

В. С. Сизиков

**УСТОЙЧИВЫЕ
МЕТОДЫ ОБРАБОТКИ
РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ**

Учебное пособие

Санкт-Петербург
Специальная Литература
1999

Сизиков В. С. Устойчивые методы обработки результатов измерений. Учебное пособие. — СПб.: «СпецЛит», 1999. — 240 с.

Изложены актуальные обратные прикладные задачи и современные (регулярные, устойчивые) численные методы обработки результатов измерений в этих задачах. Сформулированы обратные прикладные задачи компьютерной томографии, восстановления искаженных изображений, спектроскопии, диагностики плазмы, обработки сигналов, биофизики, механики, редукции измерений к идеальному измерительному устройству (радиолокатору, антенне и т. д.). Дано физическое и математическое описание задач. Изложены некоторые сведения из линейной алгебры, метод наименьших квадратов Гаусса, метод псевдообратной матрицы Мура-Пенроуза, преобразования Фурье, Хартли и Лапласа, элементы теории обобщенных функций, корректность и некорректность по Адамару, интерполяция, экстраполяция, сглаживание, аппроксимация, сплайн-функции, а также устойчивые методы регуляризации Тихонова, оптимальной фильтрации Калмана и Винера решения интегральных уравнений I рода, систем линейных алгебраических уравнений и т. д.

Для студентов, аспирантов, преподавателей и научных сотрудников, специализирующихся по современным численным (компьютерным) методам обработки результатов измерений.

Библиогр. 106 назв. Илл. 138. Табл. 3.

Р е ц е н з е н т ы: д-р техн. наук, профессор А. К. Черников
и профессор Ю. В. Юдин

Одобрено на заседании кафедры Измерительных технологий и компьютерной томографии СПбГИТМО (ТУ) 25.10.99, протокол № 3.

Рекомендовано УМО в качестве учебного пособия для подготовки специалистов по направлению (специализации) Приборостроение.

ОГЛАВЛЕНИЕ

Список сокращений	7
Введение	9
Общая схема измерений и обработки (9). Некоторые примеры прикладных задач (11). Необходимость использования устойчивых методов (12). Сравнение с другими источниками (12). Краткое содержание учебного пособия (13). Контрольные задания и вопросы (15).	
Часть I. ОБРАТНЫЕ ПРИКЛАДНЫЕ ЗАДАЧИ	16
Глава 1. Задачи компьютерной томографии	17
1.1. Рентгеновская томография	17
Недостатки обычной рентгенографии (17). Идея РТ (18). Постановка задачи (18). Закон Бера (18). Уравнение Радона (19). Историческая справка (20). Интегральное уравнение Фредгольма I рода (21). Решение уравнения методом ПФ без регуляризации и с регуляризацией (21). Численные иллюстрации (23). 5 поколений рентгеновских томографов (24). Области применения РТ (26). Снятие влияния аппаратных искажений (26). Визуализация результатов (представление слоя на дисплее) (30). Об алгоритмах и программах (31). Общая схема обработки в РТ (32). Контрольные задания и вопросы (33).	
1.2. Ядерно-магнитно-резонансная (ЯМР-) томография	33
Эффект ЯМР (33). Уравнение Лармора (34). Ансамбль протонов (36). Движение магнитных моментов изолированных протонов в постоянном и переменном магнитных полях (36). Уравнения Блоха (38). Эхо-сигнал, $\pi/2$ - и π -импульсы (41). Градиентные поля (44). Реконструкция ЯМР-изображений (45). Примеры реконструкции изображений (51). Влияние неоднородности полей на разрешающую способность томограмм (51). Математический учет технических неоднородностей полей (53). Синтез магнитного поля на оси катушки ЯМР-томографа (55). Области применения ЯМР-томографии (61). Контрольные задания и вопросы (61).	
Глава 2. Некоторые обратные задачи оптики и спектроскопии	63
2.1. Восстановление смазанных изображений	63
Постановка задачи (63). Вывод интегрального уравнения (65). Учет характеристики чувствительности пленки (67). Методы решения уравнений (68). О программах (71). Контрольные задания и вопросы (72).	
2.2. Восстановление дефокусированных изображений	72
Постановка задачи (73). Вывод основного соотношения (74). Стандартная форма уравнения (75). Решение методом двумерного ПФ (76). Применение метода регуляризации Тихонова (77). О программах (78). Контрольные задания и вопросы (79).	

2.3. Обратные задачи спектроскопии	79
Спектральный анализ и приборы (79). Области применения спектрального анализа (80). Типы спектров (80). Типы спектрального анализа (81). Экспериментальный спектр (82). Задача редукции к идеальному спектральному прибору (83). Непрерывный спектр (84). Дискретный спектр (85). Решение СЛНУ (87). Краткое изложение алгоритма интегральной аппроксимации (88). Модельный пример (88). Другие обратные задачи спектроскопии (90). О программах (90). Контрольные задания и вопросы (90).	
2.4. Обратная задача диагностики плазмы	91
Понятие плазмы (91). Характеристики плазмы (91). Спектр излучения плазмы (92). Диагностика плазмы (92). Схема пассивной диагностики плазмы (93). Методы решения уравнения (94). Случай цилиндрической симметрии (96). Случай шаровой симметрии (98). Решение уравнений (98). Контрольные задания и вопросы (99).	
Глава 3. Обобщенная формулировка обратных задач	100
3.1. Обработка сигналов	100
Постановка задачи (100). Типы обработки сигналов (101). Перечень методов вторичной обработки сигналов (103). Классические методы пеленгования (103). Методы компенсации локальных сигналов-помех (103). Методы адаптации (105). Другие методы обработки сигналов (109). Контрольные задания и вопросы (110).	
3.2. Редукция измерений к идеальному измерительному устройству	110
Введение в задачу (110). Редукционная проблема Рэлея (111). Пример 1 (редукция локальных сигналов) (112). Пример 2 (редукция протяженных сигналов) (114). Об аппаратной функции (115). Техническая реализация алгоритмов редукции (118). Контрольные задания и вопросы (118).	
Глава 4. Некоторые обратные задачи теории управления, биофизики и механики	119
4.1. Обратные задачи теории управления	119
Восстановление сигнала в динамической системе (119). Восстановление сигнала в динамической системе без обратной связи (121). Восстановление сигнала в системе, не являющейся динамической (122). Контрольные задания и вопросы (123).	
4.2. Обратные задачи биофизики	123
Восстановление искаженных изображений биологических микрообъектов (124). Томография биологических микрообъектов (125). Обратная задача речевой акустики (126). Распад клеток и радиоактивных элементов (126). Задача удельной приливной вентиляции в легких (127). Контрольные задания и вопросы (127).	
4.3. Использование преобразования Фурье в прикладных задачах механики	128
Преобразование Фурье (129). Спектральные задачи механики (131). Контрольные задания и вопросы (131).	

Часть II. УСТОЙЧИВЫЕ МЕТОДЫ РЕШЕНИЯ УРАВНЕНИЙ	132
Глава 5. Основные типы уравнений и сопутствующие им понятия	134
5.1. Основные типы уравнений	134
Интегральные уравнения (134). Системы линейных алгебраических уравнений (140). Системы линейно-нелинейных уравнений (141). Операторные уравнения (141). Контрольные задания и вопросы (142).	
5.2. Некоторые сведения из линейной алгебры	142
Система линейных алгебраических уравнений (142). Характеристическое уравнение и типы матриц (143). Нормы векторов и матриц (144). Число обусловленности (144). Умножение матриц и векторов (145). Примеры (145). Контрольные задания и вопросы (147).	
5.3. Элементы теории вероятностей	147
Основные определения (147). Программа RNDAN (150). Контрольные задания и вопросы (154).	
Глава 6. Обобщенные функции, формула Эйлера и интегральные преобразования	155
6.1. Элементы теории обобщенных функций	155
Определение обобщенной функции (155). Функция Хэвисайда (155). δ -функция Дирака (156). Контрольные задания и вопросы (157).	
6.2. Формула Эйлера	158
Определение формулы Эйлера (158). Формула Эйлера и δ -функция (158). Контрольные задания и вопросы (160).	
6.3. Интегральные преобразования	160
Непрерывное преобразование Фурье (160). Связь между прямым и обратным ПФ (164). Двухмерное ПФ (166). Дискретное ПФ (166). Об алгоритмах БПФ (170). Использование регуляризации (170). Преобразование Хартли (172). Преобразование Лапласа (175). Контрольные задания и вопросы (176).	
Глава 7. Предыстория регулярных методов	178
7.1. Корректность и некорректность по Адамару	178
Определения корректности и некорректности (178). Примеры (178).	
7.2. Классические методы решения интегральных уравнений Фредгольма I рода	180
Метод квадратур (180). Метод ПФ (182). Метод ПФ для двухмерного уравнения (185). Контрольные задания и вопросы (185).	
7.3. Метод наименьших квадратов Гаусса	186
Переопределенная СЛАУ (186). Вывод нормальной СЛАУ (186). МНК применительно к интегральному уравнению (188). Контрольные задания и вопросы (189).	
7.4. Метод псевдообратной матрицы Мура-Пенроуза	189
Недоопределенная СЛАУ (189). Нормальное решение и псевдообратная матрица (189). Пример (190). МПОМ применительно к другим уравнениям (191). Общий вывод (191). Контрольные задания и вопросы (191).	

Глава 8. Методы регуляризации, фильтрации и аппроксимации	192
8.1. Метод регуляризации Тихонова	192
Существо метода (192). Анализ метода (193). Регуляризованное интегральное уравнение (194). Способы выбора параметра регуляризации α (194). Численный алгоритм (195). Программы (196). Численные примеры (196). Метод регуляризации для уравнения типа свертки (199). Контрольные задания и вопросы (205).	
8.2. Метод оптимальной фильтрации Калмана-Бьюси	205
Одношаговый (однокрапный) фильтр Калмана (206). Сравнение одношагового фильтра Калмана с методом регуляризации Тихонова (207). Многошаговый (многократный) фильтр Калмана (208). Контрольные задания и вопросы (209).	
8.3. Метод оптимальной линейной фильтрации Винера	210
Существо метода (210). Сравнение методов Винера и Тихонова (211). Контрольные задания и вопросы (211).	
8.4. Интерполяция, экстраполяция, сглаживание и аппроксимация	212
Линейная интерполяция и экстраполяция (212). Квадратичная интерполяция и экстраполяция (213). Полином Лагранжа (214). Сплайны (215). Кубические интерполирующие сплайны (217). Линейная аппроксимация (линейное сглаживание) (221). Квадратичная аппроксимация (квадратичное сглаживание) (222). Сглаживающие (аппроксимирующие) кубические сплайны (223). Контрольные задания и вопросы (223).	
Список литературы	225
Предметный указатель	231
Contents	236

СПИСОК СОКРАЩЕНИЙ

АЦП	— аналого-цифровой преобразователь
АФ	— аппаратная функция
БПФ	— быстрое преобразование Фурье
БПХ	— быстрое преобразование Хартли
ВВ	— входные воздействия
ВЧ	— высокочастотный (импульс)
г/а	— гидроакустическая (антенна)
ДН	— диаграмма направленности
ДПФ	— дискретное преобразование Фурье
ДПХ	— дискретное преобразование Хартли
ИВК	— измерительно-вычислительный комплекс
ИзУ	— измерительное устройство
ИнУ	— интегральное уравнение
ИП	— индикаторный процесс
ИПС	— измерительная подсистема
КТ	— компьютерная томография
КТХ	— космический телескоп «Хаббл»
ЛИУ	— линейное интегральное уравнение
ЛТ	— ложная тревога
МНК	— метод наименьших квадратов (Гаусса)
МО	— математическое ожидание
МПОМ	— метод псев дообраной ма трицы (Мура-Пенроуза)
НПФ	— непрерывное преобразование Фурье
НПХ	— непрерывное преобразование Хартли
ОБПФ	— обратное быстрое преобразование Фурье
ОБПХ	— обратное быстрое преобразование Хартли
ОДПФ	— обратное дискретное преобразование Фурье
ОДПХ	— обратное дискретное преобразование Хартли
ОНПФ	— обратное непрерывное преобразование Фурье
ОНПХ	— обратное непрерывное преобразование Хартли

ОПФ	— обратное преобразование Фурье
ОПХ	— обратное преобразование Хартли
ПЛ	— преобразование Лапласа
ПФ	— преобразование Фурье
ПХ	— преобразование Хартли
РТ	— рентгеновская томография
СВУ	— специализированное вычислительное устройство
СКО	— среднеквадратическое отклонение
СЛАУ	— система линейных алгебраических уравнений
СЛНУ	— система линейно-нелинейных уравнений
СНУ	— система нелинейных уравнений
СП	— стандартная программа
СПМ	— спектральная плотность мощности
СФ	— сканирующая функция
С Ч	— спектральная чувствительность
Т л	— единица напряженности (индукции) магнитного поля
токамак	— тороидальная камера с магнитными катушками
УО	— устройство обработки
ХН	— характеристика направленности
Х Ч	— характеристика чувствительности (пленки)
ЦАП	— цифро-аналоговый преобразователь
ЭВМ	— электронно-вычислительная машина
ЯМР	— ядерный магнитный резонанс

ВВЕДЕНИЕ

Данное учебное пособие посвящено современным, устойчивым методам решения обратных прикладных задач. Это — актуальные задачи компьютерной (рентгеновской и ЯМР-) томографии, оптики (восстановление смазанных и дефокусированных изображений), спектроскопии (восстановление спектров), диагностики плазмы, обработки сигналов (радиолокационных, гидроакустических), биофизики (восстановление изображений биомикрообъектов, исследование речевого тракта, распад клеток, вентиляция в легких), механики (расчет конструкции глушителя автомобиля по спектру его шумов, определение свойств газовой, жидкой и твердой среды по спектру проходящего звука, борьба с шумами в цехах) и др.

Общая схема измерений и обработки. Суть обратных прикладных задач состоит в том, что по некоторой измеренной с погрешностями функции f (например, эхо-сигналу в ЯМР-томографии, искаженному снимку в оптике или сканирующей функции в гидроакустике), а также по аппаратной функции A (например, характеристике направленности — ХН в гидроакустике) можно определить исходную (входную) функцию y (например, плотность протонов в некотором сечении в ЯМР-томографии, неискаженный снимок в оптике, поле на входе антенны в гидроакустике) путем решения уравнения

$$Ay = f \quad (1)$$

относительно y , которое может быть интегральным уравнением (ИнУ), дифференциальным уравнением, системой линейных алгебраических уравнений (СЛАУ), системой линейно-нелинейных уравнений (СЛНУ) и т. д.

Это можно отобразить в виде двух следующих *схем* (см. рис. 1 и 2). Согласно *технической схеме* (рис. 1), на вход измерительного устройства (ИЗУ) поступает входной процесс: полезный сигнал

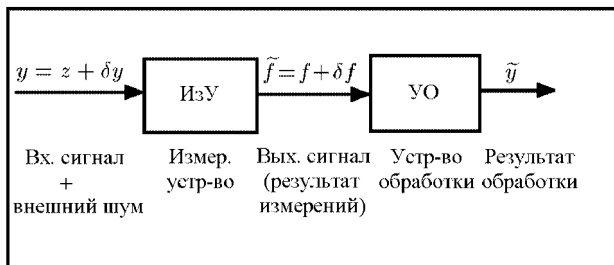


Рис. 1. Техническая схема

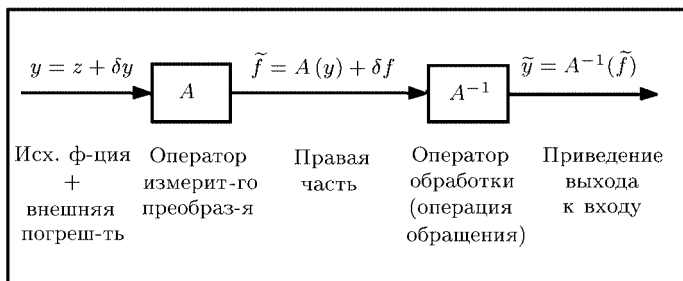


Рис. 2. Математическая схема

$z + \text{шум (помеха)}$ δy из внешней среды. Пройдя через ИзУ, характеризующееся аппаратной функцией (АФ), например, характеристикой направленности (ХН) антенны, сигнал + шум преобразуются в выходной сигнал (результат измерений), например, сканирующую функцию $\tilde{f} = f + \delta f$, где δf — аппаратная погрешность измерений. Далее следует устройство обработки (УО), цель которого — получить \tilde{y} — результат обработки, по возможности близкий к процессу $y = z + \delta y$ или даже к сигналу z . Следует отметить, что помеха δf — это мешающий фактор, с которым нужно бороться, а δy в зависимости от некоторого критерия может относиться к помехе, а может относиться и к одной из компонент входного сигнала (это характерно для адаптивных методов обработки — см. п. 3.1).

Помимо технической схемы измерений и обработки рассмотрим также соответствующую ей *математическую схему*, которая с учетом введенных обозначений имеет вид, показанный на рис. 2. В математической схеме оператор измерительного преобразования A (аналогичный аппаратной функции) преобразует входной сигнал + внешний шум $y = z + \delta y$ в так называемую правую часть (выходной сигнал) $\tilde{f} = A(y) + \delta f$, где δf — погрешность правой части. Далее с помощью обратного оператора A^{-1} вычисляется приближенное решение $\tilde{y} = A^{-1}(\tilde{f})$ и цель математических мето-

дов и алгоритмов — построить такой обратный оператор A^{-1} , чтобы он давал хорошее приближение к процессу $y = z + \delta y$ или даже к сигналу z и при этом был устойчив по отношению к погрешностям δf .

Если технически A — это аппаратная функция (АФ) измерительного устройства, например, характеристика направленности (ХН) антенны, глаза, уха или спектральная чувствительность (СЧ) спектрометра и т. д., то математически A — это интегральный, дифференциальный, алгебраический, нелинейный и т. д. оператор.

Решение уравнения (1) позволяет, в принципе, выполнить редукцию (приведение) результатов измерений к идеальному измерительному устройству (например, к фотоаппарату без смаза и ошибки фокуса или к антенне с бесконечно узкой ХН), причем выполнить это математически (с использованием компьютера), что позволит использовать даже несовершенное, недорогое измерительное устройство. Сопряжение измерительного устройства с вычислительным, обеспечивающим решение задачи редукции, равнозначно созданию нового измерительного устройства с более высокой разрешающей способностью (по углу, времени, частоте и т. д.). Еще более важным является случай, когда в силу специфики задачи даже совершенное измерительное устройство (например, томограф) не позволяет непосредственно (без математической обработки) получить искомую функцию y (см. гл. 1).

Поэтому следует считать актуальным использование математической (и компьютерной) обработки результатов измерений для определения входной информации на измерительном устройстве (например, плотности вещества в томографии) или для редукции к более совершенному измерительному устройству (прекрасный пример этого — математико-компьютерная обработка дефокусированных цветных снимков, сделанных американским космическим телескопом «Хаббл» [82]).

Некоторые примеры прикладных задач. Приведем некоторые характерные примеры обратных прикладных задач. Обработка искаженных изображений (п. 2.1, 2.2), восстановление спектра (п. 2.3), обработка сигналов (п. 3.1), редукция измерений (п. 3.2) и др. — это задачи, в которых в принципе можно получить выходной сигнал f , близкий к входному сигналу y , без математической обработки, если уровень шумов δf очень мал и аппаратная функция A очень узкая, например, снимок без смаза и в фокусе, измерение спектра спектрометром с очень узкой СЧ, прием сигналов антенной с очень узкой ХН и т. д. Однако для этого нужно иметь совершенное и дорогое измерительное устройство, например, антенну очень больших размеров, что часто невыполнимо. Математическая же обработка позволяет скомпенсировать этот недостаток. В еще большей степени необходима математическая

обработка результатов измерений в таких задачах, как рентгеновская и ЯМР-томография (гл. 1), диагностика плазмы (п. 2.4), задачи механики на основе преобразования Фурье (п. 4.3) и др., так как, например, в РТ измеряемая функция — интенсивность на детекторах $I(l, \theta)$ не имеет практически ничего общего с искомой функцией — плотностью вещества в сечении $c(x, y)$ и поэтому, сколь бы совершенен ни был томограф, без математической обработки в принципе невозможно определить плотность $c(x, y)$ по измеренной интенсивности $I(l, \theta)$ (см. п. 1.1).

Необходимость использования устойчивых методов. Задача решения уравнения типа (1) является, как правило, некорректной (сильно неустойчивой) и для эффективного ее решения нужно использовать современные, устойчивые методы. Это — методы регуляризации Тихонова, итеративной, статистической, локальной, дискриптивной регуляризации, субоптимальной фильтрации, решения на компакте и др. — первая группа методов, развитых советскими (российскими) учеными, а также методы оптимальной фильтрации Калмана-Бьюси и Винера, методы управляемой линейной фильтрации (Бэйкуса-Гильберта) и др. — вторая группа методов, развитых западными учеными. Хотя методы второй группы являются в принципе более точными, но методы первой группы (в первую очередь, метод регуляризации Тихонова) требуют гораздо меньше дополнительной информации о решении и поэтому находят более широкое применение при решении вышеперечисленных (а также других) обратных прикладных задач.

Поэтому нужно отметить большую необходимость систематического изложения (в частности, в виде учебного пособия) как актуальных прикладных задач, так и эффективных математических (и компьютерных) методов их решения.

Сравнение с другими источниками. Отметим следующие учебные пособия, а также учебники и монографии, посвященные данной тематике (в хронологическом порядке): [42] *Красильников В. А.* Звуковые и ультразвуковые волны в воздухе, воде и твердых телах (изложены прикладные задачи механики); [15] *Василенко Г. И.* Теория восстановления сигналов: О редукции к идеальному прибору в физике и технике (обработка искаженных изображений, метод регуляризации Тихонова, метод оптимальной фильтрации Винера); [14] *Вапник В. Н.* (ред.) Алгоритмы и программы восстановления зависимостей (распознавание образов, восстановление регрессии, интерпретация экспериментов, метод структурной минимизации риска, программы); [19] *Верлань А. Ф., Сизиков В. С.* Интегральные уравнения: методы, алгоритмы, программы (прикладные задачи, методы Тихонова, Калмана-Бьюси, Винера, программы); [69] *Тихонов А. Н. и др.* Математические задачи компьютерной томографии (рентгеновская, ЯМР- и ультразвуковая томография, метод регуляризации Тихонова, метод локальной регуляризации); [4] *Бакушинский А. Б., Гончарский А. В.*

Некорректные задачи. Численные методы и приложения (обработка искаженных изображений, метод регуляризации Тихонова); [49] *Наттерер Ф.* Математические аспекты компьютерной томографии (рентгеновская томография); [71] *Тихонов А. Н. и др.* Численные методы решения некорректных задач (метод регуляризации Тихонова, решение на компакте, программы); [74] *Уэбб С.* (ред.) Физика визуализации изображений в медицине (рентгеновская и ЯМР-томография); [51] *Потеев М. И., Сизиков В. С.* Повышение разрешающей способности измерительных устройств путем компьютерной обработки результатов измерений (прикладные задачи, метод регуляризации Тихонова); [93] *Cho Z. H. et al.* Foundations of medical imaging (ЯМР-томография); [16] *Васильев В. Н., Гуров И. П.* Компьютерная обработка сигналов в приложении к интерферометрическим системам (обработка сигналов, преобразования Фурье, Гильберта и Хартли, фильтрация Калмана и Стратоновича). Отметим также источники [2, 3, 8, 10, 13, 17, 18, 23, 27, 30, 32, 36, 45, 47, 48, 52, 54–57, 67, 70, 72, 79, 81, 84, 87] и др.

Анализ приведенных источников показывает, что несмотря на их обилие, практически нет источника, в котором современные прикладные задачи, методы и программы были бы изложены полно и в стиле учебного пособия. Предлагаемое учебное пособие — это попытка создания такого источника.

В заключение отметим, что в данном учебном пособии не рассматриваются такие вопросы теории измерений, как метрология (а именно, средства и методы измерений), анализ и планирование экспериментов, математическая статистика. Эти, уже ставшие классическими, вопросы достаточно полно изложены в таких книгах, как [39] *Клепиков Н. П., Соколов С. Н.* Анализ и планирование экспериментов методом максимума правдоподобия; [86] *Яноши Л.* Теория и практика обработки результатов измерений; [12] *Бурдун Г. Д., Марков Б. Н.* Основы метрологии и др. Предлагаемое пособие, по существу, продолжает эту тематику, рассматривая (еще достаточно новые) неустойчивые обратные прикладные задачи и регулярные методы их решения.

Краткое содержание учебного пособия. Учебное пособие состоит из двух частей. Часть I (содержащая главы 1–4) посвящена изложению обратных прикладных задач, а часть II (содержащая главы 5–8) — устойчивым методам решения уравнений, описывающих эти задачи.

В главе 1 изложены задачи компьютерной томографии, а именно, рентгеновской томографии и ЯМР-томографии.

В главе 2 рассмотрены некоторые обратные задачи оптики и спектроскопии: восстановление смазанных и дефокусированных изображений, обратные задачи спектроскопии и обратная задача диагностики плазмы.

В главе 3 дана обобщенная формулировка обратных задач в терминах задачи обработки сигналов и задачи редукции (приведения) измерений к идеальному измерительному устройству (редукционная проблема Рэлея).

В главе 4 сформулированы некоторые обратные задачи теории управления, биофизики и механики.

В главах 1–4 даны постановки задач, выведены основные уравнения, описывающие задачи, и даны численные иллюстрации.

В главе 5 перечислены основные типы уравнений (интегральные уравнения, системы линейных алгебраических уравнений и т. д.), описывающие прикладные задачи, и даны основные сведения из линейной алгебры и теории вероятностей.

В главе 6 даны элементы теории обобщенных функций (функция Хэвисайда, δ -функция Дирака), приведена формула Эйлера и рассмотрены интегральные преобразования (Фурье, Хартли и Лапласа) непрерывные, дискретные и быстрые, одно- и двухмерные.

В главе 7 изложена предыстория регулярных методов: корректность и некорректность по Адамару, классические методы решения интегральных уравнений, МНК Гаусса, МПОМ Мура-Пенроуза.

В главе 8 рассмотрены метод регуляризации Тихонова, методы оптимальной фильтрации Калмана-Бьюси и Винера, а также интерполяция, экстраполяция, сглаживание и аппроксимация (полином Лагранжа, сплайны и т. д.).

Многие методы доведены до рабочих алгоритмов и даны ссылки на стандартные программы, реализующие эти методы и алгоритмы.

Текст данного пособия набран в редакторе \LaTeX . Что же касается иллюстраций, то рисунки 1.1, 1.4, 1.5, 1.30, 2.2, 2.6, 2.8, 2.9, 2.18, 2.19 и 3.14 получены путем ксерокопирования из первоисточников; для получения рисунков 1.36, 1.37, 3.15, 6.14 — 6.16, 7.1, 8.1 — 8.6 выполнены расчеты на языке Fortran MS 5 (и Fortran 90), построены графики в графопостроителе Grapher и выполнена их доводка в графическо-текстовом редакторе PaintBrush; рисунки 2.3 и 2.7 рассчитаны и построены в Visual C++ с доводкой в PaintBrush'e; рисунки 2.24 и 2.25 построены в редакторе MathCad 6 с доводкой в PaintBrush'e; рисунки 1.25, 1.26, 5.3, 6.4–6.8 построены в редакторе Excell с доводкой в PaintBrush'e; остальные рисунки построены непосредственно в PaintBrush'e. Окончательная доводка рисунков выполнена в CorelDraw и PhotoShop'e. В соответствии с этим можно давать студентам задания, например: рассчитать рис. 6.16 на Фортране (или другом языке) согласно формул (6.63)–(6.65), построить графики в Grapher'e и выполнить окончательную доводку графиков в PaintBrush'e, CorelDraw или PhotoShop'e.

Данное пособие есть результат многолетней научной и педагогической работы автора. Автор искренне благодарен проф. А. Б. Бакушинскому, проф. И. В. Бойкову, prof. H. Brunner, проф. А. Ф. Верланю, проф. Ю. Е. Воскобойникову, проф. П. А. Галайдину, проф. В. А. Желудеву, проф. В. Б. Жукову, доц. А. И. Замятину, доц. В. Ф. Звягину, проф. В. А. Иванову, проф. Г. И. Мельникову, проф. М. И. Потееву, проф. А. И. Седельникову за плодотворное сотрудничество и содержательные беседы. Автор благодарит рецензентов проф. А. К. Черникова и проф. Ю. В. Юдина за труд по просмотру рукописи и ценные замечания. Автор благодарен своей жене Елене за долготерпение в период работы над рукописью. Наконец, автор благодарит студентов Д. Апполонова, И. Белова, А. Козаченко и др. за помощь при наборе пособия.

Все пожелания и замечания по поводу данного учебного пособия автор просит направлять по адресу: 197101, С-Петербург, ул. Саблинская, 14, СПб ГИТМО (ТУ), проф. В. С. Сизикову или по эл. адресу: sizikov@beam.ifmo.ru или по факсу: 812-2335952.

Контрольные задания и вопросы

1. Привести примеры обратных прикладных задач.
2. Изобразить техническую и математическую схемы измерений.
3. Перечислить устойчивые методы обработки измерений.
4. Объяснить суть обратных прикладных задач на примере уравнения $Ay = f$. Что означают A , y и f математически и технически?

Часть I

ОБРАТНЫЕ ПРИКЛАДНЫЕ ЗАДАЧИ

В данной части рассматривается ряд актуальных обратных прикладных задач томографии, оптики, спектроскопии, гидроакустики, радиолокации, биофизики, механики, медицины и т. д. Под *обратной задачей* будем подразумевать задачу, в которой по измеренному выходному сигналу (процессу) f и известной аппаратной функции A нужно определить искомый входной сигнал (процесс) y путем решения операторного уравнения

$$Ay = f$$

Пример обратной задачи — обратная задача спектроскопии (задача восстановления непрерывного спектра), которая формулируется следующим образом. Спектрометром со спектральной чувствительностью (аппаратной функцией) $K(\nu)$, где ν — частота, измеряется интенсивность спектра $u(\nu)$ в функции частоты; тогда истинный (незаглаженный аппаратной функцией и неискаженный помехами) спектр $z(\nu)$ может быть найден путем решения интегрального уравнения Фредгольма I рода:

$$\int_a^b K(\nu - \nu') z(\nu') d\nu' = u(\nu), \quad c \leq \nu \leq d,$$

где $[a, b]$ — область поиска решения $z(\nu)$, $[c, d]$ — область измерения $u(\nu)$.

Глава 1

ЗАДАЧИ К ОМПУТЕРНОЙ ТОМОГРАФИИ

В этой главе рассматриваются задачи рентгеновской томографии (РТ) и ЯМР-томографии. Существуют также ультразвуковая, протон-эмиссионная и др. томографии.

Слово «томография» происходит от греческих слов *τομή* — сечение и *γραφω* — пишу, т.е. «пишу по сечениям». Суть всех типов томографии одинакова: по суммарной информации, полученной от некоторого сечения (слоя) вещества, нужно определить локальную информацию, а именно, плотность вещества в сечении $c(x, y)$, 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 $I(l, \theta)$, в ЯМР-томографии это эхо-сигналы $s(t, g_x)$ и т.д.) и математические описания различны (в РТ это интегральное уравнение Радона или Фредгольма, в ЯМР-томографии это двухмерное ПФ и т.д.), хотя есть тенденция описать все виды томографии единым так называемым *основным уравнением компьютерной томографии* [69, с. 19, 26, 32, 38]. Однако конечная цель одинакова — получение $c(x, y)$, поэтому, например, рентгеновские и ЯМР-томограммы (отображение на прозрачных пленках плотности $c(x, y)$) выглядят практически одинаково, хотя требуемая для их получения физика, техника и математика различны.

1.1. Рентгеновская томография (РТ)

Для определенности будем рассматривать РТ [3, 49, 68, 69], [74, т. 1], [79] применительно к исследованию руки, грудной клетки или мозга, хотя область применения РТ шире (см. дальше).

Недостатки обычной рентгенографии. Обычная рентгенография (получение одного рентгеновского снимка, например, грудной клетки) имеет следующие недостатки:

1) Различение по плотности соседних участков возможно лишь в случае, когда их плотности отличаются на $\gtrsim 2\%$ (если два соседних участка отличаются по плотности, например, на 1%, то они выглядят как имеющие одинаковую плотность, что ухудшает качество медицинского анализа).

2) Неразличие пространственных структур (если один участок заслоняет другой, то последний на рентгеновском снимке не виден и т.д.).

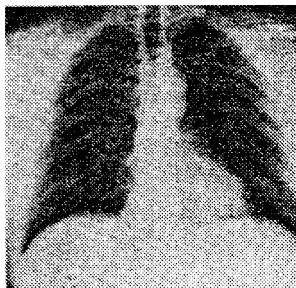


Рис. 1.1

На рис.1.1 представлена типичная рентгенограмма грудной клетки [74, т.1, с.139].

Идея РТ. Недостатки обычной рентгенографии породили идею получения не одного, а ряда снимков, выполненных под разными ракурсами, и определения по ним путем математической обработки плотностей в ряде сечений. Идею можно отобразить в виде рис.1.2, где в центре рис.1.2а — исследуемый мозг, а вокруг него — набор снимков, из которых берется лишь узкая полоска, соответствующая некоторому узкому слою мозга. После обработки интенсивностей в полосках снимков должна быть найдена плотность $c(x, y)$ в слое (сечении) мозга — см. рис.1.2б. Затем аналогичную обработку можно выполнить для других слоев (сечений) мозга и в результате получить объемную плотность $c(x, y, z)$.

Постановка задачи. Более конструктивно идея РТ выглядит следующим образом (см. рис.1.3, где представлена сканирующая схема рентгеновского томографа 2-го поколения).

На раме помещен ряд рентгеновских трубок (источников), излучающих остронаправленные параллельные лучи (пучки). Лучи проходят через сечение объекта (например, мозга) и их интенсивности фиксируются соответствующими детекторами (приемниками). Затем вся рама с источниками и приемниками поворачивается на новый угол θ и эксперимент (излучение и прием) повторяется и т. д.

Закон Бера. Согласно закону Бера [74, т.1, с.138], интенсивность рентгеновского луча, принятого на детекторе, равна

$$I(l, \theta) = I_0(l, \theta) e^{-\int_{\mathcal{L}(l, \theta)} c(x, y) ds}, \quad (1.1)$$

где l — координата детектора, θ — угол поворота рамы, $I_0(l, \theta)$ — интенсивность соответствующей излучающей трубки (обычно $I_0(l, \theta) = I_0 = \text{const}$), $c(x, y)$ — плотность вещества на луче-прямой $\mathcal{L}(l, \theta)$, уравнение которой

$$x \cos \theta + y \sin \theta = l. \quad (1.2)$$

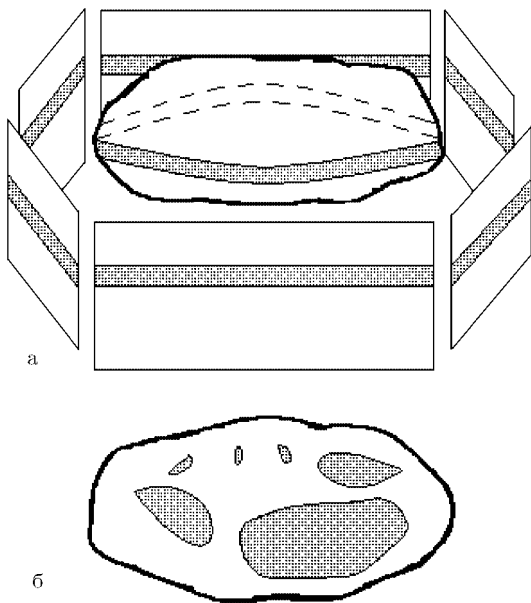


Рис. 1.2

Интегрирование в (1.1) ведется по лучу $\mathcal{L}(l, \theta)$, в результате чем больше будет $\int_{\mathcal{L}(l, \theta)} c(x, y) ds$ — масса вещества на луче, тем меньше будет принятая интенсивность $I(l, \theta)$.

Уравнение Радона относительно плотности $c(x, y)$. Запишем (1.1) иначе:

$$I(l, \theta)/I_0(l, \theta) = e^{-q(l, \theta)}, \quad (1.3)$$

где

$$q(l, \theta) = \int_{\mathcal{L}(l, \theta)} c(x, y) ds. \quad (1.4)$$

Определение. Выражение (1.4), где $\mathcal{L}(l, \theta)$ есть некоторый луч зрения, $c(x, y)$ — плотность вещества на этом луче, а s направлено вдоль этого луча, называется *преобразованием Радона* (ПР) (1917 г.).

Логарифмируя (1.3), получим:

$$q(l, \theta) = -\ln[I(l, \theta)/I_0(l, \theta)]. \quad (1.5)$$

Определение. Функция $q(l, \theta)$ называется *поглощением* и может принимать значения от 0 (среда прозрачна) до ∞ (среда абсолютно непрозрачна).

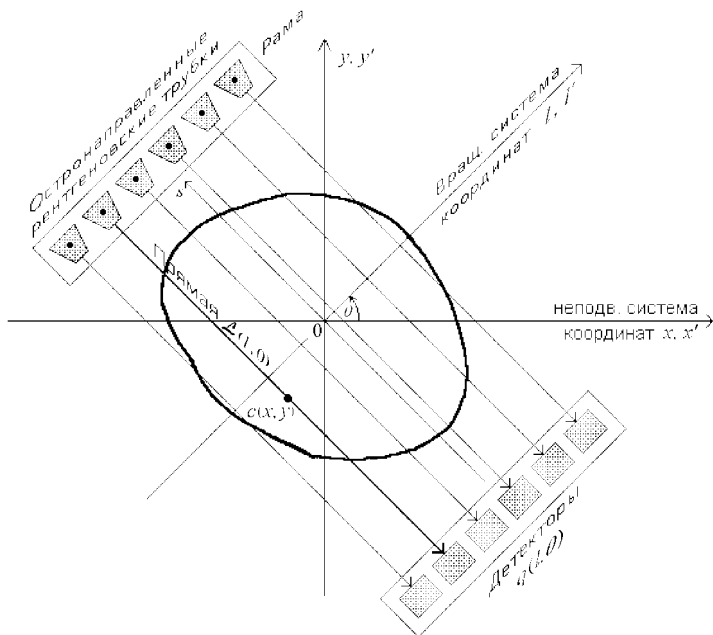


Рис. 1.3

Определение. Отношение $I(l, \theta)/I_0(l, \theta)$ называется *прозрачностью* и может принимать значения от 0 (среда совершенно непрозрачна) до 1 (среда прозрачна).

Запишем (1.4) иначе:

$$\int_{L(l, \theta)} c(x, y) ds = q(l, \theta). \quad (1.6)$$

Выражение (1.6) есть *формула*, или *уравнение Радона* (1917 г.). Здесь двухмерная функция $q(l, \theta)$ определяется согласно (1.5), где $I(l, \theta)$ получается в результате измерений. А двухмерная функция $c(x, y)$ является искомой. Поэтому (1.6) можно рассматривать как *интегральное уравнение* относительно $c(x, y)$ по измеренной правой части $q(l, \theta)$. Решение уравнения (1.6) позволяет в принципе найти плотность вещества $c(x, y)$ в некотором сечении (слое) объекта (например, мозга) по измеренной $I(l, \theta)$, а значит и $q(l, \theta)$. Такая задача называется *реконструкцией рентгеновского изображения*.

Историческая справка. Впервые задачу реконструкции изображения рассмотрел Радон в 1917 г. Он же получил решение уравнения (1.6), которое в современной литературе записывается

в виде [69, с. 40]:

$$c(x, y) = -\frac{1}{2\pi^2} \int_0^\pi d\theta \int_{-\infty}^{\infty} \frac{\partial q(l, \theta)}{\partial l} \frac{dl}{l - (x \cos \theta + y \sin \theta)}. \quad (1.7)$$

Однако анализ показал, что решение уравнения (1.6) согласно (1.7) является сильно *неустойчивым* (теперь мы говорим, что задача его решения *некорректна*) и в результате данная задача на много лет была отложена в сторону, но затем была вновь перетворена. Это сделали советские ученые Тетельбаум, Коренблюм и Тютин в 1956–58 гг., разработавшие первую систему реконструкции рентгеновских медицинских изображений. И лишь в 1973 г. на Западе Г. Хаунсфилд разработал первую коммерческую систему (сканер головного мозга английской фирмы ЕМІ), в результате в 1979 г. Г. Хаунсфилд (а также А. Кормак за лабораторные эксперименты по РТ в 1963 г.) и в 1982 г. А. Клуз за применение в био химии получили Нобелевские премии. Отметим также, что в 1980 г. был создан первый советский рентгеновский компьютерный томограф СРТ-1000 2-го поколения, в котором в качестве метода решения уравнения (1.6) заложен высокоэффективный метод локальной регуляризации Арсенина (более точный, чем метод регуляризации Тихонова) — см. *далее*.

Интегральное уравнение Фредгольма I рода относительно $c(x, y)$. Мы не будем подробно останавливаться на «западных» методах решения уравнения (1.6), а рассмотрим приведение уравнения (1.6) к стандартной форме и использование устойчивых методов его решения. Тихонов А. Н. и др. (1982 г.) привели уравнение (1.6) к стандартной форме — *двухмерному интегральному уравнению Фредгольма I рода типа свертки*:

$$\iint_{-\infty}^{\infty} \frac{c(x', y') dx' dy'}{\sqrt{(x - x')^2 + (y - y')^2}} = S(x, y), \quad (1.8)$$

где

$$S(x, y) = \frac{1}{\pi} \int_0^\pi q(x \cos \theta + y \sin \theta, \theta) d\theta. \quad (1.9)$$

Уравнение (1.8) имеет стандартную форму. В нем ядро равно $K(x - x', y - y') = 1/\sqrt{(x - x')^2 + (y - y')^2}$, искомая функция есть $c(x, y)$, а правая часть $S(x, y)$ может быть вычислена численно по известной $q(l, \theta)$ согласно (1.9).

Решение уравнения (1.8) методом ПФ без регуляризации и с регуляризацией. В п. 7.2 изложен метод ПФ решения двух-

мерного интегрального уравнения Фредгольма I рода типа свертки. Согласно нему, решение уравнения (1.8) имеет вид:

$$c(x, y) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \hat{c}(\omega_1, \omega_2) e^{-i(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2, \quad (1.10)$$

где ПФ решения (его спектр, или Фурье-образ) равно

$$\hat{c}(\omega_1, \omega_2) = \frac{\hat{S}(\omega_1, \omega_2)}{\hat{K}(\omega_1, \omega_2)} = \frac{\sqrt{\omega_1^2 + \omega_2^2}}{2\pi} \hat{S}(\omega_1, \omega_2), \quad (1.11)$$

где $\hat{S}(\omega_1, \omega_2)$ и $\hat{K}(\omega_1, \omega_2)$ — ПФ правой части и ядра, равные

$$\hat{S}(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} S(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy, \quad (1.12)$$

$$\hat{K}(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} K(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy = \frac{2\pi}{\sqrt{\omega_1^2 + \omega_2^2}}. \quad (1.13)$$

Однако решение (1.9) неустойчиво. Действительно, из-за погрешностей измерений функции $I(l, \theta)$ будут иметь погрешности также функции $q(l, \theta)$ и $S(x, y)$. При этом погрешности обычно имеют компоненту белого шума, т.е. постоянную (независящую от частот ω_1 и ω_2) компоненту. В результате $\hat{S}(\omega_1, \omega_2) \rightarrow \text{const}$ при $\omega_1, \omega_2 \rightarrow \infty$, а $\hat{c}(\omega_1, \omega_2) \rightarrow \infty$ при $\omega_1, \omega_2 \rightarrow \infty$ (см. (1.10)) и интеграл в (1.10) расходится. На практике интеграл (1.10) (т.е. ОНПФ) вычисляется через конечную сумму (ДПФ и БПФ, см. п. 6.3) до конечных максимальных значений частот ω_1 и ω_2 (т.е. имеет место усечение по частотам) и эффект неустойчивости снижается, хотя и остается.

В методе регуляризации Тихонова (см. п. 8.1) для $\hat{c}(\omega_1, \omega_2)$ вместо (1.11) используется формула [69, с. 68]:

$$\hat{c}_\alpha(\omega_1, \omega_2) = \frac{1}{2\pi} \frac{\omega}{1 + \alpha\omega^2(\omega^4 + 1)} \hat{S}(\omega_1, \omega_2), \quad \omega = \sqrt{\omega_1^2 + \omega_2^2}, \quad (1.14)$$

где $\alpha > 0$ — параметр регуляризации, способы выбора которого изложены далее (см. п. 8.1).

Для выражения (1.14) характерно следующее. Во-первых, $\hat{c}_\alpha(\omega_1, \omega_2) \rightarrow 0$ при $\omega \rightarrow \infty$ и интеграл (1.10) сходится. Во-вторых, подавление высоких частот ω в (1.14) происходит более аккуратно, чем в методах, использующих усечение по частоте. 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 и т.д.), поэтому

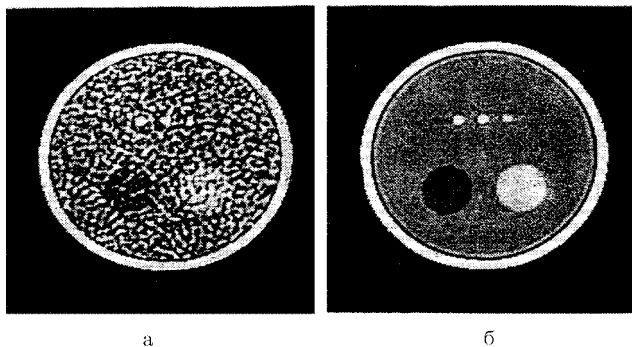


Рис. 1.4

это подавление должно быть умеренным. Обоим этим критериям (в пропорции, регулируемой параметром α) удовлетворяет метод регуляризации Тихонова.

Численные иллюстрации. Приведем примеры реконструкции рентгеновских изображений, выполненные согласно различных методик. На рис. 1.4 приведены результаты решения модельного примера [68] согласно метода Шеппа-Логана (метода «интуитивной регуляризации», использующего ПФ и прием усечения частот) — см. рис. 1.4а и согласно метода локальной регуляризации Арсенина [68, 69, 91, 104] (еще более точного, чем метод регуляризации Тихонова) — см. рис. 1.4б. Видим, каким неустойчивым, зашумленным является решение методом Шеппа-Логана и каким высокоточным — решение методом локальной регуляризации.

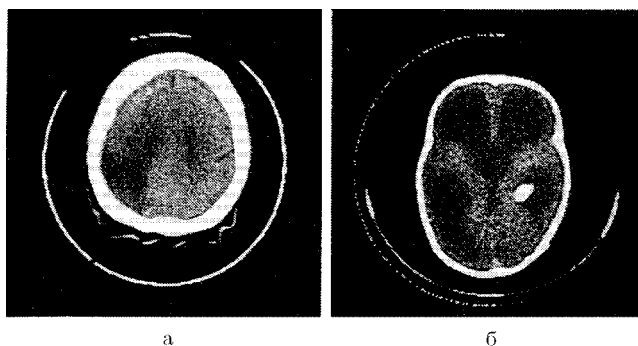


Рис. 1.5

На рис. 1.5 приведены томограммы [68] одного и того же сечения головного мозга человека, полученные по методике английской фирмы EMI (томограф СТ-1010) — см. рис. 1.5а и по методу

локальной регуляризации (томограф СРТ-1000) — см. рис.1.5б. Для этого сечения характерно то, что плотности соседних участков отличаются лишь на 0.5 % и методика фирмы ЕМІ (рис. 1.5а) их не разрешает, в то время как методом локальной регуляризации (рис. 1.5б) они разрешаются.

Примеры приведены также в [49, 69], [74, т. 1, с. 141] и др.

5 поколений рентгеновских томографов [74, т. 1, с. 142–146].

В томографах *1-го поколения* (рис. 1.6) одна остронаправленная рентгеновская трубка и один детектор передвигаются синхронно вдоль рамы для получения функции $I(l, \theta_1)$ (1-е сканирование). Затем рама поворачивается на угол θ_2 и аналогично измеряется $I(l, \theta_2)$ и т. д. Общее время измерений $T \approx 4$ мин.

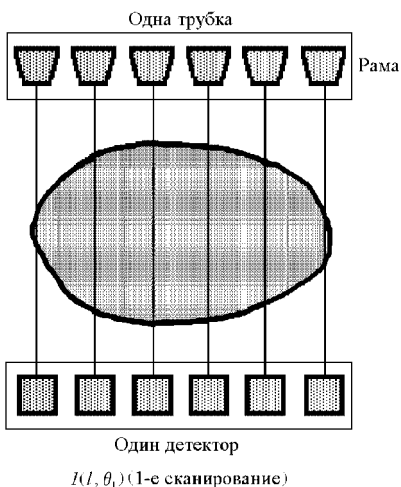


Рис. 1.6. РТ 1-го поколения

В томографах *2-го поколения* — то же, только имеется N трубок и N детекторов, работающих одновременно. Общее время измерений $T \approx 20$ сек. Первый советский рентгеновский томограф СРТ-1000 (1980 г.) [68] — это томограф 2-го поколения.

Заметим, что в томографах 1-го и 2-го поколений используется параллельное сканирование, а начиная с томографов 3-го поколения используется веерное сканирование.

В томографах *3-го поколения* (см. рис.1.7) одна трубка излучает веерный пучок, принимаемый детекторами (их количество ~ 100), расположенными по дуге — 1-е сканирование. Затем трубка и детектор поворачиваются и осуществляется 2-е сканирование и т. д. Время $T = 4$ –5 сек.

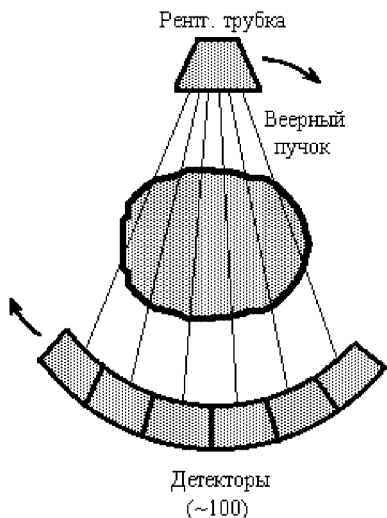


Рис. 1.7. РТ 3-го поколения

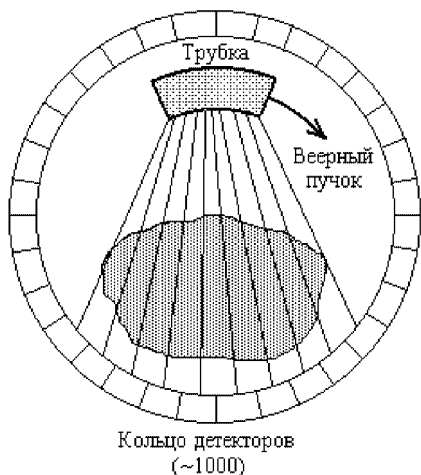


Рис. 1.8. РТ 4-го поколения

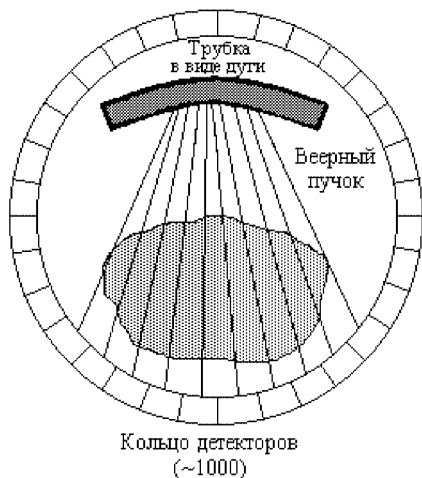


Рис. 1.9. РТ 5-го поколения

В томографах *4-го поколения* (см. рис.1.8), например, в томографе Pfizer 0450 (США) имеется сплошное неподвижное кольцо детекторов (~ 1000) и поворачивающаяся рентгеновская трубка. Время $T \sim 0.1$ сек.

В томографах *5-го поколения* (см. рис.1.9) имеется также сплошное неподвижное кольцо детекторов (~ 1000) и неподвижная

рентгеновская трубка в виде дуги окружности в $\approx 210^\circ$. Осуществляется сканирование электронными пучками по поверхности дуги. Время T — несколько мсек.

Отметим еще, что шкала современных компьютерных томографов различает ~ 1000 градаций плотностей выше и ниже плотности воды (от плотности воздуха до плотности кости), если конечно математический метод способен различать плотности соседних областей до $\approx 0.5\%$.

Как мы видим, главное в томографах — не аппаратура (она достаточно совершенна в зарубежных томографах), а математическое и программное обеспечение. В зарубежных томографах используются преобразования Радона или Фурье и прием усечения или сглаживания — «интуитивная регуляризация», суть которой составляет тайну фирмы, но которая зачастую недостаточно эффективна. В отечественных же томографах используется метод регуляризации — в результате точность восстановления $c(x, y)$ выше, чем в зарубежных томографах (см. рис. 1.4, 1.5).

Области применения РТ. Основная область применения РТ — медицина [74], а именно, исследование мозга, брюшной полости, грудной клетки, рук и т. д. с целью выявления травм, предраковых опухолей на ранней стадии их развития (размером ~ 1 мм), исследования крови в кровеносных сосудах, тонкой структуры мягких тканей, деталей анатомического строения сердца, получения динамической картины работы сердца, печени и кровотока с кинематографической регистрацией (в РТ 5-го поколения) и т. д. Другие области применения РТ: исследование трехмерной внутренней структуры технических деталей сложной формы, биологических объектов, плазмы, алмазных выработок, древесины без ее распиловки, контроль узлов реактивных двигателей, просмотр содержимого багажа без его вскрытия на таможне, просвечивание мантии Земли (геофизика), оценка распределения плотности электронов на сферических поверхностях вокруг Солнца с использованием естественного томографа Солнце–Земля (астрофизика) и т. д.

Чтобы повысить качество томограмм, следует дополнительно выполнить следующие операции (редукции): снятие влияния аппаратурных искажений и визуализация результатов $c(x, y)$ на дисплее [3].

Снятие влияния аппаратурных искажений. Аппаратурные искажения могут быть следующие.

а) Трубка излучает не бесконечно узкий луч, а узконаправленный пучок (см. рис. 1.10), вследствие чего детектор воспринимает излучение не только соответствующей ему трубки, но и других трубок (правда, более слабо).

б) По пути рентгеновский луч испытывает рассеяние (см. рис. 1.11) и попадает (в ослабленном виде) в «чужой» детектор.

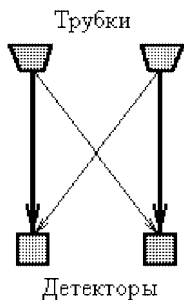


Рис. 1.10

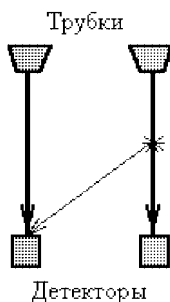


Рис. 1.11

в) Детектор воспринимает излучение, приходящее не только в него, но и (в ослабленной форме) в соседние детекторы (конструктивный эффект взаимного влияния детекторов).

Эффекты а), б), в) присущи томографам 2-го поколения, а томографам 3–5 поколений присущи эффекты б) и в).

В результате действия этих эффектов измеряется (в зависимости от l) не функция $I(l, \theta)$, а некоторая более сглаженная функция $\tilde{I}(l, \theta)$. Если $\tilde{I}(l, \theta)$ подставить в формулу (1.5) вместо $I(l, \theta)$, то в результате вместо функции $q(l, \theta)$ мы получим некоторую более сглаженную (в зависимости от l) функцию $v(l, \theta) = -\ln[\tilde{I}(l, \theta)/I_0(l, \theta)]$ (см. рис. 1.12).

Математически эффекты а), б), в) целесообразно объединить в один эффект. Тогда функции $v(l, \theta)$ и $q(l, \theta)$ будут связаны следующим соотношением:

$$v(l, \theta) = \int_{-\infty}^{\infty} q(l', \theta) K(l - l') dl', \quad (1.15)$$

где $K(l)$ — АФ системы, учитывающая эффекты а), б), в). Запишем соотношение (1.15) иначе, опустив θ , играющее роль параметра:

$$\int_{-\infty}^{\infty} K(l - l') q(l') dl' = v(l). \quad (1.16)$$

Соотношение (1.16) есть одномерное интегральное уравнение Фредгольма I рода типа свертки при каждом значении параметра θ . В нем $v(l)$ известна из измерений, $K(l)$ можно определить заранее экспериментально, а $q(l)$ есть искомая функция. Классическое решение уравнения (1.16) методом ПФ (инверсный фильтр) имеет вид (см. п. 7.2):

$$q(l') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{q}(\omega) e^{-i\omega l'} d\omega, \quad (1.17)$$

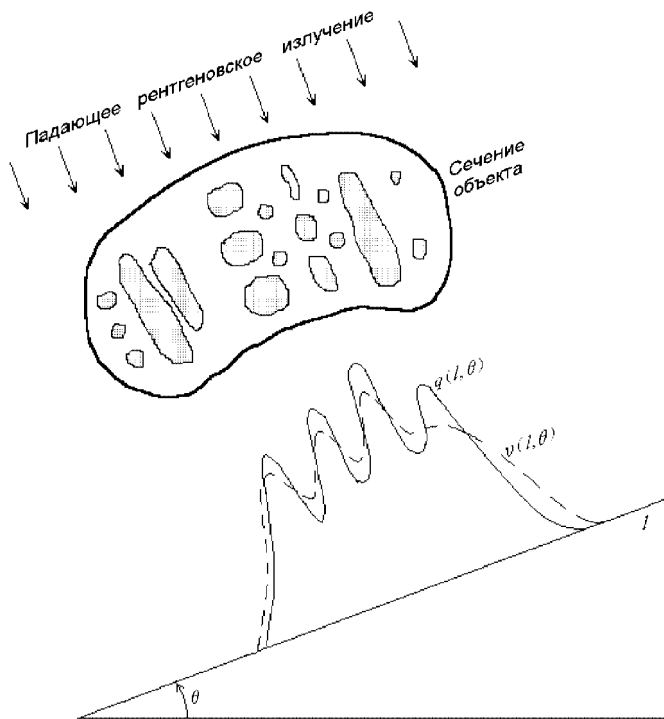


Рис. 1.12

где

$$\hat{q}(\omega) = \frac{\hat{v}(\omega)}{\hat{K}(\omega)}, \quad (1.18)$$

$$\hat{v}(\omega) = \int_{-\infty}^{\infty} v(l) e^{i\omega l} dl, \quad (1.19)$$

$$\hat{K}(\omega) = \int_{-\infty}^{\infty} K(l) e^{i\omega l} dl. \quad (1.20)$$

Однако решение (1.17) неустойчиво, так как задача решения уравнения (1.16) некорректна (см. п. 7.1). Поясним это. Вместо $v(l)$ измеряется $\tilde{v}(l) = v(l) + \delta v(l)$, где $\delta v(l)$ — погрешность измерений

(помеха). Поэтому (1.19) нужно записать в виде:

$$\widehat{v}(\omega) = \int_{-\infty}^{\infty} v(l) e^{i\omega l} dl + \int_{-\infty}^{\infty} \delta v(l) e^{i\omega l} dl, \quad (1.19')$$

где $\widehat{\delta v}(\omega) = \int_{-\infty}^{\infty} \delta v(l) e^{i\omega l} dl$ есть ПФ от помехи, которая обычно содержит компоненту белого шума ($\widehat{\delta v}(\omega) \rightarrow \text{const}$ при $\omega \rightarrow \infty$) и интеграл (1.17) расходится. На практике интегралы заменяются конечными суммами и эффект неустойчивости ослабевает, но остается. Его можно проиллюстрировать рис. 1.13, который говорит о том, что «рябь» в $\tilde{v}(l)$ (даже очень малая) приводит к следующему эффекту: $\widehat{v}(\omega) \rightarrow \text{const}$ при $\omega \rightarrow \infty$, что ведет (даже при очень малом значении const) к неустойчивому спектру $\widehat{q}(\omega)$ в виде «пилы» большой амплитуды при высоких частотах $|\omega|$, а это ведет к неустойчивому решению $\tilde{q}(l)$ в виде «пилы» большой амплитуды при всех значениях l (на рис. 1.13 $\tilde{q}(l)$ — точное решение).

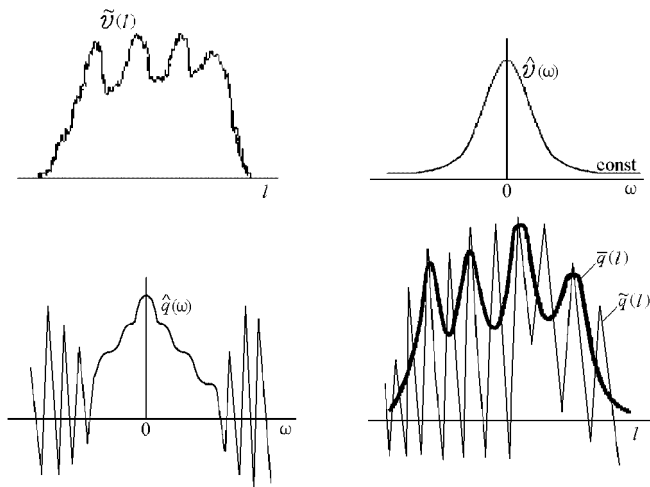


Рис. 1.13

Устойчивое решение уравнения (1.16) дает метод регуляризации Тихонова (см. п. 8.1), согласно которому

$$q_{\alpha}(l') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\widehat{K}(-\omega) \widehat{v}(\omega)}{|\widehat{K}(\omega)|^2 + \alpha |\omega|^{2p}} e^{-i\omega l'} d\omega, \quad (1.21)$$

где $\alpha > 0$ — параметр регуляризации, $p \geq 0$ — порядок регуляризации (обычно $p = 1$). Устойчивость решения $q_\alpha(l)$ (его близость к $\bar{q}(l)$) обусловлена (умеренным) подавлением высоких частот в $\hat{v}(\omega)$.

Конечно, для приближения функции v к функции q можно не использовать данную математическую методику, а использовать более жесткие рентгеновские лучи — тогда эффект а) уменьшится, но жесткие лучи вредны для пациентов и обслуживающего персонала или же использовать более совершенные (имеющие меньший эффект в)) детекторы, но это потребует больших финансовых затрат. Однако этого можно достичь и математическим путем, а именно, путем решения уравнения (1.16) согласно (1.21), (1.19), (1.20).

Визуализация результатов $s(x, y)$ (представление слоя на дисплее). Рассчитанная функция $s(x, y)$ (плотность вещества в сечении) подается на дисплей (видеомонитор, экран) компьютера для отображения (см. рис. 1.4, 1.5). Однако дисплей имеет ограниченный диапазон яркостей. Его *яркостная характеристика* имеет следующий вид (см. рис. 1.14).

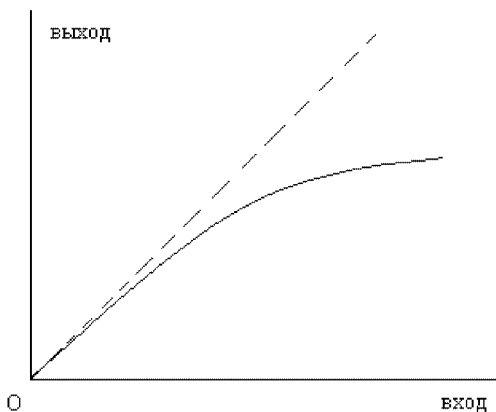


Рис. 1.14

Рис. 1.14 говорит о том, что дисплей отображает большие яркости с занижением и изображение получается недостаточно контрастным. Можно, конечно, увеличить контраст изображения, поворачивая ручку контраста дисплея, но в этом случае преумножится эффект конечного набора градаций яркости дисплея и это будет выглядеть как помеха. Выход из положения — использовать совершенный дисплей (с большим диапазоном яркостей и большим набором градаций яркости) или воспользоваться математическим решением данного вопроса.

Математически задача сводится к решению следующего уравнения:

$$\iint_G H(x - x', y - y') w(x', y') dx' dy' = c(x, y), \quad (1.22)$$

где $H(x, y)$ — АФ дисплея (учитывающая эффект ограниченности диапазона яркостей), G — граница дисплея, $c(x, y)$ — та функция, которую мы хотим отобразить на дисплее, $w(x, y)$ — искомая, более контрастная, чем $c(x, y)$, функция, которую нужно подать на полутоновый (черно-белый или серый) дисплей, чтобы он отобразил менее контрастную функцию $c(x, y)$. Соотношение (1.22) есть двухмерное интегральное уравнение Фредгольма I рода типа свертки (см. п. 5.1). Его классическое решение методом двухмерного ПФ (инверсный фильтр) имеет вид:

$$w(x, y) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \hat{w}(\omega_1, \omega_2) e^{-i(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2, \quad (1.23)$$

где

$$\hat{w}(\omega_1, \omega_2) = \frac{\hat{c}(\omega_1, \omega_2)}{\hat{H}(\omega_1, \omega_2)}, \quad (1.24)$$

$$\hat{c}(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} c(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy, \quad (1.25)$$

$$\hat{H}(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} H(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy. \quad (1.26)$$

Однако решение (1.23) неустойчиво, так как задача решения уравнения (1.22) некорректна. Устойчивое решение уравнения (1.22) методом регуляризации Тихонова имеет вид (см. п. 8.1):

$$w_\alpha(x, y) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \frac{\hat{H}(-\omega_1, -\omega_2) \hat{c}(\omega_1, \omega_2)}{|\hat{H}(\omega_1, \omega_2)|^2 + \alpha M(\omega_1, \omega_2)} e^{-i(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2, \quad (1.27)$$

$$M(\omega_1, \omega_2) = \omega_1^2 + \omega_2^2. \quad (1.28)$$

Об алгоритмах и программах. На практике изложенные методы решения интегральных уравнений (1.6), (1.8), (1.16), (1.22) реализуются в виде алгоритмов. Имеется в виду следующее. Поскольку все уравнения являются уравнениями типа свертки и их решения выражаются через преобразования Фурье, то все сводится к вычислению непрерывных преобразований Фурье (прямых и обратных, одно- и двухмерных) (1.10), (1.12), (1.17), (1.19), (1.20), (1.21), (1.23), (1.25), (1.26). Непрерывные преобразования Фурье

(НПФ) расписываются в виде дискретных преобразований Фурье (ДПФ), а ДПФ реализуются в виде быстрых преобразований Фурье (БПФ) (см. п. 6.3). Это можно отобразить схемой:

$$\begin{aligned} \text{НПФ} &\rightarrow \text{ДПФ} \rightarrow \text{БПФ}, \\ \text{ОНПФ} &\rightarrow \text{ОДПФ} \rightarrow \text{ОБПФ}. \end{aligned}$$

Для компьютерной реализации БПФ разработано много стандартных программ, см., например, FFT [56] и FTF1C [71, с. 183, 190] — программы одномерного БПФ и FTF1C [71, с. 190] — программа двухмерного БПФ. Существуют также спецпроцессоры БПФ.

Разработаны также программы для решения одно- и двухмерных интегральных уравнений Фредгольма I рода типа свертки методом регуляризации Тихонова. Это — программы РТИКР [71, с. 178], CONV1, CONV2, CONV3, CONV4, CONV5 [19, с. 379–388], CONVOL [61, пакет CONF] и др. — для решения одномерных уравнений и РТИТР [71, с. 185] и др. — для решения двухмерных уравнений.

Общая схема обработки в РТ. В итоге общая схема обработки результатов измерений в рентгеновской томографии выглядит следующим образом (рассмотрим ее на примере методики, использующей интегральное уравнение (1.8), метод ПФ и метод регуляризации Тихонова):

- 1) Измеряется $I(l, \theta)$ (см. рис. 1.3, 1.6–1.9) для дискретных значений l и θ : $l_1, l_2, \dots, \theta_1, \theta_2, \dots$
- 2) Вычисляется $q(l, \theta)$ согласно (1.5).
- 3) Вычисляется $S(x, y)$ путем численного интегрирования (1.9) для равномерных сеток дискретных значений x и y :

$$x_1, x_2, \dots, x_N, \quad y_1, y_2, \dots, y_M, \quad (1.29)$$

где N и M — целые степени числа 2, например, $N = M = 256$ (это нужно для БПФ). При этом для вычисления $q(x \cos \theta + y \sin \theta, \theta)$ используется интерполяция.

4) Вычисляется двухмерное ПФ $\hat{S}(\omega_1, \omega_2)$ согласно (1.12) на равномерных сетках дискретных значений ω_1 и ω_2 с помощью двухмерного БПФ.

5) Вычисляется двухмерное ПФ с регуляризацией $\hat{c}_\alpha(\omega_1, \omega_2)$ согласно (1.14) при α , выбранном, например, способом подбора (см. п. 8.1).

6) Вычисляется искомая плотность $c(x, y)$ согласно (1.10) на сетках узлов (1.29) с помощью двухмерного ОБПФ (где вместо $\hat{c}(\omega_1, \omega_2)$ используется $\hat{c}_\alpha(\omega_1, \omega_2)$).

При этом дополнительно могут быть подключены задачи снятия влияния аппаратных искажений (см. (1.21)) и визуализации результатов на дисплее (см. (1.27)).

Контрольные задания и вопросы

1. Дать определение обратной задачи и привести пример.
2. Объяснить суть всех типов томографии (РТ, ЯМР и др.), указав их сходства и различия.
3. Перечислить недостатки обычной рентгенографии.
4. В чем заключается идея РТ?
5. Привести схему рентгеновского томографа 2-го поколения.
6. Записать и истолковать закон Бера.
7. Записать преобразование Радона.
8. Записать и интерпретировать уравнение Радона.
9. Записать уравнение Радона в виде двухмерного интегрального уравнения Фредгольма I рода типа свертки.
10. Охарактеризовать 5 поколений рентгеновских томографов.
11. Перечислить области применения РТ.
12. Перечислить 3 типа аппаратурных искажений в РТ. Записать интегральное уравнение, описывающее эти эффекты. Привести его решение методом ПФ без регуляризации и с регуляризацией.
13. В чем заключается недостаток визуализации плотности $c(x, y)$ на дисплее?
14. Записать двухмерное интегральное уравнение Фредгольма I рода типа свертки для задачи визуализации и дать его решение методом ПФ без регуляризации и с регуляризацией.
15. В чем заключается практическая реализация методов решения основных интегральных уравнений рентгеновской томографии?
16. Дать общую схему обработки в РТ.

1.2. Ядерно-магнитно-резонансная (ЯМР-) томография

В 1960 г. В. А. Иванов [34] впервые предложил способ определения плотности вещества на основе ядерного магнитного резонанса (ЯМР) [1, 7], [74, т. 2], [85]. Подобные способы составляют суть ЯМР-томографии [1, 7, 24, 25, 34, 35], [74, т. 2], [85, 93, 96, 99, 100]. Эффект ЯМР впервые наблюдался в 1937 г. Раби на изолированных ядрах, а в 1946 г. группами физиков, руководимых Блохом и Парселлом, в конденсированных средах.

Эффект ЯМР. Как известно, вещество состоит из молекул (например, H_2O), молекулы — из атомов, а атомы имеют ядра, содержащие протоны (p) и нейтроны (n). Протон p и нейтрон n имеют собственные механические (вращательные) моменты количества движения — спины, а также магнитные моменты $\sigma^{\lambda\mu}$. У ядер, состоящих из p и n и имеющих нечетное число p и n , есть спин и магнитный момент. Это справедливо, например, для

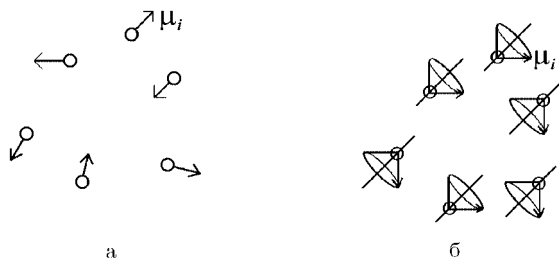


Рис. 1.15

воды, молекула которой H_2O состоит из двух атомов водорода и одного атома кислорода, а ядро водорода состоит из одного p .

В отсутствие внешнего магнитного поля магнитные моменты ядер вещества μ ориентированы случайным образом (рис. 1.15а).

При включении же магнитного поля напряженности \mathbf{H}_0 вектора магнитных моментов ядер μ_i будут прецессировать вокруг вектора \mathbf{H}_0 как по полю, так и против него (см. рис. 1.15б).

З а м е ч а н и е. В данном параграфе для обозначения магнитного поля мы будем использовать два обозначения: H (напряженность) или B (индукция).

Уравнение Лармора. Движение магнитного момента изолированного спина (протона) в постоянном магнитном поле описывается дифференциальным *уравнением Лармора* [7, с. 19], [74, т. 2, с. 109]:

$$\frac{d\boldsymbol{\mu}(t)}{dt} = \gamma [\boldsymbol{\mu}(t) \times \mathbf{H}_0(t)], \quad (1.29)$$

где γ — гиромагнитное отношение, равное $\gamma = e/2m$, причем e и m — заряд и масса протона. Введем неподвижную (*лабораторную*) систему координат xuz , причем ось z направим вдоль \mathbf{H}_0 . Спроектируем уравнение (1.29) на z , x и y , получим:

$$\left. \begin{aligned} \frac{d\mu_z(t)}{dt} &= \gamma (\mu_x H_y - \mu_y H_x) = 0, \\ \frac{d\mu_x(t)}{dt} &= \gamma (\mu_y H_z - \mu_z H_y) = \gamma \mu_y H_0, \\ \frac{d\mu_y(t)}{dt} &= \gamma (\mu_z H_x - \mu_x H_z) = -\gamma \mu_x H_0, \end{aligned} \right\} \quad (1.30)$$

так как $H_z = H_0$, $H_x = H_y = 0$. Из первого уравнения (1.30) имеем: $\mu_z(t) = \text{const}$, а значит, $\theta = \text{const}$, где θ — угол между $\boldsymbol{\mu}(t)$ и \mathbf{H}_0 . Продифференцировав второе уравнение (1.30) по t и используя третье уравнение, получим:

$$\frac{d^2\mu_x}{dt^2} + (\gamma H_0)^2 \mu_x = 0. \quad (1.31)$$

Уравнение (1.31) есть дифференциальное уравнение гармонических колебаний. Его решение

$$\mu_x(t) = A \cos(\gamma H_0 t + \varphi),$$

где A и φ — некоторые константы. Аналогично,

$$\mu_y(t) = -A \sin(\gamma H_0 t + \varphi).$$

Введем μ_{\perp} — составляющую в плоскости x, y , равную $\mu_{\perp} = \sqrt{\mu_x^2 + \mu_y^2}$. Получим:

$$\mu_{\perp}(t) = A = \text{const}.$$

В результате $\mu_z(t) = \text{const}$, а конец вектора $\mu_{\perp}(t)$ вращается в плоскости xy с угловой скоростью

$$\omega_0 = \gamma H_0, \quad (1.32)$$

т. е. в пространстве вектор $\mu(t)$ прецессирует по конусу прецессии вокруг \mathbf{H}_0 с угловой скоростью (1.32) (см. рис. 1.16).

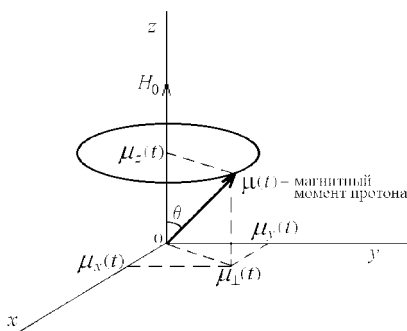


Рис. 1.16

Прецессия магнитных моментов протонов (а также электронов и атомных ядер) в магнитном поле — это основной процесс, характерный для ЯМР.

Определение. Угловая скорость $\omega_0 = \gamma H_0$ называется *угловой скоростью Ларморовой прецессии*.

Во вращающейся же вокруг оси z с угловой скоростью $\omega_0 = \gamma H_0$ системе координат имеем: $\mu_z(t) = \text{const}$, $\theta(t) = \text{const}$, $\mu_x(t) = A \cos \varphi = \text{const}$, $\mu_y(t) = -A \sin \varphi = \text{const}$.

Однако это — классический подход. С позиций же квантово-механического подхода магнитный момент μ не прецессирует по конусу с определенной частотой, а нужно говорить о вероятности его нахождения на поверхности конуса.

Ансамбль протонов. Рассмотрим некоторое большое количество протонов — ансамбль (популяцию) протонов. Практически это миллионы или миллиарды протонов. Все протоны, входящие в ансамбль, имеют лишь m конусов прецессии, причем $m = 2I + 1$, где $I = 0, \dots, 6$ — спин, направленный как вдоль поля \mathbf{H}_0 , так и против него.

Вводится в рассмотрение суммарный вектор

$$\mathbf{M} = \sum_i \boldsymbol{\mu}_i, \quad (1.33)$$

называемый *вектором ядерной намагниченности*. Казалось бы, в силу симметрии, \mathbf{M} должен быть равен нулю. Однако несимметричная квантовая механика показывает, что отношение числа протонов, прецессирующих по полю, к числу протонов, прецессирующих против него, чуть больше единицы и равно [74, т. 2, с. 118]

$$N_{\text{отн}} = e^{\Delta E/kT}, \quad (1.34)$$

где $\Delta E = \gamma \hbar H_0$, $\hbar = h/2\pi$, h и k — постоянные Планка и Больцмана, T — абсолютная температура. Если, например, $T = 25^\circ\text{C} = 298^\circ\text{K}$, $H_0 = 1$ Тл, то $N_{\text{отн}} \cong 1.000003 = 1 + 3 \cdot 10^{-6}$, т. е. из миллиона протонов лишь на 3 протона будет больше прецессировать по полю, чем против поля. В результате вектор ядерной намагниченности \mathbf{M} будет направлен по полю \mathbf{H}_0 :

$$\mathbf{M} = \chi \mathbf{H}_0, \quad (1.35)$$

где χ называется *ядерной парамагнитной восприимчивостью*. Однако \mathbf{M} очень мал, почему и не могли долго обнаружить эффект ядерной намагниченности. Более того, вся ЯМР-томография, по существу, основана на отличии χ от нуля, другими словами, на несимметричности ориентации протонов по полю \mathbf{H}_0 и против него, хотя и очень мал этот эффект. Поэтому нижеизложенные методы ЯМР-томографии связаны с очень слабыми сигналами (эхо-сигналами) и требуется охлаждение томографа жидким гелием, высокая степень однородности поля \mathbf{H}_0 ($\Delta H/H_0 \sim 10^{-6}$) и т. д.

Движение магнитных моментов изолированных протонов в постоянном и переменном магнитных полях [7, с. 23–26]. Пусть помимо постоянного магнитного поля \mathbf{H}_0 (вдоль z) действует перпендикулярное ему переменное поле, осциллирующее вдоль x : $H_x = H_x(t) = 2H_1 \cos \omega t$, причем $H_1 \ll H_0$ (на практике $\omega \sim 1\text{--}10$ МГц — радиочастотный диапазон). Можно показать, что $H_x(t)$ равно сумме двух вращающихся в разных направлениях полей $\mathbf{H}^{(1)}(t)$ и $\mathbf{H}^{(2)}(t)$ с одинаковой частотой ω и амплитудой H_1 (см. рис. 1.17).

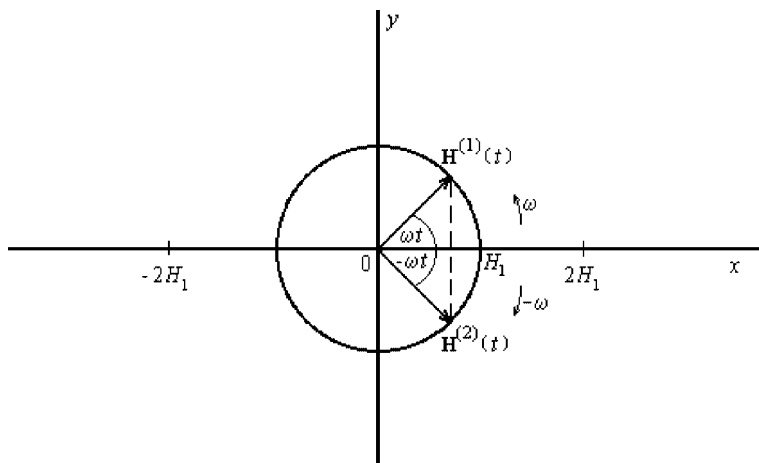


Рис. 1.17

Действительно, сумма проекций $\mathbf{H}^{(1)}(t)$ и $\mathbf{H}^{(2)}(t)$ на ось x равна

$$H_x^{(1)}(t) + H_x^{(2)}(t) = H_1 \cos \omega t + H_1 \cos \omega t = 2H_1 \cos \omega t = H_x(t), \quad (1.36)$$

а сумма проекций $\mathbf{H}^{(1)}(t)$ и $\mathbf{H}^{(2)}(t)$ на y равна

$$H_y^{(1)}(t) + H_y^{(2)}(t) = H_1 \sin \omega t - H_1 \sin \omega t = 0. \quad (1.37)$$

Постоянному магнитному полю H_0 соответствует частота $\omega_0 = \gamma H_0$, а одному из вращающихся полей $H^{(1)}(t)$ или $H^{(2)}(t)$ — частота $\omega_1 = \gamma H_1$.

Определение. Частота $\omega_0 = \gamma H_0$ называется *частотой прецессии*, частота $\omega_1 = \gamma H_1 \ll \omega_0$ — *частотой нутации*, а ω — *частотой вращения полей* $H^{(1)}(t)$ и $H^{(2)}(t)$.

Определение. Разность частот $\Delta\omega = \omega_0 - \omega$ называется *расстройкой по частоте*.

Если $\Delta\omega = 0$, т.е. одно из полей $H^{(1)}(t)$ или $H^{(2)}(t)$ вращается в такт с прецессией протона вокруг H_0 , то имеет место *резонанс* ($\omega_0 = \omega$). В результате магнитный момент $\boldsymbol{\mu}$ будет совершать одновременно две прецессии: вокруг H_0 (или оси z) с угловой скоростью прецессии $\omega_0 = \gamma H_0$ и вокруг H_1 (или оси x) с частотой нутации $\omega_1 = \gamma H_1 \ll \omega_0$, а именно, в лабораторной (неподвижной) системе координат конец вектора $\boldsymbol{\mu}$ будет выписывать спираль на поверхности шара. При этом если $\Delta\omega = 0$, то угол нутации Эйлера θ будет изменяться от некоторого минимального значения θ_1 до некоторого максимального значения θ_2 , а вектор $\boldsymbol{\mu}$ будет переходить с верхнего конуса прецессии на нижний конус нутации (см. рис. 1.18).

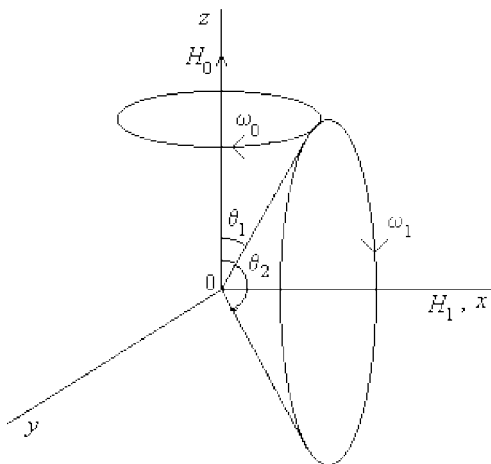


Рис. 1.18

В квантовомеханическом представлении этот переход будет скачкообразным. Если $\Delta\omega = 0$, т.е. $\omega = \omega_0$, то будет иметь место *ядерный магнитный резонанс*, или избирательное поглощение электромагнитной энергии веществом. Как показано ниже, это приведет к тому, что само вещество будет излучать электромагнитную энергию в виде так называемого *эхо-сигнала*, по которому можно судить о плотности вещества.

Уравнения Блоха. Опять рассмотрим ансамбль протонов и случай изменения поля, а именно, положим, что поле \mathbf{V}_0 (старое значение) отключено и сразу включено новое поле \mathbf{V}_z (ось z направим вдоль \mathbf{V}_z) и пусть угол между \mathbf{V}_0 и \mathbf{V}_z равен θ . Естественно, в начальный момент вектор \mathbf{M} ансамбля протонов равен $\mathbf{M}(0) = \chi\mathbf{V}_0$, а при $t \rightarrow \infty$ $\mathbf{M}(t) = M_z = \chi\mathbf{V}_z$. Спрашивается, каково будет поведение вектора $\mathbf{M}(t)$ при $t \in [0, \infty)$? Ответ на этот вопрос дают *феноменологические дифференциальные уравнения Блоха* (в лабораторной системе координат) [7, с. 42], [74, т. 2, с. 119], [85]:

$$\frac{d\mathbf{M}(t)}{dt} = \gamma [\mathbf{M}(t) \times \mathbf{V}_z] - \frac{M_z(t) - M_z}{T_1} \mathbf{e}_z - \frac{M_x(t) \mathbf{e}_x + M_y(t) \mathbf{e}_y}{T_2^*}, \quad (1.38)$$

где $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ — единичные вектора́ вдоль x, y, z .

Первое слагаемое в правой части (1.38) $\gamma[\mathbf{M}(t) \times \mathbf{V}_z]$ по аналогии с (1.29) означает прецессию вектора $\mathbf{M}(t)$ вокруг \mathbf{V}_z с угловой скоростью $\omega = \gamma V_z$, а следующие слагаемые учитывают так называемые *процессы релаксации* (ослабления, уменьшения, перехода в новое положение), а именно, спин-решетчатую и спин-спиновую релаксации.

Спин-решетчатая, или *продольная релаксация* (характеризуемая временем T_1) обусловлена энергетическими обменами между спинами (протонами) и средой, в которой они находятся. *Спин-спиновая*, или *поперечная релаксация* (характеризуемая временем T_2) обусловлена взаимодействиями между спинами. Эти взаимодействия создают локальные поля, т. е. неоднородности поля \mathbf{B} , что ведет к неодинаковости угловых скоростей ω или к *расфазированию* спинов и потере резонанса в плоскости xy (см. рис. 1.19).

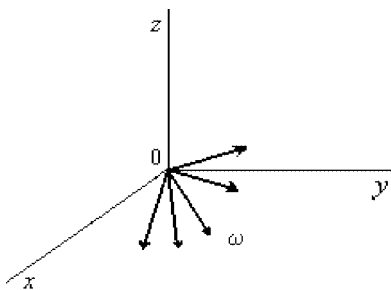


Рис. 1.19

Кроме того, имеет место *техническая неоднородность магнитного поля* ΔB , усиливающая расфазирование. Вводится новое время поперечной релаксации T_2^* [74, т. 2, с. 131]:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma \frac{\Delta B}{2},$$

в результате $T_2^* \leq T_2$, т. е. релаксация (расфазирование, потеря резонанса) в плоскости xy происходит быстрее при $\Delta B \neq 0$, чем при $\Delta B = 0$.

В лабораторной (неподвижной) системе координат решение уравнений (1.38) имеет вид:

$$\left. \begin{aligned} M_z(t) &= M_z + (M_0 \cos \theta - M_z) e^{-t/T_1}, \\ M_x(t) &= M_0 \sin \theta \cos \beta(t) e^{-t/T_2^*}, \\ M_y(t) &= M_0 \sin \theta \sin \beta(t) e^{-t/T_2^*}, \end{aligned} \right\} \quad (1.39)$$

где $\beta(t) = \beta_0 + \omega t = \beta_0 + \gamma M_z t$, или

$$\left. \begin{aligned} M_z(t) &= M_z + (M_0 \cos \theta - M_z) e^{-t/T_1}, \\ M_{\perp}(t) &= M_0 \sin \theta e^{-t/T_2^*}. \end{aligned} \right\} \quad (1.40)$$

Решение (1.40) показывает, что конец вектора \mathbf{M} будет двигаться по спирали от \mathbf{M}_0 (при $t = 0$) до \mathbf{M}_z (при $t \rightarrow \infty$), причем проекция $M_{\perp}(t)$ будет совершать круговое вращение в плоскости xy с угловой скоростью $\omega = \gamma M_z$ (см. рис. 1.20).

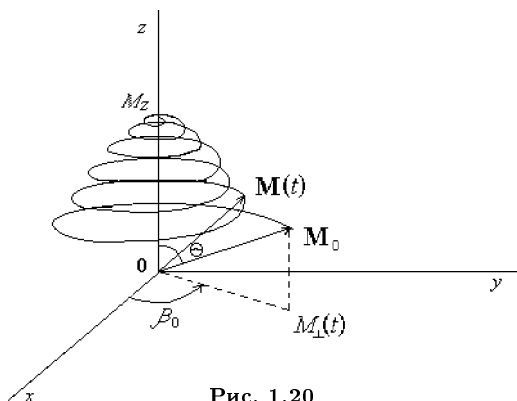


Рис. 1.20

В табл. 1.1, 1.2, 1.3 приведены значения γ , T_1 и процентное содержание воды в различных тканях организма [34, 35].

Таблица 1.1

ядро	γ , МГц/Тл	$\gamma/2\pi$
водорода H^2	267.8	42.58
углерода C^{13}	67.1	10.7
фосфора P^{31}	108.1	17.2

Таблица 1.2

орган	T_1 , сек (здоровая ткань)	T_1 , сек (рак)
легкое	0.73	1.1
кишечник	0.68	1.15
желудок	0.75	1.2
печень	0.59	0.8
селезенка	0.7	1.1

Таблица 1.3

ткань	% воды
серое вещество мозга	83
почки	81
сердце	80
мышцы, легкие	78
печень	71
кожа	68
кость	12

Табл. 1.1 показывает, что значения γ , а значит ларморовы частоты различных веществ весьма различны. Поэтому изучая, например, распределение водорода в организме в районе частоты $\omega = \gamma B$, где $\gamma = 267.8 \text{ МГц/Тл}$, мы не будем получать искажающую информацию от других веществ (углерода, фосфора и т. д.). Табл. 1.2 показывает, что время продольной релаксации T_1 заметно отличается у здоровой и раковой ткани. Поэтому, определяя T_1 некоторого органа или его части, можно поставить диагноз (здоровая ткань, предраковая ткань, рак). Время T_1 можно определить по огибающей эхо-сигнала, которая $\sim e^{-t/T_1}$ (см. рис. 1.22) или другими способами [74, т. 2, с. 129–130]. Табл. 1.3 показывает, что в тканях организма достаточно воды, а значит, протонов, поэтому реконструкцию ЯМР-изображений (см. дальше) наиболее эффективно выполнять, определяя плотность протонов $c_z(x, y)$.

Отметим также, что $e^{-1} \cong 0.37$, $e^{-2} \cong 0.13$, $e^{-3} \cong 0.05$, $e^{-4} \cong 0.02$, $e^{-5} \cong 0.007$, $e^{-6} \cong 0.0025$, $e^{-7} \cong 0.0009, \dots$ Кроме того, значение T_2 (и T_2^*) близко к T_1 , т. е. $T_2 \approx T_2^* \sim 1 \text{ сек}$. Это говорит о том, что процессы продольной релаксации (см. первое уравнение (1.40)) и поперечной релаксации (см. второе уравнение (1.40)) проходят в тканях организма практически за несколько секунд (после чего M_{\perp} практически обнуляется).

Если говорить о различных веществах, то отметим следующее. В жидкостях обычно T_1 и T_2 близки друг другу. При этом кристаллизация приводит к значительному уменьшению T_2 . В чистых диамагнитных (намагничивающихся в магнитном поле в противоположном ему направлении) кристаллах $T_1 \sim 1\text{--}10 \text{ час}$. А парамагнитные примеси (намагничивающиеся по направлению поля) уменьшают T_1 : например, для парамагнитных жидких растворов $T_1 \sim 10^{-3} \text{--} 10^{-4} \text{ сек}$. В металлах (при температуре $1\text{--}10 \text{ К}$) $T_1 \sim 10^{-3} \text{--} 10 \text{ сек}$.

Эхо-сигнал, $\pi/2$ - и π -импульсы. Важной особенностью ЯМР-томографии является то, что ансамбль протонов излучает так называемый (электромагнитный радиочастотный) *эхо-сигнал* и это имеет место лишь при наличии M_{\perp} компоненты (поперечной намагниченности, см. (1.40) и рис. 1.20). Однако $M_{\perp}(t)$ убывает с ростом t , как это видно из (1.40), и это убывание обусловлено расфазированием протонов и неоднородностью поля. А в пределе (при $t \rightarrow \infty$) наступает равновесное состояние, когда \mathbf{M} направлен вдоль \mathbf{M}_z и $M_{\perp} = 0$, т. е. не будет прецессии \mathbf{M} вокруг \mathbf{M}_z и эхо-сигнала от системы. Чтобы получить эхо-сигнал, нужно отклонить вектор \mathbf{M} от \mathbf{M}_z — в этом случае $M_{\perp} \neq 0$. Это — основа ЯМР-измерений.

Конкретно данную процедуру можно выполнить с помощью так называемого *импульсного метода*. Он заключается в следующем. Создается короткий мощный ВЧ (радиочастотный) импульс в направлении, не параллельном \mathbf{B}_z , лучше всего под углом 90° к \mathbf{B}_z

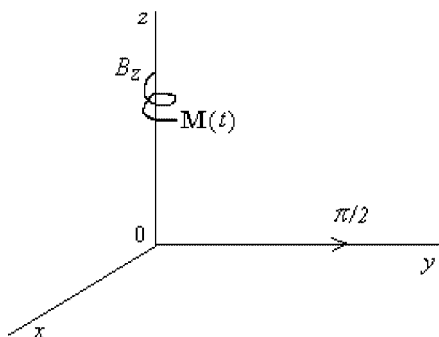


Рис. 1.21

($\pi/2$ -импульс Карра-Парселла резонансной частоты $\omega = \gamma B_z$), например, вдоль y (без отключения B_z). В этом случае вектор \mathbf{M} будет совершать спиралевидное отклонение от оси z с появлением компоненты M_{\perp} (см. рис.1.21).

От ансамбля протонов будет исходить эхо-сигнал, или *сигнал спада свободной индукции* (ССИ). Однако этот сигнал будет слабым и затухающим из-за расфазирования протонов — см. рис.1.22а, где отображены две компоненты эхо-сигнала — по y и по x (но обычно отображают лишь одну компоненту, причем нередко в виде огибающей).

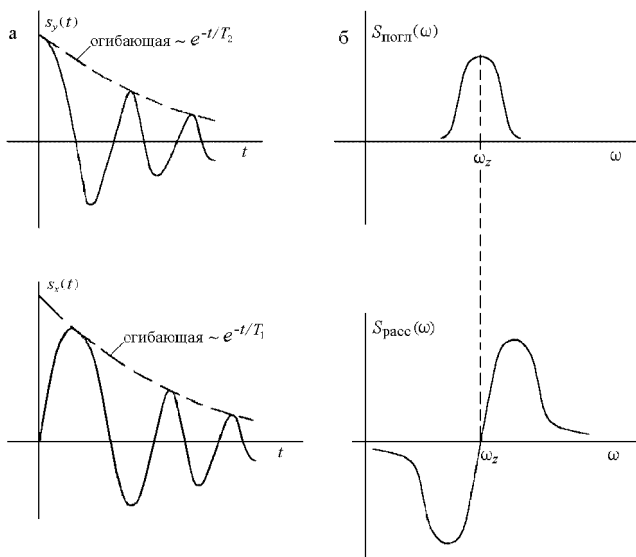


Рис. 1.22

Спектр Фурье эхо-сигнала, отображенного на рис. 1.22а, имеет вид, представленный на рис. 1.22б, где даны вещественная часть спектра — *сигнал поглощения* $S_{\text{погл}}(\omega)$ и мнимая часть — *сигнал дисперсии*, или *рассеяния* $S_{\text{расс}}(\omega)$ [74, т. 2, с. 121]. Для $S_{\text{погл}}(\omega)$ и $S_{\text{расс}}(\omega)$ получены следующие формулы [7, с. 51]:

$$S_{\text{погл}}(\omega) = |\gamma| H_1 T_2 M_z / (1 + \Delta\omega^2 T_2^2 + \gamma^2 H_1^2 T_1 T_2),$$

$$S_{\text{расс}}(\omega) = -|\gamma| H_1 T_2 M_z \Delta\omega T_2 / (1 + \Delta\omega^2 T_2^2 + \gamma^2 H_1^2 T_1 T_2),$$

где $\Delta\omega = \omega_z - \omega$, $\omega_z = \gamma B_z$.

Эхо-сигнал, изображенный на рис. 1.22а, практически нельзя зарегистрировать. Чтобы получить достаточно сильный эхо-сигнал, создается (по истечении некоторого времени T) еще один импульс, а именно, π -импульс Карра-Парселла — импульс, противоположный B_z . Он возвращает спины в фазу и ведет к тому, что по прошествии еще времени T появится сильный (регистрируемый) эхо-сигнал (см. рис. 1.23).

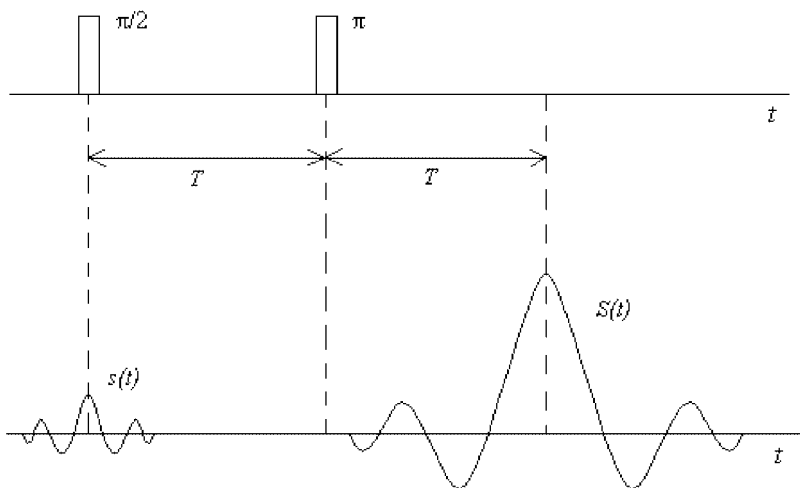


Рис. 1.23

Данная процедура называется *методом эха Ганна* (существует также метод градиентного эха [24, 25]). Метод эха Ганна можно объяснить по аналогии с бегунами на дорожке [72, т. 2, с. 127]. Стартер дает выстрелом старт ($\pi/2$ -импульс) и бегуны (спины) разбегаются вдоль дорожки, но бегут с разными скоростями (расфазирование) и от их ног исходит слабый сигнал $s(t)$. Затем через $t = T$ стартер дает второй выстрел (π -импульс), бегуны поворачивают, бегут обратно со своими скоростями и «медленные» бегуны оказываются впереди «быстрых». В результате при $t = 2T$ все

бегуны будут на старте и одновременно (в фазе) ударят ногами по дорожке — возникнет сильный эхо-сигнал $S(t)$.

При этом отметим следующее:

- 1) мощность эхо-сигнала $S(t)$ несет информацию о плотности вещества,
- 2) частота ларморовой прецессии ω — о типе вещества (см. табл. 1.1),
- 3) время продольной релаксации T_1 — о фазовом состоянии вещества (см. табл. 1.2).

Градиентные поля. Предположим, что исследуется, например, часть руки от кисти до локтя. Направим ось z вдоль ее длины (см. рис. 1.24, где изображен ряд сечений руки).

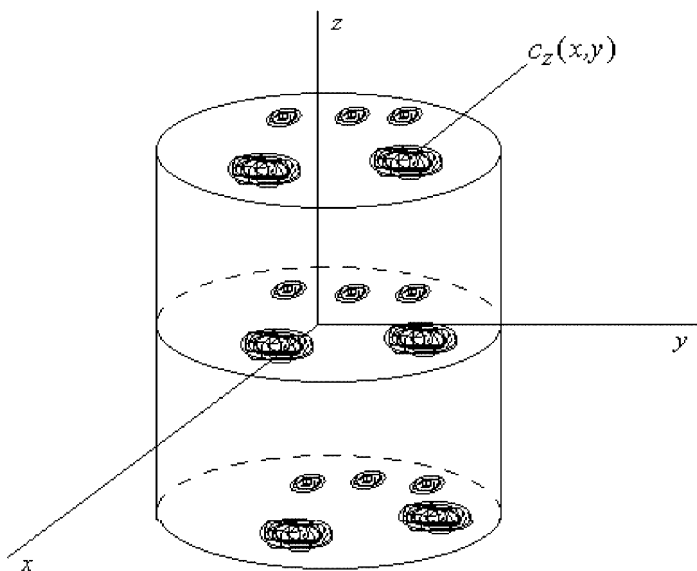


Рис. 1.24

Пусть в ее объеме создано постоянное магнитное поле B_0 , направленное вдоль z . Временно будем полагать, что расфазирование спинов-протонов отсутствует и что $\Delta B = 0$. В этом случае все протоны будут прецессировать с одинаковой ларморовой частотой $\omega = \gamma B_0$ и если приложить $\pi/2$ - и π -импульсы, то мы получим ответный эхо-сигнал $S(t)$ (см. рис. 1.23). Недостатком такого эксперимента является то, что информация в виде эхо-сигнала на частоте ω будет поступать от протонов всего исследуемого объема, а идея томографии в том и состоит, чтобы получать информацию от отдельных сечений (или тонких слоев) — тогда можно

решать задачу реконструкции (см. дальше). Отделить же различные сечения друг от друга можно, сопоставив им различные значения индукции поля B , а значит, ларморовой частоты ω . Тогда, принимая эхо-сигнал на различных частотах ω , мы получим информацию отдельно для каждого сечения.

Чтобы каждому значению z соответствовало свое и только свое значение B , функция $B(z)$ должна быть монотонно возрастающей (или убывающей). Проще (математически и технически) закон изменения поля $B(z)$ полагать линейным — этому соответствует понятие градиентного поля.

Определение. *Градиентным z -полем* называется поле, напряженность (или индукция) которого $G(z)$ изменяется по линейному закону:

$$G_z \equiv G(z) = g_z z, \quad (1.41)$$

где $g_z = \text{const}$ — некоторая постоянная величина, означающая: $g_z = \partial G(z)/\partial z$.

Аналогично вводятся *градиентные поля по x и по y* :

$$\left. \begin{aligned} G_x &= g_x x, \\ G_y &= g_y y. \end{aligned} \right\} \quad (1.42)$$

Теперь мы можем обратиться к основному вопросу — реконструкции ЯМР-изображений.

Реконструкция ЯМР-изображений. Под реконструкцией (формированием) ЯМР-изображений подразумевается определение плотности протонов $c_z(x, y)$ в сечениях по измеренным на разных частотах $\omega = \omega(z)$ эхо-сигналам $S_{\omega(z)}(t)$.

Основной принцип реконструкции ЯМР-изображений заключается в так называемом *пространственном кодировании частоты резонанса* [21, 22]. Он состоит в том, что на статическое поляризующее однородное магнитное поле B_0 накладываются градиентные поля, индукции которых зависят от координат x, y, z , а значит и наоборот каждой точке (x, y, z) соответствует определенная индукция, а следовательно, частота и/или фаза. Другими словами, точки пространства (x, y, z) кодируются под частоту и/или фазу и наоборот под частотой и/или фазой закодированы координаты x, y, z .

Предположим для простоты, что поле B_0 отсутствует, а включены сразу три градиентных поля G_x, G_y и G_z , тогда в некоторой точке (x, y, z) индукция поля будет равна [24, с. 25], [25] $G_x + G_y + G_z = g_x x + g_y y + g_z z$, а резонансная частота (частота ларморовой прецессии) будет равна $\omega(x, y, z) = \gamma(g_x x + g_y y + g_z z) = \gamma \mathbf{G} \mathbf{r}$, где \mathbf{G} — вектор суммы градиентных полей, а \mathbf{r} — радиус-вектор точки (x, y, z) . Введение градиентных полей позволяет решать обратную задачу — определять координаты x, y, z точки, из которой исходит сигнал некоторой частоты ω . Это — основа методики реконструкции ЯМР-изображений.

Множество точек, соответствующих некоторому фиксированному значению ω , образуют в пространстве плоскость:

$$\gamma(g_x x + g_y y + g_z z) = \omega. \quad (1.43)$$

Если измерения выполнять на этой частоте ω , то эхо-сигнал будет равен

$$S_\omega(t) = A \iiint c(x, y, z) e^{i\gamma(g_x x + g_y y + g_z z)t} dx dy dz, \quad (1.44)$$

где A — некоторая константа. Интегрирование в (1.44) ведется по плоскости (1.43) с учетом фазы, поскольку $S(t)$ — волновой процесс, поэтому вводится множитель $\exp[i\gamma(g_x x + g_y y + g_z z)t]$, учитывающий фазовые различия интегрируемых точек.

Однако (1.44) нельзя рассматривать как интегральное уравнение относительно $c(x, y, z)$, поскольку искомая функция $c(x, y, z)$ — трехмерная функция, а измеренная функция $S_\omega(t)$ — одномерная. Поэтому вводится ряд практических схем, в которых это противоречие разрешено. Некоторые из них мы рассмотрим. Во всех схемах полагается, что $g_z \gg g_x$ и $g_z \gg g_y$ — это создает так называемую *селекцию по z* , в результате которой, во-первых, плоскости (1.43) получаются практически перпендикулярными z , а во-вторых, получается четкое разграничение по частотам ω вдоль z .

Рассмотрим идею селекции плоскости подробнее [24, с. 27–31], [25]. Если возбуждать ансамбль протонов с частотой $\omega_0 = \gamma B_0$ и одновременно прикладывать градиентное поле $G_z = g_z(z - z_0)$, так что

$$B_z = B_0 + g_z(z - z_0), \quad (1.45)$$

то будут возбуждаться только протоны в плоскости $z = z_0$. Однако исходящий из этой плоскости эхо-сигнал будет бесконечно слаб. Поэтому нужно возбуждать не плоскость, а слой некоторой толщины δ_z , которая должна быть достаточно «большой», чтобы получить надежно детектируемый эхо-сигнал, и достаточно «малой», чтобы частоты протонов в пределах слоя не сильно отличались.

Обычно $\pi/2$ -импульс имеет следующий вид в функции t (см. рис. 1.25).

Его огибающая хорошо представляется гауссианой. Обозначим через 2τ ширину гауссианы на полувысоте (эффективную ширину гауссианы). Преобразование Фурье от этой гауссианы (спектр огибающей импульса) есть также гауссиана (см. рис. 1.26) с шириной на полувысоте (эффективной шириной) $\delta\omega = 4 \ln 2/\tau$.

Если включить поле (1.45), то резонансная частота спинов будет изменяться по закону

$$\omega(z) = \omega_0 + \gamma g_z(z - z_0). \quad (1.46)$$

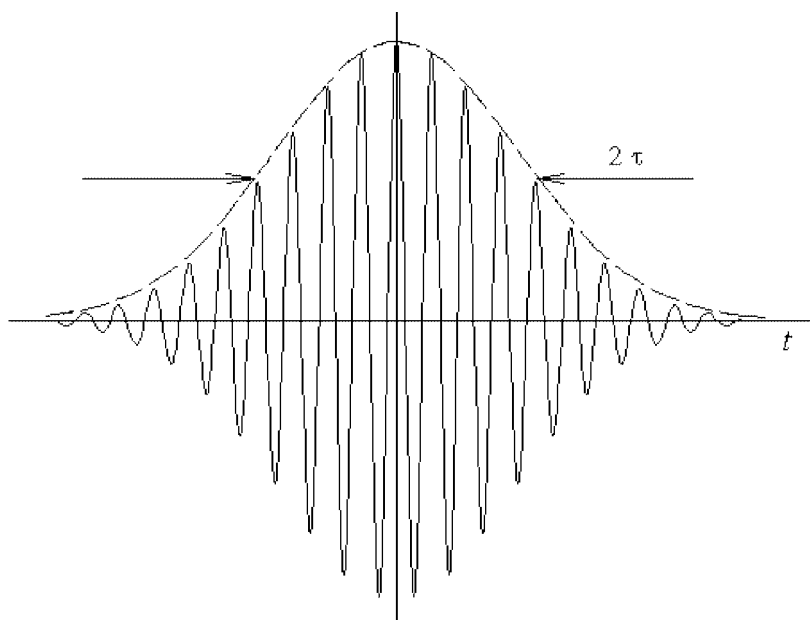


Рис. 1.25

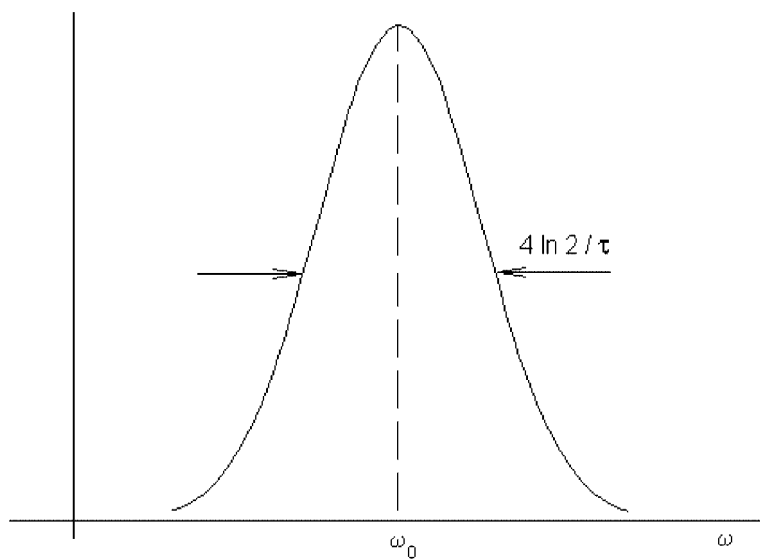


Рис. 1.26

И если теперь облучать вещество $\pi/2$ -импульсом (рис. 1.25, 1.26), то будут возбуждаться протоны в слое с эффективной толщиной

$$\delta_z = \frac{\delta\omega}{\gamma g_z} = \frac{4 \ln 2}{\gamma g_z \tau}. \quad (1.47)$$

Из (1.46) и (1.47) следует, что можно изменять положение возбуждаемого слоя (его среднюю z -координату), меняя ω_0 , и его эффективную толщину δ_z , меняя g_z и/или τ . Однако в пределах слоя (1.47) протоны будут иметь различные частоты (1.46) и будет иметь место расфазирование вдоль z (в отличие от расфазирования в плоскости x, y , изображенного на рис. 1.19), в результате чего суммарный эхо-сигнал от слоя будет очень слабым. Для устранения этого эффекта Хоулт в 1977 г. предложил после действия $\pi/2$ -импульса изменять направление градиентного поля G_z на противоположное, что ведет к рефазированию протонов и к сильному эхо-сигналу (см. рис. 1.27).

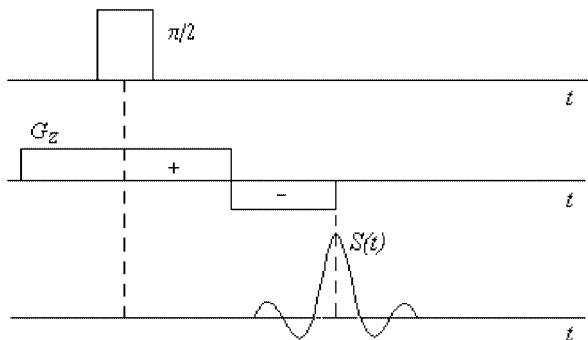


Рис. 1.27

Это напоминает действие π -импульса, описанное выше.

Заметим, что часто $\pi/2$ - и π -импульсы, а также градиентные поля изображаются прямоугольниками (см., например, рис. 1.27). Однако это условность. В действительности, например, под $\pi/2$ -импульсом некоторой ширины 2τ подразумевается импульс типа изображенного на рис. 1.25 с шириной 2τ на полувысоте.

1-я практическая схема [93, с. 272] (см. рис. 1.28).

Согласно этой схеме, для селекции включается градиентное поле G_z с большим значением g_z . Вдоль x включаются поочередно несколько градиентных полей G_x (выполняется несколько экспериментов) с различными g_x продолжительностью T_x , т. е. при каждом значении g_x создается *фазовое кодирование*, поскольку в каждом эксперименте эхо-сигналу соответствует постоянство фазы $\gamma g_x x T_x$. Вдоль y включается одно градиентное поле G_y в принципе бесконечной длительности, т. е. создается *частотное*

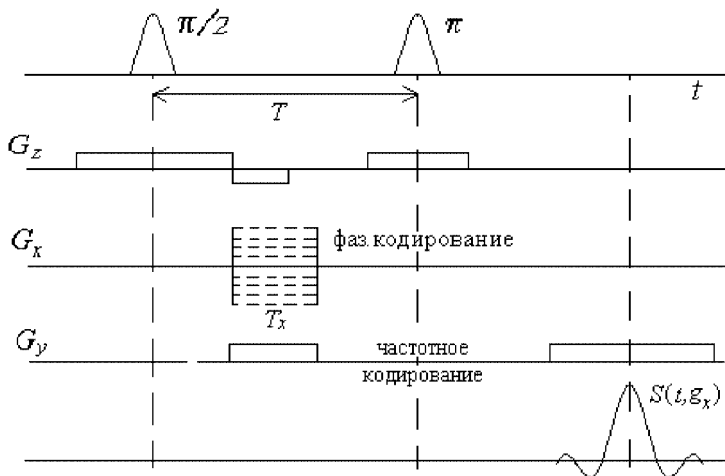


Рис. 1.28

кодирование, поскольку этому соответствует постоянство лишь частоты $\gamma g_y y$. В результате эхо-сигнал может быть записан в виде:

$$S(g_x, t) = A \iint c(x, y) e^{i\gamma(g_x x T_x + g_y y t)} dx dy. \quad (1.48)$$

В соотношении (1.48) g_x и t переменны, а T_x и g_y постоянны (имеют по одному значению). Технически это означает, что задается ряд значений g_x и при каждом значении g_x измеряется эхо-сигнал $S(t)$. В результате будет измерена двумерная функция $S(g_x, t)$. Следовательно, соотношение (1.48) можно рассматривать как двумерное интегральное уравнение относительно $c(x, y)$:

$$A \iint_{-\infty}^{\infty} c(x, y) e^{i\gamma(g_x x T_x + g_y y t)} dx dy = S(g_x, t), \quad (1.49)$$

соответствующее некоторому селективному сечению (слою) с координатой z и, следовательно, частотой $\omega_z = \gamma g_z z$. Это означает, что измерения правой части $S(g_x, t)$ выполнены на некоторой средней частоте ω_z , а искомая функция $c(x, y)$ соответствует координате $z = \omega_z / \gamma g_z$. Следовательно, уравнение (1.49) можно записать более точно:

$$A \iint_{-\infty}^{\infty} c_z(x, y) e^{i\gamma(g_x x T_x + g_y y t)} dx dy = S_{\omega_z}(g_x, t). \quad (1.50)$$

Однако для простоты мы будем пользоваться записью типа (1.49).

Решение уравнения (1.49) с помощью обратного преобразования Фурье (ОПФ) имеет вид (ср. (6.46), п. 6.3):

$$c(\xi/\gamma T_x, \omega_y/\gamma g_y) = \frac{\gamma^2 T_x g_y}{4\pi^2 A} \iint_{-\infty}^{\infty} S(g_x, t) e^{-i(\xi g_x + \omega_y t)} dg_x dt. \quad (1.51)$$

Итак, плотность c определяется через двумерное ОПФ от набора эхо-сигналов S .

2-я практическая схема [93, с.272] (см. рис.1.29).

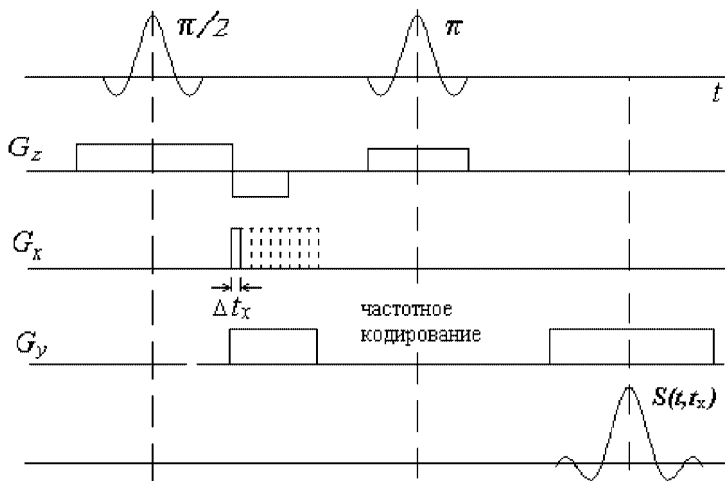


Рис. 1.29

В этой схеме задается набор градиентных полей по x с одинаковыми g_x , но с разной продолжительностью действия t_x . А по y , как в 1-й схеме, создается частотная кодировка. В результате получается набор эхо-сигналов, отличающихся разными значениями t_x :

$$S(t_x, t) = A \iint_{-\infty}^{\infty} c(x, y) e^{i\gamma(g_x x t_x + g_y y t)} dx dy, \quad (1.52)$$

где $t_x = n \cdot \Delta t_x$, $n = 1, 2, \dots$

Существует еще ряд похожих схем обработки двумерных эхо-сигналов [74, т. 2, с.142–145]. Отметим при этом, что обработку типа схем 1 и 2 нужно выполнять на ряде частот ω_z — этому будет соответствовать определение плотности $c(x, y)$ в ряде сечений с различными z , т. е. $c_z(x, y)$.

Рассмотрена также трехмерная Фурье-обработка [93, с. 273–274].

Все виды обработок (одно-, двух- и трехмерная, т. е. по линии, по сечению и по объему) можно записать как следующее *соответствие между эхо-сигналом S и плотностью* (точнее, плотностью намагниченности) c [24, с. 27], [25]:

$$S(\mathbf{k}) = \int c(\mathbf{r}) e^{-i\mathbf{k}\mathbf{r}} d\mathbf{r}, \quad (1.53)$$

$$c(\mathbf{r}) = \int S(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} d\mathbf{k}, \quad (1.54)$$

где \mathbf{r} — радиус-вектор, а \mathbf{k} — волновой вектор, равный

$$\mathbf{k} = \gamma \int_0^t \mathbf{G}(t') dt',$$

причем \mathbf{G} — суммарный вектор градиентных полей, вообще говоря, зависящий от t . Таким образом, реконструкцию изображений можно трактовать как описание пространства $\{\mathbf{k}\}$ и последующий переход в пространство $\{\mathbf{r}\}$ посредством преобразования Фурье.

Отметим, что метод преобразования Фурье наиболее часто используется для решения задачи реконструкции ЯМР-изображений. Этот метод впервые сформулировали Кумар, Велти и Эрнст в 1975 г. [85]. Помимо него, используются также следующие методы: метод «чувствительной точки» Хиншоу [35], [74, т. 2, с. 147], методы линейного сканирования [74, т. 2, с. 147–149], методы «быстрой» визуализации [74, т. 2, с. 149–154], методы «объемной» визуализации [74, т. 2, с. 154–157] и др.

Примеры реконструкции изображений. На рис. 1.30а приведено изображение $c(x, y)$ сечения головного мозга пациента, страдающего болезнью Ходжкина, полученное с помощью схемы типа 1-й практической схемы при $T_E = 2T = 0.012$ сек, $\delta_z = 7$ мм, 256 точек в фазовом кодировании и 512 точек в частотном кодировании. На рис. 1.30б и в — вещественная и мнимая части двухмерного ПФ $\hat{c}(\omega_x, \omega_y)$ изображения $c(x, y)$ (по горизонтальной и вертикальной осям отложены частоты Фурье ω_x и ω_y) [74, т. 2, с. 141].

Влияние неоднородности полей на разрешающую способность томограмм. Пусть создано статическое поляризующее поле $B(x) = B_0 + \Delta B_0(x)$, где $B_0 = \text{const}$, а $\Delta B_0(x)$ — неоднородность, а также градиентное поле $G_x(x) = g_x x$. Тогда, как показано в [24, 25], неоднородность ΔB_0 ведет к тому, что в ЯМР-изображении будут разрешаться детали с расстоянием

$$\delta x \geq \frac{B_0}{g_x} \frac{\Delta B_0}{B_0}, \quad (1.55)$$

т. е. разрешающая способность томограмм ухудшается (δx увеличивается) пропорционально относительной неоднородности поляризующего поля $\Delta B_0/B_0$ и обратно пропорционально g_x . Можно

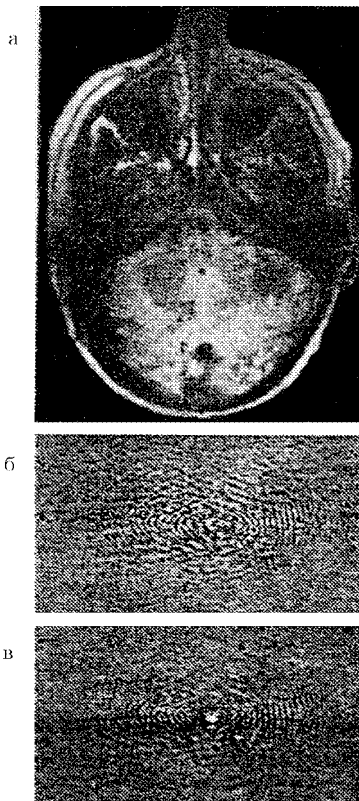


Рис. 1.30

записать оценку относительной неоднородности поляризующего поля, соответствующей заданному разрешению δx :

$$\frac{\Delta B_0}{B_0} \leq \frac{g_x}{B_0} \delta x. \quad (1.56)$$

Если же созданы поля B_0 и $g_x x + \Delta g_x x$, где Δg_x — отклонение от линейности градиентного x -поля, то относительное отклонение от линейности градиентного x -поля, соответствующее δx ,

$$\frac{\Delta g_x}{g_x} \leq \frac{\delta x}{x}. \quad (1.57)$$

Аналогичны оценки для $\Delta g_y/g_y$ и $\Delta g_z/g_z$.

Рассмотрен пример [24, 25]: Пусть $B_0 = 0.1$ Тл, $g_x = 5$ мТл/м, $x = 20$ см, требуемая разрешающая способность $\delta x = 0.2$ мм, тогда

$$\frac{\Delta B_0}{B_0} \leq 10^{-5}, \quad \frac{\Delta g_x}{g_x} \leq 10^{-3}, \quad (1.58)$$

т.е. для обеспечения высокой разрешающей способности томограмм (в доли мм) необходима высокая степень однородности поляризующего поля B_0 и линейности градиентных полей G_x , G_y , G_z .

Определение. *Рабочим объемом* называется та область зоны доступа томографа, в которой выполняются требования по разрешающей способности и относительным отклонениям полей типа (1.55)–(1.58).

В оценке (1.57) x — линейный размер рабочего объема.

Обычно диаметр рабочего объема не превышает 40–50% диаметра зоны доступа. Однако П.А. Галайдиным и А.И. Замятыми [25] разработана методика увеличения рабочего объема за счет введения корректирующих катушек.

Математический учет технических неоднородностей полей. Гладкие технические неоднородности полей (в отличие от локальных негладких полей релаксации) можно не компенсировать, как в методике Галайдина-Замятина или в задаче синтеза магнитного поля (см. дальше), а учесть математически, как это сделано в работе [96]. Рассматривается задача реконструкции изображения, решаемая по схеме типа 1-й практической схемы (ср. рис. 1.28) — см. рис. 1.31.

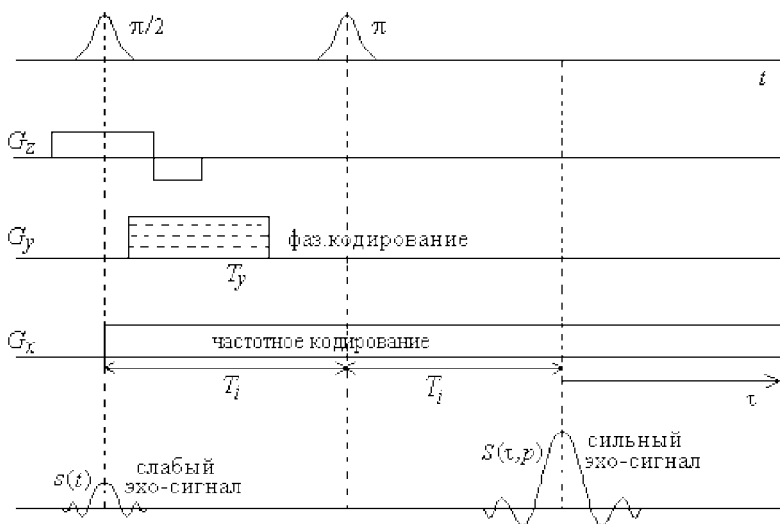


Рис. 1.31

Математически задача описывается следующим двумерным интегральным уравнением (ср. (1.49)) [60, 96, 100]:

$$A \iint_{-\infty}^{\infty} c(x, y) e^{i[\Omega(x, y)\tau + P(x, y)p]} dx dy = s(\tau, p), \quad (1.59)$$

$$-\infty < \tau, \quad p < \infty,$$

где $\Omega(x, y) = \gamma[B_0 + \Delta B(x, y) + G_x + \Delta G_x(x, y)]$, $P(x, y) = \gamma[G_y + \Delta G_y(x, y)]T_y$, причем T_y — продолжительность действия поля G_y , а $\Delta B(x, y)$, $\Delta G_x(x, y)$ и $\Delta G_y(x, y)$ — неоднородности полей (известные гладкие монотонные функции). Имеются в виду неоднородности, обусловленные, в первую очередь, техническими особенностями. Например, если статическое поле B создается одной катушкой-соленоидом с намоткой постоянной толщины, то на ее оси поле будет монотонно убывать от центра к ее краям.

Рассматривается ряд градиентных полей по y , равных pG_y , где $p \in (-\infty, \infty)$. Правая часть $s(\tau, p)$ есть совокупность эхо-сигналов в функции τ при каждом p . На практике $\tau \in [-\tau_{\max}, \tau_{\max}]$, $p \in [-p_{\max}, p_{\max}]$.

Решение уравнения (1.59) посредством двумерного ПФ имеет вид [96, 60, 100]:

$$c(x(\xi, \eta), y(\xi, \eta)) = Q(\xi, \eta) \iint_{-\infty}^{\infty} s(\tau, p) e^{-i(\xi\tau + \eta p)} d\tau dp, \quad (1.60)$$

где $Q(\xi, \eta) = 1/4\pi^2 A |J(\xi, \eta)|$, а $x = x(\xi, \eta)$ и $y = y(\xi, \eta)$ определяются из системы нелинейных уравнений (СНУ):

$$x = \frac{\xi - \omega_0}{\gamma g_x} - \frac{\Delta B(x, y) + \Delta G_x(x, y)}{g_x},$$

$$y = \frac{\eta}{\gamma g_y T_y} - \frac{\Delta G_y(x, y)}{g_y},$$

где $\omega_0 = \gamma B_0$, а

$$J(\xi, \eta) = \begin{vmatrix} \partial x / \partial \xi & \partial x / \partial \eta \\ \partial y / \partial \xi & \partial y / \partial \eta \end{vmatrix}$$

— якобиан преобразования.

Для повышения устойчивости решения (1.60) применим метод регуляризации Тихонова n -го порядка [60]. Решение (1.60) с регуляризацией будет иметь вид:

$$c_\alpha(x(\xi, \eta), y(\xi, \eta)) = Q(\xi, \eta) \iint_{-\infty}^{\infty} \frac{s(\tau, p)}{1 + \alpha M(\tau, p)} e^{-i(\xi\tau + \eta p)} d\tau dp, \quad (1.61)$$

где $M(\tau, p)$ — функция вида

$$M(\tau, p) = \tau^{2n} + qp^{2n}$$

или

$$M(\tau, p) = (\tau/T)^{2n} + (p/P)^{2n},$$

причем $n \in \mathbb{N}$, т. е. $n = 1, 2, 3, \dots$ (обычно $n = 1$), $\alpha > 0$ — параметр регуляризации, $q > 0$ — некоторый множитель такой, чтобы слагаемые τ^{2n} и qp^{2n} были одного порядка, а $T, P > 0$, в частности, $T = \tau_{\max}$, $P = p_{\max}$. Параметр регуляризации α может быть выбран различными способами, например, способом подбора контраста изображения (чем меньше α , тем выше контраст и наоборот — аналогично настройке контраста телеизображения) или способом невязки из уравнения [60]

$$\iint_{-\infty}^{\infty} |\tilde{s}(\tau, p)|^2 \left[\frac{\alpha M(\tau, p)}{1 + \alpha M(\tau, p)} \right]^2 d\tau dp = \delta^2$$

при условии

$$\iint_{-\infty}^{\infty} |\tilde{s}(\tau, p)|^2 d\tau dp > \delta^2,$$

где

$$\|\tilde{s} - s\|_{L_2}^2 \leq \delta^2,$$

т. е. δ — среднеквадратическая погрешность измерения эхо-сигнала $s(\tau, p)$, полагаемая известной (s — точный эхо-сигнал, а \tilde{s} — измеренный эхо-сигнал).

Решение примеров согласно (1.60), приведенное в [96], показывает, что можно математически учесть неоднородности полей. В этих примерах $\Delta B/B_0 \approx 3.6 \cdot 10^{-5}$, $\Delta G_x/g_x \approx 3.4$ мм, $\Delta G_y/g_y \approx 4.1$ мм или $\Delta g_x/g_x \approx 2.8 \cdot 10^{-2}$, $\Delta g_y/g_y \approx 3.3 \cdot 10^{-2}$. Однако данная методика может быть использована и в случае, когда $\Delta B/B_0$, $\Delta g_x/g_x$ и $\Delta g_y/g_y$ заметно больше этих значений. Главное, нужно знать функции $\Delta B(x, y)$, $\Delta G_x(x, y)$ и $\Delta G_y(x, y)$. А они могут быть измерены в томографе или рассчитаны по формулам [24, 25].

Что касается регуляризации, то как показывает решение примеров [60], использование регуляризации (см. (1.61)) понижает отношение $\Delta c_\alpha/c_\alpha$ (где Δc_α — погрешность решения c_α) в 2–3 раза, т. е. регуляризация повышает качество томограмм (снижает их зашумленность).

Синтез магнитного поля на оси катушки ЯМР-томографа. Рассматривается следующая задача ЯМР-томографии: определение распределения плотности тока (точнее, ампер-витков или магнитодвижущей силы) $J(a)$ вдоль обмотки цилиндрической катушки по заданному полю (напряженности) $H(z)$ на ее оси — *задача синтеза* магнитного поля на оси катушки. В случае катушки с бесконечно тонкой обмоткой (см. рис. 1.32)

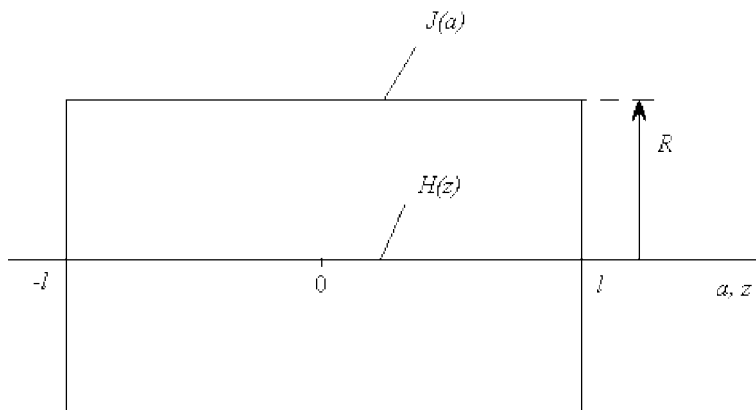


Рис. 1.32

имеет место интегральное уравнение Фредгольма I рода (в безразмерных переменных) [88]:

$$\int_{-s_0}^{s_0} K(x, s) J(s) ds = H(x), \quad -s_0 \leq x \leq s_0, \quad (1.62)$$

где

$$K(x, s) = K(x - s) = \frac{0.5}{\sqrt{[1 + (x - s)^2]^3}},$$

$s = a/R$, $x = z/R$, $s_0 = l/R$, R — радиус катушки, l — ее полудлина, a — расстояние вдоль обмотки катушки, z — расстояние от центра катушки на ее оси, $H(x)$ — заданное магнитное поле на оси катушки, $J(s)$ — искомый ток в изолированных друг от друга витках обмотки.

Если положить, например, $H(x) = \text{const}$, $x \in [-s_0, s_0]$, то решение уравнения (1.62) и его техническая реализация позволят создать (в принципе) постоянное статическое поляризующее поле на оси катушки.

Замечание. Нельзя говорить о соленоиде, как в работе [88], так как соленоид — это единый намотанный на катушку провод с $J(s) = \text{const}$, а следует говорить о катушке, например, следующего типа (см. рис. 1.33).

Имеется один общий источник напряжения U . От него делается N отводов с сопротивлениями $\rho(s) = U/J(s)$, где $J(s)$ — решение уравнения (1.62). Каждый отвод передает ток $J(s)$ только на один, соответствующий ему виток обмотки, изолированный от других витков. В результате мы получаем обмотку из N изолированных друг от друга витков, в каждом из которых течет

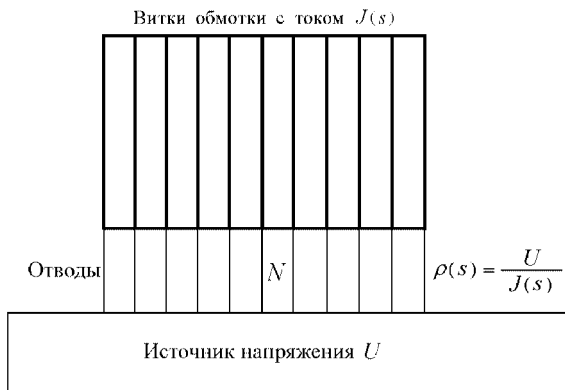


Рис. 1.33

свой ток $J(s)$. В идеале, когда $N \rightarrow \infty$, получим гладкие функции $J(s)$ и $H(x)$.

Можно предложить и другое техническое решение данного вопроса, например, в виде соленоида с единым намотанным на цилиндр проводом, имеющим на каждом витке (длиной $2\pi R$) свое сопротивление $\rho(s) = U/J(s)$, где U — подаваемое на провод напряжение.

Рассмотрим вопрос о решении уравнения (1.62) при $H(x) = \text{const}$. В работе [88] использован метод регуляризации Тихонова в соединении с методом квадратур (см. п. 8.1). Решен пример со следующими данными: $s_0 = 1$, шаг дискретизации $\Delta s = \Delta x = 0.1$, число шагов дискретизации $n = 20$ (число витков $N = n + 1 = 21$). Наиболее важным оказался вопрос выбора значения α . Дело в том, что поле $H(x)$ задается точно (без погрешностей) и один из наиболее эффективных способов выбора α — способ невязки (см. п. 8.1) для данной задачи практически не подходит. Поэтому в работе [88] выбран другой способ выбора α (назовем его *способом минимальной невязки*, или минимального отклонения). На рис. 1.34 даны результаты решения уравнения (1.62) методом регуляризации Тихонова и методом квадратур при различных значениях α согласно [88].

Видно, что с уменьшением α флуктуации, а значит, неустойчивость $J(s)$ возрастают, ток $J_\alpha(s)$ имеет как положительные, так и отрицательные значения, т.е. даже направление тока в витках различно. На рис. 1.35 даны процентные отклонения (невязки)

$$\delta_\alpha(x), \% = \left[\int_{-s_0}^{s_0} K(x, s) J_\alpha(s) ds - H(x) \right] / H(x) \cdot 100.$$

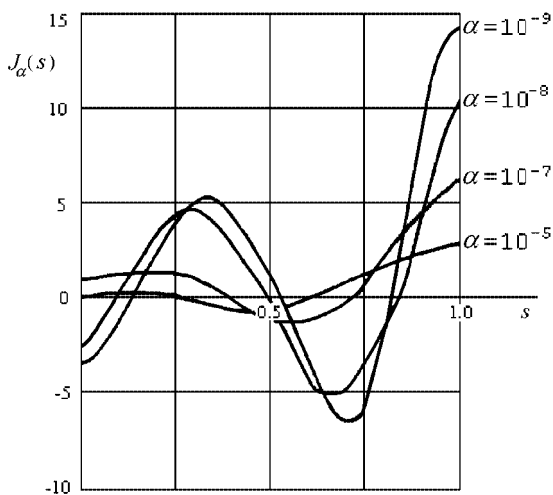


Рис. 1.34

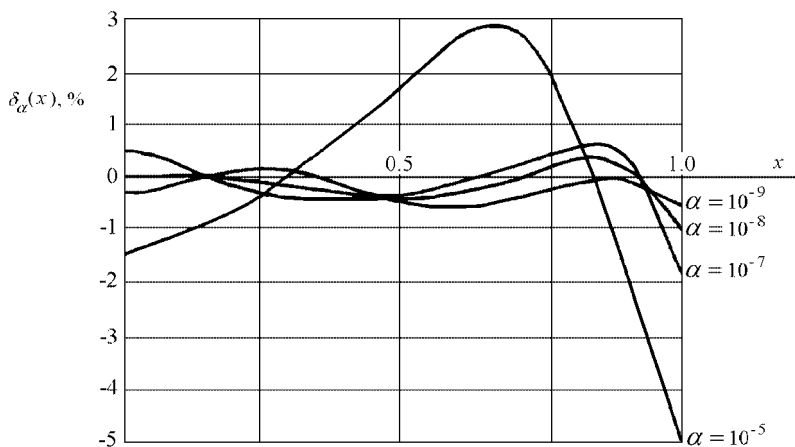


Рис. 1.35

Видим, что с уменьшением α невязка $\delta_\alpha(x)$ уменьшается, т.е. казалось бы, точность решения $J_\alpha(s)$ повышается. Однако это — хорошо известный эффект (см., например, [67, с.162]), когда при малых α решение получается в виде так называемой «пилы большой амплитуды» (т.е. решение очень неустойчиво и совершенно не похоже на точное решение), хотя при подстановке пилообразного решения в (1.62) получается небольшая невязка $\delta_\alpha(x)$ и тем

меньшая, чем меньше α . В статье же [88] сделан именно такой некорректный шаг: α выбирается исходя из условия:

$$\min_{\alpha} \max_{x \in [-s_0, s_0]} |\delta_{\alpha}(x)| . \quad (1.63)$$

Обозначим его через α_{\min} . В результате для α_{\min} получено очень малое значение: $\alpha_{\min} \approx 10^{-13}$. Этому соответствует крайне неустойчивое распределение тока $J_{\alpha_{\min}}(s)$ — решение уравнения (1.62) методом квадратур практически без регуляризации.

В работе [99] предложен более эффективный способ выбора α . Этот способ основан на следующих физико-технических предпосылках. Если $J(s) = \text{const}$, то $H(x)$ будет уменьшаться от центра ($x = 0$) к краям катушки ($x = \pm s_0$). Если же $J(s)$ монотонно и гладко возрастает от центра к краям, то при некотором законе возрастания (его нужно найти) будет $H(x) = \text{const}$. Поэтому в качестве способа выбора α при $H(x) = \text{const}$ взят следующий способ (назовем его *способом монотонного решения*): выбирается минимальное α , при котором $J_{\alpha}(s)$ монотонно возрастает от центра ($s = 0$) к краям катушки ($s = \pm s_0$). Обозначим такое α через $\alpha_{\text{мон}}$.

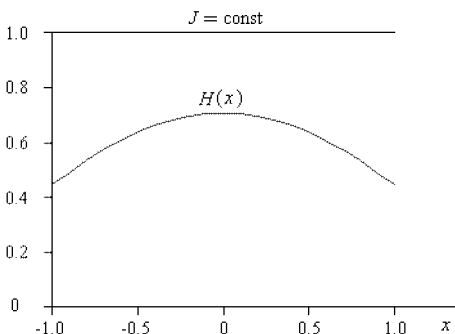


Рис. 1.36

Разработан пакет программ MFS (the Magnetic Field Synthesis) на Fortran MS, ver.5 и Fortran 90. С помощью него для данного примера найдено $\alpha_{\text{мон}}$ равным $5 \cdot 10^{-4}$. На рис.1.36 представлено поле $H(x)$ при $J = \text{const}$, а на рис.1.37 — численное решение $J_{\alpha}(s)$ при $\alpha = \alpha_{\text{мон}}$ и $N = 21$, а также поле $H_{\alpha}(x) = \int_{-s_0}^{s_0} K(x, s) J_{\alpha}(s) ds$.

Проанализируем результаты, отображенные на рис.1.37. Видим, что решение $J_{\alpha_{\text{мон}}}(s)$ примерно в 10 раз возрастает к краям катушки по сравнению с центром, что технически вполне реализуемо не в пример решению $J_{\alpha_{\min}}(s)$. При этом $|\delta_{\alpha_{\text{мон}}}(x)| \approx 0.1\%$, т.е. относительная неоднородность поля $H(x)$ равна $\Delta H/H \approx 10^{-3}$.

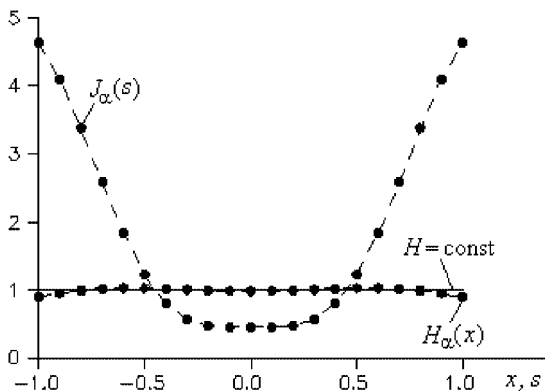


Рис. 1.37

Чтобы уменьшить это значение, нужно увеличить число витков n . Например, при $n \approx 40$ имеем $\Delta H/H \approx 10^{-4}$ и т. д.

В работах [88, 99] рассмотрен также случай катушки с конечной (постоянной) толщиной изолированных витков — получается уравнение типа (1.62), но с иным ядром $K(x, s)$. Все вышеизложенное в принципе подходит к этому случаю.

Наконец, в работе [99] рассмотрен случай соленоида (т. е. единого провода) с однородным током в проводе $J(s) = J = \text{const}$ и переменной (искомой) толщиной намотки провода $y = y(s)$, другими словами, с переменным (искомым) распределением ампер-витков $J \cdot y(s)$ вдоль цилиндра катушки (см. рис. 1.38).

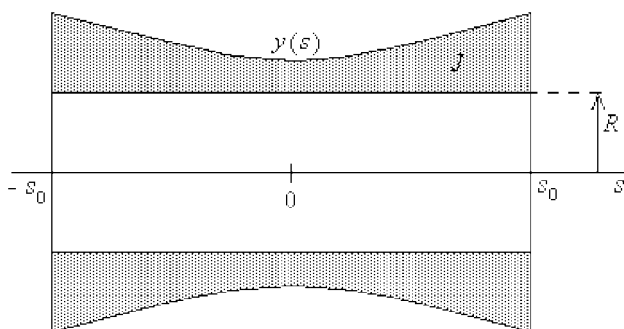


Рис. 1.38

В этом случае имеет место нелинейное интегральное уравнение Урысона I рода:

$$J \int_{-s_0}^{s_0} K[x, s, t(s)] ds = H(x), \quad -s_0 \leq x \leq s_0, \quad (1.64)$$

где

$$K[x, s, t(s)] = K[x - s, t(s)] = \frac{1}{\sqrt{1 + (x - s)^2}} - \frac{1 + t(s)}{\sqrt{[1 + t(s)]^2 + (x - s)^2}} + \frac{\ln \left\{ 1 + t(s) + \sqrt{[1 + t(s)]^2 + (x - s)^2} \right\}}{1 + \sqrt{1 + (x - s)^2}},$$

$t(s) = y(s)/R$ — искомая функция.

Данный подход (см. рис. 1.38) напоминает использование катушек седловидной формы для обеспечения однородности поля [74, т. 2, с. 211].

Области применения ЯМР-томографии. Основная область применения ЯМР-томографии, как и РТ, — медицина [74, т. 2]. При этом ЯМР-томография «вытесняет» РТ из медицины, поскольку рентгеновское излучение гораздо вреднее магнитных полей для пациентов и обслуживающего персонала. Переходить же на более мягкие (т.е. низкочастотные) рентгеновские лучи нежелательно, так как это ведет к понижению разрешающей способности томограмм.

ЯМР-томография также применяется в физике, химии, биологии, технике.

Контрольные задания и вопросы

1. В чем заключается эффект ЯМР (объяснить понятия: спин, магнитный момент, вектор ядерной намагниченности, прецессия, гиромангнитное отношение, ядерная восприимчивость, частота Ларморовой прецессии)?

2. Объясните с помощью уравнения Лармора характер движения магнитного момента протона в постоянном магнитном поле.

3. Что такое вектор ядерной намагниченности ансамбля протонов?

4. Опишите движение магнитного момента протона в постоянном и переменном магнитных полях (прецессия, нутация, расстройка, резонанс).

5. Запишите и истолкуйте уравнения Блоха, описывающие переориентацию вектора \mathbf{M} в результате изменения поля \mathbf{H} с учетом продольной и поперечной релаксации.

6. Выведите решения (1.39) или (1.40) уравнений Блоха (1.38) (задание повышенной трудности).

7. Что такое эхо-сигнал? Условие его появления и причина его затухания.
8. Для чего нужны: постоянное магнитное поле B_0 , переменное электро-магнитное поле с частотой $\Omega = \omega = \gamma B_0$, короткий импульс резонансной частоты?
9. Информацию о каких свойствах вещества несут: мощность эхо-сигнала $s(t)$, частота ларморовой прецессии ω и время продольной релаксации T_1 ?
10. Объясните $\pi/2$ - и π -импульсы Карра-Парселла, эхо-сигналы, расфазирование. Приведите аналогию с бегунами.
11. Что такое расфазирование протонов? Его причины.
12. Градиентные поля по z , x и y . Для чего они вводятся?
13. Реконструкция ЯМР-изображений. Определение и основной принцип реконструкции.
14. Описать 1-ю и 2-ю практические схемы реконструкции ЯМР-изображений.
15. Что такое фазовое и частотное кодирование пространства?
16. Разрешающая способность томограмм. Что это и от чего она зависит?
17. Опишите технически и математически реконструкцию изображений с учетом неоднородности полей $\Delta B(x, y)$, $\Delta G_x(x, y)$, $\Delta G_y(x, y)$.
18. Сформулируйте задачу синтеза магнитного поля на оси катушки ЯМР-томографа (три варианта обмотки).
19. Два способа выбора параметра регуляризации α в задаче синтеза. В чем они заключаются и каковы их особенности?
20. Перечислите области применения ЯМР-томографии.

Глава 2

НЕКОТОРЫЕ ОБРАТНЫЕ ЗАДАЧИ ОПТИКИ И
СПЕКТРОСКОПИИ

В этой главе рассмотрим задачу обработки искаженных изображений как одну из обратных задач оптики [4, 55, 84, 101]. При этом под *изображением* будем подразумевать фотоснимок человека, текста, объекта, природы, телескопический снимок или оптико-электронное изображение космического объекта и т. д. Однако для определенности под изображениями будем подразумевать фотографии. Будем полагать, что выполнена предварительная обработка изображений, а именно, устранены царапины на снимке, подобрана его контрастность и т. п. (операции, не требующие математической обработки). А мы остановимся на наиболее трудной задаче — на обработке (восстановлении, реконструкции) изображений, искаженных в результате смаза (сдвига, смещения) и дефокусировки.

2.1 Восстановление смазанных изображений

Рассмотрим данную задачу на примере смазанного фотоснимка [4, 55, 70, 101, 105, 106].

Постановка задачи. Считаем, что фотографируемый объект (для простоты полагаемый плоским) и фотопленка фотоаппарата расположены параллельно апертуре линзы фотоаппарата по разные стороны от нее на расстояниях соответственно f_1 и f_2 от линзы, причем [46, 77]

$$\frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f}, \quad (2.1)$$

где f — фокусное расстояние линзы и $f_1 \geq f$. В результате на фотопленке возникнет перевернутое изображение (см. рис. 2.1).

Введем в плоскости объекта прямоугольную систему координат $\xi'o'\eta'$, а на фотопленке $\xi\eta$. Возьмем на объекте некоторую точку $A'(\xi', \eta')$ с интенсивностью излучения $w'(\xi', \eta')$. Лучи, исходящие из нее и прошедшие через линзу, пересекутся в некоторой точке $A(\xi, \eta)$. Из подобия треугольников $A'CO'$ и ACO следует:

$$\frac{\overrightarrow{A'O'}}{f_1} = \frac{\overrightarrow{OA}}{f_2}$$

или в проекциях:

$$\frac{\xi'}{f_1} = -\frac{\xi}{f_2}, \quad \frac{\eta'}{f_1} = -\frac{\eta}{f_2}.$$

В результате точка $A(\xi, \eta)$ на фотопленке (соответствующая точке $A'(\xi', \eta')$ на объекте) будет иметь следующие интенсивность w и

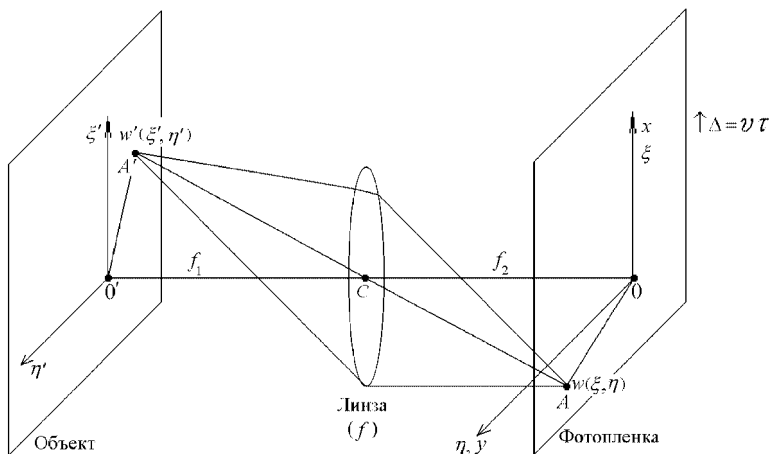


Рис. 2.1

координаты ξ, η (прямая задача):

$$w(\xi, \eta) = w'(\xi', \eta'), \quad \xi = -\xi'/q, \quad \eta = -\eta'/q, \quad (2.2)$$

где $q = f_1/f_2$, причем f_2 определяется как

$$f_2 = \left(\frac{1}{f} - \frac{1}{f_1} \right)^{-1}. \quad (2.3)$$

Таким образом, каждой точке A' на объекте будет соответствовать точка A на фотопленке с той же интенсивностью, но с уменьшенными (и перевернутыми) в q раз координатами (см. (2.2)).

Пример. Пусть $f_1 = 5$ м, $f = 4$ см, тогда $f_2 = 4.04$ см (согласно (2.3)), $q = 123$, т. е. изображение будет в 123 раза меньше объекта.

По фотоснимку можно восстановить объект (обратная задача):

$$w'(\xi', \eta') = w(\xi, \eta), \quad \xi' = -q\xi, \quad \eta' = -q\eta. \quad (2.4)$$

Из (2.1) имеем формулу для f_1 :

$$f_1 = \left(\frac{1}{f} - \frac{1}{f_2} \right)^{-1}. \quad (2.5)$$

Далее полагаем, что за время экспозиции τ фотопленка совершила прямолинейный и равномерный сдвиг со скоростью v вдоль оси ξ , т. е. на величину $\Delta = v\tau$. Можно также считать, что сдвиг совершила не фотопленка, а объект (движущаяся цель) на величину $-\Delta \cdot q$. В результате изображение на фотопленке будет смазанным (сдвинутым, смещенным) вдоль ξ — см. рис. 2.2а [55] и 2.3а [105, 106].

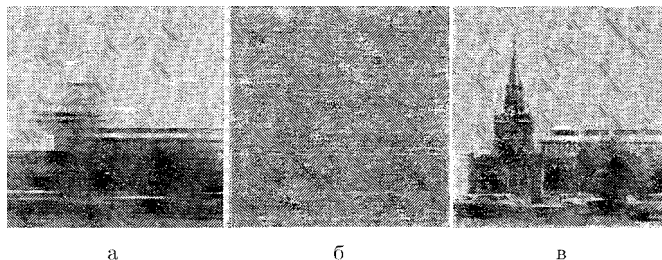


Рис. 2.2

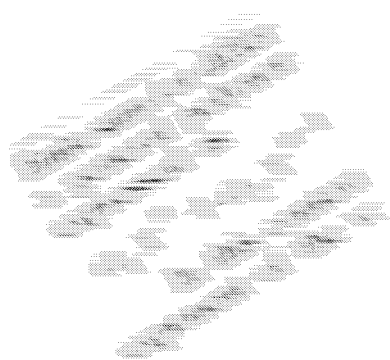
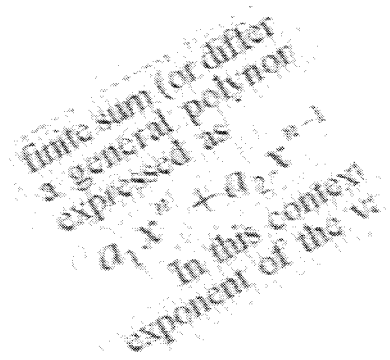
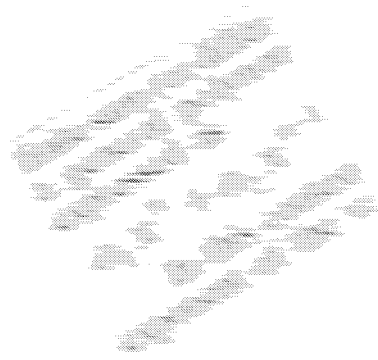
а) Смазанное изображение ($\Delta=20$)б) Восстановленное изображение при $\alpha = 0$ в) Восстановленное изображение при $\alpha = 2.5 \cdot 10^{-3}$ г) Восстановленное изображение при $\alpha = 0.1$

Рис. 2.3

Вывод интегрального уравнения. Опишем математически данную задачу. Введем наряду с абсолютно неподвижной системой координат $\xi\eta$ на рис. 2.1 систему координат xoy , неподвижную

относительно пленки и совпадающую с $\xi\eta$ лишь в начальный момент ($\tau = 0$). На точку (x, y) фотопленки за время τ будут проецироваться точки A с абсциссами от $\xi = x$ до $\xi = x + \Delta$ и с интенсивностями $w(\xi, y)$, т. е. результирующая интенсивность (обозначим ее через g) в некоторой точке (x, y) фотопленки будет равна сумме (точнее, интегралу) интенсивностей $w(\xi, y)$, $\xi \in [x, x + \Delta]$:

$$g(x, y) = \frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi. \quad (2.6)$$

В (2.6) перед интегралом поставлен множитель $1/\Delta$. Объясним, почему это сделано. При $\Delta \rightarrow 0$ (отсутствие смаза) в правой части (2.6) имеем неопределенность типа $\frac{0}{0}$, так как интеграл $\int \rightarrow 0$. Раскрывая ее по правилу Лопиталя, имеем:

$$g(x, y)|_{\Delta \rightarrow 0} = \frac{\partial[\int_x^{x+\Delta} w(\xi, y) d\xi]/\partial\Delta}{\partial\Delta/\partial\Delta} \Big|_{\Delta=0} = w(x + \Delta, y)|_{\Delta=0} = w(x, y), \quad (2.7)$$

как и должно быть при $\Delta \rightarrow 0$. Если же в (2.6) перед интегралом поставить другой множитель, например, $1/\Delta^2$, то равенство (2.7) не будет иметь места.

Запишем (2.6) иначе:

$$\frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi = g(x, y). \quad (2.8)$$

Соотношение (2.8) является основным в задаче восстановления смазанных изображений. В нем $g(x, y)$ — измеренная интенсивность на фотопленке (смазанное изображение — см. рис. 2.2а и 2.3а), Δ — величина смаза, полагаемая известной, а $w(\xi, y)$ — истинная неискаженная интенсивность — то изображение, которое было бы на фотопленке в отсутствие сдвига (смаза).

Соотношение (2.8) есть одномерное интегральное уравнение типа Вольтерры I рода относительно $w(\xi, y)$ при каждом фиксированном значении y , играющем роль параметра, другими словами, (2.8) есть *совокупность одномерных уравнений*.

Отметим, что в ряде работ ([4, 70] и др.) рассматриваются более сложные формулировки данной задачи: неравномерный и/или непрямолинейный сдвиг пленки (или объекта), непараллельность плоскостей объекта и пленки и т. д.

Отметим также, что значение Δ (или v) часто априори неизвестно и его обычно определяют путем подбора на основе визуальной оценки решения $w(\xi, y)$ [4, с.164]. Что же касается направления смаза (вдоль которого направляется ось x), то его можно определить по штрихам на снимке (см. рис. 2.2а и 2.3а).

Итак, в принципе, правильно выбрав направление оси x (вдоль смаза) и значение смаза Δ , можно, решив уравнение (2.8) (точнее, совокупность уравнений), восстановить неискаженный снимок $w(x, y)$ по искаженному снимку $g(x, y)$. А искаженный снимок может быть, например, очень старым снимком, на котором заложена очень ценная, но нераспознаваемая информация (фотография знатного человека, исторического задания, важного текста и т. д.), и только математическим путем (с использованием компьютера) можно будет извлечь эту информацию.

Учет характеристики чувствительности пленки. Чтобы обработка снимка была более точной, нужно учесть характеристику чувствительности пленки.

Определение. *Характеристикой чувствительности (ХЧ) пленки* называется зависимость реакции пленки p от падающего на нее излучения g (в некоторую точку пленки).

Обычно ХЧ пленки имеет следующий вид (см. рис. 2.4):

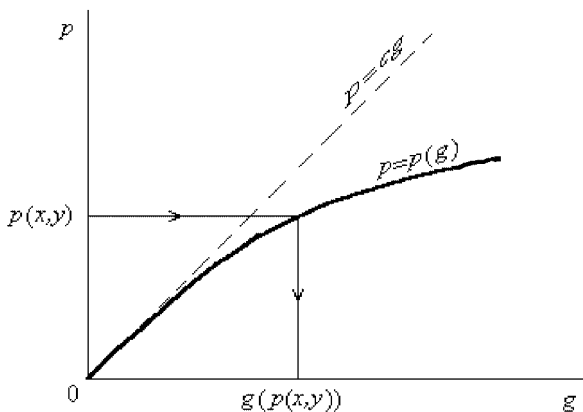


Рис. 2.4

ХЧ пленки напоминает яркостную характеристику дисплея (см. рис. 1.14).

Реакция пленки p выражается в виде степени ее почернения, если это негатив (и в виде «побеления», если это позитив), которая, в свою очередь, отображается количеством прореагировавшего серебра. При малых падающих интенсивностях g реакция пленки p будет линейна: $p = cg$, где c — некоторый коэффициент, но по мере увеличения g будет все в большей степени сказываться дефицит серебра и реакция p все сильнее будет отличаться от $p = cg$ (см. рис. 2.4). Поэтому в действительности в каждой точке пленки (x, y) отображено не значение $g(x, y)$, а некоторое меньшее значение $p(x, y)$. Однако, зная ХЧ пленки $p = p(g)$ (она

должна быть заложена в паспорте пленки или может быть получена от завода-изготовителя или же измерена экспериментально), можно по значениям $p(x, y)$ найти значения $g(x, y)$, т. е. $g(p(x, y))$ (идя по стрелкам на рис. 2.4), другими словами, используя функцию, обратную ХЧ. Тогда уравнение (2.8) можно записать в виде:

$$\frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi = g(p(x, y)). \quad (2.9)$$

Использование вместо p более точных значений g должно повысить контрастность обработки пленки.

Мы будем пользоваться как выражением (2.9), так и (2.8), а также еще более простым выражением:

$$\frac{1}{\Delta} \int_x^{x+\Delta} w(\xi) d\xi = g(x), \quad (2.10)$$

справедливым при каждом y .

Рассмотрим

Методы решения уравнений (2.8), (2.9), (2.10).

Первый метод — *метод дифференцирования* [63, 105, 106]. Продифференцируем по x соотношение (2.9), получим:

$$\frac{1}{\Delta} [w(x + \Delta, y) - w(x, y)] = \frac{\partial}{\partial x} g(p(x, y)),$$

откуда

$$w(x + \Delta, y) = w(x, y) + \Delta \frac{\partial}{\partial x} g(p(x, y)). \quad (2.11)$$

Соотношение (2.11) есть *рекуррентное соотношение*, означающее, что по предыдущему значению $w(x, y)$ можно определить последующее значение $w(x + \Delta, y)$. Практически это означает следующее. Пусть априори известно $w(x_0, y)$ при некотором $x = x_0$ (например, в местах «провалов» в изображении на рис. 2.3а можно положить $w(x_0, y) = 0$). Тогда с помощью (2.11) можно найти $w(x_0 + \Delta, y)$, $w(x_0 + 2\Delta, y)$, $w(x_0 + 3\Delta, y)$, ...

Запишем (2.11) иначе:

$$w(x, y) = w(x + \Delta, y) - \Delta \frac{\partial}{\partial x} g(p(x, y))$$

или (заменив x на $x - \Delta$)

$$w(x - \Delta, y) = w(x, y) - \Delta \frac{\partial}{\partial x} g(p(x - \Delta, y)). \quad (2.12)$$

Тогда, используя (2.12), можно найти $w(x_0 - \Delta, y)$, $w(x_0 - 2\Delta, y)$, ... При этом производную $\frac{\partial}{\partial x} g(p(x, y))$ следует определять путем *численного дифференцирования*, например, по формуле:

$$\frac{\partial}{\partial x} g(p(x, y)) \cong \frac{g(p(x + h, y)) - g(p(x - h, y))}{2h}, \quad (2.13)$$

где h — шаг численного дифференцирования (вообще говоря, отличный от Δ). Однако задача численного дифференцирования функции, измеренной с погрешностями (а $g(p(x, y))$ измерена, конечно, с погрешностями) является некорректной (неустойчивой) и необходимо предварительно сгладить функцию $g(p(x, y))$, например, с помощью аппроксимирующих сплайнов (см. п. 8.4) или использовать метод регуляризации Тихонова для устойчивого дифференцирования зашумленных функций [67, с. 18–19, 158–159].

Второй метод — *метод приведения к интегральному уравнению типа свертки* [4, 70, 105, 106]. Этот метод является более эффективным и распространенным. В данном методе уравнение (2.10) (а также (2.8) и (2.9)) преобразуется к стандартной форме. Дело в том, что запись (2.10) не является стандартной записью интегрального уравнения Вольтерры или Фредгольма (см. п. 5.1), поэтому непосредственно к уравнению (2.10) нельзя применять традиционные методы решения интегральных уравнений (см. п. 7.2). Уравнение (2.10) запишем в виде

$$\int_x^{x+\Delta} \frac{1}{\Delta} w(\xi) d\xi = g(x) \quad (2.14)$$

и приведем к стандартному виду — одномерному интегральному уравнению Фредгольма I рода типа свертки

$$\int_{-\infty}^{\infty} k(x - \xi) w(\xi) d\xi = g(x), \quad -\infty < x < \infty, \quad (2.15)$$

где

$$k(x - \xi) = \begin{cases} 1/\Delta, & -\Delta \leq x - \xi \leq 0, \\ 0, & \text{иначе,} \end{cases} \quad (2.16)$$

или

$$k(x) = \begin{cases} 1/\Delta, & -\Delta \leq x \leq 0, \\ 0, & \text{иначе.} \end{cases} \quad (2.17)$$

Уравнение (2.15) обычно решается *методом преобразования Фурье* (ПФ) (см. п. 7.2). Согласно нему, решение имеет вид (ОПФ):

$$w(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) e^{-i\omega\xi} d\omega, \quad (2.18)$$

где ПФ (спектр) решения

$$W(\omega) = \frac{G(\omega)}{K(\omega)}, \quad (2.19)$$

а $G(\omega)$ и $K(\omega)$ — ПФ (спектры) правой части $g(x)$ и ядра $k(x)$ уравнения (2.15), равные

$$G(\omega) = \int_{-\infty}^{\infty} g(x) e^{i\omega x} dx, \quad (2.20)$$

$$K(\omega) = \int_{-\infty}^{\infty} k(x) e^{i\omega x} dx. \quad (2.21)$$

Ядро $k(x)$ имеет аналитическое представление (см. (2.17)), поэтому $K(\omega)$ может быть найдено аналитически согласно (2.21). Имеем:

$$K(\omega) = \int_{-\infty}^{\infty} k(x) e^{i\omega x} dx = \frac{1}{\Delta} \int_{-\Delta}^0 e^{i\omega x} dx = \frac{\sin(\omega\Delta)}{\omega\Delta} + \frac{\cos(\omega\Delta) - 1}{\omega\Delta} i. \quad (2.22)$$

А $G(\omega)$ может быть найдено численно по некоторой стандартной программе вычисления ДПФ (обычно в виде БПФ).

Проанализируем поведение спектров $G(\omega)$, $K(\omega)$ и $W(\omega)$. При $|\omega| \rightarrow \infty$ спектр $G(\omega)$ правой части $g(x)$ с учетом ее зашумленности стремится к некоторой константе (уровню «белого шума»), а спектр $K(\omega)$, точнее, огибающая $K(\omega)$ при $|\omega| \rightarrow \infty$ ведет себя как $\sim 1/\omega$ (см. (2.22)), т.е. $K(\omega) \rightarrow 0$ при $|\omega| \rightarrow \infty$ и, следовательно, $W(\omega) \rightarrow \infty$ при $|\omega| \rightarrow \infty$ и интеграл в (2.18) расходится. Таким образом, решение согласно (2.18)–(2.22) хотя и выглядит изящно, но строго говоря, не может быть использовано на практике из-за неустойчивости. Это демонстрируют рис. 2.26 и 2.36, на которых представлены решения согласно (2.18)–(2.22), причем интегралы в (2.18), (2.20), (2.21) заменялись суммами, т.е. НПФ заменялось на ДПФ (и на БПФ), в результате чего имело место усечение спектра частот (см. п. 7.2), поэтому выражения (2.18) и (2.19) в бесконечность не обращались, но, тем не менее, имела место сильная неустойчивость решения.

Итак, на рис. 2.26 и 2.36 представлены результаты восстановления смазанных изображений методом ПФ без регуляризации. Видим, что изображения не улучшились, а наоборот ухудшились. Поэтому необходимо использование некоторого устойчивого метода. В качестве такового рассмотрим метод регуляризации Тихонова.

Использование метода регуляризации Тихонова [4, 70, 105, 106]. Решение уравнения (2.15) методами ПФ и регуляризации Тихонова имеем вид (см. п. 8.1):

$$w_\alpha(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_\alpha(\omega) e^{-i\omega\xi} d\omega, \quad (2.23)$$

где

$$W_\alpha(\omega) = \frac{K(-\omega)G(\omega)}{L(\omega) + \alpha M(\omega)}, \quad (2.24)$$

$L(\omega) = |K(\omega)|^2 = K(\omega)K(-\omega)$, $M(\omega) = \omega^2$, $\alpha > 0$ — параметр регуляризации.

Для выбора α разработан ряд способов, например, способ невязки (см. [4, 70, 71] и п. 8.1). Однако для задачи восстановления изображения более эффективен способ выбора α , называемый *способом подбора* [63, 105, 106]. Он заключается в следующем. С уменьшением α контраст восстанавливаемого изображения увеличивается, но уменьшается устойчивость, а с увеличением α , наоборот, контраст изображения уменьшается, а его устойчивость увеличивается. Следовательно, должно быть выбрано некоторое умеренное значение α . При этом при выборе α нужно привлекать не столько математические, сколько физиологические критерии. Этот эффект аналогичен настройке контраста телеизображения.

Формально способ подбора состоит в следующем. Задается ряд значений α :

$$\alpha = \alpha_1, \alpha_2, \dots, \alpha_n.$$

При каждом α вычисляется $w_\alpha(\xi) \equiv w_\alpha(\xi, y)$ согласно (2.23), выводится на дисплей решение $w_\alpha(\xi, y)$ и выбирается то α , при котором решение $w_\alpha(\xi, y)$ является наиболее разрешимым, устойчивым и/или правдоподобным, как, например, на рис. 2.2в и 2.3в.

В работе [63] для решения уравнения типа свертки типа (2.15) использован не метод ПФ, а метод преобразования Хартли (см. [9] и п. 6.3). Это позволяет осуществлять обработку вещественных функций k , w и g в области вещественных чисел в отличие от ПФ, отображающего вещественные функции в комплексную область, и тем самым экономить компьютерную память и время (приблизительно в 2 раза).

О программах. Задача реконструкции смазанного изображения сводится к многократному решению интегрального уравнения (2.8), (2.9) или (2.15) методом регуляризации Тихонова согласно формулам (2.23), (2.24) при некотором значении α , причем нужно получить столько решений $w_\alpha(\xi) \equiv w_\alpha(\xi, y)$, сколько задано значений y . Если дисплей имеет $m \times n$ пикселей (например, 640×480), то это означает, что нужно получить $\leq n$ решений. Можно использовать следующие программы на Фортране: РТИКР [71, с. 124–130, 178–179], CONV3 [19, с. 384–385], CONVOL [61, пакет CONF] и др., которые, в свою очередь, обращаются к программам вычисления БПФ — можно использовать программы FFT [56], FTF1C [71, с. 183–184, 190–192] и др. Еще более эффективным является пакет программ IMAGE на языке Visual C++ [102, 105, 106]. Пакет IMAGE позволяет решать как прямую задачу (моделирование смазанного изображения), так и обратную задачу (реконструкция изображения методом регуляризации

Тихонова с подбором значений α и Δ). Пакет IMAGE предназначен для обработки как черно-белых изображений (в этом случае для получения большой градации яркостей используется серый цвет — смесь красного, зеленого и синего цветов в одинаковой пропорции), так и цветных изображений (в этом случае используется раздельная обработка в трех цветах и последующее наложение трех изображений).

Контрольные задания и вопросы

1. Сформулируйте задачу восстановления смазанных изображений.

2. Используя формулы (2.5) и (2.3), покажите, что $f_1 \geq f$ и $f_2 \geq f$.

3. Используя формулу (2.3), вычислите f_2 (расстояние от линзы до фотопленки) при $f_1 = \infty$, $f_1 = 3f$ и $f_1 = f$.

4. Используя общую формулу дифференцирования по параметру:

$$\frac{\partial}{\partial x} \int_{\varphi(x)}^{\psi(x)} f(\xi, x, y) d\xi = \int_{\varphi(x)}^{\psi(x)} \frac{\partial}{\partial x} f(\xi, x, y) d\xi + f(\psi(x), x, y) \psi'(x) - f(\varphi(x), x, y) \varphi'(x),$$

запишите выражение

$$\text{для } \frac{\partial}{\partial \Delta} \int_x^{x+\Delta} w(\xi, y) d\xi \text{ (см. (2.6), (2.7))}$$

$$\text{и для } \frac{\partial}{\partial x} \left[\frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi \right] \text{ (см. (2.9)).}$$

5. Классифицируйте уравнение (2.8) (одномерное или двумерное и т. д.).

6. Что такое ХЧ пленки?

7. Обоснуйте переход от (2.14) к (2.15)–(2.16).

8. Дайте более подробный вывод формулы (2.22).

9. В чем причина неустойчивости решения (2.18)–(2.22)?

10. В чем причина устойчивости решения (2.23)–(2.24)?

11. Сформулируйте способ подбора параметра регуляризации α . Опишите поведение решения $w_\alpha(\xi)$ при уменьшении/увеличении α .

2.2. Восстановление дефокусированных изображений

Рассмотрим еще одну обратную задачу оптики — задачу обработки (восстановления, реконструкции) дефокусированных изображений (фотоснимков человека, текста, космического объекта и т. д.) [4, 15, 16, 28, 62, 70, 82, 84, 105, 106].

Эту задачу рассмотрим на примере дефокусированного фотоснимка.

Постановка задачи. Данная задача имеет много общего с предыдущей задачей (восстановление смазанных изображений), но имеет и существенные отличия. Считаем, что снимаемый объект (полагаемый плоским) и фотопленка расположены параллельно линзе (по разные стороны от нее) на расстояниях от линзы f_1 и $f_2 + \delta$ соответственно, где δ — погрешность фокусировки изображения (установления расстояния f_2) — см. рис. 2.5.

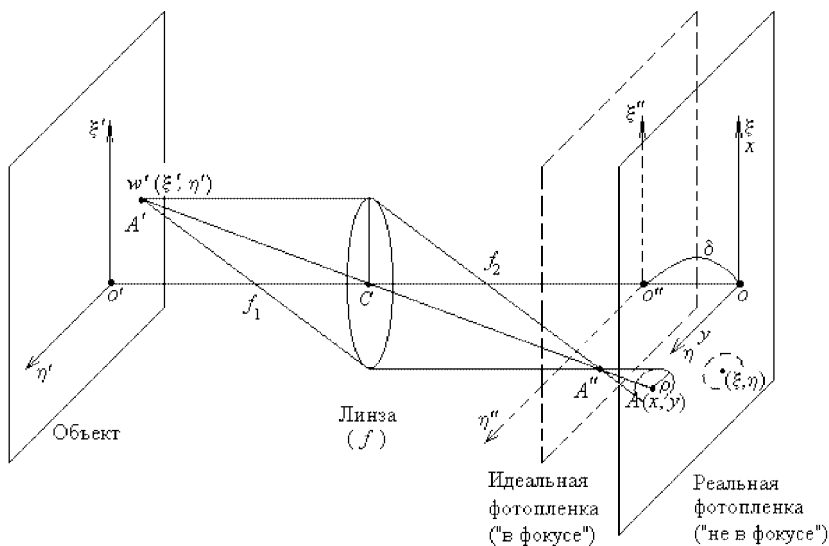


Рис. 2.5

При этом, как и в предыдущей задаче, имеет место соотношение (2.1), 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 $\xi \eta$, 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. (2.2)), причем f_2 определяется согласно (2.3).

На реальной же фотопленке точка A' отобразится не в точку, а в дифракционный круг радиуса

$$\rho = \frac{a\delta}{f_2}, \quad (2.25)$$

где a — радиус апертуры линзы, с центром в точке $A(x, y)$, причем

$$x = -\frac{f_2 + \delta}{f_1} \xi', \quad y = -\frac{f_2 + \delta}{f_1} \eta' \quad (2.26)$$

(ср. (2.2)).

Вывод основного соотношения. Опишем математически задачу дефокусировки. Рассмотрим, помимо дифракционного круга с центром в точке $A(x, y)$, также некоторый другой круг с центром в точке (ξ, η) (см. рис. 2.5). Радиусы этих (а также других) кругов одинаковы и равны ρ (см. (2.25)), а площади кругов равны $S = \pi\rho^2$. В результате некоторая интенсивность $w(\xi, \eta)$, соответствующая точке (ξ, η) , будет «размазана» по кругу радиуса ρ и площади $S = \pi\rho^2$ с плотностью интенсивности $w(\xi, \eta)/\pi\rho^2$ (постоянной, в первом приближении, в пределах дифракционного круга).

Интенсивность в точке $A(x, y)$ будет результатом суммирования (интегрирования) по всем 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

$$\sqrt{(x - \xi)^2 + (y - \eta)^2} \leq \rho. \quad (2.27)$$

В результате интенсивность в точке $A(x, y)$ будет равна

$$g(x, y) = \iint_{\sqrt{(x-\xi)^2 + (y-\eta)^2} \leq \rho} \frac{w(\xi, \eta)}{\pi\rho^2} d\xi d\eta. \quad (2.28)$$

Соотношение (2.28) является основным в задаче реконструкции дефокусированных изображений.



Рис. 2.6

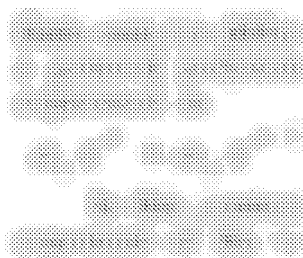
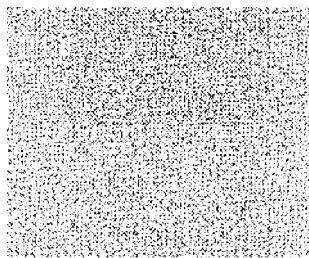
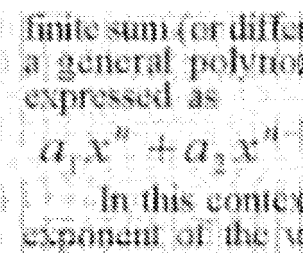
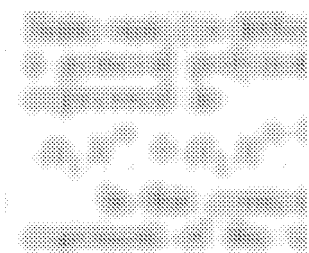
а) Дефокусированное изображение ($\rho=10$)б) Восстановленное изображение при $\alpha=0$ в) Восстановленное изображение при $\alpha=5 \cdot 10^{-4}$ г) Восстановленное изображение при $\alpha=5 \cdot 10^{-2}$

Рис. 2.7

На рис. 2.6а и 2.7а приведены примеры [15, с. 171], [105, 106] дефокусированных изображений $g(x, y)$.

Стандартная форма уравнения. Запишем (2.28) в виде

$$\iint_{\sqrt{(x-\xi)^2+(y-\eta)^2} \leq \rho} \frac{w(\xi, \eta)}{\pi \rho^2} d\xi d\eta = g(x, y). \quad (2.29)$$

Соотношение (2.29) есть двухмерное интегральное уравнение I рода относительно $w(\xi, \eta)$. Однако оно записано не в стандартной форме. Преобразуем его к стандартной форме. Запишем (2.29) в виде [62]:

$$\iint_{-\infty}^{\infty} k(x-\xi, y-\eta) w(\xi, \eta) d\xi d\eta = g(x, y), \quad -\infty < x, y < \infty, \quad (2.30)$$

где

$$k(x-\xi, y-\eta) = \begin{cases} \frac{1}{\pi \rho^2}, & \sqrt{(x-\xi)^2 + (y-\eta)^2} \leq \rho, \\ 0, & \text{иначе,} \end{cases} \quad (2.31)$$

или

$$k(x, y) = \begin{cases} \frac{1}{\pi \rho^2}, & \sqrt{x^2 + y^2} \leq \rho, \\ 0, & \text{иначе.} \end{cases} \quad (2.32)$$

Соотношение (2.30) есть двумерное интегральное уравнение Фредгольма I рода типа свертки. В нем $g(x, y)$ — интенсивность в плоскости реальной (расположенной «не в фокусе») фотопленки, которая может быть записана как $g(p(x, y))$, где $p(g)$ — характеристика чувствительности (ХЧ) фотопленки (см. рис. 2.4), а $g(p)$ — ее обратная характеристика, причем $p(x, y)$ — измеренное потемнение пленки с учетом ее ХЧ, меньшее, чем $g(x, y)$. Далее, $k(x, y)$ есть ядро интегрального уравнения, причем ρ определяется согласно (2.25), где a и f_2 известны, а δ (или ρ) может быть определено путем подбора (аналогично Δ в предыдущей задаче). Ядро интегрального уравнения $k(x, y)$ называется *функцией рассеяния точки* [81, с. 34]. Наконец, $w(\xi, \eta)$ есть искомая интенсивность, которая была бы на снимке при $\delta = 0$ (неискаженное изображение в плоскости $\xi''\eta''$).

Отметим, что в работе [70] рассмотрена также задача дефокусировки для случая непараллельности плоскости объекта и плоскости пленки.

После решения уравнения (2.30) можно восстановить исходное изображение в плоскости объекта (*обратная задача*, ср. (2.4)):

$$w'(\xi', \eta') = w(\xi, \eta), \quad \xi' = -\frac{f_1}{f_2 + \delta} \xi, \quad \eta' = -\frac{f_1}{f_2 + \delta} \eta. \quad (2.33)$$

Решение методом двумерного ПФ. Уравнение (2.30) как двумерное интегральное уравнение Фредгольма I рода типа свертки может быть решено методом двумерного преобразования Фурье — ПФ (инверсная фильтрация). Решение записывается в виде двумерного ОПФ (см. п. 7.2):

$$w(\xi, \eta) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} W(\omega_1, \omega_2) e^{-i(\omega_1 \xi + \omega_2 \eta)} d\omega_1 d\omega_2, \quad (2.34)$$

где ПФ (спектр) решения

$$W(\omega_1, \omega_2) = \frac{G(\omega_1, \omega_2)}{K(\omega_1, \omega_2)}, \quad (2.35)$$

а $G(\omega_1, \omega_2)$ и $K(\omega_1, \omega_2)$ — преобразования Фурье (спектры) правой части $g(x, y)$ и ядра интегрального уравнения (2.30), равные

$$G(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} g(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy, \quad (2.36)$$

$$K(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} k(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy. \quad (2.37)$$

Ядро $k(x, y)$ выражается в виде аналитической формулы (2.32), поэтому в принципе $K(\omega_1, \omega_2)$ может быть найдено аналитически согласно (2.37) (предоставляем читателю возможность сделать это). А $G(\omega_1, \omega_2)$ (а также $K(\omega_1, \omega_2)$) должно быть найдено численно по стандартной программе двумерного ДПФ (обычно в виде БПФ).

Однако задача решения уравнения (2.30) является некорректной [4, 15, 19, 23, 45, 48, 67]. Это связано с тем, что функция $p(x, y)$, а значит и $g(p(x, y))$ измеряется с погрешностью и это ведет к чрезвычайно большим (в принципе, бесконечно большим) погрешностям решения $w(\xi, \eta)$. Поэтому формулы (2.34)–(2.37) не годятся для устойчивого решения уравнения (2.30). Используем метод регуляризации Тихонова для этой цели. Отметим, что в работе [4, с.164–177] используются также методы итеративной регуляризации, а в работе [62] — метод регуляризации Тихонова в совокупности с преобразованием Хартли (вместо ПФ).

Применение метода регуляризации Тихонова. Решение уравнения (2.30) методами двумерного ПФ и регуляризации Тихонова имеет вид (см. п. 8.1):

$$w_\alpha(\xi, \eta) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} W_\alpha(\omega_1, \omega_2) e^{-i(\omega_1\xi + \omega_2\eta)} d\omega_1 d\omega_2, \quad (2.38)$$

где

$$W_\alpha(\omega_1, \omega_2) = \frac{K(-\omega_1, -\omega_2) G(\omega_1, \omega_2)}{L(\omega_1, \omega_2) + \alpha M(\omega_1, \omega_2)}, \quad (2.39)$$

$L(\omega_1, \omega_2) = |K(\omega_1, \omega_2)|^2 = K(\omega_1, \omega_2) K(-\omega_1, -\omega_2)$, $M(\omega_1, \omega_2) = (\omega_1^2 + \omega_2^2)^2$, $\alpha > 0$ — параметр регуляризации. Решение (2.38)–(2.39) при правильно выбранных значениях α и δ (или ρ) обладает устойчивостью и достаточной разрешающей способностью.

Обычно для выбора α используется способ невязки (см. п. 8.1). Однако для задачи реконструкции дефокусированных изображений (как и для задачи реконструкции смазанных изображений) более эффективен способ подбора [62]. Алгоритмически он состоит в следующем. Задаем ряд значений α :

$$\alpha_1 < \alpha_2 < \alpha_3 < \dots < \alpha_n,$$

причем $\alpha_{\min} = \alpha_1$ и $\alpha_{\max} = \alpha_n$ задаем на основе дополнительной информации о решении $w(\xi, \eta)$, например, используя опыт обработки других дефокусированных изображений. При каждом значении α рассчитываем решение $w_\alpha(\xi, \eta)$ согласно (2.38)–(2.39) и выводим его, как изображение, на экран компьютера. На рис. 2.6 в качестве примера приведены: на рис. 2.6а — исходное дефокусированное изображение $g(x, y)$ (точнее, $p(x, y)$), на рис. 2.6б —

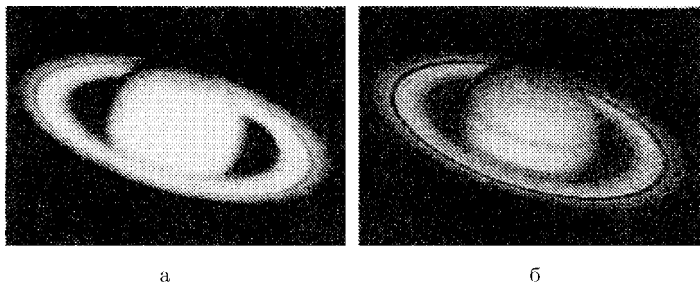


Рис. 2.8

решение $w_\alpha(\xi, \eta)$ при заниженном α , а именно $\alpha_1 = 10^{-5}$ (решение с повышенным контрастом), на рис. 2.6в — решение $w_\alpha(\xi, \eta)$ при большем α , а именно, $\alpha_2 = 10^{-3}$ (контраст уменьшился и устойчивость повысилась), а на рис. 2.6г — решение $w_\alpha(\xi, \eta)$ при еще большем (завышенном) α , а именно, $\alpha_3 = 10^{-2}$ (устойчивость еще повысилась, но контраст, а значит и разрешающая способность, снизились). Видим, что наиболее приемлемое значение α — это некоторое «умеренное» значение, обеспечивающее умеренный контраст и умеренную устойчивость. В данном случае это $\alpha = \alpha_2 = 10^{-3}$ (см. рис. 2.6в). Данный способ выбора α аналогичен настройке (выбору) контраста телеизображения, когда используются не столько математические, сколько физиологические критерии. Аналогичные результаты приведены на рис. 2.7 [105, 106], полученные с помощью пакета программ IMAGE [102].

На рис. 2.8 и 2.9 представлены еще более показательные примеры. Дело в том, что в 1990 г. в космос был запущен американский космический телескоп «Хаббл» (КТХ) [82] диаметром 2.4 м с целью получения снимков космических объектов с исключительно высоким разрешением (из-за отсутствия искажающего влияния атмосферы на орбите). Однако из-за ошибки радиуса кривизны зеркала КТХ (при его изготовлении) из космоса стали поступать дефокусированные снимки (типа рис. 2.8а и 2.9а — снимки Сатурна и пылевого облака звезды η Киля).

Применение же математической (и компьютерной) обработки позволило восстановить неискаженные изображения (см. рис. 2.8б и 2.9б), однако при этом был использован не метод типа метода регуляризации Тихонова, а метод типа усечения спектра частот и сглаживающих окон.

О программах. Задача реконструкции дефокусированных изображений сводится к методу двумерного ПФ и методу регуляризации Тихонова согласно формулам (2.36)–(2.39). Можно использовать следующие программы на Фортране: PTITR [71, с. 130–136, 185–186], CON2 [61, пакет CONF] и др., которые используют программы двумерного ПФ (можно использовать программу FTFTC

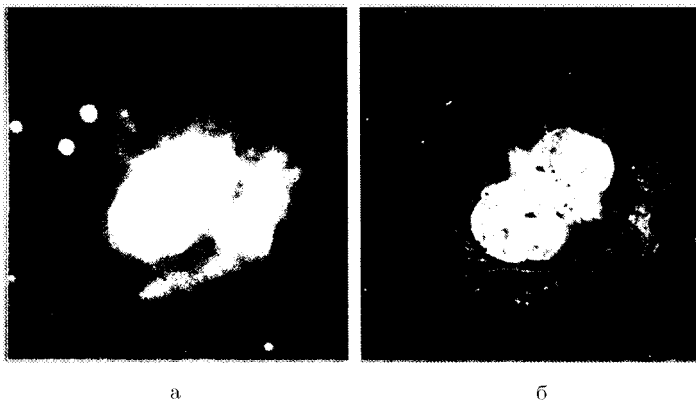


Рис. 2.9

[71, с. 190] и др.). Можно воспользоваться также пакетом программ IMAGE [102] на языке Visual C++ (решение прямой и обратной задач, использование серого цвета для обработки черно-белых изображений и раздельная обработка в трех цветах цветных изображений).

Контрольные задания и вопросы

1. Сформулируйте задачу восстановления дефокусированных изображений.
2. Докажите равенства (2.25), (2.26) и (2.33).
3. Классифицируйте уравнения (2.29) и (2.30) (одномерные или двумерные и т. д.).
4. К чему придет левая часть уравнения (2.30) при $\rho \rightarrow 0$?
5. Используя (2.32), выведите аналитическое выражение для $K(\omega_1, \omega_2)$ согласно (2.37) (задание повышенной трудности).
6. В чем причина некорректности (неустойчивости) решения (2.34)–(2.35) ?
7. В чем причина устойчивости решения (2.38)–(2.39) ?
8. Сформулируйте способ подбора параметра регуляризации α . Как ведет себя решение $w_\alpha(\xi, \eta)$ при уменьшении/увеличении α ?
9. Опишите сходства и различия задач восстановления смазанных и дефокусированных изображений (исходные данные, искомые решения, размерности задач, типы уравнений и т. д.).

2.3. Обратные задачи спектроскопии

Спектральный анализ и приборы. Для качественного и количественного исследования веществ широко используется *спектральный анализ*. Он основан на изучении спектров излучения

(испускания), поглощения, отражения, комбинационного рассеяния света и люминесценции. Для разложения излучения в спектр и его регистрации используются *оптические спектральные приборы*. Такой прибор состоит из трех основных частей: осветительной (необязательная часть), спектральной и регистрирующей. В зависимости от способа регистрации спектра различают следующие приборы: *спектроскопы* (с визуальной регистрацией), *спектрографы* (с фоторегистрацией), *спектрометры* (например, интерферометр Фабри-Перо) и *спектрофотометры* (с фотоэлектрической регистрацией).

Области применения спектрального анализа. Это — физика (изучение спектров газов, жидкостей, металлов и плазмы, исследование атомной сенсibilизированной флуоресценции смесей и паров металлов по контурам спектральных линий), астрофизика (изучение спектров звезд, планет, галактик, туманностей, комет, квазаров), металлургия (определение по спектру состояния расплавленного металла), химия (определение по спектру химического состава вещества), геофизика (разведка руд, минералов) и т. д. Отметим еще такое применение спектрального анализа, как определение магнитного (или электрического) поля по сверхтонкой структуре линии на основе эффекта Зеемана (или Штарка). Говоря далее о спектральном анализе, мы будем иметь в виду, в основном, спектры излучения и их изучение с помощью спектрометров, хотя нижеизложенное справедливо и для иных типов спектров и приборов.

Под *спектром* будет подразумевать зависимость интенсивности излучения от частоты ν . При этом используется следующая терминология: спектральный контур, спектральный профиль, контур спектра, профиль спектра.

Типы спектров. По виду спектры бывают [46], [53, с. 712–713] *непрерывные*, или *сплошные* (примеры: спектры расплавленного металла, Солнца и т. д.) — см. рис. 2.10, 2.11, 2.13–2.15, *дискретные*, или *линейчатые*, состоящие из отдельных спектральных линий, соответствующих дискретным значениям частоты ν (примеры: спектры атомарного водорода, натрия, меди [53, с. 528]) — см. рис. 2.16, 2.17 и *полосатые*, состоящие из отдельных полос, каждая из которых охватывает некоторый интервал частот ν (пример: спектр испускания паров иода [53, с. 528]). Строго говоря, отдельная спектральная линия также не соответствует вполне определенному значению ν , поскольку имеет, во-первых, минимальную (естественную, радиационную) ширину, обусловленную квантовыми эффектами, во-вторых, ширину, обусловленную эффектами Доплера (тепловое уширение), Зеемана (магнитное уширение), Штарка (электрическое уширение) и т. д. Однако если линия имеет лишь естественную ширину, то в большинстве исследований линию считают дискретной (монокроматической). Пример:

линии спектров межзвездных туманностей; они имеют практически лишь естественную ширину, так как тепловые, магнитные и электрические эффекты чрезвычайно малы.

Типы спектрального анализа. Спектральный анализ можно разделить на широкополосный и узкополосный. *Широкополосная спектрометрия* — это изучение спектра в широкой области частот, например, изучение спектра звезды во всем видимом диапазоне (от красного до фиолетового) — см. рис. 2.10.

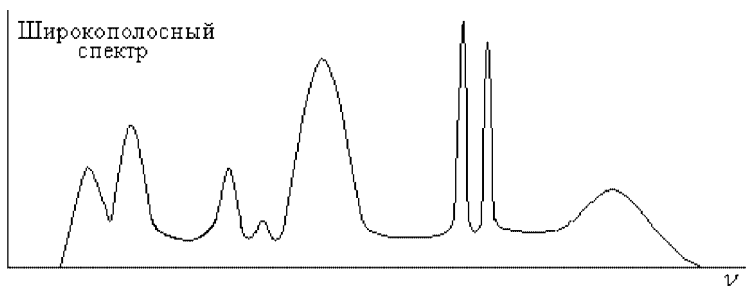


Рис. 2.10

Узкополосная спектрометрия — это изучение спектра в узкой полосе частот, например, изучение сверхтонкой структуры мессбауэровской линии, обусловленной магнитными или электрическими полями и тепловыми эффектами [53, с. 407] — см. рис. 2.11.



Рис. 2.11

Однако деление на широко- и узкополосную спектрометрию зачастую условно.

Экспериментальный спектр. Введем следующее

Определение. *Аппаратной функцией* — $A\Phi$ (или *функцией пропускания, частотной характеристикой* — $\mathcal{C}\mathcal{X}$, *спектральной чувствительностью* — $\mathcal{C}\mathcal{Ч}$) спектрометра [53, с. 704] $K(\nu, \nu')$ называется реакция спектрометра (в виде измеренной интенсивности) на дискретную линию единичной интенсивности и частоты ν' при настройке спектрометра на частоту ν — см. рис. 2.12.

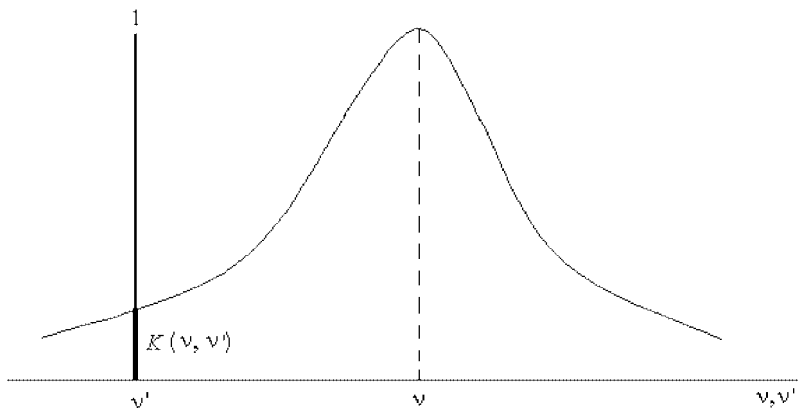


Рис. 2.12

Пример $A\Phi$, аппроксимированной аналитической формулой в виде гауссианы:

$$K(\nu, \nu') = \exp\left(-\frac{(\nu - \nu')^2}{2(a - b\nu)^2}\right)$$

с максимумом при $\nu' = \nu$ и с эффективной шириной, обратно пропорциональной ν ($a > 0$ и $b > 0$ — некоторые константы).

Если форма $K(\nu, \nu')$ не меняется при изменении ν , то $K(\nu, \nu') = K(\nu - \nu')$, т.е. $A\Phi$ (или $\mathcal{C}\mathcal{Ч}$) является разностной функцией.

$A\Phi$ (или $\mathcal{C}\mathcal{Ч}$) спектрометра полагается известной. Она должна быть представлена в паспорте спектрометра или определена, например, с помощью одночастотного лазера, стабилизированного по провалу Лэмба [44].

Будем полагать, что экспериментальный (измеренный) спектр (обозначим его через $u(\nu)$) — это спектр, измеренный путем сканирования по частоте реальным спектрометром, а именно, с учетом его $\mathcal{C}\mathcal{Ч}$ $K(\nu, \nu')$ и при наличии различных шумов (погрешностей измерений и т.д.), а истинный спектр (обозначим его через $z(\nu)$) — это спектр, который был бы измерен при бесконечно узкой $\mathcal{C}\mathcal{Ч}$ и в отсутствие шумов (но при наличии магнитных и электрических полей, тепловых движений атомов, самопоглощения и т.д.).

Измеренный спектр $u(\nu)$ отличается от истинного $z(\nu)$. Это проявляется, во-первых, в большей сглаженности $u(\nu)$ по сравнению с $z(\nu)$ (неразрешены близкие линии, заглажена микроструктура спектра) и, во-вторых, в зашумленности $u(\nu)$ (слабые линии «тонут» в шуме). На рис. 2.13а изображен экспериментальный спектр без шума $u(\nu)$, а на рис. 2.13б — экспериментальный зашумленный спектр

$$\tilde{u}(\nu) = u(\nu) + \delta u(\nu) + U, \quad (2.40)$$

где $\delta u(\nu)$ — случайная шумовая компонента (обусловленная погрешностями измерений и внешними помехами), а U — детерминированная шумовая компонента (фон), полагаемая постоянной.

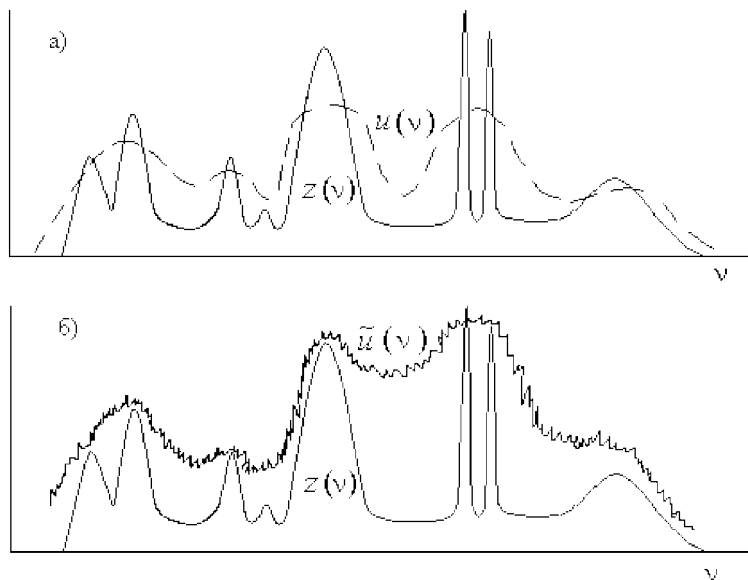


Рис. 2.13

Задача редукции к идеальному спектральному прибору.

Естественно возникает следующая задача: по экспериментальному спектру $\tilde{u}(\nu)$ и СЧ $K(\nu, \nu')$ путем математической обработки восстановить истинный спектр $z(\nu)$. Успешное решение этой задачи позволит повысить разрешающую способность спектрометра, а значит, качество спектрального анализа (например, более точно определить фазовое состояние расплавленного металла в домне или химический состав космического объекта — туманности, кометы, поверхности звезды и т. д.).

Задача определения истинного спектра $z(\nu)$ по экспериментальному спектру $\tilde{u}(\nu)$ и СЧ $K(\nu, \nu')$ называется *задачей редукции*

(или приведения) *профиля спектра к идеальному спектральному прибору* [44]. Отметим, что восстанавливаемая функция $z(\nu)$ может обладать сложной структурой (см. рис. 2.13, 2.14), причем ширина ее отдельных пиков часто соизмерима с шириной СЧ или даже меньше ее (см. рис. 2.13). Это — одна из обратных задач спектроскопии [17, с.130 — 133]. Рассмотрим две ее формулировки — применительно к непрерывному и дискретному спектру.

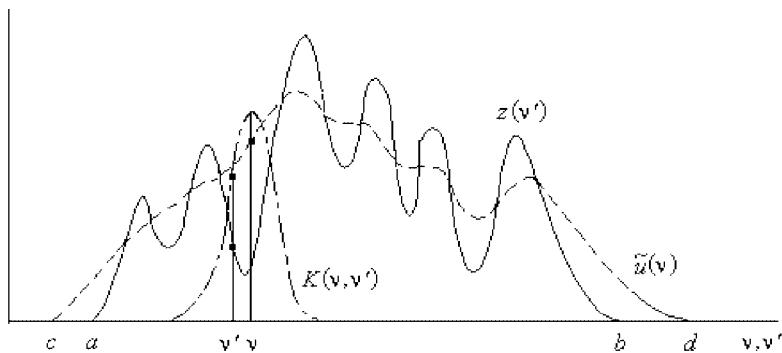


Рис. 2.14

Непрерывный спектр. Рассмотрим случай *непрерывного спектра*, когда искомым спектром $z(\nu')$, $\nu' \in [a, b]$, есть кусочно-непрерывная функция (см. рис. 2.13 или 2.14). Этот случай имеет место обычно для веществ с повышенной плотностью (жидкий металл) или при изучении сверхтонкой структуры линии, когда пределы $[a, b]$ узкие. В этом случае измеренное значение интенсивности $u(\nu)$ при настройке спектрометра на частоту ν равно интегралу по всем интенсивностям $z(\nu')$ с весовой функцией, равной K , т. е.

$$u(\nu) = \int_a^b z(\nu') K(\nu, \nu') d\nu', \quad (2.41)$$

где $[a, b]$ — пределы изменения ν' . Из (2.41), варьируя ν (т. е. выполняя сканирование с помощью сканирующей системы) и учитывая зашумленность $u(\nu)$, имеем:

$$\int_a^b K(\nu, \nu') z(\nu') d\nu' = \tilde{u}(\nu), \quad c \leq \nu \leq d, \quad (2.42)$$

где $[c, d]$ — пределы изменения ν (более широкие, чем $[a, b]$) — см. рис. 2.15.

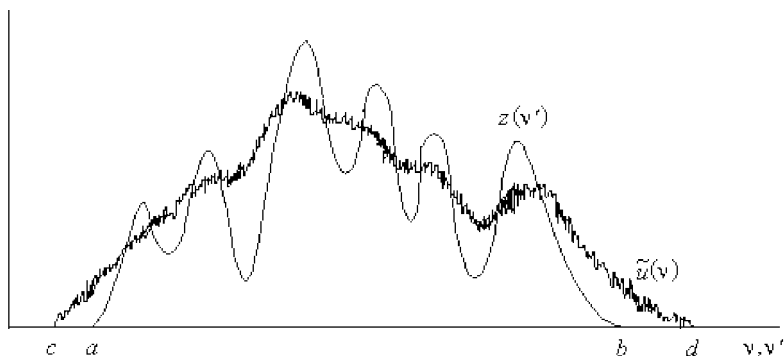


Рис. 2.15

В соотношении (2.42) известны (измерены или заданы) $\tilde{u}(\nu)$, $K(\nu, \nu')$, a , b , c , d , а $z(\nu')$ является искомой. Соотношение (2.42) есть интегральное уравнение Фредгольма I рода относительно $z(\nu')$. Если $K(\nu, \nu') = K(\nu - \nu')$, то (2.42) обычно записывают в виде:

$$\int_0^{\infty} K(\nu - \nu') z(\nu') d\nu' = \tilde{u}(\nu), \quad 0 \leq \nu < \infty. \quad (2.43)$$

Соотношение (2.43) есть интегральное уравнение Фредгольма I рода типа свертки на полуоси. Задача решения уравнений (2.42) и (2.43) является некорректной (см. п. 7.1). Поэтому для их устойчивого решения необходимо применение регулярных (устойчивых) методов, например, метода регуляризации Тихонова (см. п. 8.1) аналогично решению уже рассмотренных (одномерных) уравнений (1.16), (1.62), (2.15).

Если пределы $[a, b]$ в (2.42) положены широкими, то мы имеем дело с обработкой спектра в широкой полосе частот. А если пределы $[a, b]$ взяты узкими (вплоть до рассмотрения структуры лишь одной линии), то это будет задача обработки спектра в узкой полосе частот (или задача восстановления тонкой структуры линии).

Результаты решения примеров мы приводить не будем — они типа изображенных на рис. 3.15, 8.3, 8.4, 8.6 (см. дальше).

Дискретный спектр. Теперь рассмотрим случай *дискретного (линейчатого) спектра*, когда искомый спектр $z(\nu')$ состоит из отдельных дискретных (мономатических) линий (см. дальше рис. 2.21), характеризующихся их частотами и амплитудами. Такой спектр имеют, например, туманности и низкотемпературная, в частности, газоразрядная плазма, а вот в рекламных люминесцентных лампах и лампах «дневного света» имеют место процессы

ионизации и рекомбинации, в результате чего их спектр — это набор полос, т. е. нечто среднее между непрерывным и дискретным спектром — такой спектр называется *полосатым* [53, с. 712–713] (см. дальше рис. 2.22). В случае дискретного спектра измеренное значение интенсивности $u(\nu)$ при настройке спектрометра на частоту ν равно сумме интенсивностей (амплитуд) всех линий с весовой функцией K (см. рис. 2.16),

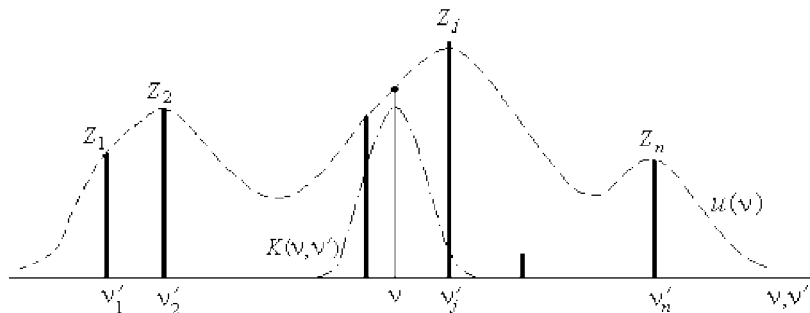


Рис. 2.16

т. е.

$$u(\nu) = \sum_{j=1}^n z_j K(\nu, \nu'_j), \quad (2.44)$$

где z_j — амплитуда (интенсивность) j -й линии, ν'_j — ее частота, n — число линий. Из (2.44), варьируя ν и учитывая зашумленность измерений, получим:

$$\sum_{j=1}^n K(\nu_i, \nu'_j) z_j + F = \tilde{u}(\nu_i), \quad i = \overline{1, m}, \quad c \leq \nu_i \leq d, \quad (2.45)$$

где ν_i — дискретный отсчет ν , m — число таких отсчетов, $[c, d]$ — границы отсчетов, $\tilde{u}(\nu_i) = u(\nu_i) + \delta u(\nu_i)$, δu — случайная погрешность, F — детерминированная погрешность (фон) (см. рис. 2.17).

В (2.45) известны (измерены или заданы) $\tilde{u}(\nu_i)$, $K(\nu_i, \nu'_j)$, ν_i , c , d , m , а искомыми являются z_j , ν'_j , n , F (амплитуды и частоты линий, их число, а также детерминированная составляющая шумов). Соотношение (2.45) есть система линейно-нелинейных уравнений (СЛНУ) (нест андартный термин), поскольку часть неизвестных (z_j и F) входит линейно, а часть (ν'_j) — нелинейно.

Отметим некоторые особенности записи (2.45). Неизвестное F специально выделено в отдельное слагаемое. Во-первых, как показало решение модельных примеров [59, ч. III], это повышает

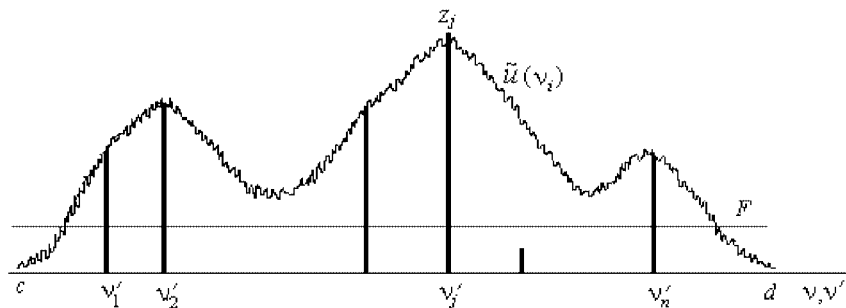


Рис. 2.17

точность решения СЛНУ, а во-вторых, определение F , как фона, также полезно, как и определение z_j и ν'_j линий. Кроме того, введена дискретизация по ν . Это сделано для того, чтобы использовать известные методы решения СЧУ и СЛНУ.

Решение СЛНУ. Рассмотрим вопрос о решении СЛНУ (2.45) (см. также п. 3.2). Система уравнений (2.45) может рассматриваться как система нелинейных уравнений (СНУ) относительно $2n$ неизвестных z_j и ν'_j (а также F) при некотором n . Ее можно решать известными методами решения СЧУ без ограничений на решение: методами градиента, Ньютона, хорд и др. [5, 11, 71] или методами решения СЧУ с ограничениями на решение (методами нелинейного программирования): проекции градиента, оврагов и др. [80]. При этом более эффективно применение методов нелинейного программирования, так как на неизвестные z_j, ν'_j и F мы можем наложить следующие ограничения:

$$z_j > 0, \quad a \leq \nu'_j \leq b, \quad F \geq 0, \quad (2.46)$$

где $[a, b]$ — некоторая область, более узкая, чем $[c, d]$.

Однако эти методы не учитывают специфики системы (2.45), кроме того, они оставляют открытым вопрос о n . Поэтому решение системы (2.45) будет более эффективным (потребуется меньше компьютерного времени и памяти, будет меньше вероятность ложных корней нелинейной системы и т. д.), если мы учтем специфику системы (2.45), а именно, то, что половина неизвестных (z_j) в нее входит линейно, а половина (ν'_j) — нелинейно (если не считать F). Для решения подобных СЛНУ весьма эффективен, например, метод Прони [38], однако он подходит лишь для СЛНУ с матрицей Вандермонда (когда $K(\nu_i, \nu'_j)$ изменяется вдоль строки по геометрической прогрессии), а матрица в (2.45) таковой, вообще говоря, не является. Используется еще алгоритм Пиблза-Берковича [50], однако он весьма неточен, а также алгоритм Фальковича-Коновалова [76], но это слишком громоздкий

алгоритм. Поэтому для решения (2.45) можно воспользоваться алгоритмом интегральной аппроксимации [59, ч. I, III].

Краткое изложение алгоритма интегральной аппроксимации. В этом алгоритме:

1. Решается интегральное уравнение (2.42) или (2.43) методом регуляризации Тихонова с заниженным значением параметра регуляризации α . Получается решение $z_\alpha(\nu')$.

2. В решении $z_\alpha(\nu')$ выделяются $L \leq N$ наиболее мощных максимумов, где N задается на основе дополнительной информации так, чтобы $N \geq n$, и фиксируются их углы $\tilde{\nu}'_j$, $j = \overline{1, L}$.

3. Решается уточняющая СЛАУ

$$\sum_{j=1}^L K(\nu_i, \tilde{\nu}'_j) \tilde{z}_j + \tilde{F} = \tilde{u}(\nu_i), \quad i = \overline{1, m}, \quad c \leq \nu_i \leq d, \quad (2.47)$$

относительно \tilde{z}_j и \tilde{F} .

4. Оставляются лишь те \tilde{z}_j и \tilde{F} , для которых справедливо:

$$\tilde{z}_j \geq Z, \quad j = \overline{1, k}, \quad \tilde{F} > 0, \quad (2.48)$$

где $Z > 0$ — некоторый априори заданный барьер, а $k \leq L$ — количество \tilde{z}_j , преодолевших барьер Z . При этом барьер Z можно определить, используя соотношение [72, с. 126]:

$$Z = \sigma_u \sqrt{-2 \ln F_{\text{лт}}}, \quad (2.49)$$

где σ_u — среднеквадратическое значение случайной составляющей помехи $\delta u(\nu)$, полагаемое известным, а $F_{\text{лт}} \in [0, 1]$ — задаваемая условная вероятность ложной тревоги.

Достоинством этого алгоритма является то, что наиболее трудная часть задачи — определение нелинейно входящих значений ν'_j , а также n — их числа — решается линейно, а именно, путем решения линейного интегрального уравнения (2.42) или (2.43).

Модельный пример. На рис. 2.18 представлен следующий модельный пример (типа примера 2 из [59, ч. III]): 1 — линейчатый спектр, состоящий из 6 линий, 2 — экспериментальный (измеренный) спектр $u(\nu)$ без погрешностей, 3 — зашумленный экспериментальный (измеренный) спектр $\tilde{u}(\nu)$ — см. (2.40), причем $U = 0.3$, $\sigma_u = 0.2$, 4 — разностная АФ (СЧ) $K(\nu)$.

Видим, что в измеренном спектре близкие линии (две слева и три справа) неразрешены. Задача редукции решалась с использованием алгоритма интегральной аппроксимации. При этом число дискретных отсчетов при решении интегрального уравнения (2.42) полагалось равным $m = 481$, $a = c$, $b = d$, $N = 10$, $\alpha = 10^{-4}$, барьер Z определялся по формуле (2.49) при $F_{\text{лт}} = 10^{-4}$.

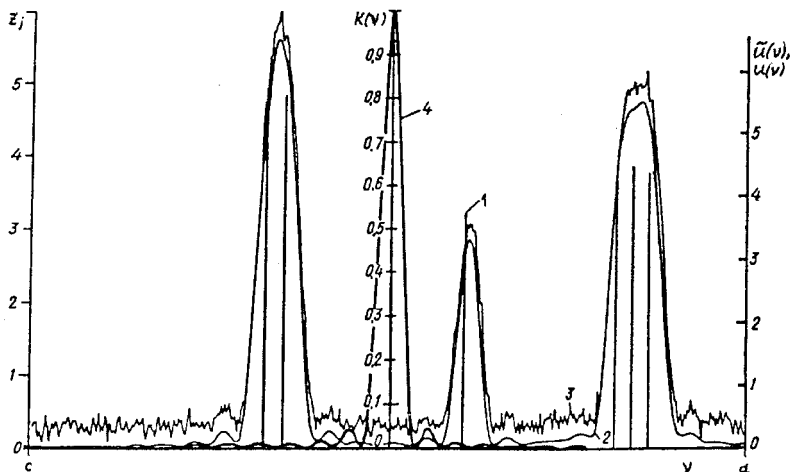


Рис. 2.18

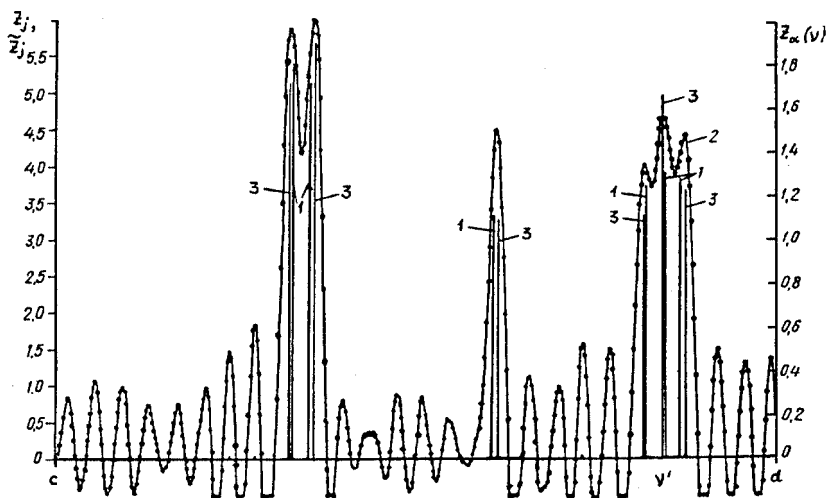


Рис. 2.19

На рис.2.19 — решение данного примера: 1 — точные интенсивности линий z_j , 2 — решение $z_\alpha(\nu')$ интегрального уравнения (2.42) методом регуляризации Тихонова при $\alpha = 10^{-4}$, 3 — решение \tilde{z}_j уточняющей СЛАУ (2.47). Видим, что в результате решения уравнения (2.42) методом регуляризации Тихонова (кривая 2) разрешились все истинные линии, однако появилось много

ложных линий (максимумов в решении $z_\alpha(\nu')$). Но последующее решение уточняющей СЛАН (2.47) позволило существенно (на 1–3 порядка) снизить интенсивности ложных линий, сделав их меньше порога Z , и приблизить вычисленные значения интенсивностей истинных линий к их точным значениям.

В результате все шесть спектральных линий разрешились и с хорошей точностью определились их частоты и интенсивности, а также фон F , причем ни одна линия не потерялась и ни одна ложная не появилась, хотя помехо-сигнальная ситуация в этом примере выбрана специально сложная, чтобы продемонстрировать возможности алгоритма интегральной аппроксимации.

Другие обратные задачи спектроскопии. Перечислим другие обратные задачи спектроскопии [44] (которые должны выполняться в порядке их перечисления): редукция к оптически тонкому слою (редукция за самопоглощение в линии, или за реабсорбцию), сепарация перекрывающихся компонентов сверхтонкой (мультиплетной) структуры линии (разложение сложного суммарного контура $z(\nu)$ — см. рис. 2.11 на контуры отдельных компонентов линии, например, в виде гауссиан), восстановление функции распределения по скоростям движущихся атомов, вносящих вклад в контур некоторого компонента линии и т. д.

О программах. Задача редукции к идеальному спектральному прибору в случае непрерывного спектра сводится к решению интегрального уравнения Фредгольма I рода (2.42) или типа свертки (2.43). Для их решения методом регуляризации Тихонова (а также методами локальной регуляризации, субоптимальной фильтрации и оптимальной фильтрации Винера) можно использовать следующие программы на Фортране: PTIMR, PTIZR, PTIPR, PTIKR [71], TIKH1, TIKH2, TIKH3, TIKH4, TIKH5, CONV1, CONV2, CONV3, CONV4, CONV5 [19], CONVOL, LOCAL0, LOCALINF, LOCALN, SUBOPT, OPT [61, пакет CONF].

В случае же дискретного спектра задача сводится к решению СЛНУ (2.45), для решения которой можно использовать программы на Фортране [80] или [61, пакет SLNE].

Контрольные задания и вопросы

1. Назовите типы спектров (излучения и т. д.).
2. Перечислите оптические спектральные приборы.
3. Назовите области применения спектрального анализа.
4. Какие бывают спектры по их виду (непрерывные и т. д.).
5. Что такое широкополосный и узкополосный спектральный анализ? Приведите примеры.
6. В чем состоят отличия экспериментального спектра от истинного?

7. Дайте определение аппаратной функции (спектральной чувствительности) спектрометра.
8. Сформулируйте задачу восстановления непрерывного и дискретного спектров.
9. Какими уравнениями описывается задача восстановления спектра?
10. Распишите подробно СЛНУ (2.45) при $n = 2$, $m = 5$.
11. Перечислите методы решения СНУ без ограничений, СНУ с ограничениями и СЛНУ.
12. В чем состоит алгоритм интегральной аппроксимации решения СЛНУ?

2.4. Обратная задача диагностики плазмы

Рассмотрим обратную задачу вычислительной диагностики плазмы.

Понятие плазмы. Сформулируем понятие плазмы в виде определения.

Определение. *Плазма* [53, с. 536] — это частично или полностью ионизованный газ, в котором плотности положительных и отрицательных зарядов (в основном ионов и электронов) практически одинаковы.

Большая часть вещества во Вселенной находится в состоянии плазмы (звезды, звездные атмосферы, туманности и межзвездная среда). Около Земли плазма существует в космосе в виде солнечного ветра, радиационных поясов и ионосферы. В лабораторных условиях и в промышленности мы имеем плазму, например, в виде электрического разряда в газах, в плазменных ускорителях и т. д.

Характеристики плазмы. Плазму характеризуют следующие *параметры*:

— *плотность* n (она заключена в очень широком диапазоне: от $\sim 10^{-6} \text{ см}^{-3}$ в межгалактическом пространстве, $\sim 10 \text{ см}^{-3}$ в солнечном ветре и до $\sim 10^{22} \text{ см}^{-3}$ в центре звезд),

— *степень ионизации* α (отношение числа ионизованных атомов к полному их числу),

— *температура* T (низкотемпературная плазма имеет $T \lesssim 10^5 \text{ К}$, а высокотемпературная плазма имеет $T \gtrsim 10^6 \text{ К}$, например, на поверхности Солнца $T = 6 \cdot 10^3 \text{ К}$, а внутри него $T = 20 \cdot 10^6 \text{ К}$),

— *локальная излучательная способность* ε (связана с n , α , T) и т. д.

Термин «плазма» впервые введен в физику в 1929 г. американскими учеными Ленгмюром и Тонксом.

Плазма во многом отличается от нейтрального газа. Для плазмы (в отличие от нейтрального газа) характерны ионизация, кулоновские силы между частицами, взаимодействие с электромагнит-

ными полями. Это позволяет рассматривать плазму как особое, четвертое состояние вещества.

В установках (системах) *токамак* (сокращение от «тороидальная камера с магнитными катушками»), предназначенных для создания и удержания высокотемпературной плазмы, за счет винтовых силовых линий магнитных полей создаются *плазменные шнуры* (см. рис. 2.20).

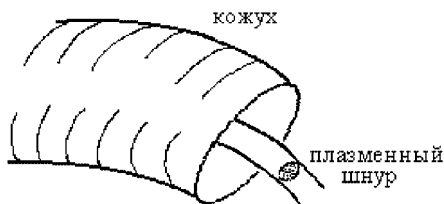


Рис. 2.20

Спектр излучения плазмы. Спектр излучения низкотемпературной (например, газоразрядной) плазмы состоит из отдельных спектральных линий (см. рис. 2.21).



Рис. 2.21

В рекламных лампах и лампах «дневного света» имеют место процессы ионизации и рекомбинации, в результате чего их спектры являются полосатыми в виде широких полос (см. рис. 2.22).



Рис. 2.22

А для высокотемпературной плазмы со значительной степенью ионизации характерно тормозное излучение с непрерывным спектром.

Диагностика плазмы.

Определение. *Диагностика плазмы* — это определение ее параметров (плотности n , температуры T , локальной излучательной способности ε , электрических и магнитных полей и т.д.) в функции координат x , y , z (т.е. локальных параметров) и времени t .

Диагностика может быть *активной* и *пассивной*. *Активная диагностика* — это определение параметров плазмы с помощью

электрического или магнитного зонда, путем просвечивания плазмы СВЧ излучением, пучками заряженных и/или нейтральных частиц, путем лазерного просвечивания и т. д. *Пассивная диагностика* — это определение параметров плазмы спектроскопическими методами, путем фотографирования, с помощью измерения электромагнитных полей и т. д.

Мы остановимся на некоторых вопросах пассивной диагностики плазмы. Дело в том, что, например, [30] в термоядерных установках необходимо знать параметры плазмы. Но из-за высокой температуры (десятки миллионов градусов) прямые измерения внутри нее (т. е. активная диагностика) невозможны. Поэтому определить внутренние параметры плазмы можно лишь на основании косвенных измерений вне плазмы и последующей их математической обработки. Это же характерно и для плазмы во вневидном пространстве, где возможна в основном пассивная диагностика с математической обработкой, к тому же здесь наши возможности ограничены (мы не можем выполнять измерения под разными ракурсами, а именно, под разными углами θ — см. далее рис. 2.23).

Схема пассивной диагностики плазмы. Рассмотрим следующую схему косвенной (пассивной) диагностики некоторого сечения плазмы (см. рис. 2.23).

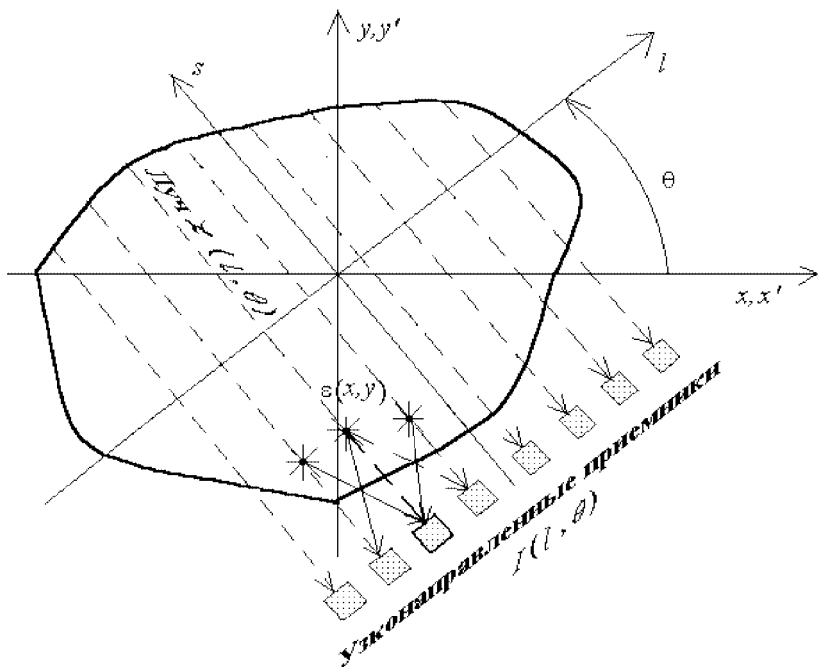


Рис. 2.23

Суть ее заключается в следующем. У зконаправленные приемники принимают (каждый) интегральное излучение $I(l, \theta)$, идущее с соответствующего луча $\mathcal{L}(l, \theta)$. За тем система приемников поворачивается на другой угол θ и измеряется новая функция $I(l, \theta)$ и т. д. В результате будет получена двухмерная функция $I(l, \theta)$.

Пусть $\varepsilon(x, y)$ — локальная излучательная способность плазмы из некоторой точки (x, y) . Тогда, если приемники обладают бесконечно узкой направленностью, каждый из них принимает излучение лишь с соответствующего луча $\mathcal{L}(l, \theta)$.

Излучение может приниматься в полосе частот — в этом случае ставится задача: по измеренной $I(l, \theta)$ определить $\varepsilon(x, y)$. Если же излучение принимается на некоторой частоте ν (или на ряде частот), то ставится задача: по измеренной (измеренным) $I_\nu(l, \theta)$ определить $\varepsilon_\nu(x, y)$. Математически обе задачи формулируются одинаково (см. дальше). Поэтому мы рассмотрим лишь одну из них — в полосе частот. Данная задача весьма напоминает задачу рентгеновской томографии (см. п. 1.1 и рис. 1.3).

Функции $I(l, \theta)$ и $\varepsilon(x, y)$ связаны следующим интегральным уравнением (ср. (1.6)):

$$\int_{\mathcal{L}(l, \theta)} \varepsilon(x, y) ds = I(l, \theta). \quad (2.50)$$

Уравнение (2.50) есть уравнение Радона относительно $\varepsilon(x, y)$. Задача решения уравнения (2.50) называется задачей определения локальных характеристик плазмы по ее интегральным характеристикам. В математической постановке эта задача восходит к работам Радона (1917 г.), а применительно к практическим задачам диагностики она появилась в шестидесятые годы почти одновременно в медицинской рентгеновской томографии (работы Кормака 1963–1964 г.) и в физике плазмы (работы Днестровского, Костомарова и др. 1966–1968 г.).

Методы решения уравнения (2.50). Уравнение (2.50) (как и уравнение (1.6)) есть интегральное уравнение относительно $\varepsilon(x, y)$ по измеренной правой части $I(l, \theta)$. Его классическое решение — это решение Радона типа (1.7).

Можно уравнение (2.50) привести к стандартной форме ср. (1.8)):

$$\iint_{-\infty}^{\infty} \frac{\varepsilon(x', y') dx' dy'}{\sqrt{(x-x')^2 + (y-y')^2}} = S(x, y), \quad (2.51)$$

где

$$S(x, y) = \frac{1}{\pi} \int_0^\pi I(x \cos \theta + y \sin \theta, \theta) d\theta, \quad (2.52)$$

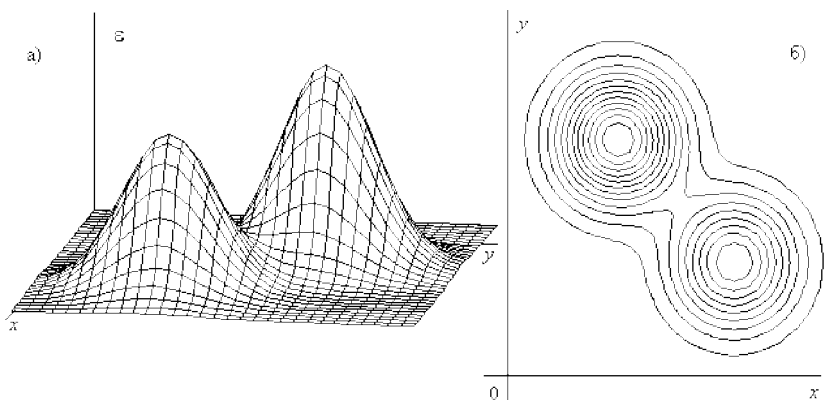


Рис. 2.24

причем

$$x \cos \theta + y \sin \theta = l \quad (2.53)$$

— уравнение прямой $\mathcal{L}(l, \theta)$. Тогда можно использовать метод ПФ для решения уравнения (2.51) и получить решение типа (1.10)–(1.13).

Однако метод Радона, а также метод ПФ сильно неустойчив, так как задача решения уравнения (2.50) и (2.51) некорректна. В результате вместо гладкой функции $\varepsilon(x, y)$ (см. рис. 2.24а, а также рис. 2.24б, где $\varepsilon(x, y)$ представлено в виде изолиний) получим неустойчивое решение в виде так называемой «пилы» (см. рис. 2.25а и 2.25б; ср. рис. 1.13).

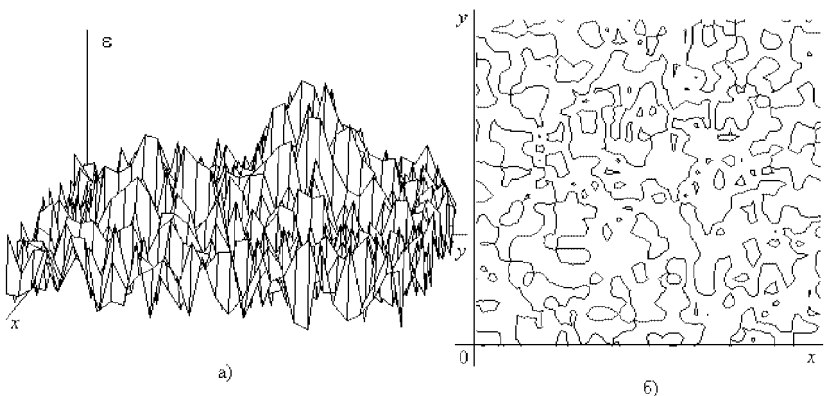


Рис. 2.25

Ясно, что такое решение неприемлемо. Устойчивое решение уравнения (2.51) дает, например, метод регуляризации Тихонова (с использованием двухмерного ПФ). Согласно ему (см. п. 8.1) регуляризованное (устойчивое) решение равно (ср. (1.10), (1.12)–(1.14))

$$\varepsilon_\alpha(x, y) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \widehat{\varepsilon}_\alpha(\omega_1, \omega_2) e^{-i(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2,$$

где

$$\widehat{\varepsilon}_\alpha(\omega_1, \omega_2) = \frac{1}{2\pi} \frac{\omega}{1 + \alpha\omega^2(\omega^4 + 1)} \widehat{S}(\omega_1, \omega_2),$$

$$\widehat{S}(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} S(x, y) e^{i(\omega_1 x + \omega_2 y)} dx dy,$$

$\omega = \sqrt{\omega_1^2 + \omega_2^2}$, $\alpha > 0$ — параметр регуляризации. Решение $\varepsilon_\alpha(x, y)$ при правильно выбранном α весьма близко к точному решению $\varepsilon(x, y)$.

В книге [52, с. 163–175] изложены также другие методы решения уравнения (2.50): представление двухмерной функции $\varepsilon(x, y)$ в виде одномерной функции $\varepsilon(z)$, где z — параметр, характеризующий вид изолиний (например, в виде эллипсов); применение отрезка двухмерного ряда Фурье; сведение двухмерного интегрального уравнения (2.50) к системе одномерных интегральных уравнений типа Абеля с использованием полиномов Чебышева, Цернике, Лагерра, Эрмита (по работам Кормака, Мальдонадо и др.); использование двухмерных сглаживающих сплайнов для численной реализации решения Радона типа (1.7) и др.

Случай цилиндрической симметрии. Направим (криволинейную) ось z вдоль плазменного шнура (см. рис. 2.26).

Рассмотрим случай, когда локальная излучательная способность ε зависит лишь от расстояния r от оси z (оси симметрии), т. е. $\varepsilon_z(x, y) = \varepsilon(r)$, $r = \sqrt{x^2 + y^2}$. Это — случай *цилиндрической симметрии* плазменного шнура. Еще реальнее случай *круговой симметрии*, когда в каждом сечении плазменного шнура ε зависит лишь от r , но при этом зависит также от z , т. е. $\varepsilon_z(x, y) = \varepsilon_z(r)$. Рассмотрим оба эти случая в одном ключе, опуская индекс z . В обоих случаях диагностику плазмы достаточно выполнять узконаправленными приемниками, не поворачивая их на разные углы θ (см. рис. 2.27).

Тогда $I(l, \theta) = I(l)$, $\varepsilon(x, y) = \varepsilon(r)$. Измеренная интенсивность $I(l)$ равна интегралу по лучу зрения, т. е.

$$I(l) = \int_G \varepsilon(r) ds, \quad (2.54)$$

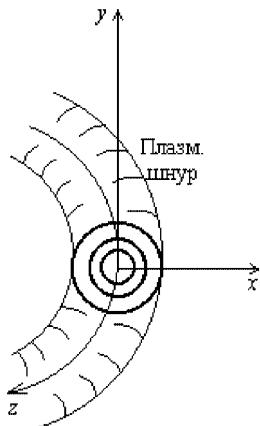


Рис. 2.26

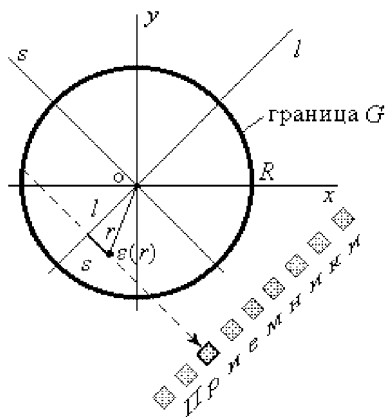


Рис. 2.27

где G — граница сечения (окружность радиуса R). Преобразуем (2.54), учитывая, что $s = \sqrt{r^2 - l^2}$, $ds = r dr / \sqrt{r^2 - l^2}$; получим:

$$I(l) = 2 \int_l^R \varepsilon(r) \frac{r dr}{\sqrt{r^2 - l^2}}$$

или

$$\int_l^R \frac{r}{\sqrt{r^2 - l^2}} \varepsilon(r) dr = \frac{I(l)}{2}, \quad -R \leq l \leq R. \quad (2.55)$$

Соотношение (2.55) есть интегральное уравнение I рода с переменным нижним пределом, поэтому это — интегральное уравнение типа Вольтерры, но поскольку при $r = l$ ядро $K(l, r) = r / \sqrt{r^2 - l^2}$ обращается в бесконечность, то это также сингулярное интегральное уравнение. Его название — интегральное уравнение Абеля [30, 52] (точнее, интегральное уравнение Цейпеля [19, с.109]). В нем

$I(l)$ — измеренная функция (излучение, принятое приемником),
 $\varepsilon(r)$ — искомая функция (локальная излучательная способность плазмы),

$K(l, r) = r / \sqrt{r^2 - l^2}$ — ядро со слабой сингулярностью.

Итак, в случае круговой (или цилиндрической) симметрии плазмы двухмерное интегральное уравнение (2.50) или (2.51) преобразуется в одномерное интегральное уравнение (2.55).

Случай шаровой симметрии. Этот случай имеет место, например, для шаровой молнии, а также (в первом приближении) для звезд. Полагаем, что плазма — это шар и $\varepsilon = \varepsilon(\rho)$, где ρ — расстояние от центра шара (см. рис. 2.28).

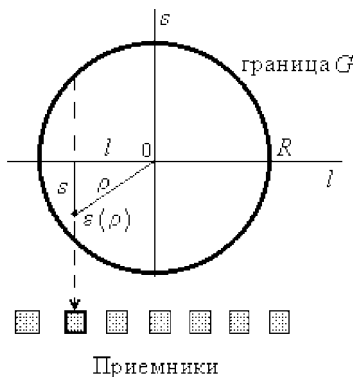


Рис. 2.28

Математическая задача описывается уравнением типа (2.55), а именно:

$$\int_l^R \frac{\rho}{\sqrt{\rho^2 - l^2}} \varepsilon(\rho) d\rho = \frac{I(l)}{2}, \quad -R \leq l \leq R. \quad (2.56)$$

Решение уравнений (2.55) и (2.56). Уравнения (2.55) и (2.56) принадлежат к тому редкому типу интегральных уравнений, которые имеют аналитическое решение. Например, решение уравнения (2.56) имеет вид

$$\varepsilon(\rho) = -\frac{1}{\pi} \int_{\rho}^R \frac{dI(l)/dl}{\sqrt{l^2 - \rho^2}} dl, \quad 0 \leq \rho \leq R. \quad (2.57)$$

Однако при всем изяществе решения (2.57) оно имеет тот недостаток, что интеграл в (2.57) является сингулярным (подынтегральная функция обращается в бесконечность при $l = \rho$). А поскольку функция $I(l)$, а значит и $dI(l)/dl$ — экспериментальная (численная, табличная) функция, то интеграл в (2.57) нужно вычислять численно по некоторой квадратурной формуле. Однако, если в качестве квадратурной формулы взять, например, формулу трапеций, то вычисления будут невозможны из-за обращения функции $1/\sqrt{l^2 - \rho^2}$ в бесконечность при $l = \rho$. Чтобы обойти это, можно использовать специальные квадратурные формулы для сингулярных интегралов [6]. Но еще удобнее воспользоваться следующей

модификацией формулы (2.57), которую можно получить с помощью интегрирования по частям:

$$\varepsilon(\rho) = \frac{1}{\pi} \int_{\rho}^R \sqrt{l^2 - \rho^2} \frac{d}{dl} \left[\frac{1}{l} \frac{dI(l)}{dl} \right] dl, \quad 0 \leq \rho \leq R. \quad (2.58)$$

В формуле (2.58) подынтегральная функция не обращается в бесконечность (кроме случая $l = \rho = 0$), правда, требуется двукратное дифференцирование экспериментальных данных $I(l)$.

Контрольные задания и вопросы

1. Дать определение плазмы, привести примеры.
2. Какие параметры характеризуют плазму?
3. Какие типы спектров бывают у плазмы?
4. Что такое диагностика плазмы? Пояснить понятия активной и пассивной диагностики.
5. Привести и пояснить схему пассивной диагностики плазмы.
6. Записать и истолковать уравнение Радона относительно локальной излучательной способности плазмы $\varepsilon(x, y)$.
7. Почему решения уравнений (2.50) методом Радона и (2.51) методом ПФ неустойчивы и почему устойчиво решение уравнения (2.51) методом регуляризации Тихонова?
8. Вывести решение (2.57) уравнения (2.56) (задание повышенной трудности; подсказка: выполнить в (2.56) интегрирование по частям и дифференцирование по параметру l , умножить на $1/\sqrt{l^2 - x^2}$, проинтегрировать по l от x до ∞ и изменить порядок интегрирования).
9. Преобразовать (2.57) в (2.58) с помощью интегрирования по частям.

Глава 3

ОБОБЩЕННАЯ ФОРМУЛИРОВКА ОБРАТНЫХ ЗАДАЧ

В этой главе мы дадим более общую, чем в предыдущих главах, постановку обратных задач.

3.1. Обработка сигналов

Рассмотрим широко распространенную задачу — *обработку сигналов* (*signal processing*) [14–16, 59]. Она имеет много общего с задачами, изложенными в гл. 2 и 4, но и имеет свою специфику.

Постановка задачи. В радиолокации, радиоастрономии, биологической микроскопии (п. 4.2), оптике (в задаче восстановления изображений — п. 2.1, 2.2), спектроскопии (п. 2.3), томографии (гл. 1), теории управления (п. 4.1), механике (п. 4.3), гидроакустике, гравиметрии и т. д. большое значение имеет *интерпретация косвенных измерений* [14], сводящаяся обычно к решению обратной задачи, которая математически формулируется в виде операторного, интегрального, дифференциального уравнений, СЛАУ, СНУ, СЛНУ и т. д.

Пусть на *вход измерительной подсистемы* (ИПС) (антенны, телескопа, микроскопа, фотоаппарата, спектроскопа, томографа, системы управления, гравиметра и т. д.) поступают *входные воздействия* (ВВ) p — входные сигналы, лучи, помехи, поток излучения и т. д. Независимой переменной может являться угловая координата (прием ВВ с разных направлений в радиолокации, радиоастрономии, гидроакустике), линейная координата (на фотоснимке, на томограмме, в микроскопе), время (в ЯМР-томографии, в теории управления, в механике, в речевой акустике), частота (в спектроскопии) и т. д.

ВВ могут быть стохастическими (случайными) или детерминированными, протяженными (распределенными) или локальными (дискретными), тональными (монохроматическими) или широкополосными.

Полагаем, что ВВ проходят через тракт ИПС, содержащий частотный фильтр, квадратичный детектор и накопитель по времени, причем за время накопления (усреднения) сигнала возможен сдвиг (смаз) ИПС (аналогичный смазу в задаче восстановления смазанных изображений - см. п. 2.1). На *выходе ИПС* формируется результат измерений f — *выходной сигнал* (сканирующая функция — СФ, индикаторный процесс — ИП и т. д.).

Связь между p и f запишем в операторной форме:

$$Rp = f, \quad p \in P, \quad f \in F, \quad (3.1)$$

где R — аппаратная функция (АФ) ИПС. Рассматривается линейная ИПС; в этом случае R есть некоторый линейный оператор, а P и F — некоторые гильбертовы пространства (например, W_2^1 и L_2). На практике соотношение (3.1) есть линейное интегральное или дифференциальное уравнение, СЛАУ и т. д. Полагаем, что вместо точных R и f известны \tilde{R} и \tilde{f} такие, что $\|\tilde{R} - R\| \leq \xi$, $\|\tilde{f} - f\| \leq \delta$, где ξ и δ — верхние оценки погрешностей R и f . В результате вместо (3.1) рассматривается уравнение

$$\tilde{R}\tilde{p} = \tilde{f}, \quad \tilde{p} \in P, \quad \tilde{f} \in F. \quad (3.2)$$

Далее для определенности будем рассматривать задачу обработки г/а сигналов. В этом случае R — это ХН антенны, p — г/а поле на входе антенны, f — измеренное г/а поле (СФ, если измеряется давление, или ИП, если измеряется мощность).

Нужно различать 3 типа обработки.

Типы обработки сигналов:

1) *Первичная обработка* — выделение сигнала из шума. Обычно в этом случае обрабатывается временной процесс $S(t)$ — см. рис. 3.1 лишь с целью определения, присутствует ли в $S(t)$ полезный сигнал (с некоторой вероятностью) [72].

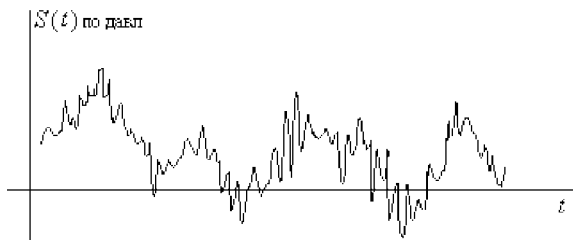


Рис. 3.1

2) *Вторичная обработка* — определение параметров сигналов: углов прихода, амплитуд и частот в случае дискретных сигналов (плоских волн) или входного процесса и его спектра в случае непрерывных сигналов, а также параметров помехи. На рис. 3.2 представлен случай дискретных сигналов, где ψ_j — углы прихода сигналов, A_j — их амплитуды, N — их число, $\tilde{f}(\psi)$ — зашумленный ИП, равный (ср. (2.45))

$$\tilde{f}(\psi) = \sum_{j=1}^N K(\psi, \psi_j) A_j + \delta f \quad (3.3)$$

в случае измерений и обработки в полосе частот, или

$$\tilde{f}_\nu(\psi) = \sum_{j=1}^N K_\nu(\psi, \psi_j) A_{j\nu} + \delta f_\nu \quad (3.4)$$

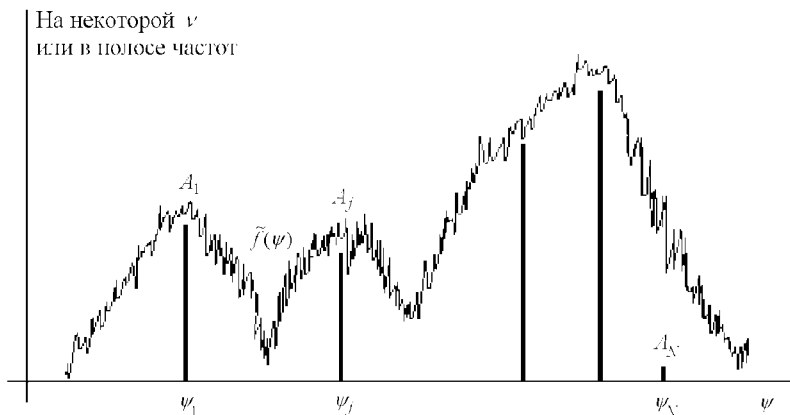


Рис. 3.2

в случае измерений и обработки на частоте ν , где K — ХН антенны по мощности, δf — шум (помеха из среды + погрешность измерений f). Задача заключается в том, что путем решения СЛНУ (3.3) или (3.4) (или другим способом — см. дальше методы Андерсона, Фроста и др.) нужно определить углы ψ_j , амплитуды A_j и число сигналов N в полосе частот или на ряде частот ν (тогда для каждого сигнала будет получен спектр $A_{j\nu} \equiv S_j(\nu)$ — см. рис. 3.3).

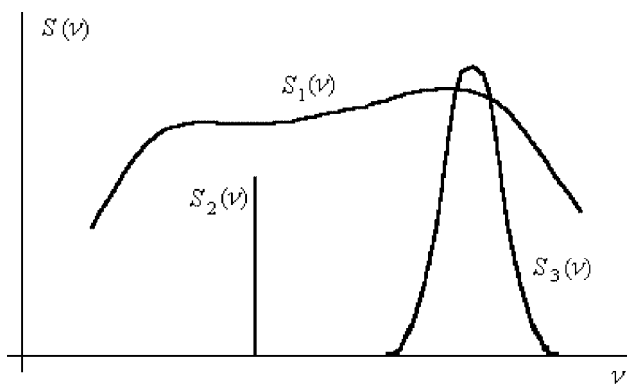


Рис. 3.3

3) *Третьичная обработка**) — классификация, распознавание целей. Например, по результатам вторичной обработки на ряде

*) нестандартный термин

частот (см. рис. 3.3) нужно определить, что 1 — это мощная широкополосная помеха, 2 — монохроматическая помеха, 3 — полезный сигнал и т. д.

Однако далее мы будем в основном рассматривать лишь вторичную обработку сигналов.

Перечень методов вторичной обработки сигналов. Для вторичной обработки сигналов применяются следующие методы: классические методы пеленгования, методы компенсации локальных сигналов-помех, методы адаптации (приспособления, подстройки под помехо-сигнальную ситуацию), методы использующие собственные значения и вектора спектрально-ковариационной матрицы, теоретико-информационные методы, обобщенный способ редукции (приведения). Изложим некоторые из них.

Классические методы пеленгования. Это — максимальный, фазовый и др. методы [73, с. 178–190]. Например, в *максимальном методе* сигналы (а именно, их параметры: углы, амплитуды и их число) определяются по максимумам в СФ $\tilde{f}(\psi)$ (см. рис. 3.4).

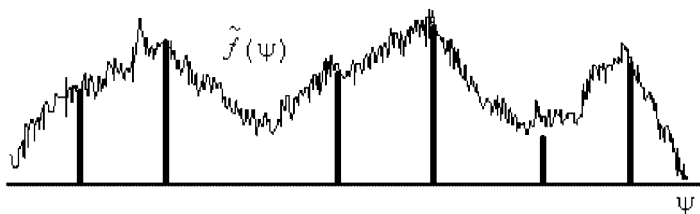


Рис. 3.4

Однако этот метод имеет следующие недостатки: 1) неразрешимость (даже в отсутствие шумов) близких сигналов; 2) потеря слабых сигналов в шуме. Впрочем, он имеет и следующие достоинства: 1) простота обработки; 2) минимум информации. Поэтому его часто используют в начале обработки для выделения наиболее сильных сигналов. А затем используют более точные, ниже следующие методы.

Методы компенсации локальных сигналов-помех. Это — методы, использующие информацию о локальных сигналах, которые или являются помехами, или мы их таковыми условно считаем. К этим методам относятся методы Андерсона и когерентной компенсации.

Изложим *метод Андерсона* [32, 47, 89]. Он состоит в компенсации (исключении) математическим путем локальных помех на выходах сформированных пространственных каналов и заключается в следующем. Пусть *антенна* имеет произвольную форму, *сигналы* являются локальными (этому соответствуют плоские

волны), узкополосными, взаимно некоррелированными и имеется *дополнительная информация* — координаты (параметры) $(\psi', A)_m$ некоторого сигнала. *Выходной сигнал* (измеренный сигнал, сканирующая функция) равен

$$f(\psi) = \sum_{j=1}^N K(\psi, \psi'_j) A_j$$

— см. рис. 3.5, из которого видно, что в $f(\psi)$ четко выделяется наибольший (m -й) максимум.

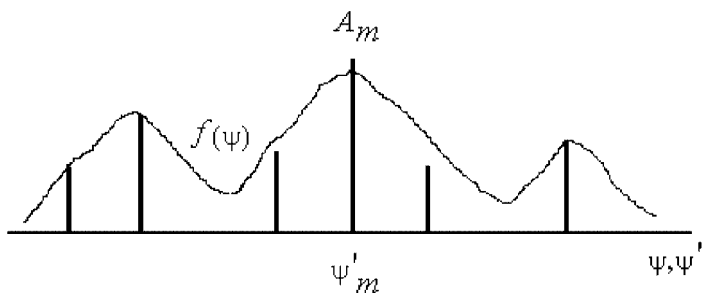


Рис. 3.5

Поэтому в качестве дополнительной информации можно использовать параметры m -го сигнала: ψ'_m и $A_m \approx f(\psi'_m)$. Тогда может быть рассчитана новая сканирующая функция (без m -го сигнала, который будем считать локальной помехой, или мешающим сигналом):

$$\underline{f}(\psi) = f(\psi) - K(\psi, \psi'_m) A_m \quad (3.5)$$

(см. рис. 3.6).

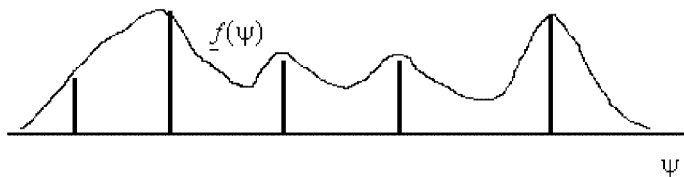


Рис. 3.6

Из функции $\underline{f}(\psi)$ также можно (если это возможно) выделить наибольший максимум с его параметрами, т. е. определить параметры еще одного сигнала, рассчитать следующую новую функцию и т. д.

Метод Андерсона — это метод компенсации локального сигнала-помехи (или нескольких таких) на выходе антенны путем математического вычитания из сканирующей функции вклада этого сигнала. Его *достоинство* — простота (метод сводится к операции вычитания согласно (3.5)). Однако он имеет *недостатки*: 1) Из графика типа рис. 3.5 параметры ψ'_m и A_m определяются приближенно, тем более при наличии случайной помехи — см. рис. 3.4. 2) Близкие сигналы (даже мощные) могут не разрешаться (два левых сигнала на рис. 3.4–3.6), а слабые могут «тонуть» в шуме.

К методу Андерсона примыкает *метод когерентной компенсации* [32, 47]. Его отличия от метода Андерсона состоят в том, что компенсация локального сигнала производится не математическим, а техническим путем и не на выходах пространственных каналов (т. е. из сканирующей функции), а на выходах отдельных приемников (элементов, преобразователей) антенны.

Методы адаптации. Метод Андерсона — это преддверие методов адаптации. После него появилось большое количество публикаций по различным вариантам методов адаптации. Это — методы Хоуэлса-Эплбаума, Уидроу, Гриффитса, Фроста и др. Они частично отражены в книгах [32, 47].

Основная идея методов адаптации состоит в том, что нужно сформировать такую АФ приемного устройства (ХН антенны или ЧХ фильтра и т. д.), чтобы в направлении (или на частоте) некоторого сигнала, который мы считаем (часто условно) мешающим, или помехой, был сформирован «прокол» (глубокий минимум) в АФ (ХН, ЧХ и т. д.). Поясним это на примере.

Пример (частотная избирательность телевизора). Пусть (см. рис. 3.7а) телевизор с ЧХ (реальным фильтром) $S_T(\nu)$ принимает полезный сигнал (телепрограмму) с центральной частотой ν_0 и спектром $S_C(\nu)$.

Пусть, кроме того, на телепрограмму воздействует шум (внешний и внутренний) со спектром $S_{\Pi} = \text{const}$ («белая помеха») и некая промышленная помеха с центральной частотой ν_{Π} и спектром $S_{\Pi}(\nu)$, гораздо более узким, чем $S_T(\nu)$. Особенность ситуации состоит в том, что хотя ν_{Π} заметно отлична от ν_0 (попадая лишь на боковое поле спектра $S_T(\nu)$), но $S_{\Pi}(\nu_{\Pi})$ в несколько раз больше $S_C(\nu_0)$ и поэтому промышленная помеха оказывает заметное мешающее воздействие на телеприем. Необходимо так перестроить ЧХ (реальный фильтр) $S_T(\nu)$, чтобы на спектр промышленной помехи $S_{\Pi}(\nu)$ приходился практически ноль («прокол») в ЧХ. Наилучшим решением была бы ЧХ в виде прямоугольника, вырезающего лишь область частот спектра $S_C(\nu)$ (идеальный фильтр, см. рис. 3.7б). Однако такую ЧХ создать и технически, и математически сложно (точнее говоря, невозможно). Гораздо реальнее создать такую ЧХ (адаптивный фильтр, см. рис. 3.7б), которая мало отличается

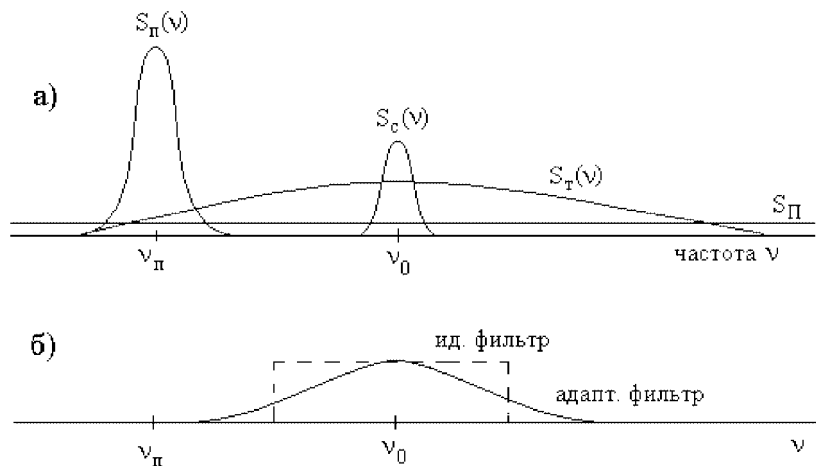


Рис. 3.7

от $S_T(\nu)$ за исключением области частот в районе ν_{Π} , где она имеет «прокол». В результате практически исключится влияние промышленной помехи на телеприем.

Особенностью задачи является то, что мешающий источник (в данном примере промышленная помеха) может менять свой спектр $S_{\Pi}(\nu)$, в результате форма «прокола» должна подстраиваться (адаптироваться) под $S_{\Pi}(\nu)$.

Из методов адаптации изложим *адаптивный алгоритм Фроста* [32, 47, 78]. Данный алгоритм рассмотрим применительно к задаче адаптации (подстраивания) ХН антенны под помехо-сигнальную ситуацию.

Рассмотрим дискретную линейную приемную антенну из M элементов (преобразователей) (см. рис. 3.8).

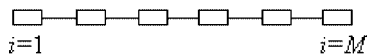


Рис. 3.8

Они формируют ХН (по мощности) [31, 65, 73]. Она будет иметь различный вид в зависимости от весовых коэффициентов преобразователей w_1, w_2, \dots, w_M (комплексные величины). Если все весовые коэффициенты одинаковы, то вводя задержки для электрического поворота ХН [31, 65, 73], мы получим ХН $K(\psi, \psi')$ следующего вида (см. рис. 3.9, где ψ — угол компенсации, а ψ' — текущий угол).

Если же весовые коэффициенты падают к краям антенны, то получим ХН типа изображенной на рис. 3.10.

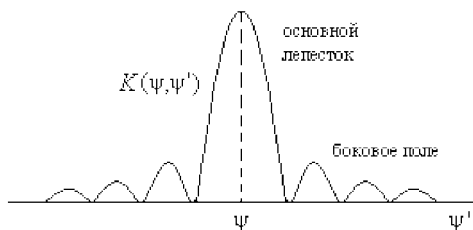


Рис. 3.9

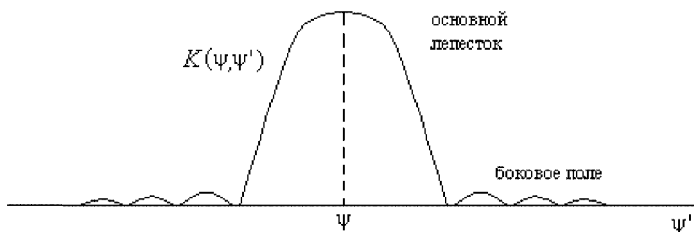


Рис. 3.10

В этом случае основной лепесток ХН будет шире, а боковое поле слабее, чем в случае одинаковости весовых коэффициентов. При некотором, более сложном, распределении коэффициентов w можно получить ХН с проколом в направлении $\bar{\psi}$ некоторого мешающего сигнала, который мы хотим подавить (см. рис. 3.11).

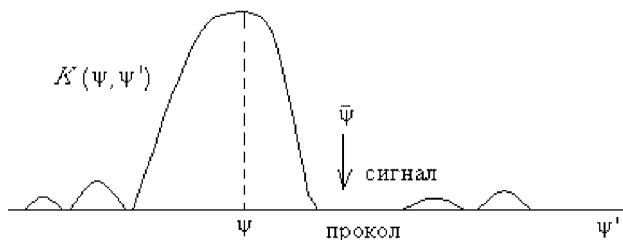


Рис. 3.11

Если же мы меняем задержки и тем самым создаем другой угол компенсации (см. рис. 3.12), то чтобы сформировать прокол в ХН в направлении $\bar{\psi}$ того же мешающего сигнала, нужно сформировать некое другое распределение коэффициентов w . В этом состоит идея алгоритма Фроста. Правда, в данном алгоритме ХН в явном виде не формируется, но неявно она присутствует.

В алгоритме Фроста полагается, что антенная решетка имеет M элементов (преобразователей). На этой основе создан процессор

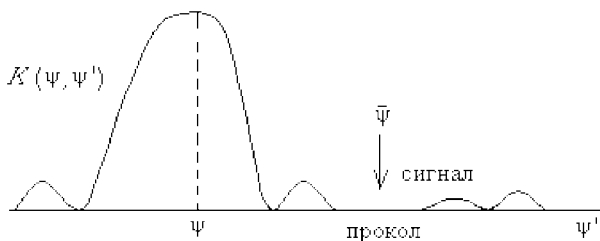


Рис. 3.12

с M элементами, у каждого из которых имеется J отводов, соответствующих J направлениям. Должна быть определена матрица W размера $M \times J$ весовых коэффициентов w . В алгоритме Фроста эта матрица записывается в виде следующего вектора размера $N \times 1$ (где $N = MJ$):

$$\begin{aligned}
 W^T = & (w_1, w_2, \dots, w_M, & - & \text{для 1-го направления } \psi_1, \\
 & w_{M+1}, \dots, w_{2M}, & - & \text{для 2-го направления } \psi_2, \\
 & \dots \dots \dots & & \\
 & w_{N-M+1}, \dots, w_N) & - & \text{для } J\text{-го направления } \psi_J.
 \end{aligned}$$

Полагается, что на решетку (и процессор) поступают с разных направлений сигналы и шумы. Каждые Δ сек происходит выборка значений напряжений в отводах решетки. Под выборкой подразумевается реализация измерений, другими словами, выполняется не одно измерение на каждом отводе, а ряд измерений (реализаций) и затем производится усреднение по ансамблю реализаций (см. ниже).

Вектор напряжений (размера $N \times 1$) в отводах при k -й выборке равен

$$X^T(k) = [x_1(k\Delta), x_2(k\Delta), \dots, x_N(k\Delta)], \quad k = 1, 2, 3, \dots,$$

причем $X(k) = L(k) + \mathcal{N}(k)$ есть сумма $L(k)$ — напряжения сигнала с заданного (j -го) направления и $\mathcal{N}(k)$ — напряжения шумов с других направлений. Полагается, что сигналы и шумы — случайные процессы с нулевыми матожиданиями и следующими корреляционными функциями:

$$\begin{aligned}
 R_{XX} &= E[X(k) X^T(k)], \\
 R_{\mathcal{N}\mathcal{N}} &= E[\mathcal{N}(k) \mathcal{N}^T(k)], \\
 R_{LL} &= E[L(k) L^T(k)],
 \end{aligned} \tag{3.6}$$

где $E[\cdot]$ — символ усреднения по ансамблю реализаций, причем вектор сигнала $L(k)$ с заданного направления и вектор шумов $\mathcal{N}(k)$ с других направлений некоррелированы: $E[\mathcal{N}(k) L^T(k)] = 0$.

Суммарный *выходной сигнал* всей решетки при k -й выборке с учетом всех J направлений равен (скаляр):

$$y(k) = W^T X(k) = X^T(k) W. \quad (3.7)$$

Матожидание *мощности сигнала* на выходе решетки равно

$$E[y^2(k)] = E[W^T X(k) X^T(k) W] = W^T R_{XX} W.$$

В алгоритме вводятся следующие ограничения: сумма весовых коэффициентов в j -м направлении равна заданному числу f_j :

$$c_j^T W = f_j, \quad j = \overline{1, J}, \quad (3.8)$$

где c_j — j -й столбец *матрицы ограничений* C (размера $N \times J$). Эти ограничения зависят от заданного (j -го) направления, а поскольку задается несколько (J) направлений, то в алгоритме Фроста, по-существу, каждое j -е направление объявляется полезным и выделяется сигнал с каждого j -го направления с подавлением шумов с других направлений. Эта задача решается как *задача оптимизации*. При этом в качестве основного критерия оптимизации используется *минимизация суммарной выходной мощности* $W^T R_{XX} W$ при ограничениях (3.8). Это осуществляется методом неопределенных множителей Лагранжа. В итоге оптимальный вектор весовых коэффициентов получается в виде

$$W_{\text{opt}} = R_{XX}^{-1} C (C^T R_{XX}^{-1} C)^{-1} \mathcal{F}, \quad (3.9)$$

причем (3.8) записывается в виде:

$$C^T W = \mathcal{F},$$

где $\mathcal{F}^T = (f_1, f_2, \dots, f_J)$.

Процессор, построенный на основании данного алгоритма, называется *оптимальным процессором* НСКОО (наименьшего среднего квадрата ошибки с ограничениями), являющимся пространственно-частотным фильтром. Подставляя (3.9) в (3.7), получим оптимальную в смысле НСКОО оценку сигнала с заданного направления:

$$y_{\text{opt}}(k) = W_{\text{opt}}^T X(k). \quad (3.10)$$

Другие методы обработки сигналов. Перечислим другие методы вторичной обработки сигналов:

— методы, использующие собственные значения и вектора спектрально-ковариационной матрицы (методы Писаренко, Редди, Прони, Шмидта, Джонсона и др.) [13, 38, 97];

— теоретико-информационные методы (максимальной энтропии Берга, максимального правдоподобия в трактовке Кейпона и др.) [16, 38];

— методы редукции (обобщенный метод редукции измерений [59], см. п. 3.2) и др.

Контрольные задания и вопросы

1. Сформулируйте постановку задачи обработки сигналов.
2. Что такое первичная, вторичная и третичная обработка сигналов?
3. Перечислите методы вторичной обработки сигналов.
4. Сформулируйте максимальный метод пеленгования.
5. Сформулируйте метод адаптации Андерсона. В чем отличие метода когерентной компенсации от метода Андерсона?
6. В чем состоит основная идея методов адаптации?
7. Изложите пример частотной избирательности телевизора.
8. Сформулируйте адаптивный алгоритм Фроста.
9. В формулах (3.6)–(3.10) проставьте размеры матриц и векторов R_{XX} , X и т. д.

3.2. Редукция измерений к идеальному измерительному устройству

Введение в задачу. Рассмотрим ряд *измерительных устройств*: радиолокатор, телескоп, радиотелескоп, гидроакустическая антенна, спектрометр, микроскоп, фотоаппарат, телевизор, система управления, томограф и т. д. Для всех них характерно то, что измеренный сигнал (например, индикаторный процесс — ИП \tilde{f} в гидроакустике — см. [рис. 3.2](#) или экспериментальный спектр \tilde{u} в спектроскопии — см. [рис. 2.13, 2.15, 2.17, 2.18](#)) отличается от истинного сигнала (поля на входе гидроакустической антенны y или истинного спектра z), а именно, в \tilde{f} или \tilde{u} не разрешены близкие максимумы, слабые максимумы «тонут» в шуме и т. д., т. е. измеренный сигнал часто обладает недостаточной разрешающей способностью. Это отличие измеренного сигнала от истинного обусловлено, во-первых, шумами различной природы (внешними шумами, шумами в измерительном устройстве и погрешностями измерений) и во-вторых, тем, что аппаратная функция — АФ K (ХН в гидроакустике — см. [рис. 3.9–3.12](#) или СЧ в спектроскопии — см. [рис. 2.12, 2.18](#)) не является бесконечно узкой, а имеет некоторую ширину.

Есть два способа приближения измеренного сигнала к истинному.

Первый способ заключается в совершенствовании измерительной аппаратуры, например, в увеличении размеров гидроакустической антенны (в этом случае становится уже ХН) или в создании спектрометров высокого разрешения (длиннофокусные монохроматоры, помещенные в вакуумные корпуса в виброзащищенных и термостабилизированных помещениях с использованием интерферометров Фабри-Перо [53, с. 705]) и т. д. Однако этот способ связан с высокой сложностью и стоимостью аппаратуры.

Кроме того, в ряде случаев этот способ принципиально не может (без использования математической обработки) привести к цели. Примеры: ограниченность размеров гидроакустической антенны длиной по дво днойю дки; старые, искаженные фотографии, которые нельзя повторить; задачи томографии (см. г л. 1), в которых по результатам измерений $q(l, \theta)$ — см. рис. 1.12 или $S(t)$ — см. рис. 1.23, 1.28, 1.29 нельзя в принципе без математической обработки определить плотность вещества $c(x, y)$ и т. д.

Второй способ заключается в использовании математической обработки результатов измерений с целью устранения искажающих факторов, а именно, сглаживающего (уширяющего) эффекта АФ и шумов, т. е. с целью приведения (редукции) измерений к идеальному измерительному устройству. Другими словами, этот способ состоит в повышении разрешающей способности измерительных устройств за счет математической (и компьютерной) обработки результатов измерений [17, с. 129], [51].

Определение. *Идеальным измерительным устройством* назовем устройство, измеряющее непосредственно искомую функцию или величину, причем без погрешностей.

Примеры: гидроакустическая антенна с бесконечно узкой ХН в отсутствие погрешностей измерений; спектрометр с бесконечно узкой АФ в отсутствие помех; томограф, измеряющий непосредственно плотность вещества $c(x, y)$ (таковым, в принципе, является, например, ЯМР-томограф на основе метода Хиншоу — см. п. 1.2) и т. д.

Определение идеального измерительного устройства в такой форме несколько шире прежнего определения, как устройства с бесконечно узкой АФ и в отсутствие шумов [15]. Оно включает в себя и прежнее определение, и имеет в виду также те задачи (томографии и т. д.), в которых АФ как таковая отсутствует и/или измеряются косвенные величины (не плотность $c(x, y)$, а например, эхо- сигнал $S(t)$ в ЯМР-томографии).

Р едукционная проблема Рэля. Еще в 1871 г. Рэлей сформулировал следующую задачу (в связи с задачами спектроскопии): обработать математически некоторую измеренную прибором функцию (выходной сигнал) так, чтобы восстановить истинный сигнал, поступивший на вход прибора (входной сигнал), учитывая его АФ. Сейчас эта задача называется *редукционной проблемой Рэля* [15, 27] и она, по- существу, говорит о редукции (приведении) измерений к идеальному измерительному устройству.

Во времена Рэля сначала казалось, что эта блестящая идея вполне осуществима и тогда мы будем иметь возможность чисто математическим путем устранять технические несовершенства различных измерительных устройств. Однако затем выяснилось (в начале XX века, во времена А дамара), что очень многие прикладные задачи в математическом отношении являются некорректными — чрезвычайно чувствительными к даже очень малым

погрешностям измерений, аппроксимации оператора (например, замене интеграла конечной суммой) и т. д. В результате редукционная проблема Рэлея в математическом отношении оказалась значительно сложнее, чем представлялось раньше.

Сформулируем математически редукционную проблему Рэлея. Пусть f — результат измерений, или выходная функция (интенсивность звукового поля в функции направления в гидроакустике, спектр — распределение энергии по частоте ν в спектроскопии, искаженное изображение в функции линейной координаты в задаче восстановления изображений, эхо-сигнал в функции времени t в томографии и т. д.), y — входная (искомая, неизвестная) функция и пусть y и f связаны математически (а также физически и технически) соотношением:

$$Ay = f, \quad (3.10)$$

где A — некоторый математический известный оператор (интегральный, дифференциальный, алгебраический и т. д.), подразумевающий АФ, ХН, ЧХ и т. д. Тогда решение уравнения (3.10) можно рассматривать как *обратную задачу* относительно y , т. е. как задачу редукции. Формально решение можно записать в виде

$$y = A^{-1}f, \quad (3.11)$$

где A^{-1} — обратный оператор (например, ОПФ, обратная матрица и т. д.)

Пример 1 (редукция локальных сигналов). Проиллюстрируем возможное решение редукционной проблемы Рэлея на примере редукции локальных (дискретных) сигналов и использования для этой цели *обобщенного метода редукции измерений* [59]. Пусть (см. пример 3 в [59, ч. III]) на вход антенны поступает $N = 4$ локальных некоррелированных сигналов с амплитудами $A_1 = 7.8$, $A_2 = 9.1$, $A_3 = 0.8$, $A_4 = 8.4$ и углами прихода $\psi'_1 = 8^\circ.1$, $\psi'_2 = 12^\circ.3$, $\psi'_3 = 22^\circ.2$, $\psi'_4 = 32^\circ.1$. Пусть сформирован веер из M характеристик направленности (M пространственных каналов, см. рис. 3.13).

Пусть углы каналов равны $\psi_i, i = \overline{1, M}$, а именно, $\psi_1 = 0$, $\psi_2 = 0^\circ.5$, ..., $\psi_M = 40^\circ$, т. е. $\Delta\psi = \text{const} = 0^\circ.5$, $M = 81$. Тогда измеренный сигнал в i -м канале (индикаторный процесс) равен (с учетом шумов)

$$\tilde{U}(\psi_i) = \sum_{j=1}^N A_j R(\psi_i, \psi'_j) + F + \delta U, \quad i = \overline{1, M}, \quad (3.12)$$

где $F = 10$ — детерминированная часть помехи, а $\delta U \in \mathcal{N}(0, 3)$ — случайная часть помехи, $R(\psi, \psi')$ — ХН антенны, равная

$$R(\psi, \psi') = e^{-\frac{15(\psi - \psi')^2}{0.1(\psi - 20)^2 + 100}}, \quad (3.13)$$

где ψ и ψ' — в град.

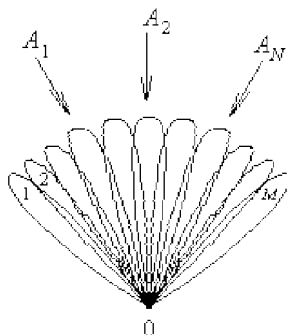


Рис. 3.13

Проанализируем ХН (5.13) (практическое занятие). Если $\psi = 20^\circ$, то

$$R(20^\circ, \psi') \cong e^{-\frac{(20-\psi')^2}{2.6^2}}. \quad (3.14)$$

Если же $\psi = 0$ или 40° , то

$$R(\psi, \psi') \cong e^{-\frac{(\psi-\psi')^2}{3.1^2}}. \quad (3.15)$$

В качестве ширины ХН ΔR будем рассматривать ее ширину по уровню $e^{-1} \cong 0.37$. Тогда видно, что ΔR при $\psi = 20^\circ$ равна $5^\circ.2$, а при $\psi = 0$ или 40° равна $6^\circ.2$, т.е. в вереве ХН, изображенном на рис. 3.13, наиболее узкая (наиболее острая) ХН — средняя ($\psi = 20^\circ$ или $i = 41$), а наиболее широкие ХН — крайние ($\psi = 0$ или $i = 1$ и $\psi = 40^\circ$ или $i = M = 81$).

Запишем (3.12) в виде

$$\sum_{j=1}^N R(\psi_i, \tilde{\psi}'_j) \tilde{A}_j + F = \tilde{U}(\psi_i), \quad i = \overline{1, M}. \quad (3.16)$$

Соотношение (3.16) есть система M линейно-нелинейных уравнений относительно неизвестных \tilde{A}_j , $\tilde{\psi}'_j$, F и N (\tilde{A}_j и F входят линейно, а $\tilde{\psi}'_j$ — нелинейно). СЛНУ (3.16) решалась обобщенным методом редукции измерений (с использованием алгоритма интегральной аппроксимации и метода регуляризации Тихонова). При этом полагалось: $a = c = 0$, $b = d = 40^\circ$. На рис. 3.14 представлены: 1 — точные значения амплитуд сигналов A_j , $j = \overline{1, N}$; 2 — точный ИП $U(\psi)$; 3 — дискретизированный и зашумленный ИП $\tilde{U}(\psi_i)$, $i = \overline{1, M}$ (см. (3.12)); 4 — решение $P_\alpha(\psi')$ интегрального уравнения

$$\int_a^b R(\psi, \psi') P(\psi') d\psi' = \tilde{U}(\psi), \quad c \leq \psi \leq d,$$

методом регуляризации Тихонова ($\alpha = 0.6$); 5 — решение уточняющей СЛАУ

$$\sum_{j=1}^L R(\psi_i, \tilde{\psi}'_j) \tilde{P}_j + F = \tilde{U}(\psi_i), \quad i = \overline{1, M},$$

относительно \tilde{P}_j и F при $L = 6$, $\tilde{\psi}'_1 = 12^\circ.5$, $\tilde{\psi}'_2 = 8^\circ.5$, $\tilde{\psi}'_3 = 32^\circ$, $\tilde{\psi}'_4 = 22^\circ$, $\tilde{\psi}'_5 = 18^\circ$, $\tilde{\psi}'_6 = 26^\circ$. Результат решения: $\tilde{P}_1 = 12.9$, $\tilde{P}_2 = 10.8$, $\tilde{P}_3 = 6.1$, $\tilde{P}_4 = 1.6$, $\tilde{P}_5 = -0.05$, $\tilde{P}_6 = -0.08$, $F = 10.7$.

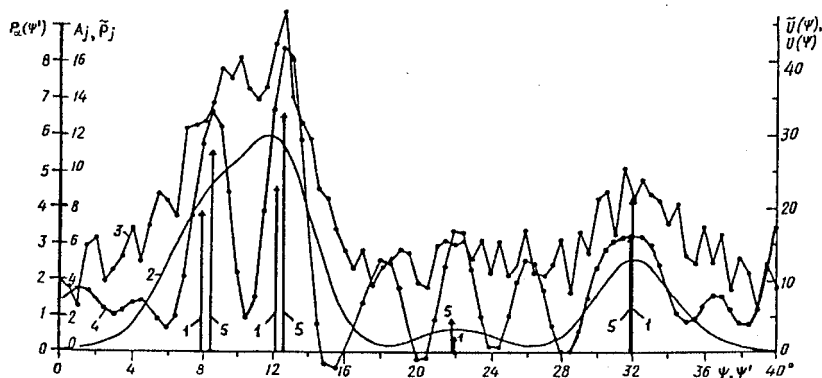


Рис. 3.14

Рис. 3.14 показывает, что в данном примере имеет место сложная помехо-сигнальная ситуация, а именно, два левых сигнала являются близкими и даже в незашумленном $U(\psi)$ (кривая 2) не разрешаются, а третий (слабый) сигнал в зашумленной $\tilde{U}(\psi)$ совершенно не просматривается. И тем не менее, все сигналы (5 на рис. 3.14) разрешились, а ложные отфильтровались (подробнее алгоритм интегральной аппроксимации см. в п. 2.3).

Пример 2 (редукция протяженных сигналов). Пусть со всех направлений из сектора $[a, b]$ поступают на антенну звуковые сигналы и их интенсивность P зависит от направления ψ следующим образом (см. сплошную линию на рис. 3.15, типа рис. 1 в [59, ч. III]).

Будем принимать их антенной, поворачивая (сканируя) ее ХН. Пусть ХН антенны (по мощности) равна

$$R(\psi, \psi') = \sqrt{\frac{Q}{\pi}} e^{-\frac{Q(\psi - \psi')^2}{1 + \psi'^2}},$$

где ψ — угол компенсации (угол наведения ХН), а ψ' — текущий угол, $Q = 59.924$.

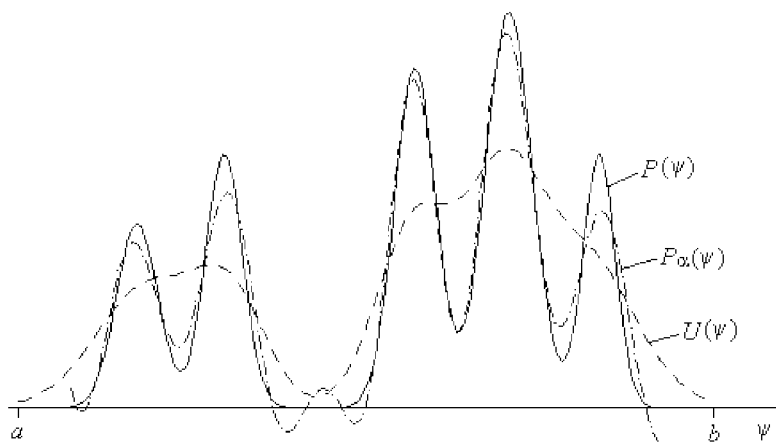


Рис. 3.15

На рис. 3.15 пунктиром отображен незашумленный индикаторный процесс $U(\psi) = \int_a^b R(\psi, \psi') P(\psi') d\psi'$. К функции $U(\psi)$ добавлялись $\approx 1\%$ погрешности. На рис. 3.15 приведено решение интегрального уравнения методом регуляризации Тихонова $P_\alpha(\psi)$ ($\alpha = 10^{-3.5}$). Видим, что все флуктуации в решении восстановились, появился лишь один небольшой ложный максимум (подробности — в [59, ч. III]; здесь обозначения изменены по сравнению с [59, ч. III]).

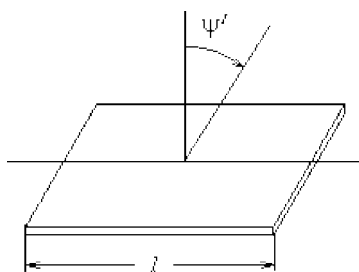


Рис. 3.16

Об аппаратной функции. Остановимся еще раз на вопросе об АФ в связи с редуционной проблемой. Применительно к антеннам АФ называется ХН или ДН [31, 65, 73]. Для прямоугольного поршневого излучателя или приемника (см. рис. 3.16) ХН по давлению в функции угла ψ' , отсчитанного от нормали,

записывается в виде [73, с. 91–100]:

$$R(\psi') = \frac{\sin\left(\frac{\pi l}{\lambda} \sin \psi'\right)}{\frac{\pi l}{\lambda} \sin \psi'}, \quad (3.17)$$

где l — длина излучателя, λ — длина волны, и графически выглядит следующим образом (см. рис. 3.17а в декартовых координатах и рис. 3.17б в полярных координатах).

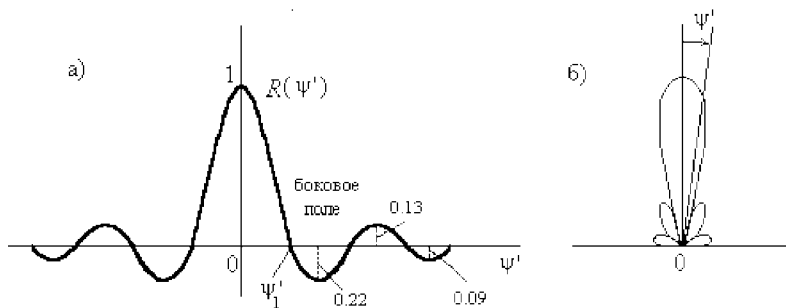


Рис. 3.17

На рис. 3.17 $2\psi'_1$ — угол раствора главного максимума ХН, причём ψ'_1 находится из условия: $\sin\left(\frac{\pi l}{\lambda} \sin \psi'_1\right) = 0$ или $\frac{\pi l}{\lambda} \sin \psi'_1 = \pi$, откуда

$$\psi'_1 = \arcsin \frac{\lambda}{l}. \quad (3.18)$$

Например, при $\lambda = 0.1l$ имеем: $\psi'_1 = \arcsin 0.1 \cong 6^\circ$. Формула (3.18) показывает, что ψ'_1 уменьшается (ХН обостряется) с уменьшением длины волны λ (т.е. с переходом на более высокие частоты) и с увеличением размеров излучателя. Иначе говоря, ψ'_1 уменьшается с увеличением так называемого *волнового размера излучателя* l/λ . Однако большую роль играет и боковое поле ХН (см. рис. 3.17).

ХН рассматривают также по модулю давления:

$$|R(\psi')| = \left| \frac{\sin\left(\frac{\pi l}{\lambda} \sin \psi'\right)}{\frac{\pi l}{\lambda} \sin \psi'} \right| \quad (3.19)$$

и по мощности:

$$R^2(\psi') = \left[\frac{\sin\left(\frac{\pi l}{\lambda} \sin \psi'\right)}{\frac{\pi l}{\lambda} \sin \psi'} \right]^2 \quad (3.20)$$

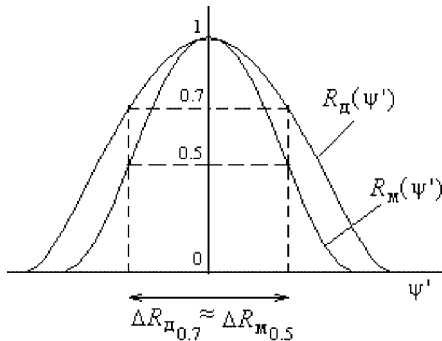


Рис. 3.18

(в этом случае резко снижается уровень бокового поля). Часто рассматривают ХН в полосе частот — тогда она будет гораздо более гладкой функцией, чем (3.17)–(3.20) (см. рис. 3.18).

Если ХН измеряется по давлению ($R_{\text{д}}$), то используется ширина ХН по уровню 0.7: $\Delta R_{\text{д}0.7}$, а если ХН измеряется по мощности ($R_{\text{м}}$), то используется ширина ХН по уровню 0.5: $\Delta R_{\text{м}0.5}$. Поскольку $0.7^2 \approx 0.5$, то $\Delta R_{\text{д}0.7} \approx \Delta R_{\text{м}0.5}$.

При механическом вращении (повороте) поршня ХН записывается в функции двух переменных:

$$R(\psi, \psi') = \frac{\sin \left[\frac{\pi l}{\lambda} \sin(\psi - \psi') \right]}{\frac{\pi l}{\lambda} \sin(\psi - \psi')} = R(\psi - \psi'), \quad (3.21)$$

где ψ — направление нормали к поршню, $\psi - \psi'$ — угол, отсчитанный от нормали. Видим, что ХН в этом случае является разностной функцией, т. е. поворачивается вместе с поршнем без деформации. Если же поворот ХН осуществляется электрически (за счет внесения задержек), что имеет место для антенн подводных лодок, то

$$R(\psi, \psi') = \frac{\sin \left[\frac{\pi l}{\lambda} (\sin \psi - \sin \psi') \right]}{\frac{\pi l}{\lambda} (\sin \psi - \sin \psi')}, \quad (3.22)$$

где ψ и ψ' отсчитываются от нормали к борту. Первый нуль в ХН вида (3.22) имеет место при $\frac{\pi l}{\lambda} (\sin \psi'_1 - \sin \psi) = \pi$, т. е. при

$$\psi'_1 = \arcsin \left(\frac{\lambda}{l} + \sin \psi \right). \quad (3.23)$$

Формула (3.23) отличается от (3.18). При механическом повороте угол раствора $2\psi'_1$ главного максимума ХН (см. (3.18)) не

зависит от ψ , а при электрическом повороте $2\psi'_1$ растет с ростом ψ (см. (3.23)). Например, если $\lambda = 0.1l$, то при $\psi = 0$ имеем по обеим формулам: $\psi'_1 = \arcsin 0.1 \approx 6^\circ$, а при $\psi = 30^\circ$ формула (3.18) дает $\psi'_1 \approx 6^\circ$, а формула (3.23) дает $\psi'_1 = \arcsin 0.6 \approx 37^\circ$, т. е. ХН (3.22) при $\psi = 30^\circ$ в 6 раз (!) шире, чем при $\psi = 0$.

Техническая реализация алгоритмов редукции. Итак, решение редукционной проблемы Рэлея сводится к адекватному математическому описанию процессов и эффективному решению соответствующих математических соотношений (интегральных, дифференциальных, трансцендентных уравнений, систем линейных, нелинейных или линейно-нелинейных уравнений и т. д.). Под эффективным решением подразумевается не только устойчивое и высокоточное решение, но и решение, требующее по возможности малых затрат компьютерного времени и памяти. Эти вычисления могут выполняться с помощью цифровых или аналоговых ЭВМ, преобразователей (ЦАП и АЦП), фильтров, а также специализированных вычислительных устройств (СВУ).

Соединение измерительного устройства с вычислительным с целью решения задачи редукции измерений ведет к созданию так называемых *измерительно-вычислительных комплексов* (ИВК). Это равнозначно созданию новых измерительных устройств с более высокой разрешающей способностью (по углу, времени, частоте и т. д.). Такие устройства способны восстанавливать микроструктуру измеряемого поля, разрешать близкие сигналы, выделять слабые сигналы из шума и т. д. Если под измерительным устройством подразумевать, например, радиоантенну или телескоп, то создание ИВК будет равносильно созданию антенны или телескопа больших размеров. Можно сказать, что создание ИВК, т. е. решение задач редукции измерений может дать значительный технический и финансово-экономический эффект.

Контрольные задания и вопросы

1. В чем состоят два способа приближения измеренного сигнала к истинному? Примеры.
2. Сформулируйте редукционную проблему Рэлея.
3. На основе рис. 3.13 сформулируйте задачу редукции сигналов.
4. Выведите (подробно) соотношения (3.14) и (3.15) из (3.13).
5. На основе рис. 3.14 опишите последовательность операций в обобщенном методе редукции измерений.
6. Нарисуйте графики ХН по модулю давления $|R(\psi')|$ (см. (3.19)) и по мощности $R^2(\psi')$ (см. (3.20)) типа рис. 3.17а и укажите значения максимумов в боковом поле ХН.
7. Опишите техническую реализацию методов редукции с использованием измерительно-вычислительного комплекса.

Глава 4

НЕКОТОРЫЕ ОБРАТНЫЕ ЗАДАЧИ ТЕОРИИ УПРАВЛЕНИЯ,
БИОФИЗИКИ И МЕХАНИКИ

4.1. Обратные задачи теории управления

Рассмотрим кратко ряд обратных задач современной теории управления, описываемых интегральными уравнениями [10, 57, 87].

Восстановление сигнала в динамической системе. Обширной областью приложения интегральных уравнений типа Вольтерры являются задачи анализа процессов в непрерывных динамических системах управления с обратной связью или без нее.

Пусть $y(t)$ — входной сигнал в систему управления (управляющее воздействие на систему), $f(t)$ — выходной сигнал (отклик, реакция системы), где t — время (поэтому система называется *динамической*). Тогда, если система линейна, связь между входным и выходным сигналами выражается в виде следующего общего соотношения [10, с. 198]:

$$l(t)y(t) + \int_{t_0}^t L(t, \tau)y(\tau) d\tau = F(t), \quad t \in [t_0, T], \quad (4.1)$$

где

$$F(t) = k(t)f(t) + \int_{t_0}^t K(t, \tau)f(\tau) d\tau - v(t, t_0), \quad (4.2)$$

$L(t, \tau)$ и $K(t, \tau)$ — импульсные переходные функции, определяющие инерционность (последствие) системы; $l(t)$ и $k(t)$ определяют обратные связи в системе: чем больше l и k , тем сильнее обратная связь и тем выше порядок системы [10, с. 62], если $l = k = 0$, то система не имеет обратных связей; t_0 — момент начала функционирования системы, в частности, $t_0 = -\infty$ (система с бесконечной памятью, или инерционностью) или $t_0 = t - \Delta$ (система с конечной памятью $\Delta > 0$) или $t_0 = 0$ (система в состоянии покоя при $t < 0$); $v(t, t_0)$ характеризует влияние предыстории системы до момента t_0 на выходной сигнал в момент t (другими словами, начальный запас энергии, накопленный системой до начала функционирования).

Соотношение (4.1) есть интегральное уравнение Вольтерры II рода относительно входного сигнала $y(t)$, если $l(t) \neq 0$ при $t \in [t_0, T]$, или уравнение Вольтерры I рода, если $l(t) \equiv 0$ (система без обратной связи), или уравнение Вольтерры III рода, если $l(t) = 0$ при некоторых, но не всех значениях $t \in [t_0, T]$.

Методы решения уравнений Вольтерры II и I рода существенно различны. Задача решения уравнения Вольтерры II рода является корректной (устойчивой) и для его решения весьма эффективны такие классические методы, как методы квадратур, итераций, резольвенты, коллокации, сплайнов, кусочно-гладких полиномов и др. [19, с. 24–94]. Задача же решения уравнения Вольтерры I рода, строго говоря, некорректна (неустойчива), хотя «степень некорректности» ниже, чем уравнения Фредгольма I рода. Поэтому для решения уравнения Вольтерры I рода наряду с методами квадратур, коллокации, сплайнов и кусочно-гладких полиномов [19, с. 120–126, 134–137] применяют методы регуляризации (Тихонова, Денисова, Апарцина, Сергеева, Магницкого) [19, с. 126–134]. Что же касается редкого уравнения Вольтерры III рода, то оно занимает промежуточное положение между уравнениями Вольтерры II и I рода.

Соотношение (4.1) часто рассматривается также как уравнение Вольтерры относительно выходного сигнала $f(t)$ и записывается в виде:

$$k(t)f(t) + \int_{t_0}^t K(t, \tau) f(\tau) d\tau = Y(t), \quad t \in [t_0, T], \quad (4.3)$$

где

$$Y(t) = l(t)y(t) + \int_{t_0}^t L(t, \tau) y(\tau) d\tau + v(t, t_0). \quad (4.4)$$

Если $L(t, \tau) = L(t - \tau)$, $K(t, \tau) = K(t - \tau)$, $l(t) = l = \text{const}$, $k(t) = k = \text{const}$, то система называется *стационарной*, или системой с постоянными параметрами, и для нее имеет место одностороннее (при $t_0 = 0$) или двустороннее (при $t_0 = -\infty$) *уравнение Вольтерры II рода типа свертки*:

$$ly(t) + \int_{t_0}^t L(t - \tau) y(\tau) d\tau = F(t), \quad t \in [t_0, T], \quad (4.5)$$

или

$$kf(t) + \int_{t_0}^t K(t - \tau) f(\tau) d\tau = Y(t), \quad t \in [t_0, T]. \quad (4.6)$$

Если система состоит из n элементов, то имеет место (при $l = k = 1$) следующая *система n уравнений Вольтерры II рода* [87, с. 13–15]:

$$y_i(t) + \sum_{j=1}^n \int_{t_0}^t L_{ij}(t, \tau) p_{ij}(\tau) y_j(\tau) d\tau = F_i(t), \quad i = \overline{1, n}, \quad t \in [t_0, T], \quad (4.7)$$

или

$$f_i(t) + \sum_{j=1}^n \int_{t_0}^t K_{ij}(t, \tau) q_{ij}(\tau) f_j(\tau) d\tau = Y_i(t), \quad (4.8)$$

где i или j — номер элемента, p_{ij} и q_{ij} — известные интенсивности структурных связей по каналу $j \rightarrow i$.

Если система *нелинейна*, то она описывается *нелинейным интегральным уравнением Вольтерры-Урысона II рода* [10, с. 29], [87, с. 16]:

$$y(t) - \int_{t_0}^t A[t, \tau, y(\tau), f(\tau)] d\tau = y(t_0), \quad t \in [t_0, T], \quad (4.9)$$

или

$$f(t) - \int_{t_0}^t B[t, \tau, f(\tau), y(\tau)] d\tau = f(t_0), \quad t \in [t_0, T], \quad (4.10)$$

где A и B — некоторые известные нелинейные функции.

Замечание. Поскольку переменная t (и τ) является временем, то условие физической реализуемости системы (выходной сигнал не может появиться раньше входного) требует, чтобы

$$L(t, \tau) = K(t, \tau) = 0 \quad \text{при } t < \tau \quad (4.11)$$

или для стационарной системы

$$L(t) = K(t) = 0 \quad \text{при } t < 0. \quad (4.12)$$

Поэтому верхние пределы интегралов в (4.1)–(4.10) переменны и равны t .

Восстановление сигнала в динамической системе без обратной связи. Специально рассмотрим уравнение (4.1) при $l = K = v = 0$, $k = 1$:

$$\int_{t_0}^t L(t, \tau) y(\tau) d\tau = f(t), \quad t \in [t_0, T]. \quad (4.13)$$

Уравнение (4.13) есть *интегральное уравнение Вольтерра I рода* и оно описывает восстановление сигнала в упрощенной динамической системе без обратной связи. В нем $y(t)$ — *входной сигнал* в систему (фильтр), $f(t)$ — *выходной сигнал*, или отклик, а $L(t, \tau)$ — функция, называемая в теории автоматического управления и радиотехнике *импульсной переходной функцией* (или импульсной реакцией, или весовой функцией). Если $L(t, \tau) = L(t - \tau)$, то система с такой импульсной функцией в теории автоматического управления называется *стационарной*, в теории систем —

однородной, или инвариантной к сдвигу, в оптике — изопланатической и т. д. В этом случае

$$\int_{t_0}^t L(t - \tau) y(\tau) d\tau = f(t), \quad t \in [t_0, T]. \quad (4.14)$$

Уравнение (4.14) есть интегральное уравнение Вольтерры I рода с разностным ядром, а при $t_0 = 0$ — одностороннее или при $t = -\infty$ двустороннее интегральное уравнение Вольтерры I рода типа свертки. Для решения уравнения (4.14) типа свертки удобным является метод ПФ (с регуляризацией для устойчивости) [19, с. 110–114, 126–134].

Частные случаи данной задачи (уравнение (4.14) при $t_0 = 0$) [67, с. 29, 164]: определение радиоимпульса $y(t)$, излученного источником, по принятому сигналу $f(t)$ и известной импульсной переходной функции среды $L(t)$; определение электрического импульса на входе кабеля $y(t)$ по измеренному сигналу на выходе кабеля $f(t)$ и известной импульсной функции кабеля $L(t)$.

Замечание. Для решения уравнения (4.14) требуется знание импульсной переходной функции L . Поэтому должна быть поставлена следующая (предшествующая) задача: по заданным входному y и выходному f сигналам определить импульсную переходную функцию L . Математически эта задача для линейной стационарной системы при $t_0 = 0$ описывается следующим односторонним уравнением Вольтерры I рода типа свертки [67, с. 30]:

$$\int_0^t y(t - \tau) L(\tau) d\tau = f(t), \quad t \geq 0. \quad (4.15)$$

Для решения такой задачи нужен следующий дополнительный специальный эксперимент: задаем $y(t)$, измеряем $f(t)$, после чего находим $L(t)$ путем решения уравнения (4.15).

Если динамическая система нелинейна (в частности, следящая система), то она описывается нелинейным уравнением Вольтерры–Урысона I рода [87, с. 11]:

$$\int_{t_0}^t H[t, \tau, y(\tau)] d\tau = f(t), \quad (4.16)$$

где H — некоторая известная нелинейная функция.

Восстановление сигнала в системе, не являющейся динамической. Если t (и τ) в уравнениях (4.1)–(4.16) не является временем (а является, например, угловой координатой, линейной координатой, расстоянием, частотой, энергией и т. д.), то система

не является динамической и ограничения (4.11), (4.12) неприменимы. Тогда имеют место интегральные уравнения Фредгольма II рода [10, с. 81]:

$$y(x) + \int_a^b h(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (4.17)$$

или Фредгольма I рода [19, с. 144]:

$$\int_a^b h(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (4.18)$$

в зависимости от характера задачи. В уравнениях (4.17) и (4.18) обозначения изменены по сравнению с (4.1)–(4.16). В (4.17) и (4.18) $h(x, s)$ — аппаратная функция (функция отклика на единичный импульс, весовая функция и т. д.), $y(s)$ — искомый входной сигнал, $f(x)$ — измеренный выходной сигнал. Сама задача в данном случае называется задачей восстановления сигнала (см. п. 3.1), а также задачей редукции измерений к идеальному прибору (см. п. 3.2) и т. д.

Контрольные задания и вопросы

1. Сформулируйте задачу анализа процессов в непрерывных динамических системах с обратной связью и без нее.
2. Как Вы понимаете физико-технически (а не математически) обратную связь?
3. Почему в соотношениях (4.1)–(4.16) верхний предел интегрирования равен t , а не постоянной величине, как в (4.17), (4.18)?
4. Каково Ваше физико-техническое (а не математическое) понимание стационарности системы?
5. Как Вы понимаете фразу: «Система состоит из n элементов»?
6. Используя соотношения (4.1) и (4.2), объясните, какими должны быть входящие в них функции $l(t)$, $L(t, \tau)$, $k(t)$, $K(t, \tau)$ и $v(t, t_0)$, чтобы выходной сигнал $f(t)$ был близок (а в пределе равен) входному сигналу $y(t)$.

4.2. Обратные задачи биофизики

Описанные в п. 4.1 задачи теории управления применимы также и к исследованию процессов в биофизических динамических (и не динамических) системах, например, к анализу передачи биотоков (как сигналов) в мозгу человека с учетом сложных связей, запаздываний (задержек), импульсных переходных функций,

возможных нелинейных связей и т.д. В этом случае возможны применения уравнений (4.1)–(4.16) для исследования данных процессов.

Однако есть и другие, более специфические для биофизики задачи.

Восстановление искаженных изображений биологических микрообъектов. При наблюдении биологических микрообъектов (вирусов, белков, аминокислот и т.д.) с помощью некоторой системы (оптического или электронного микроскопа и т.д.) могут иметь место *следующие факторы*, снижающие качество их изображений: 1) *смаз (сдвиг) изображения*, 2) *ошибка в установлении фокуса* и 3) *недостаточная разрешающая способность системы*.

В результате *смаза* (см. рис. 4.1) истинное четкое изображение микрообъектов, описываемое функцией (интенсивностью) $w(x, y)$ (см. рис. 4.1а), превратится в нечеткое фотоизображение, описываемое интенсивностью $P(x, y)$ (см. рис. 4.1б).

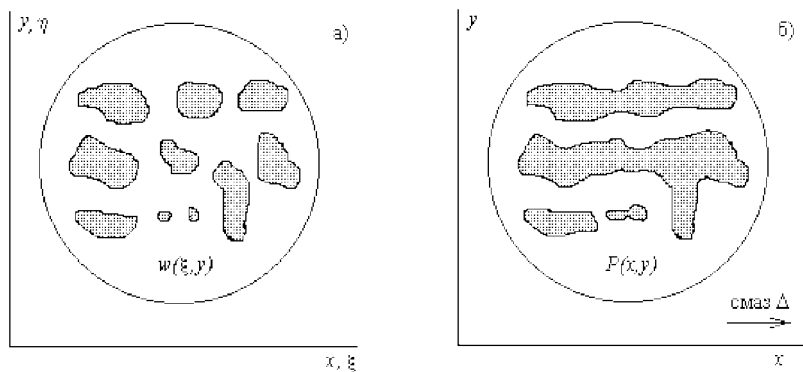


Рис. 4.1

Функции $w(x, y)$ (искомая) и $P(x, y)$ (измеренная) связаны следующим одномерным интегральным уравнением Фредгольма I рода типа свертки с параметром:

$$\int_{-\infty}^{\infty} K(x - \xi) w(\xi, y) d\xi = P(x, y), \quad (4.19)$$

где

$$K(x) = \begin{cases} \frac{1}{\Delta}, & -\Delta \leq x \leq 0, \\ 0, & \text{иначе,} \end{cases} \quad (4.20)$$

Δ — величина смаза (сдвига), а y играет роль параметра. Подробно эта задача рассмотрена в п. 2.1.

В результате *ошибки в установлении фокуса* будет получено дефокусированное (размытое) изображение. Имеет место двухмерное интегральное уравнение Фредгольма I рода типа свертки:

$$\iint_{-\infty}^{\infty} K(x - \xi, y - \eta) w(\xi, \eta) d\xi d\eta = P(x, y), \quad (4.21)$$

где

$$K(x, y) = \begin{cases} \frac{1}{\pi\rho^2}, & \sqrt{x^2 + y^2} \leq \rho, \\ 0, & \text{иначе,} \end{cases} \quad (4.22)$$

$\rho = a\delta/f_2$ — радиус дифракционного круга, в который превращается каждая точка истинного изображения, a — радиус апертуры линзы системы (микроскопа), δ — погрешность фокусировки изображения, f_2 — расстояние от линзы до зрачка наблюдателя. Поскольку в микроскопе несколько линз с разными a , δ и f_2 , то целесообразнее не оперировать величинами a , δ и f_2 , а использовать лишь величину ρ , которую следует рассматривать как некую среднюю величину для всех линз и ее можно найти по дбором.

Подробно данная задача (применительно к одной линзе) рассмотрена в п. 2.2.

Наконец, *недостаточная разрешающая способность системы* порождает *задачу редукции наблюдений биологических микрообъектов за аппаратную функцию системы*, которая описывается двухмерным интегральным уравнением Фредгольма I рода типа свертки:

$$\iint_{-\infty}^{\infty} h(x - x', y - y') w(x', y') dx' dy' = P(x, y), \quad (4.23)$$

где $w(x, y)$ — искомое истинное распределение интенсивности микрообъектов (типа рис. 4.1а), $P(x, y)$ — наблюдаемое распределение интенсивности, например, по оболочке вируса белковых молекул при увеличении в 640000 раз (типа рис. 4.1б), не обладающее должным разрешением, $h(x, y)$ — функция рассеяния (аппаратная функция) системы. Эффективное (достаточно точное и устойчивое) решение уравнения (4.23) даст возможность приблизить $P(x, y)$ к $w(x, y)$, т.е. математическим (и компьютерным) путем повысить разрешающую способность системы (например, микроскопа). Подобная задача рассмотрена в п. 3.2.

Томография биологических микрообъектов. Для восстановления 3-мерных структур вирусов, белков, аминокислот и т.д. может быть использована рентгеновская томография, изложенная в п. 1.1 (см., например, [рис. 1.12](#)), а также ЯМР-томография (см. п. 1.2).

Обратная задача речевой акустики [98]. Речевой процесс может быть описан интегральным уравнением Фредгольма I рода типа свертки:

$$\int_{-\infty}^{\infty} h(x - \tau) x(\tau) d\tau = y(t), \quad -\infty < t < \infty, \quad (4.24)$$

где $x(t)$ — возбуждение в голосовой щели как в источнике (входной процесс), $h(t)$ — импульсная характеристика голосового (вокального) тракта, $y(t)$ — речевой выход (речь). В уравнении (4.24) искомой является функция $x(t)$ при известной $h(t)$ и измеренной $y(t)$ — все величины суть акустические давления.

Рассматривается также другая задача: по известным $x(t)$ и $y(t)$ определить $h(t)$. Эта задача описывается уравнением:

$$\int_{-\infty}^{\infty} x(t - \tau) h(\tau) d\tau = y(t), \quad -\infty < t < \infty. \quad (4.25)$$

Вторая задача является более важной, так как определение импульсной характеристики голосового тракта связано с определением его спектра Фурье $H(\omega)$ (вопрос о решении уравнений типа (4.24), (4.25) см. в п. 7.2 и 8.1), а знание $H(\omega)$ (например, типа изображенного на рис. 4.2) позволяет поставить медицинский диагноз относительно состояния голосового тракта.

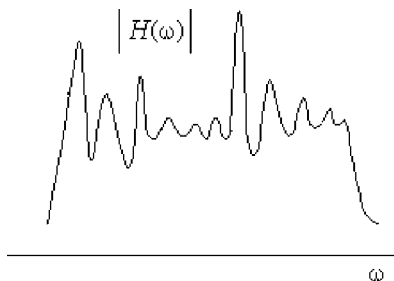


Рис. 4.2

Распад клеток и радиоактивных элементов [92]. В биологии в задаче *распада клеток* и в физике в задаче *распада радиоактивных элементов* полагается, что вещество состоит из клеток или элементов, распадающихся по одинаковому закону $\sim e^{-t/T}$, где t — текущее время, а T — эффективное время распада, но с разными значениями T . В этом случае имеет место интеграль-

ное уравнение Фредгольма I рода:

$$\int_0^{\infty} e^{-t/T} \rho(T) dT = M(t)/M(0), \quad t \geq 0, \quad (4.26)$$

где $\rho(T)$ — искомая функция распределения частиц вещества по эффективному времени распада T (по закону $\sim e^{-t/T}$), причем $\int_0^{\infty} \rho(T) dT = 1$, а $M(t)$ — измеренная в функции времени t общая масса нераспавшегося вещества.

Задача удельной приливной вентиляции в легких [92]. В физиологии имеет место задача, аналогичная последней, а именно, задача удельной приливной вентиляции в легких с вымыванием азота N_2 при многократном дыхании (в модели, не учитывающей поступления нового N_2). Справедливо уравнение (4.26), в котором $\rho(T)$ — искомая функция распределения по T — эффективному времени вымывания N_2 из легких, а $M(t)$ — общее количество невымытого N_2 в функции времени t , связанного с процессом дыхания.

В более общем случае рассматривается некоторый закон распада или вымывания иной, нежели $\sim e^{-t/T}$, а именно, $k(t, T)$, причем $k(t, T)$ — некоторая убывающая функция времени t , а производная $|\partial k(t, T)/\partial t|$ обратно пропорциональна T , т.е. чем больше T , тем медленнее происходит распад или вымывание. Тогда справедливо следующее интегральное уравнение Фредгольма I рода:

$$\int_{D_1} k(t, T) \rho(T) dT = g(t), \quad t \in D_2, \quad (4.27)$$

где $g(t)$ — наблюдаемые данные, $\rho(T)$ — искомая функция распределения частиц вещества по эффективному времени распада T по закону $k(t, T)$, а D_1 и D_2 — заданные области изменения T и t .

В заключение этого пункта отметим еще две интересные задачи математической биологии — *задача кинетики печени* [95] и *задача моделирования зрения лягушки* [94]. Обе задачи используют аппарат интегральных уравнений.

Контрольные задания и вопросы

1. В уравнении (4.19) как направляются оси x , ξ , y и η ?
2. Дайте классификацию уравнения (4.19). Это одно- или двух-мерное уравнение или совокупность уравнений или же система уравнений? Чем отличается совокупность уравнений от системы уравнений?

3. В уравнении (4.23) какой должна быть функция $h(x, y)$, чтобы $P(x, y) = w(x, y)$?

4. Сделать переход от (4.24) к (4.25) (используя замену переменной).

5. Как преобразуется (4.26), если все частицы вещества будут иметь одинаковое эффективное время распада, например T_0 ? Подсказка: использовать δ -функцию (задание повышенной трудности).

6. Для общего случая (см. (4.27)) придумать примеры функции $k(t, T)$, убывающей с ростом t , причем закон убывания обратно пропорционален T (т. е. чем больше T , тем медленнее убывает $k(t, T)$).

4.3. Использование преобразования Фурье в прикладных задачах механики

Задачи механики — это обширная область приложений различных разделов математики. Например, многие нестационарные задачи механики описываются дифференциальными уравнениями в частных производных. Это (см., например, [64]) — продольные волны в стержне, продольный удар стержня при его падении на жесткую горизонтальную преграду, волны в бесконечной цепочке из шариков, расположенных вдоль прямой и соединенных пружинами, продольные волны в пластине и круговой цилиндрической оболочке, изгибные волны (динамический изгиб стержня под действием единичной сосредоточенной силы, поля напряжений при движении поршня с жестким фланцем, задача о штампе), некоторые линейные задачи динамики сжимаемой жидкости и ее взаимодействие с упругими телами (дифракция плоской акустической волны около жесткого кругового цилиндра, акустические волны в жидкости при ее падении на цилиндр, удар тела о поверхность сжимаемой жидкости, волны в оболочке, окруженной идеальной сжимаемой жидкостью, изгибные волны в пластине, погруженной в сжимаемую жидкость) и др. Многие из этих задач решаются с помощью преобразования Фурье.

Отметим следующую характерную особенность перечисленных задач: уравнения практически всех этих задач непосредственно не имеют в качестве правой части функцию, получаемую в результате измерений, а значит, с погрешностями (такая функция присутствует лишь в граничных условиях). Исключение составляет, например, задача о распространении продольных волн в стержне. Она описывается следующим дифференциальным уравнением [64, с. 180]

$$\rho F \frac{\partial^2 u}{\partial t^2} - EF \frac{\partial^2 u}{\partial x^2} = Q(t, x), \quad (4.28)$$

где $Q(t, x)$ — задаваемая внешняя продольная нагрузка на единицу длины стержня, ρ — плотность стержня, F — площадь его поперечного сечения, E — модуль Юнга, $u = u(t, x)$ — искомое продольное перемещение точек стержня. Поскольку задаваемая и практически реализуемая нагрузки $Q(t, x)$ отличаются (вследствие погрешности исходных данных), то и решение $u(t, x)$ помимо методической погрешности будет содержать погрешность, обусловленную погрешностью задания $Q(t, x)$, а значит такая задача для своего решения требует устойчивых математических методов.

В механике есть еще целый класс задач, имеющих дело с погрешностями исходных данных (измерений). Это — задачи, связанные с вычислением и последующим анализом спектра Фурье [33, 42, 56].

Преобразование Фурье. Пусть дана некоторая непрерывная функция $y(t)$, например, звуковой процесс (см. рис. 4.3).

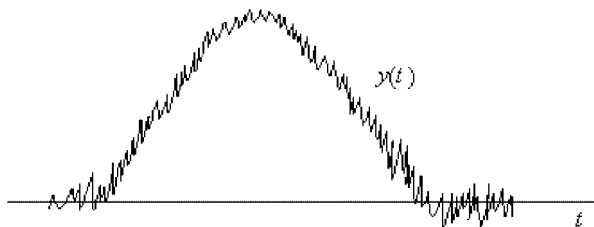


Рис. 4.3

В этом случае очень важную информацию о процессе дает НПФ (спектр Фурье, Фурье-образ):

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{i\omega t} dt. \quad (4.29)$$

При этом для анализа используют вещественную и мнимую части НПФ:

$$\operatorname{Re} Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos \omega t dt, \quad (4.30)$$

$$\operatorname{Im} Y(\omega) = \int_{-\infty}^{\infty} y(t) \sin \omega t dt \quad (4.31)$$

или же модуль от ПФ:

$$|Y(\omega)| = \sqrt{\operatorname{Re}^2 Y(\omega) + \operatorname{Im}^2 Y(\omega)} \quad (4.32)$$

(см. рис. 4.4).

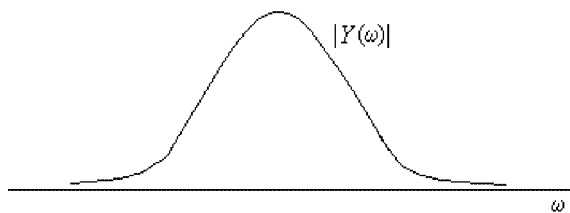


Рис. 4.4

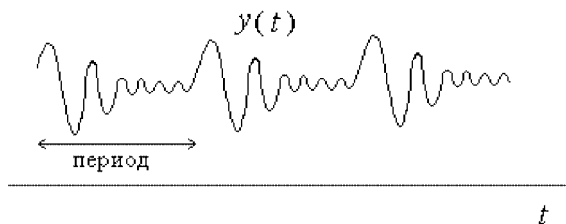


Рис. 4.5

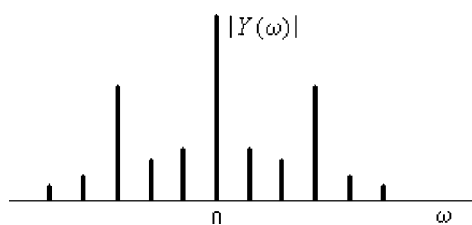


Рис. 4.6

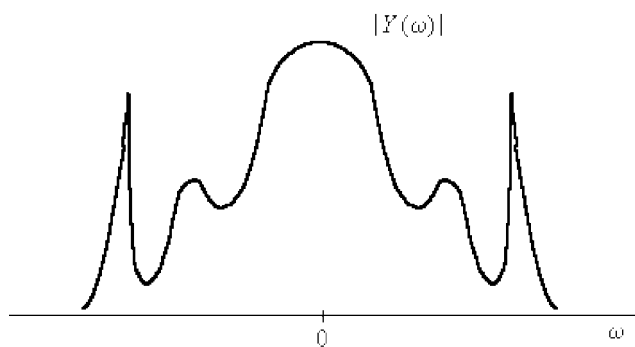


Рис. 4.7

Если временной процесс $y(t)$ — непериодические колебания типа рис. 4.3, то спектр непрерывен (типа рис. 4.4). Если же временной процесс $y(t)$ — периодически повторяющиеся колебания (см. рис. 4.5), то спектр — дискретный (см. рис. 4.6).

А практически спектр обычно сложнее (рис. 4.7).

Подробнее о ПФ см. п. 6.3.

Спектральные задачи механики. Перечислим прикладные задачи механики, в которых использование ПФ, а именно, частотный анализ звука, позволяет делать важные заключения о свойствах изучаемого объекта:

— определение свойств газовой, жидкой и твердой среды по спектру прошедшего звука,

— расчет конструкции глушителя двигателя автомобиля или самолета по спектру его шумов (выхлопов),

— расчет усталостной прочности конструкции ракеты и предотвращение ее разрушения под действием шумов двигателей на основе знания их спектров,

— борьба с шумами в цехах, где много станков, на основании анализа их спектров и т. д.

Определение, например, звукового спектра станков (типа рис. 4.6) даст возможность заглушить некоторые частоты с большими амплитудами, присутствующие в цеховом шуме. А определение, например, спектра шумов двигателя автомобиля (типа рис. 4.7) позволит рассчитать такую конструкцию глушителя, которая будет ослаблять лишь некоторые «неприятные» частоты.

Контрольные задания и вопросы

1. Как называется дифференциальное уравнение (4.28)?
2. Используя формулы (4.29)–(4.31), найти $Y(\omega)$ при
 - а) $y(t) = \cos(ct)$, б) $y(t) = e^{-a^2 t^2}$.
3. Доказать, что если $y(t) = \sum_j b_j \cos(c_j t)$ (периодическая функция), то $Y(\omega)$ — сумма δ -функций, т. е. спектр — дискретный (задание повышенной трудности).
4. Доказать, что если $y(t) = e^{-a^2 t^2} + \int_{-\infty}^{\infty} b(c) \cos(ct) dc$ (непериодический процесс типа рис. 4.3), то $Y(\omega)$ — непрерывная функция типа рис. 4.4 (задание повышенной трудности).
5. Какие частоты нужно подавить в спектрах на рис. 4.6 и 4.7?

Часть II

УСТОЙЧИВЫЕ МЕТОДЫ РЕШЕНИЯ УРАВНЕНИЙ

В данной части рассматриваются современные методы (регуляризации, фильтрации, сглаживания, аппроксимации) решения тех уравнений, которые приведены в части I. Следует отметить, что в части I приведены в основном задачи, которые описываются интегральными уравнениями Фредгольма I рода. Этому нужно дать объяснение. Во-первых, если говорить лишь о тех прикладных задачах, которым соответствуют интегральные уравнения, то как показывает богатая практика [4, 10, 14–19, 23, 45, 48, 52, 54, 55, 59, 67–71, 87, 90–92, 94–96, 98–100, 104, 105], большинство прикладных задач описываются уравнениями Фредгольма I рода. Во-вторых, имеется, конечно, множество задач, которым соответствуют дифференциальные уравнения (задачи механики, акустики, теплопроводности и т. д.), но как уже было отмечено в п. 4.3, в большинстве этих задач дифференциальное уравнение непосредственно не содержит экспериментально измеряемую правую часть. Например, задача определения потенциала φ некоторого (например, звукового) поля описывается уравнением Гельмгольца

$$\Delta\varphi + k^2\varphi = 0,$$

в котором непосредственно нет экспериментальной функции. Однако она присутствует в граничных условиях (на некоторой замкнутой поверхности S): $\varphi = \Phi$ (задача Дирихле) и $\frac{\partial\varphi}{\partial n} = N$ (задача Неймана), где k — волновое число, а n — внешняя нормаль к S . Решение дается *формулой Грина*:

$$\varphi(p) = \frac{1}{4\pi} \iint_S \left\{ N(S) \frac{e^{ikr(p,S)}}{r(p,S)} - \Phi(S) \frac{\partial}{\partial n} \left[\frac{e^{ikr(p,S)}}{r(p,S)} \right] \right\} dS,$$

где p — некоторая точка внутри S , r — расстояние между p и некоторой точкой поверхности S . При этом Φ и N задаются (или измеряются) с погрешностями, но вычисление по формуле Грина есть прямая, а не обратная задача. Если же задана лишь

одна функция Φ или N , то мы приходим через посредство метода граничных интегральных уравнений [19, 37, 83] к необходимости решать интегральное уравнение относительно другой функции, т. е. к обратной задаче с измеряемой с погрешностями функцией.

В-третьих, как показано в [18, 19, 43] и других работах, любое дифференциальное уравнение в обыкновенных или частных производных или система таковых уравнений (как линейных, так и нелинейных) может быть приведено к некоторому интегральному уравнению, но не всякое интегральное уравнение может быть приведено к дифференциальному, т. е. аппарат интегральных уравнений является более общим, чем аппарат дифференциальных уравнений.

В-четвертых, ряд задач (например, задача восстановления дискретного спектра — см. п. 2.3 или задача редукции локальных сигналов — см. п. 3.2) описывается системами линейно-нелинейных уравнений (СЛНУ) и для их решения весьма эффективным является алгоритм интегральной аппроксимации [59, ч. I, III] (см. п. 2.3), в основе которого лежит применение аппарата интегральных уравнений.

Таким образом, интегральным уравнениям должно быть уделено повышенное внимание, что и сделано в данном учебном пособии.

Глава 5

ОСНОВНЫЕ ТИПЫ УРАВНЕНИЙ И СОПУТСТВУЮЩИЕ ИМ ПОНЯТИЯ

5.1. Основные типы уравнений

Перечислим основные типы уравнений, которыми описываются прикладные задачи, рассмотренные в части I.

Интегральные уравнения. *Одномерное (линейное) интегральное уравнение Фредгольма I рода:*

$$\int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (5.1)$$

где $K(x, s)$ — математическое ядро (а физически и технически АФ, ХН и т. д.), $f(x)$ — правая часть (выходной сигнал, СФ, ИП и т. д.) — известные функции, $y(s)$ — искомая функция (входной сигнал и т. д.), x и s — линейные или угловые координаты, время, частота, температура, энергия и т. д., причем x — наружная переменная, а s — внутренняя переменная, $[c, d]$ — заданная область измерения $f(x)$, $[a, b]$ — задаваемая область поиска $y(s)$. Практически вместо точной $f(x)$ задается экспериментальная зашумленная правая часть $\tilde{f}(x) = f(x) + \delta f(x)$, где $\delta f(x)$ — аддитивная погрешность, вместо точного ядра $K(x, s)$ также часто известно приближенное $\tilde{K}(x, s)$, а вместо точной $y(s)$ получаем приближенное решение $\tilde{y}(s) = y(s) + \delta y(s)$, причем $\delta y(s)$ обусловлена не только погрешностями $\delta f(x)$ и $\delta K(x, s)$, но и методом решения.

Прикладные задачи, описываемые уравнением (5.1): восстановление непрерывного спектра в обратной задаче спектроскопии (уравнение (2.42)), редукция протяженных сигналов (п. 3.2), восстановление сигналов в системе, не являющейся динамической (уравнение (4.18)), распад клеток и радиоактивных элементов (уравнения (4.26), (4.27)), вентиляция в легких (уравнения (4.26), (4.27)).

Пример [71, с. 43, 127], [91, 104] — см. рис. 5.1.

Одномерное интегральное уравнение Фредгольма I рода с разностным ядром:

$$\int_a^b K(x - s) y(s) ds = f(x), \quad c \leq x \leq d. \quad (5.2)$$

Прикладная задача: синтез магнитного поля на оси катушки ЯМР-томографа (уравнение (1.62)).

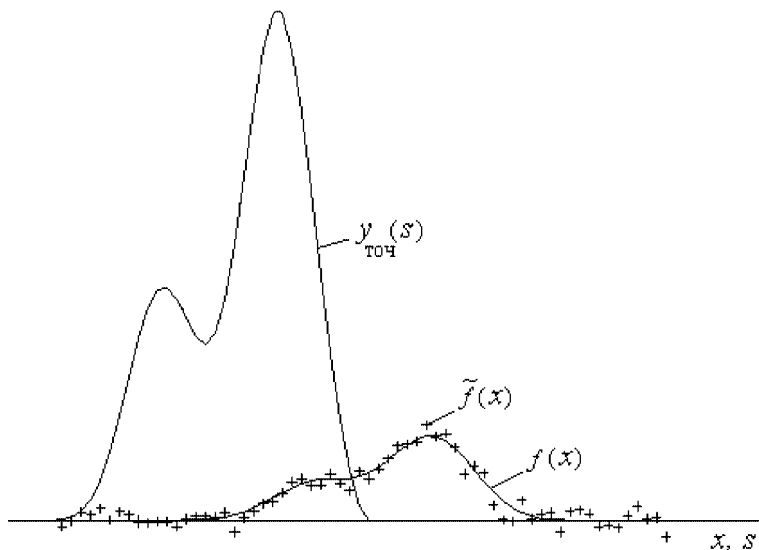


Рис. 5.1

Одномерное интегральное уравнение Фредгольма I рода типа свертки:

$$\int_{-\infty}^{\infty} K(x-s) y(s) ds = f(x), \quad -\infty < x < \infty. \quad (5.3)$$

Прикладные задачи: снятие влияния аппаратных искажений (уравнение (1.16)), восстановление смазанных фотоизображений (уравнение (2.15)), восстановление смазанных изображений биологических микрообъектов (уравнение (4.19)), обратная задача речевой акустики (уравнения (4.24), (4.25)).

Одномерное интегральное уравнение Фредгольма I рода типа свертки на полуоси:

$$\int_0^{\infty} K(x-s) y(s) ds = f(x), \quad 0 \leq x < \infty, \quad (5.4)$$

Прикладная задача: восстановление непрерывного спектра (уравнение (2.43)).

Двухмерное интегральное уравнение Фредгольма I рода типа свертки:

$$\iint_{-\infty}^{\infty} K(x_1-s_1, x_2-s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2), \quad (5.5)$$

$$-\infty < x_1, x_2 < \infty.$$

Прикладные задачи: определение плотности вещества $c(x, y)$ в РТ (уравнение (1.8)), визуализация результатов $c(x, y)$ на дисплее (уравнение (1.22)), восстановление дефокусированных фотоизображений (уравнение (2.30)), обратная задача диагностики плазмы (уравнение (2.51)), восстановление дефокусированных изображений биологических микрообъектов (уравнение (4.21)).

Одномерное интегральное уравнение Фредгольма II рода:

$$y(x) + \int_a^b K(x, s) y(s) ds = f(x), \quad a \leq x \leq b. \quad (5.6)$$

Прикладная задача: восстановление сигнала в системе, не являющейся динамической (уравнение (4.17)).

Двухмерное интегральное уравнение Фредгольма II рода:

$$y(x_1, x_2) + \int_{a_1}^{b_1} \int_{a_2}^{b_2} K(x_1, s_1, x_2, s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2), \quad (5.7)$$

$$a_1 \leq x_1 \leq b_1, \quad a_2 \leq x_2 \leq b_2.$$

Прикладная задача: определение потенциала на поверхности в месте де граничных интегральных уравнений [19, с. 146].

Двухмерное интегральное уравнение Радона (уравнение интегральной геометрии):

$$\int_{\mathcal{L}(l, \theta)} c(x, y) ds = q(l, \theta), \quad (5.8)$$

где $\mathcal{L}(l, \theta)$ — луч зрения (прямая или кривая, уравнение которой задается в функции l и θ), $c(x, y)$ — искомая функция, $q(l, \theta)$ — измеренная правая часть (поглощение). Уравнение (5.8) может быть приведено к уравнению типа (5.5) (точнее, к уравнению (1.8) или (2.51)).

Прикладные задачи: определение плотности вещества в РТ (уравнение (1.6)), обратная задача диагностики плазмы (уравнение (2.50)).

Нелинейное одномерное интегральное уравнение Урысона I рода:

$$\int_a^b K[x, s, y(s)] ds = f(x), \quad c \leq x \leq d, \quad (5.9)$$

где $K[x, s, y(s)]$ — заданная нелинейная функция.

Прикладная задача: синтез магнитного поля на оси катушки ЯМР-томографа (уравнение (1.64)).

Линейное одномерное интегральное уравнение Вольтерра I рода:

$$\int_a^x K(x, s) y(s) ds = f(x), \quad a \leq x \leq b. \quad (5.10)$$

Прикладная задача: восстановление сигнала в динамической системе без обратной связи (уравнение (4.13)).

Линейное одномерное интегральное уравнение Вольтерра I рода с разностным ядром:

$$\int_a^x K(x-s) y(s) ds = f(x), \quad a \leq x \leq b. \quad (5.11)$$

Прикладная задача: восстановление сигнала в стационарной динамической системе без обратной связи (уравнение (4.14)).

Линейное одномерное интегральное уравнение Вольтерра I рода типа свертки одностороннее (или на полуоси):

$$\int_0^x K(x-s) y(s) ds = f(x), \quad x \geq 0. \quad (5.12)$$

Прикладная задача: определение импульсной переходной функции в теории управления (уравнение (4.15)).

Линейное одномерное интегральное уравнение типа Вольтерры I рода с параметром:

$$\frac{1}{\Delta} \int_x^{x+\Delta} w(\xi, y) d\xi = g(x, y), \quad -\infty < x, y < \infty, \quad (5.13)$$

где $w(\xi, y)$ — искомая функция, $\Delta > 0$ — некоторое число, $g(x, y)$ — заданная правая часть, y — параметр, т. е. (5.13) можно рассматривать и как совокупность одномерных уравнений (столько уравнений, сколько задано значений y). Уравнение (5.13) может быть приведено к уравнению типа (5.3) (см. п. 2.1).

Прикладная задача: восстановление смазанных фотоизображений (уравнение (2.8)).

Интегральное уравнение Цейделя:

$$\int_l^R \frac{r}{\sqrt{r^2 - l^2}} \varepsilon(r) dr = \frac{I(l)}{2}, \quad -R \leq l \leq R, \quad (5.14)$$

где $\varepsilon(r)$ — искомая функция, $R > 0$ — некоторое известное число (в частности, $R = \infty$, если (5.14) применять к шаровым звездным скоплениям [19, с. 109]), $I(l)$ — измеренная функция. Уравнение (5.14) можно классифицировать как *линейное одномерное интегральное уравнение типа Вольтерры I рода сингулярное*.

Прикладная задача: обратная задача диагностики плазмы (случай цилиндрической, круговой или шаровой симметрии, уравнения (2.55), (2.56)).

Линейное одномерное интегральное уравнение Вольтерра:

$$g(x) y(x) + \int_a^x K(x, s) y(s) ds = f(x), \quad a \leq x \leq b, \quad (5.15)$$

где $g(x)$ — заданная функция. Если $g(x) \equiv 0$, то уравнение (5.15) есть уравнение Вольтерры I рода (см. (5.10)); если $g(x) \neq 0 \forall x \in [a, b]$, то уравнение (5.15) можно записать в виде

$$y(x) + \int_a^x \frac{K(x, s)}{g(x)} y(s) ds = \frac{f(x)}{g(x)}, \quad a \leq x \leq b, \quad (5.16)$$

т. е. в виде *линейного одномерного интегрального уравнения Вольтерры II рода*; если же $g(x) = 0$ при некоторых, но не все x значениях $x \in [a, b]$, то (5.15) есть *линейное одномерное интегральное уравнение Вольтерра III рода*.

Прикладная задача: восстановление сигнала в линейных динамических системах управления с различной «силой» обратной связи (уравнения (4.1), (4.3), (4.5), (4.6), (4.13)–(4.15)).

Нелинейное одномерное интегральное уравнение Вольтерры-Урысона I рода:

$$\int_a^x K[x, s, y(s)] ds = f(x), \quad a \leq x \leq b, \quad (5.17)$$

где $K[x, s, y(s)]$ — заданная нелинейная функция.

Прикладная задача: восстановление сигнала в нелинейной динамической системе управления без обратной связи (уравнение (4.16)).

Нелинейное одномерное интегральное уравнение Вольтерры-Урысона II рода:

$$y(x) + \int_a^x K[x, s, y(s)] ds = f(x), \quad a \leq x \leq b, \quad (5.18)$$

где $K[x, s, y(s)]$ — заданная нелинейная функция.

Прикладная задача: восстановление сигнала в нелинейной динамической системе управления с обратной связью (уравнения (4.9), (4.10)).

К линейным интегральным уравнениям Фредгольма I рода также следует отнести *непрерывные интегральные преобразования*

[9, 11, 33, 40, 43, 56, 64]. Это — *одномерное преобразование Фурье (НПФ)*:

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{i\omega t} dt, \quad -\infty < \omega < \infty, \quad (5.19)$$

которое можно записать в виде:

$$\int_{-\infty}^{\infty} e^{i\omega t} y(t) dt = Y(\omega), \quad -\infty < \omega < \infty, \quad (5.20)$$

где $e^{i\omega t}$ можно считать ядром уравнения, $Y(\omega)$ — заданной правой частью, а $y(t)$ — искомой функцией. Тогда соотношение (5.20) можно рассматривать как *линейное одномерное комплексное интегральное уравнение Фредгольма I рода* относительно $y(t)$. Далее, *двухмерное преобразование Фурье*:

$$Y(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} y(t_1, t_2) e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2, \quad -\infty < \omega_1, \omega_2 < \infty, \quad (5.21)$$

или

$$\iint_{-\infty}^{\infty} e^{i(\omega_1 t_1 + \omega_2 t_2)} y(t_1, t_2) dt_1 dt_2 = Y(\omega_1, \omega_2), \quad -\infty < \omega_1, \omega_2 < \infty. \quad (5.22)$$

Соотношение (5.22) можно рассматривать как *линейное двухмерное комплексное интегральное уравнение Фредгольма I рода*. Имеют место также *косинус-преобразование Фурье*:

$$Y_c(\omega) = \int_{-\infty}^{\infty} y(t) \cos(\omega t) dt, \quad (5.23)$$

синус-преобразование Фурье:

$$Y_s(\omega) = \int_{-\infty}^{\infty} y(t) \sin(\omega t) dt, \quad (5.24)$$

преобразование Хартли:

$$Y_H(\omega) = \int_{-\infty}^{\infty} y(t) \operatorname{cas}(\omega t) dt, \quad (5.25)$$

где $\operatorname{cas} x \stackrel{\text{def}}{=} \cos x + \sin x$, *одностороннее преобразование Лапласа*:

$$\Psi(p) = \int_0^{\infty} y(t) e^{-pt} dt, \quad (5.26)$$

где $p = \lambda + i\sigma$ — комплексная переменная, $y(t)$ — оригинал, $\Psi(p)$ — изображение, а также интегральные преобразования Радона (см. (1.4)), Меллина, Ганкеля, Бесселя [19, 40] и др., которые также можно рассматривать как интегральные уравнения Фредгольма I рода относительно $y(t)$.

В уравнениях (5.1)–(5.7), (5.9)–(5.12), (5.15)–(5.18) $y \in Y$, $f \in F$, где Y и F — некоторые гильбертовы пространства. Обычно $Y = W_2^1$ (пространство Соболева) или $Y = L_2$ (пространство квадратично суммируемых функций), а $F = L_2$.

Определение *норм элементов* (функций) в этих пространствах:

$$\|y\|_{W_2^1} = \sqrt{\int_a^b y^2(s) ds + \int_a^b y'^2(s) ds}, \quad (5.27)$$

$$\|y\|_{L_2} = \sqrt{\int_a^b y^2(s) ds}, \quad (5.28)$$

$$\|f\|_{L_2} = \sqrt{\int_c^d f^2(x) dx}. \quad (5.29)$$

Следует отметить, что ряд интегральных уравнений решается аналитически. Это — уравнения (5.3)–(5.5), (5.8), (5.12), (5.14), (5.19)–(5.26). Остальные уравнения решаются, вообще говоря, лишь численно.

Методы решения интегральных уравнений изложены в гл. 7, 8.

Системы линейных алгебраических уравнений. Система *линейных алгебраических уравнений* (СЛАУ) записывается в виде:

$$Ay = f, \quad (5.30)$$

где A — заданная $m \times n$ матрица элементов a_{ij} , $i = \overline{1, m}$, $j = \overline{1, n}$ (m — число строк или уравнений, n — число столбцов или неизвестных), y — искомый (неизвестный) вектор-столбец $n \times 1$ элементов y_j , $j = \overline{1, n}$, f — заданная (измеренная) правая часть — вектор-столбец $m \times 1$ элементов f_i , $i = \overline{1, m}$, или

$$\sum_{j=1}^n a_{ij} y_j = f_i, \quad i = \overline{1, m}. \quad (5.31)$$

СЛАУ используется в данном учебном пособии, главным образом, при решении интегральных уравнений методом квадратур (см. п. 7.2, 8.1) и при решении СЛНУ алгоритмом интегральной аппроксимации (см. (2.47)).

Системы линейно-нелинейных уравнений. Система линейно-нелинейных уравнений (СЛНУ) записывается в виде:

$$\sum_{j=1}^n K(\nu_i, \nu'_j) z_j + F = u(\nu_i), \quad i = \overline{1, m}, \quad (5.32)$$

или более кратко

$$\sum_{j=1}^n K_{ij} z_j + F = u_i, \quad i = \overline{1, m}, \quad (5.33)$$

где m — число уравнений, n — число линейно входящих неизвестных z_j (а также F) и нелинейно входящих неизвестных ν'_j (т. е. всего $2n + 1$ неизвестных), $u_i = u(\nu_i)$ — заданные (измеренные) правые части, $K_{ij} = K(\nu_i, \nu'_j)$ — некоторая нелинейная функция.

Прикладные задачи: восстановление дискретного спектра в обратной задаче спектроскопии (уравнение (2.45)), редукция локальных сигналов (уравнение (3.16)).

Операторные уравнения. Все вышеприведенные уравнения (5.1)–(5.26), (5.30)–(5.33) можно записать в виде единого *операторного уравнения*:

$$Ay = f, \quad y \in Y, \quad f \in F, \quad (5.34)$$

где A — заданный оператор (линейный или нелинейный, интегральный, дифференциальный или алгебраический и т. д.), часто называемый измерительным преобразованием, f — измеренная правая часть, y — искомое решение, Y и F — некоторые гильбертовы пространства, например, W_2^1 , L_2 (см. (5.27)–(5.29)).

Определение. *Норма оператора* определяется следующим образом:

$$\|A\| = \sup_y \frac{\|Ay\|_F}{\|y\|_Y}. \quad (5.35)$$

Определение. Измерения называются *однократными* (или *одношаговыми*), когда дана одна реализация f , и *многократными* (или *многошаговыми*), когда измерено много реализаций f (см. п. 8.2).

Определения. Задача решения уравнения (5.34) (а значит, и уравнений (5.1)–(5.26), (5.30)–(5.33)) называется *статической*, когда A , y и f не зависят от времени, *динамической*, когда A , y и f зависят от времени, и *квазистатической*, когда A , y

и f зависят от номера эксперимента (реализации), но за время эксперимента практически не изменяются.

Определения. Задачу отыскания y в виде: $y = Vf$ будем называть *задачей приведения выхода ко входу*, а оператор V , равный A^{-1} , $(A^*A)^{-1}A^*$, A^+ , $(\alpha E + A^*A)^{-1}$ и т. д. (см. п. 5.2, 7.3, 7.4, 8.1), — *оператором обработки*.

Контрольные задания и вопросы

1. В уравнении (5.1) переменные x и s обязательно должны иметь одинаковый физический смысл (например, линейные координаты) или могут иметь разный физический смысл (например, x — линейная координата, а s — время)? А сами функции y и f ?

2. Ответить на предыдущий вопрос применительно к уравнениям (5.6), (5.10), (5.15).

3. Что означает разностность ядра $K(x - s)$ математически (в смысле его геометрической формы) и физико-технически (на примере ХН антенны)?

4. Если в уравнениях (5.3), (5.5) ядро K стремится к δ -функции, то в какие соотношения переходят эти уравнения?

5. Если в уравнении (5.13) $\Delta \rightarrow 0$, то в какое соотношение переходит это уравнение?

6. Почему (5.14) есть сингулярное уравнение?

7. Запишите уравнение (1.16), восстановив переменную θ , и классифицируйте полученное уравнение.

8. Приведите примеры статической, динамической и квазистатической задач.

5.2. Некоторые сведения из линейной алгебры

Многие из численных методов решения интегральных уравнений (методы квадратур, аппроксимации полиномами), дифференциальных уравнений (методы конечных разностей) и т. д. приводят к необходимости решать системы линейных алгебраических уравнений (СЛАУ). Поэтому необходимо привести некоторые понятия из линейной алгебры [19, с. 504–509], [20, 21, 26, 75].

Система линейных алгебраических уравнений (СЛАУ).

Система m линейных алгебраических уравнений относительно n неизвестных записывается в виде (5.30) или (5.31), где A — матрица $m \times n$, y — искомый вектор-столбец $n \times 1$, f — заданный вектор-столбец $m \times 1$ — правая часть (A , y и f — вообще говоря, комплексные) или подробнее

$$\left. \begin{aligned} a_{11} y_1 + a_{12} y_2 + \dots + a_{1n} y_n &= f_1, \\ a_{21} y_1 + a_{22} y_2 + \dots + a_{2n} y_n &= f_2, \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots & \\ a_{m1} y_1 + a_{m2} y_2 + \dots + a_{mn} y_n &= f_m. \end{aligned} \right\} \quad (5.36)$$

Будем полагать, что строки матрицы A линейно независимы. В этом случае если $m = n$, то A — квадратная матрица, а СЛАУ имеет одно и только одно решение, равное

$$y = A^{-1}f, \quad (5.37)$$

где A^{-1} — обратная матрица, определение которой:

$$A^{-1}A = E, \quad (5.38)$$

где E — единичная (квадратная) матрица, элементы e_{ij} которой определяются как

$$e_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (5.39)$$

На практике решение СЛАУ обычно находится не по формуле (5.37), а по правилу Крамера, методами Гаусса, Краута-Холецкого и др. [11, 20, 21, 26, 75].

Характеристическое уравнение и типы матриц. *Характеристическое уравнение* для квадратной матрицы (при $m = n$) имеет вид:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0, \quad (5.40)$$

где $|\cdot|$ есть *определитель*. Корни уравнения (5.40) называются *собственными значениями*, или *собственными числами* λ_i , $i = \overline{1, n}$. Собственные значения произвольной квадратной комплексной матрицы, вообще говоря, комплексны. Величины $1/\lambda_i$, $i = \overline{1, n}$, называются *характеристическими числами*.

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 отличных от нуля миноров. Вводится в рассмотрение также $\rho = \text{rang}(A|f)$ — *ранг расширенной матрицы*. С использованием понятия ранга можно следующим образом изложить вопрос о решении СЛАУ (5.36) [11, с. 162]. Если $\rho > r$ (см. (7.2), где $\rho = 3$, а $r = 2$), то СЛАУ не имеет решения и является переопределенной. Если $\rho = r$, то при $\rho = n$ (см. (7.4), где $\rho = r = n = 2$) СЛАУ имеет единственное решение, а при $\rho < n$ (см. (7.3), где $\rho = r = 1 < n = 2$) СЛАУ имеет множество решений и является недоопределенной.

Матрица $A^* = \bar{A}^T = (b_{ij})$, $i = \overline{1, n}$, $j = \overline{1, m}$, называется *сопряженной с A* (или *эрмитово сопряженной*, или *комплексно сопряженной и транспонированной*), если $b_{ij} = \bar{a}_{ji}$ (знак $\bar{\cdot}$ означает комплексное сопряжение, T — транспонирование, а $*$ — эрмитово сопряжение). Квадратная комплексная матрица A называется

эрмитовой (самосопряженной), если $A = A^*$ или $a_{ij} = \bar{a}_{ji}$. Если такая матрица вещественна, то она называется *симметричной*: $A = A^T$ или $a_{ij} = a_{ji}$. У эрмитовой и симметричной матриц все λ_i вещественны (но любого знака). Квадратная матрица A называется *положительно определенной*, если $\sum_{i,j=1}^n a_{ij} y_i y_j > 0$ при любых вещественных y_i . Примеры положительно определенной матрицы: A^*A , AA^* , $A^T A$, AA^T , E (где A , вообще говоря, прямоугольна). У положительно определенной матрицы все λ_i вещественны и неотрицательны.

Для прямоугольной $m \times n$ -матрицы A вместо собственных значений используются *сингулярные числа* — это вещественные неотрицательные числа $\mu_i(A) = \sqrt{\lambda_i(A^*A)}$, $i = \overline{1, n}$, обычно располагаемые в порядке убывания: $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > \mu_{r+1} = \dots = \mu_n = 0$, где r — ранг матрицы. Если $r < n$, то матрица A есть *вырожденная*, или *особенная матрица*; 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, $|A| = \sqrt{|A^*A|} = \mu_1 \mu_2 \dots \mu_n$, а обратные матрицы A^{-1} (при $m = n$) и $(A^*A)^{-1}$ (в общем случае) существуют. При этом в случае $m = n$

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \quad (5.41)$$

где A_{ij} — алгебраические дополнения.

Если A — произвольная (комплексная прямоугольная) матрица, то ее *норма* $\|A\| = \mu(A)_{\max} = \sqrt{\lambda(A^*A)_{\max}}$, а норма обратной матрицы $\|A^{-1}\| = 1/\mu(A)_{\min} = 1/\sqrt{\lambda(A^*A)_{\min}}$. Если A — эрмитова или симметричная матрица, то $\|A\| = \mu(A)_{\max} = |\lambda(A)|_{\max}$, $\|A^{-1}\| = 1/\mu(A)_{\min} = 1/|\lambda(A)|_{\min}$. Если A — положительно определенная матрица, то $\|A\| = \mu(A)_{\max} = \lambda(A)_{\max}$, $\|A^{-1}\| = 1/\mu(A)_{\min} = 1/\lambda(A)_{\min}$.

Нормы векторов и матриц. Нормы комплексных векторов y и f и комплексной квадратной матрицы A определяются как

$$\|y\| = \sqrt{\sum_{j=1}^n |y_j|^2}, \quad \|f\| = \sqrt{\sum_{i=1}^m |f_i|^2}, \quad \|A\| = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}. \quad (5.42)$$

Эти нормы называются *эрмитовыми*. Если же y , f и A вещественны, то нормы называются *евклидовыми*.

Число обусловленности. Пусть вместо точных f и A заданы \tilde{f} и \tilde{A} такие, что $\|\tilde{f} - f\| \leq \delta$, $\|\tilde{A} - A\| \leq \xi$, где δ и ξ — погрешности задания правой части и матрицы. Тогда относительная погрешность решения СЛАУ (5.36) (по формуле (5.37) или др.) может быть оценена в виде следующего неравенства:

$$\frac{\|\delta y\|}{\|y\|} \leq \text{cond}(A) \left(\frac{\delta}{\|f\|} + \frac{\xi}{\|A\|} \right), \quad (5.43)$$

где

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\| = \mu(A)_{\max} / \mu(A)_{\min} \geq 1 \quad (5.44)$$

— *число обусловленности* матрицы A .

Если $\text{cond}(A)$ относительно мало (обычно $\lesssim 10^3$), то матрица A (и СЛАУ) называется *хорошо обусловленной*. Если же $\text{cond}(A)$ относительно велико (обычно $\gtrsim 10^4$), то матрица A (и СЛАУ) называется *плохо обусловленной*. Заметим, что малость (по сравнению с единицей) определителя $|A|$, вообще говоря, не есть критерий плохой обусловленности.

Умножение матриц и векторов. Умножение двух прямоугольных матриц:

$$Z = X \times Y \quad (5.45)$$

$$m \times l \quad m \times n \quad n \times l$$

или подробнее

$$z_{ik} = \sum_{j=1}^n x_{ij} y_{jk}, \quad i = \overline{1, m}, \quad k = \overline{1, l}. \quad (5.46)$$

Умножение матрицы на вектор:

$$f = A \times y \quad (5.47)$$

$$m \times 1 \quad m \times n \quad n \times 1$$

или подробнее

$$f_i = \sum_{j=1}^n a_{ij} y_j, \quad i = \overline{1, m}. \quad (5.48)$$

Примеры. Рассмотрим два примера.

Пример 1. Рассмотрим *систему* двух уравнений с двумя неизвестными:

$$\left. \begin{aligned} y_1 + 2y_2 &= 3, \\ -y_1 + 3y_2 &= 7. \end{aligned} \right\} \quad (5.49)$$

Ее матрица

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad (5.50)$$

является квадратной, поэтому можно рассматривать следующее *характеристическое уравнение*:

$$\begin{vmatrix} 1 - \lambda & 2 \\ -1 & 3 - \lambda \end{vmatrix} = 0, \quad (5.51)$$

откуда $\lambda^2 - 4\lambda + 5 = 0$ и корни равны $\lambda_{1,2} = 2 \pm i$, т. е. *собственные значения* комплексные, что возможно для матрицы (5.50),

не являющейся симметричной или положительно определенной.

Определитель

$$|A| = \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} = 5. \quad (5.52)$$

Обратная матрица, в соответствии с (5.41), равна

$$A^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3/5 & -2/5 \\ 1/5 & 1/5 \end{pmatrix}. \quad (5.53)$$

Проверка (см. (5.38)) с использованием правила (5.46) умножения матриц:

$$A^{-1}A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5.54)$$

Решение СЛАУ (5.49) согласно (5.37) с использованием (5.48) равно

$$y = A^{-1}f = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \quad (5.55)$$

или по правилу Крамера [11, с. 163]

$$y_1 = \frac{\begin{vmatrix} 3 & 2 \\ 7 & 3 \end{vmatrix}}{5} = -1, \quad y_2 = \frac{\begin{vmatrix} 1 & 3 \\ -1 & 7 \end{vmatrix}}{5} = 2. \quad (5.56)$$

Пример 2. Рассмотрим *прямоугольную матрицу* 3×2 :

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \\ 1 & 4 \end{pmatrix}. \quad (5.57)$$

Эрмитово сопряженная матрица A^* равна

$$A^* = \begin{pmatrix} 2 & -1 & 1 \\ -3 & 2 & 4 \end{pmatrix}. \quad (5.58)$$

Произведение этих матриц (согласно (5.46)) дает положительно определенную квадратную матрицу 2×2 :

$$B = A^*A = \begin{pmatrix} 6 & -4 \\ -4 & 29 \end{pmatrix}. \quad (5.59)$$

Характеристическое уравнение для матрицы B :

$$\begin{vmatrix} 6 - \lambda & -4 \\ -4 & 29 - \lambda \end{vmatrix} = 0, \quad (5.60)$$

откуда $\lambda^2 - 35\lambda + 158 = 0$ и *корни* (*собственные значения матрицы* B) равны $\lambda_1(A^*A) \cong 29.675$, $\lambda_2(A^*A) \cong 5.325$, т.е. корни вещественны и неотрицательны, как и должно быть для положительно определенной матрицы A^*A . *Сингулярные числа* матрицы A равны $\mu_1(A) = \sqrt{\lambda_1(A^*A)} \cong 5.47 = \mu(A)_{\max} = \|A\|$,

$\mu_2(A) = \sqrt{\lambda_2(A^*A)} \cong 2.32 = \mu(A)_{\min} = \|A^{-1}\|^{-1}$. Число обусловленности матрицы A равно $\text{cond}(A) = \mu(A)_{\max}/\mu(A)_{\min} \cong 2.36$, т. е. матрица (5.57) является хорошо обусловленной.

Контрольные задания и вопросы

1. Как Вы понимаете линейную независимость строк матрицы?
2. Дайте определения матриц: A^T — транспонированной по отношению к матрице A и \bar{A} — комплексно сопряженной с A .
3. Вспомните определения минора и алгебраического дополнения.
4. Вспомните правило Крамера решения СЛАУ (5.36) при $m = n$.
5. Глядя на формулы (5.45) и (5.47), ответьте, допустимо ли умножение $X_{7 \times 5}$ на $Y_{4 \times 6}$ или $A_{8 \times 9}$ на $y_{11 \times 1}$?
6. Решите следующий пример (СЛАУ):

$$\left. \begin{aligned} 2y_1 - 3y_2 &= -4, \\ -y_1 + 2y_2 &= 3 \end{aligned} \right\} \quad (5.61)$$

(найдите собственные значения матрицы и решение СЛАУ по правилу Крамера и по формуле (5.37)).

7. Используя формулы (5.46) и (5.48), выполните умножение двух матриц и матрицы на вектор:

$$\begin{pmatrix} -2 & 5 & -1 & 0 \\ 0 & 3 & 4 & 3 \\ 1 & 2 & -3 & -4 \end{pmatrix} \begin{pmatrix} 3 & -4 \\ -1 & 0 \\ 0 & 1 \\ 2 & -3 \end{pmatrix} \quad (5.62)$$

и

$$\begin{pmatrix} 2 & -3 \\ -1 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \quad (5.63)$$

5.3. Элементы теории вероятностей

Основные определения. Дадим несколько определений.

Определение 1. Действительная переменная, которая в зависимости от исхода опыта, т. е. в зависимости от случая принимает различные значения, называется *случайной величиной*.

Пусть X — некоторая случайная величина. Тогда имеет место

Определение 2. *Функцией распределения* $F(x)$ случайной величины X называется функция

$$F(x) = P(X < x),$$

где $P(A) \in [0, 1]$ — вероятность события A .

Случайные величины могут быть дискретными и непрерывными. Рассмотрим лишь непрерывные случайные величины.

Определение 3. Случайная величина называется *непрерывной*, если функция распределения может быть представлена в виде:

$$F(x) = \int_{-\infty}^x f(t) dt,$$

где $f(x)$ — плотность распределения.

При этом $f(x) = dF(x)/dx$, $F(\infty) = \int_{-\infty}^{\infty} f(x) dx = 1$, $P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$, $P(X = x_0) = 0$.

Определение 4. *Равномерное распределение* случайной величины x — это распределение с плотностью

$$f(x) = \begin{cases} 1/(b-a) = \text{const}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases} \quad (5.64)$$

В случае равномерного распределения *математическое ожидание* (МО, или первый центральный момент) случайной величины x равно

$$m_0 \equiv MX = \int_{-\infty}^{\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}, \quad (5.65)$$

а *дисперсия* (второй центральный момент) равна

$$\sigma^2 \equiv DX = \int_{-\infty}^{\infty} (x - m_0)^2 f(x) dx = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{dx}{b-a} = \frac{(b-a)^2}{12}, \quad (5.66)$$

откуда *среднеквадратическое отклонение* (СКО) равно

$$\sigma = \frac{b-a}{2\sqrt{3}} \approx 0.577 \cdot \frac{b-a}{2}$$

(см. рис. 5.2).

Определение 5. *Нормальное распределение* (распределение Гаусса) случайной величины x — это распределение с плотностью

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \quad (5.67)$$

(см. рис. 5.3).

Закона (5.67) соответствует МО

$$m_0 = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = a \quad (5.68)$$

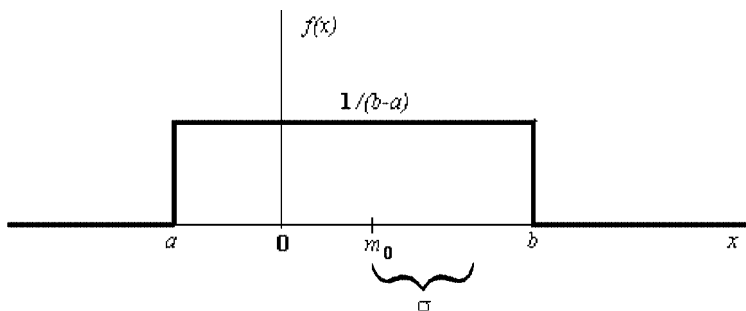


Рис. 5.2. Равномерное распределение

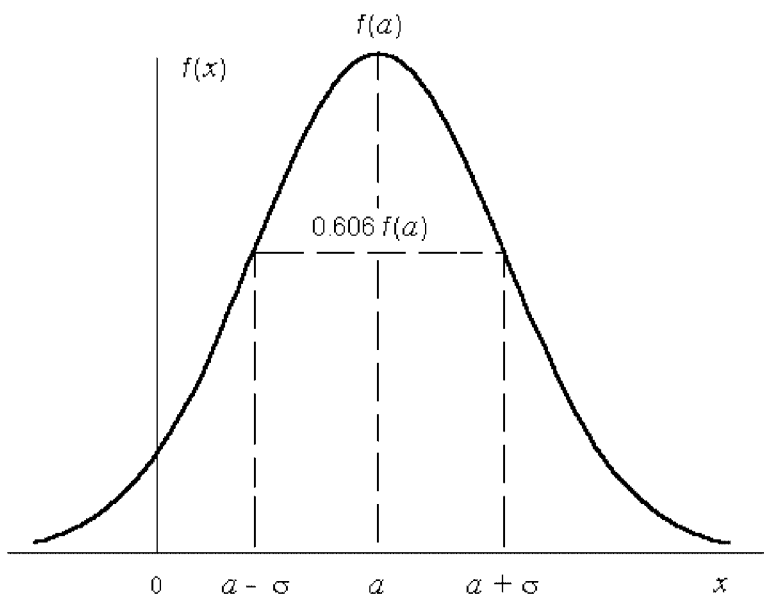


Рис. 5.3. Нормальное распределение

и дисперсия

$$\sigma^2 = \int_{-\infty}^{\infty} \frac{(x-a)^2}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma^2. \quad (5.69)$$

Имеем (см. рис. 5.3): $f(a + \sigma) \approx 0.606 f(a)$.

Если случайная величина X имеет нормальное распределение с параметрами a и σ , то говорят, что X распределена нормально согласно закону $\mathcal{N}(x; a, \sigma)$ или $\mathcal{N}(a, \sigma)$ и пишут: $X \in \mathcal{N}(x; a, \sigma)$ или $X \in \mathcal{N}(a, \sigma)$.

Функция

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

т. е. функция $f(x)$ (см. (5.67)) при $a = 0$, $\sigma = 1$ называется *плотностью нормированного центрированного нормального распределения*. Плотность $\varphi(x)$ и соответствующая ей функция распределения

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

затабулированы. Функцию $\Phi(x)$ часто называют *гауссовым интегралом ошибок*. Иногда задается функция

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt,$$

причем $\Phi(x)$ может быть выражена через $\Phi_0(x)$ посредством соотношения

$$\Phi(x) = \begin{cases} \Phi_0(x) + 0.5, & x \geq 0, \\ -\Phi_0(|x|) + 0.5, & x \leq 0 \end{cases} \quad (5.70)$$

или

$$\Phi(x) = \text{sign}(x) \cdot \Phi_0(|x|) + 0.5. \quad (5.71)$$

Связь между $f(x)$ и $\varphi(x)$, а также $F(x)$ и $\Phi(x)$ имеет вид:

$$f(x) = \frac{1}{\sigma} \varphi\left(\frac{x-a}{\sigma}\right), \quad (5.72)$$

$$F(x) = \Phi\left(\frac{x-a}{\sigma}\right) = \text{sign}(x-a) \cdot \Phi_0\left(\frac{|x-a|}{\sigma}\right) + 0.5. \quad (5.73)$$

Программа RNDAN. На Фортране (и других языках программирования) разработаны стандартные программы, позволяющие имитировать случайные процессы, а именно, выдавать последовательности случайных чисел, подчиненных какому-то закону распределения. Такие программы обычно называют *датчиками случайных чисел*. Одной из наиболее распространенных программ является подпрограмма RNDAN.

В настоящее время наиболее употребимыми для решения прикладных задач языками программирования являются Fortran, C, Pascal, Basic и др. Наиболее совершенными версиями языка Fortran являются MS Fortran, ver. 5.0, 5.1 и Fortran 90 (только в этих версиях имеет место, например, такой исключительно удобный прием, как динамическое распределение памяти в виде операторов ALLOCATE и др.), язык C имеет версии QuickC, TurboC, C++ и т. д. В данном пособии мы будем демонстрировать

результаты, полученные на MS Fortran'e 5.0, 5.1 (совпадающие с результатами на Fortran'e 90), а также на QuickC, ver. 4.2.

Ниже приведены тексты головной программы EXAMPLE1.for (пример обращения к RNDAN.for), подпрограммы RNDAN.for и файла результатов решения EXAMPLE1.dat на Fortran'e, а также головной программы EXAMPLE1.c (пример обращения к RNDAN.c), функции без возвращаемого значения RNDAN.c и файла результатов EXAMPLE1.dat на QuickC. В программах EXAMPLE1.for и EXAMPLE1.c выполнено обращение к RNDAN.for и RNDAN.c при $l=0$ (равномерный закон распределения), $mo=0$, $sig=1$, $ur=0.37843$. Отметим следующее. Параметру ur нужно обязательно присваивать некоторое начальное значение от 0 до 1 перед получением последовательности случайных чисел. При этом вид последовательности случайных чисел будет зависеть от начального значения ur . Далее, сравнение файлов EXAMPLE1.dat на Fortran'e и QuickC показывает, что различные языки программирования выдают несколько различные последовательности случайных чисел (если использовать двойную точность, то различия уменьшаются), однако это не существенно, так как при большом количестве случайных чисел (большой выборке) они подчиняются выбранному закону распределения и его параметрам вне зависимости от языка программирования.

C EXAMPLE1.for:

```

program EXAMPLE1
integer n,l,i
parameter (n=12)
real ur,mo,sig,f(n),df(n),f1(n)
common /urn/ur
ur=.37843 !или другое число от 0 до 1
OPEN(1,file='EXAMPLE1.dat')
WRITE(1,'(//3x,''EXAMPLE1.dat:''')')
PRINT *,' Введите l<1 - закон равномерный',
*      ' или l>=1 - закон норм.:'
READ *,l
WRITE(1,'('' l='',i)')l
PRINT *,' Введите mo>=0 и sig>0:'
READ *,mo,sig
WRITE(1,'('' mo='',f8.4,'' , sig='',f8.4)')mo,sig

```

C Формируем f, df и f1:

```

do 10 i=1,n
  f(i)=3.141593+.5*float(i**2) !или другая формула

```

C Обращение к датчику случайных чисел:

```

  call RNDAN(l,mo,sig,df(i))
10  f1(i)=f(i)+df(i) !зашумленная функция
  WRITE(1,*)f=' !точная функция
  WRITE(1,20)f
20  format(7f9.4)

```

```

WRITE(1,*)'df=' !случайные числа (погрешности, шумы)
WRITE(1,20)df
WRITE(1,*)'f1=' !зашумленная функция
WRITE(1,20)f1
CLOSE(1)
STOP
end ! EXAMPLE1.for

```

```

subroutine RNDAN(l,mo,sig,r) !подпрограмма на Fortran'e

```

```

C*****
C Датчик случайных чисел (модификация Сизикова В.С.) *
C Входные параметры: *
C l<1 - равномерное распределение, *
C l>=1 - нормальное распределение, *
C mo>=0 - математическое ожидание, *
C sig>0 - среднеквадратическое отклонение, *
C Выходной параметр: r - случайное число *
C (от -0.5sig+mo до 0.5sig+mo при l<1) *
C*****
integer l,k,j
real mo,sig,r,ur
common /urn/ur
r=0.
if (l .lt. 1) then
k=1
else
k=12
endif
do 5 j=1,k
f=37.*ur ! от 0 до 37
ur=f-aint(f) ! от 0 до 1
5 r=r+ur-.5 ! от -0.5 до 0.5 при k=1 (l<1)
r=mo+r*sig ! от -0.5sig+mo до 0.5sig+mo при k=1
RETURN
end ! RNDAN.for

```

EXAMPLE1.dat:

```

l=0
mo= .0000, sig= 1.0000
f=
 3.6416  5.1416  7.6416  11.1416  15.6416  21.1416  27.6416
35.1416  43.6416  53.1416  63.6416  75.1416
df=
-.4981  -.4293  .1152  .2639  -.2361  .2627  -.2811
-.3993  .2248  .3179  -.2393  .1441
f1=
 3.1435  4.7123  7.7568  11.4055  15.4054  21.4043  27.3605
34.7423  43.8664  53.4594  63.4022  75.2857

```



```

//EXAMPLE1.c:
#include <stdio.h>
#define N 12
#define NEWLINE    if (i%7==0) fprintf(F1,"\n ")
extern float ur=.37843; //или другое число от 0 до 1
void RNDAN(int l,float mo,float sig,float *r); //прототип
main() //головная программа на QuickC
{ int l,i1; float mo,sig,f[N],df[N],f1[N]; register int i; FILE *F1;
if ((F1=fopen("EXAMPLE1.dat","w")) == NULL)
{ printf("\n Файл F1=EXAMPLE1.dat не открылся на запись"); exit(1);}
fprintf(F1,"\n\n  EXAMPLE1.dat:");
printf("\n Введите l<1 - закон равномерный");
printf(" или l>=1 - закон нормальный: ");
scanf("%i",&l); fprintf(F1,"\n l=%i",l);
printf("\n Введите mo>=0 и sig>0: "); scanf("%f%f",&mo,&sig);
fprintf(F1,"\n mo=%.4f, sig=%.4f",mo,sig);
// Формируем f,df и f1:
for (i=0; i<N; i++) { i1=i+1;
f[i]=3.141593+.5*(float)(i1*i1); //или другая формула
RNDAN(l,mo,sig,&df[i]); //обращение к датчику случайных чисел
f1[i]=f[i]+df[i]; //зашумленная функция
} //end i
fprintf(F1,"\n f[%d]=",N); //точная функция
for (i=0; i<N; i++) { NEWLINE; fprintf(F1,"%9.4f",f[i]); } //end i
fprintf(F1,"\n df[%d]=",N); //случайные числа (погрешности, шумы)
for (i=0; i<N; i++) { NEWLINE; fprintf(F1,"%9.4f",df[i]); } //end i
fprintf(F1,"\n f1[%d]=",N); //зашумленная функция
for (i=0; i<N; i++) { NEWLINE; fprintf(F1,"%9.4f",f1[i]); } //end i
fclose(F1); return 0; } //end MAIN38.c

```

```

//RNDAN.c:
extern float ur;
void RNDAN(int l,float mo,float sig,float *r)
/*Датчик случайных чисел на QuickC (модификация Сизикова В.С.)
Вход: l<1 - закон равномерный, l>=1 - закон нормальный,
      mo>=0 - математическое ожидание,
      sig>0 - СК0 (среднеквадратическое отклонение)
Выход: r - случайное число (от -0.5sig+mo до 0.5sig+mo при l<1)*/
{ int k; register int j; float f;
*r=0.; k = (l<1)? 1 : 12;
for (j=1; j<=k; j++) { f=37.*ur; //от 0 до 37
ur=f-(float)(int)f; //от 0 до 1
*r += ur - .5; //от -0.5 до 0.5 при k=1 (l<1)
} //end j
*r = mo + (*r) * sig; //от -0.5sig+mo до 0.5sig+mo при k=1
} //end RNDAN.c

```

EXAMPLE1.dat:

l=0

mo=0.0000, sig=1.0000

f[12]=

3.6416	5.1416	7.6416	11.1416	15.6416	21.1416	27.6416
35.1416	43.6416	53.1416	63.6416	75.1416		

df[12]=

-0.4981	-0.4293	0.1151	0.2579	-0.4595	-0.0025	-0.0912
-0.3736	0.1761	-0.4845	0.0728	-0.3078		

f1[12]=

3.1435	4.7123	7.7567	11.3994	15.1821	21.1391	27.5504
34.7680	43.8177	52.6571	63.7144	74.8338		

Контрольные задания и вопросы

1. Дать определения случайной величины, функции ее распределения и плотности распределения.
2. Что такое равномерное и нормальное распределения случайной величины?
3. Доказать соотношения (5.65) и (5.66).
4. Доказать соотношение (5.68), используя табличный интеграл:

$$\int_0^{\infty} e^{-p^2 x^2} dx = \sqrt{\pi}/2p \quad (p > 0).$$

5. Доказать соотношение (5.69), используя табличный интеграл:

$$\int_0^{\infty} x^2 e^{-p^2 x^2} dx = \sqrt{\pi}/4p^3 \quad (p > 0).$$

6. Доказать соотношения (5.70)–(5.73).
7. Дать характеристику программы RNDAN — датчика случайных чисел (разъяснить ее входные параметры l, mo, sig и выходной параметр r).

Глава 6

ОБОБЩЕННЫЕ ФУНКЦИИ, ФОРМУЛА ЭЙЛЕРА И
ИНТЕГРАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ

6.1. Элементы теории обобщенных функций

Определение обобщенной функции. В современном математическом аппарате широко используются так называемые *обобщенные функции* [103]. Необходимость их введения была связана с тем, что в классическом математическом анализе постоянно нужно делать оговорки о свойствах функций: при дифференцировании функции нужно требовать, чтобы она была $n \geq 1$ раз дифференцируема, при интегрировании — чтобы она была кусочно-непрерывна и суммируема и т.д. Вместе с тем очень часто реальные функции (в первую очередь, полученные в результате измерений) такими свойствами не обладают. Поэтому возникла необходимость распространить математический анализ (дифференциальное и интегральное исчисления) на такие функции.

Определение. *Обобщенной функцией* называется функция, для дифференцирования и интегрирования которой не требуется выполнения классических свойств дифференцируемости, кусочной непрерывности, суммируемости, интегрируемости и т.д.

Функция Хэвисайда. Исторически первой обобщенной функцией является *функция Хэвисайда* (или ступенчатая функция) [40, с. 676]:

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases} \quad (6.1)$$

(см. рис. 6.1).

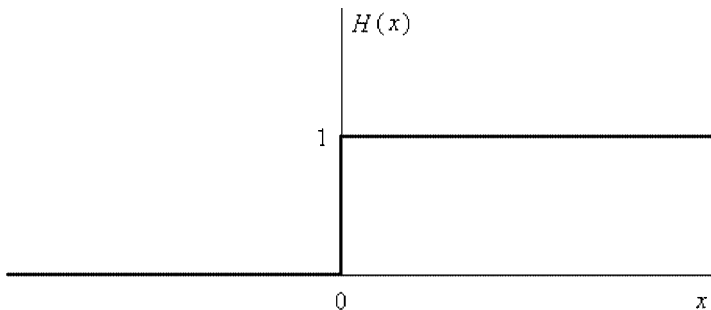


Рис. 6.1

δ -функция Дирака. Исторически следующая обобщенная функция — это δ -функция Дирака (или симметричная единичная импульсная функция) [40, с. 678–680]. Одномерная δ -функция Дирака определяется следующими тремя пунктами:

$$1) \quad \delta(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \neq 0; \end{cases} \quad (6.2)$$

$$2) \quad \int_{-\infty}^{\infty} \delta(x) dx = 1; \quad (6.3)$$

$$3) \quad \int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0), \quad (6.4)$$

где $f(x)$ — некоторая (классическая) функция, а x_0 — некоторое значение абсциссы x .

Пункту 1) (см. (6.2)) соответствует рис. 6.2а, а пункту 3) (см. (6.4)) — рис. 6.2б :

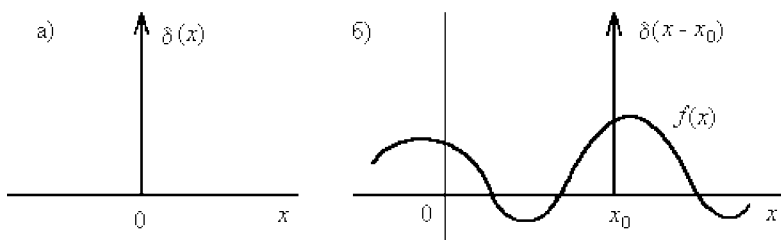


Рис. 6.2

Из (6.3) и (6.4) следуют, в частности, следующие соотношения:

$$\int_{-\infty}^{\infty} \delta(ax - x_0) dx = \frac{1}{|a|}, \quad (6.5)$$

$$\int_{-\infty}^{\infty} f(x) \delta(ax - x_0) dx = \frac{1}{|a|} f\left(\frac{x_0}{a}\right). \quad (6.6)$$

Функция Хэвисайда и δ -функция Дирака связаны соотношениями:

$$\delta(x) = H'(x), \quad (6.7)$$

$$H(x) = \int_{-\infty}^x \delta(t) dt. \quad (6.8)$$

Видим, что эти функции можно дифференцировать и интегрировать (но не в классическом смысле).

Используется также двухмерная δ -функция Дирака:

$$\delta(x, y) = \begin{cases} \infty, & x = y = 0, \\ 0, & \text{иначе,} \end{cases} \quad (6.9)$$

$$\iint_{-\infty}^{\infty} \delta(x, y) dx dy = 1, \quad (6.10)$$

$$\iint_{-\infty}^{\infty} f(x, y) \delta(x - x_0, y - y_0) dx dy = f(x_0, y_0). \quad (6.11)$$

Из (6.10) и (6.11) следуют следующие соотношения:

$$\iint_{-\infty}^{\infty} \delta(ax - x_0, by - y_0) dx dy = \frac{1}{|ab|}, \quad (6.12)$$

$$\iint_{-\infty}^{\infty} f(x, y) \delta(ax - x_0, by - y_0) dx dy = \frac{1}{|ab|} f\left(\frac{x_0}{a}, \frac{y_0}{b}\right). \quad (6.13)$$

Двухмерную δ -функцию можно представить как произведение одномерных:

$$\delta(x, y) = \delta(x) \delta(y). \quad (6.14)$$

Контрольные задания и вопросы

1. Напишите формулу для $\delta(x - x_0)$ типа 6.2 и нарисуйте график.

2. Напишите формулы для $\delta(ax)$ и $\delta(ax - x_0)$ типа 6.2.

3. Чему равны интегралы

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx \quad \text{и} \quad \int_{-\infty}^{\infty} \delta(ax) dx?$$

4. Докажите соотношения (6.5) и (6.6) (используйте замену переменной: $ax = u$).

5. Нарисуйте график $\delta(x, y)$ (см. (6.9)).

6. Напишите формулу для $\delta(x - x_0, y - y_0)$ типа (6.9) и нарисуйте график.

7. Чему равны интегралы

$$\iint_{-\infty}^{\infty} \delta(x - x_0, y - y_0) dx dy \quad \text{и} \quad \iint_{-\infty}^{\infty} \delta(ax, by) dx dy?$$

8. Докажите соотношения (6.12), (6.13) (используйте замены переменных: $ax = u$, $by = v$).

9. Покажите, что при использовании формулы (6.14) останутся в силе формулы (6.9)–(6.13).

6.2. Формула Эйлера

Определение формулы Эйлера. Формула Эйлера имеет вид:

$$Ae^{i\varphi} = A(\cos \varphi + i \sin \varphi). \quad (6.15)$$

Ее можно проиллюстрировать графически. Пусть задана комплексная плоскость xy и на ней окружность радиуса A (см. рис. 6.3). Пусть точка (x, y) на этой окружности задается полярным углом (аргументом) φ . Тогда $x = A \cos \varphi$, $y = A \sin \varphi$, а комплексное число z , соответствующее точке (x, y) , будет равно $z = A(\cos \varphi + i \sin \varphi)$. Итак, формула (6.15) дает связь между показательным и тригонометрическим представлениями комплексного числа (в дополнение к алгебраическому представлению: $z = x + iy$).

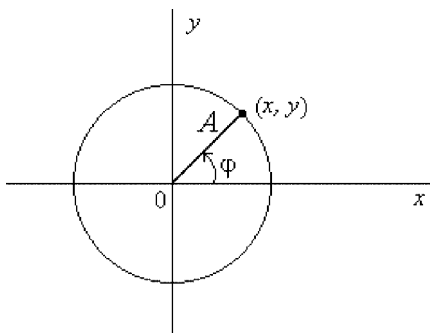


Рис. 6.3

Формула Эйлера записывается также в виде:

$$e^{i\varphi} = \cos \varphi + i \sin \varphi. \quad (6.16)$$

Она находит широкое применение в различных областях математики. В частности, нам далее (см. п. 6.3) потребуется выражение интеграла от формулы Эйлера. Покажем, что он связан с δ -функцией.

Формула Эйлера и δ -функция. Итак, нужно найти интеграл (интеграл Фурье):

$$I(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} dt. \quad (6.17)$$

Используя (6.16), запишем:

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = \int_{-\infty}^{\infty} \cos \omega t dt + i \int_{-\infty}^{\infty} \sin \omega t dt. \quad (6.18)$$

Учтем, что если некоторая функция $f(x)$ является четной, т. е. $f(-x) = f(x)$, то $\int_{-\infty}^{\infty} f(x) dx = 2 \int_0^{\infty} f(x) dx$, а если $f(x)$ является нечетной функцией: $f(-x) = -f(x)$, то $\int_{-\infty}^{\infty} f(x) dx = 0$. В правой части (6.18) первый интеграл — это интеграл от четной функции, а второй — от нечетной функции. Поэтому

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = 2 \int_0^{\infty} \cos \omega t dt. \quad (6.19)$$

При $\omega = 0$ интеграл $\int_0^{\infty} \cos \omega t dt = \infty$, а при $\omega \neq 0$

$$\int_0^{\infty} \cos \omega t dt = \frac{1}{\omega} \lim_{t \rightarrow \infty} \sin \omega t. \quad (6.20)$$

Однако, строго говоря, предел $\lim_{t \rightarrow \infty} \sin \omega t$ не существует. Но $\sin \omega t$ — это функция, осциллирующая вокруг нуля, и ее среднее значение (математическое ожидание) равно нулю. Поэтому если отойти от классического математического анализа, то можно считать, что

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = \begin{cases} \infty, & \omega = 0, \\ 0, & \omega \neq 0, \end{cases} \quad (6.21)$$

т. е. интеграл $\int_{-\infty}^{\infty} e^{i\omega t} dt$ пропорционален δ -функции $\delta(\omega)$. Более точное обоснование показывает, что [40, с. 681]:

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = \int_{-\infty}^{\infty} \cos \omega t dt = 2\pi \delta(\omega). \quad (6.22)$$

Из (6.22) следует интегральное представление δ -функции (ее выражение через интеграл Фурье):

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} dt.$$

Аналогичная формула имеет место в двухмерном случае:

$$\iint_{-\infty}^{\infty} e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2 = 4\pi^2 \delta(\omega_1, \omega_2). \quad (6.23)$$

Полученные формулы (6.22) и (6.23) часто используются в математике (см., например, п. 6.3).

Контрольные задания и вопросы

1. Запишите формулу Эйлера (6.15) или (6.16) при $\varphi = 0, \pi/2, \pi, 3\pi/2, 2\pi$.
2. Обоснуйте (исходя из геометрического смысла интеграла), что в случае четной функции $\int_{-\infty}^{\infty} f(x) dx = 2 \int_0^{\infty} f(x) dx$, а в случае нечетной функции $\int_{-\infty}^{\infty} f(x) dx = 0$.
3. Почему при $\omega = 0$ интеграл $\int_0^{\infty} \cos \omega t dt = \infty$?
4. Обоснуйте геометрически, что при $\omega \neq 0$ интеграл $\int_0^{\infty} \cos \omega t dt$ в среднем равен нулю.
5. Исходя из (6.22) и используя (6.14), докажите (6.23).

6.3. Интегральные преобразования

Рассмотрим интегральные, или непрерывные (а также дискретные и быстрые) преобразования (5.19)–(5.26), которые находят широкое применение в математике и прикладных задачах [9, 11, 15, 16, 19, 33, 40–42, 56, 60, 64, 71, 91, 93, 104–106].

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

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{i\omega t} dt, \quad -\infty < \omega < \infty, \quad (6.24)$$

называется (одномерным) *прямым непрерывным преобразованием Фурье (НПФ)* или *преобразованием Фурье (ПФ)*, *Фурье-образом*, *спектральной функцией*, *спектром*, *изображением по Фурье* 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, поскольку $Y(\omega)$ может быть записана в виде (на основании формулы Эйлера (6.16)):

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos \omega t dt + i \int_{-\infty}^{\infty} y(t) \sin \omega t dt \quad (6.25)$$

или

$$Y(\omega) = \operatorname{Re} Y(\omega) + i \operatorname{Im} Y(\omega), \quad (6.26)$$

где

$$\operatorname{Re} Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos \omega t dt, \quad (6.27)$$

$$\operatorname{Im} Y(\omega) = \int_{-\infty}^{\infty} y(t) \sin \omega t dt. \quad (6.28)$$

Соотношение (6.27) называется *косинус-преобразованием Фурье*, а (6.28) — *синус-преобразованием Фурье*. Используют часто также квадрат модуля ПФ:

$$|Y(\omega)|^2 = \operatorname{Re}^2 Y(\omega) + \operatorname{Im}^2 Y(\omega). \quad (6.29)$$

Рассмотрим несколько примеров.

Пример 1: $y(t) = \cos ct$ (исходный процесс — одна гармоника). Учитывая, что интеграл от нечетной функции равен нулю и заменяя произведение косинусов их суммой, получим:

$$\begin{aligned} Y(\omega) &= \int_{-\infty}^{\infty} \cos ct \cos \omega t dt + i \int_{-\infty}^{\infty} \cos ct \sin \omega t dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \cos(\omega + c)t dt + \frac{1}{2} \int_{-\infty}^{\infty} \cos(\omega - c)t dt. \end{aligned} \quad (6.30)$$

Используя формулу (6.22), окончательно имеем:

$$Y(\omega) = \operatorname{Re} Y(\omega) = \pi \delta(\omega + c) + \pi \delta(\omega - c). \quad (6.31)$$

Результат — на рис. 6.4.

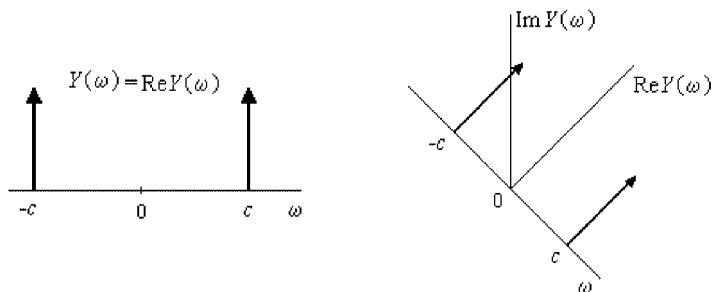


Рис. 6.4

Пример 2: $y(t) = e^{-a^2 t^2}$ (гауссиана). Имеем:

$$\begin{aligned}
 Y(\omega) &= \int_{-\infty}^{\infty} e^{-a^2 t^2} e^{i\omega t} dt = \int_{-\infty}^{\infty} e^{-a^2 t^2} \cos \omega t dt + i \int_{-\infty}^{\infty} e^{-a^2 t^2} \sin \omega t dt = \\
 &= 2 \int_0^{\infty} e^{-a^2 t^2} \cos \omega t dt = \frac{\sqrt{\pi}}{a} e^{-\omega^2/4a^2} = \operatorname{Re} Y(\omega).
 \end{aligned}$$

Здесь мы использовали табличный интеграл $\int_0^{\infty} e^{-a^2 t^2} \cos \omega t dt$ [11, с. 110]. Результат при различных значениях a — на рис. 6.5 и на рис. 6.6 (при некотором одном значении a).

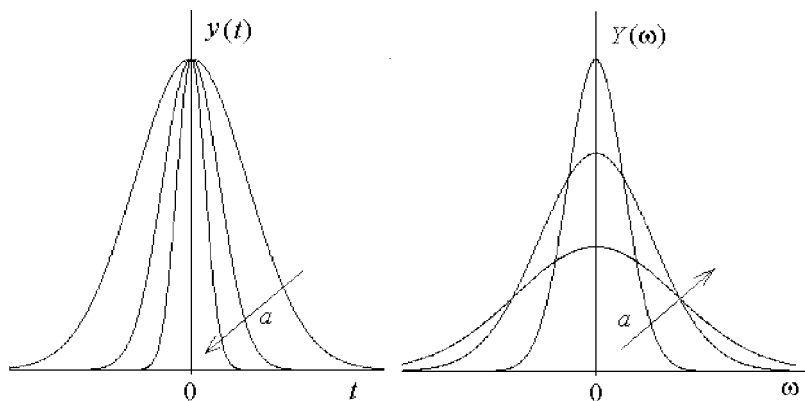


Рис. 6.5

Мы видим, что $y(t)$ и $Y(t)$ — гауссианы, причем чем уже $y(t)$ (чем больше a), тем шире спектр $Y(\omega)$. Это — общая закономерность.

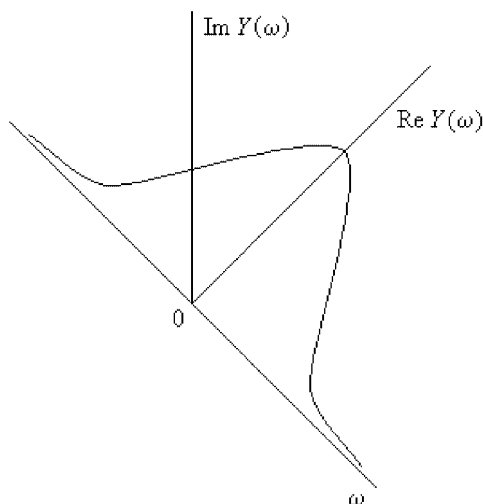


Рис. 6.6

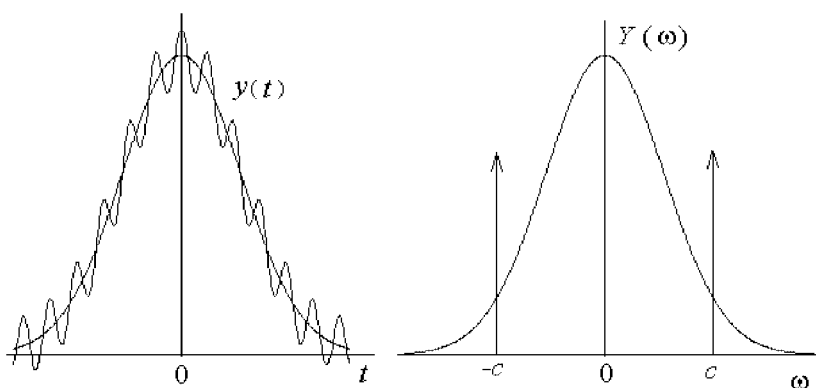


Рис. 6.7

Пример 3: $y(t) = e^{-a^2 t^2} + b \cos ct$ (гауссиана + одна гармоника). Имеем:

$$Y(\omega) = \frac{\sqrt{\pi}}{a} e^{-\omega^2/4a^2} + \pi b \delta(\omega + c) + \pi b \delta(\omega - c) \quad (6.32)$$

(см. рис. 6.7).

Пример 4: $y(t) = e^{-a^2 t^2} + \int_{-\infty}^{\infty} b(c) \cos(ct) dc$ (гауссиана + бесконечный набор гармоник, т.е. гауссиана + случайный процесс — см. рис. 6.8а)).

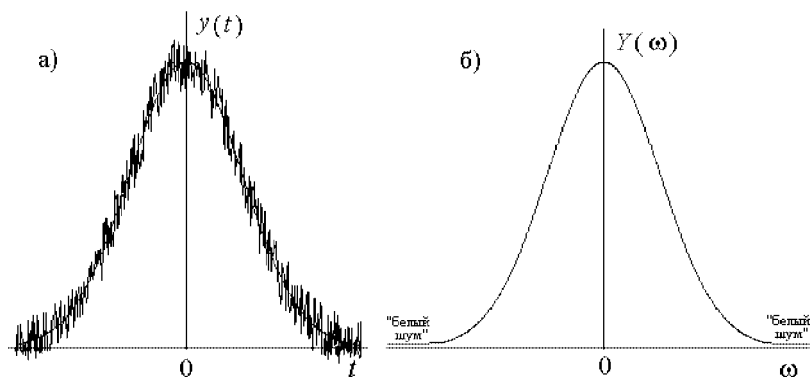


Рис. 6.8

Спектр $Y(\omega)$ равен:

$$Y(\omega) = \frac{\sqrt{\pi}}{a} e^{-\omega^2/4a^2} + \pi [b(-\omega) + b(\omega)]. \quad (6.33)$$

Если $b(\omega) \rightarrow \text{const}$ («белый шум») при $\omega \rightarrow \infty$, как это часто характерно для зашумленных процессов, то $Y(\omega)$ будет иметь вид, изображенный на рис. 6.8б.

Теперь рассмотрим вопрос об обращении формулы (6.24), т. е. по-существу, о решении интегрального уравнения Фредгольма I рода (6.24) относительно $y(t)$.

Связь между прямым и обратным ПФ. Запишем (6.24) в виде

$$Y(\omega) = \int_{-\infty}^{\infty} y(t') e^{i\omega t'} dt', \quad (6.34)$$

умножим (6.34) на $e^{-i\omega t}$, проинтегрируем по ω от $-\infty$ до ∞ и изменим порядок интегрирования. Получим:

$$\int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega = \int_{-\infty}^{\infty} y(t') \left[\int_{-\infty}^{\infty} e^{i(t'-t)\omega} d\omega \right] dt'. \quad (6.35)$$

Из формулы (6.22) следует, что

$$\int_{-\infty}^{\infty} e^{i(t'-t)\omega} d\omega = 2\pi \delta(t' - t), \quad (6.36)$$

и тогда, используя (6.4), получим:

$$\int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega = 2\pi \int_{-\infty}^{\infty} y(t') \delta(t' - t) dt' = 2\pi y(t), \quad (6.37)$$

откуда окончательно

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega. \quad (6.38)$$

Итак, исходя из прямого ПФ (6.24), мы получили выражение для *обратного преобразования Фурье (ОПФ)* $y(t)$ и это можно рассматривать как результат аналитического решения интегрального уравнения Фредгольма I рода (6.24).

Отметим при этом, что часто в (6.24) ставят знак «-» перед i . В этом случае изменится знак перед i в (6.38). Кроме того, в (6.24) перед интегралом нередко ставят множитель $1/\sqrt{2\pi}$. В этом случае такой же множитель будет перед интегралом в (6.38). Другими словами, (6.24) и (6.38) можно записывать в виде:

$$Y(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(t) e^{\pm i\omega t} dt, \quad (6.39)$$

$$y(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Y(\omega) e^{\mp i\omega t} d\omega. \quad (6.40)$$

Однако различие в записях не ведет к различию вычисляемых $Y(\omega)$ и $y(t)$. Поэтому можно пользоваться любой записью. В данном учебном пособии используются соотношения (6.24) и (6.38).

Помимо комплексного ПФ используются также:
косинус-преобразование Фурье (прямое и обратное):

$$Y_c(\omega) = \int_{-\infty}^{\infty} y(t) \cos(\omega t) dt, \quad (6.41)$$

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_c(\omega) \cos(\omega t) d\omega \quad (6.42)$$

и *синус-преобразование Фурье* (прямое и обратное):

$$Y_s(\omega) = \int_{-\infty}^{\infty} y(t) \sin(\omega t) dt, \quad (6.43)$$

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_s(\omega) \sin(\omega t) d\omega. \quad (6.44)$$

Двухмерное ПФ. Пусть задана двухмерная функция $y(t_1, t_2)$. Тогда двойной интеграл

$$Y(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} y(t_1, t_2) e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2 \quad (6.45)$$

есть *двухмерное прямое непрерывное преобразование Фурье*, а формула обращения, аналогичная (6.38):

$$y(t_1, t_2) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} Y(\omega_1, \omega_2) e^{-i(\omega_1 t_1 + \omega_2 t_2)} d\omega_1 d\omega_2, \quad (6.46)$$

дает *двухмерное обратное непрерывное преобразование Фурье*.

Преобразование Фурье (одномерное и двухмерное) широко используется для спектрального анализа временных процессов (см. п. 4.3), для решения интегральных уравнений Вольтерры и Фредгольма I, II и III рода типа свертки одномерных и двухмерных (см. п. 1.1, 2.1, 2.2, 2.4), для решения дифференциальных уравнений [64] и т.д.

Дискретное ПФ. На практике значения функций y и Y задаются и/или вычисляются не непрерывно, а на дискретных сетках узлов, и интегралы в вышеприведенных соотношениях заменяются конечными суммами, т.е. непрерывное преобразование Фурье (НПФ) реализуется в виде дискретного преобразования Фурье (ДПФ). Рассмотрим это на примере вычисления преобразования Фурье согласно (6.24). При этом мы будем пользоваться двойной терминологией (и обозначениями): согласно [41, 56], а также (в скобках) терминологией в стиле данного учебного пособия.

Пусть снимаются дискретные отсчеты $y(t)$ на равномерной сетке узлов:

$$t_k = kh, \quad k = \overline{0, N-1}, \quad (6.47)$$

где $h = \Delta t = \text{const}$ — интервал (шаг) дискретизации по t , а N — число отсчетов.

Справедлива *теорема Котельникова* [41], в силу которой:

1) линейная частота дискретизации (максимальная линейная частота) в ДПФ равна

$$f_g \equiv f_{\max} = \frac{1}{h} \quad (6.48)$$

(заметим, что ω — это круговая частота, связанная с линейной частотой f соотношением: $\omega = 2\pi f$);

2) интервал (шаг) дискретизации по частоте равен

$$\Delta f = \frac{f_{\max}}{N} = \frac{1}{h \cdot N} = \frac{1}{t_{\max}}, \quad (6.49)$$

где $t_g \equiv t_{\max} = t_N = h \cdot N$ — длина выборки;

$$3) \quad \omega_l = 2\pi f_l, \quad l = \overline{0, N-1}, \quad f_l = l \cdot \Delta f = \frac{l}{h \cdot N}. \quad (6.50)$$

Интеграл в (6.24) заменяем конечной суммой по формуле левых прямоугольников с шагом h (см. рис. 6.9) и в результате вместо непрерывного соотношения (6.24) получим дискретное:

$$Y_l \equiv Y(\omega_l) = h \sum_{k=0}^{N-1} y_k e^{i 2\pi f_l k h}, \quad (6.51)$$

откуда

$$Y_l = h \sum_{k=0}^{N-1} y_k e^{i 2\pi l k / N}, \quad l = \overline{0, N-1}, \quad (6.52)$$

где $y_k = y(t_k)$.

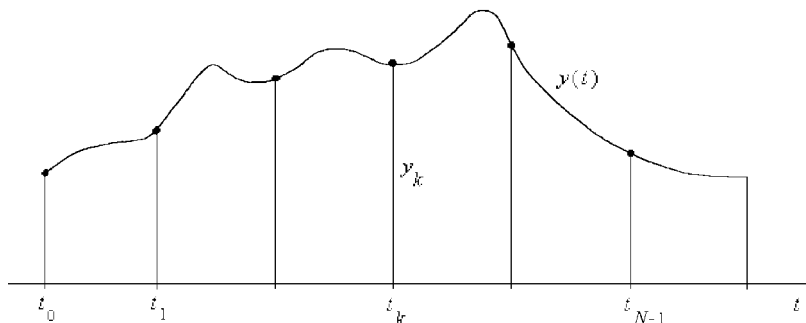


Рис. 6.9

Получили формулу (6.52) для *дискретного преобразования Фурье (ДПФ)* в случае, когда шаги дискретизации по t и ω постоянны и число отсчетов N по t и ω одинаково.

Аналогично обратному непрерывному преобразованию Фурье (ОПФ) (6.38) будет соответствовать *обратное дискретное преобразование Фурье (ОДПФ)*:

$$y_k \equiv y(t_k) = \frac{1}{h \cdot N} \sum_{l=0}^{N-1} Y_l e^{-i 2\pi l k / N}, \quad k = \overline{0, N-1}. \quad (6.53)$$

Исследуем формулу (6.52) и сравним ее с формулой (6.24). Результаты будем формулировать в виде выводов.

Вывод 1. В формуле (6.52) не фигурируют значения узлов ω_l и t_k , а фигурируют лишь номера узлов l и k . Это значительно упрощает и убыстряет вычисления, а в конечном итоге позволило

создать алгоритм быстрого преобразования Фурье (БПФ) (Куль и Тьюки, 1956 [16, 33, 56]), см. дальше.

Далее, в (6.51) к f_l добавим $n \geq 0$ значений $f_g = 1/h$, получим:

$$e^{i2\pi(f_l+n/h)kh} = e^{i2\pi f_l kh} e^{i2\pi nk} = e^{i2\pi lk/N}, \quad n = 0, 1, 2, \dots \quad (6.54)$$

Вывод 2. Y_l - периодическая функция с периодом $f_g = 1/h$.

Вывод 2 порождает следующие следствия. Пусть $Y(\omega)$ или $Y(f)$ — финитная функция, равная нулю при $|\omega| > \omega_v$ или $|f| > f_v$, где f_v — верхняя частота (см. рис. 6.10, где под $Y(f)$ подразумевается или $\text{Re}Y(f)$, или $|Y(f)|$ и т.д.).

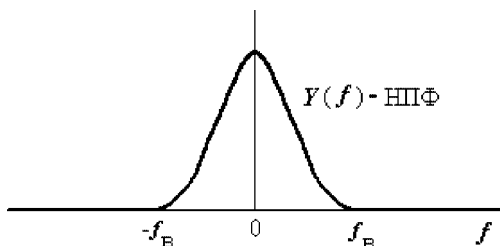


Рис. 6.10

Тогда при $f_g > 2f_v$ или $h < 1/2f_v$ получим вместо НПФ (рис. 6.10) ДПФ согласно (6.52) — см. рис. 6.11, где непрерывная линия соответствует значениям $l \in [0, N-1]$, а пунктир — значениям $l \notin [0, N-1]$. Видим, что ДПФ заметно отличается от НПФ.

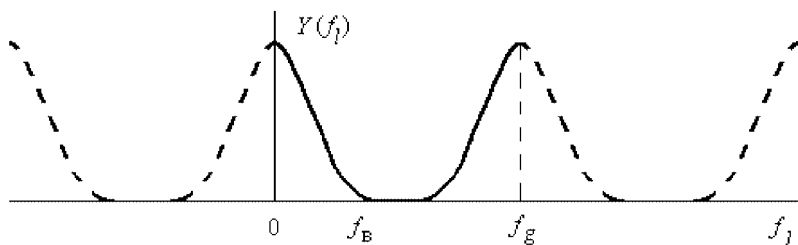


Рис. 6.11

При $f_g = 2f_v$ или $h = 1/2f_v$ имеем (см. рис. 6.12):

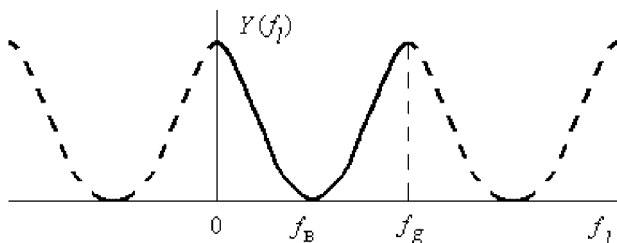


Рис. 6.12

А при $f_g < 2f_B$ или $h > 1/2f_B$ имеем так называемый *эффект наложения*, искажающий спектр (см. рис. 6.13):

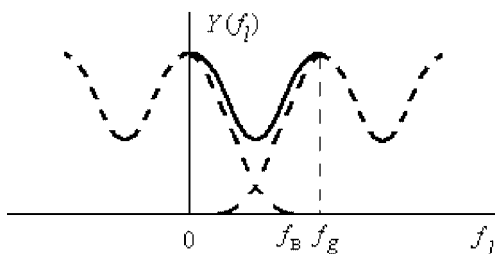


Рис. 6.13

Поскольку при $h > 1/2f_B$ имеет место эффект наложения, то шаг h дискретизации по t нужно брать в соответствии с неравенством:

$$h \leq \frac{1}{2f_B},$$

оценив при этом каким-то образом f_B .

Если добавляются отсчеты t_k изнутри области $[0, t_g)$, т.е. уменьшается h при неизменном t_g , то f_g увеличивается (см. (6.48)) и эффект наложения снижается, однако шаг дискретизации по частоте $\Delta f = 1/t_g$ не изменяется и, следовательно, не изменяется разрешение по частоте f .

Если же добавляются отсчеты t_k снаружи области $[0, t_g)$, т.е. при неизменном h увеличивается t_g , то $f_g = 1/h$ не изменяется и, следовательно, сохраняется эффект наложения, но уменьшается $\Delta f = 1/t_{\max}$, т.е. повышается разрешение по f .

Итак, *основные свойства ДПФ*:

- 1) Периодичность с периодом f_g .
- 2) Эффект наложения (при $f_g < 2f_B$).
- 3) Изменение эффекта наложения и разрешения по f при добавлении новых отсчетов по t .

Об алгоритмах БПФ. На практике ДПФ обычно реализуется в виде алгоритма быстрого преобразования Фурье (БПФ) (Fast Fourier Transform - FFT). Впервые такой алгоритм предложили Куль и Тьюки; в этом алгоритме число отсчетов по t и по f равно $N = 2^m$, где $m \in \mathbb{N}$, например, $N = 1024$; если обычное ДПФ требует N^2 операций, то БПФ требует $N \log_2 N$ операций (если $N = 1024$, то соответственно $\approx 10^6$ и $\approx 10^4$ операций — разница существенная).

Затем ряд быстрых алгоритмов предложил Виноград; в его алгоритмах N — произведение взаимно простых чисел.

К настоящему времени разработано много *стандартных программ* (СП) для БПФ, например, FFT [56], FTFC [71, с. 183, 190].

Что касается *двухмерного НПФ* (см. (6.45)), то оно на практике реализуется в виде двухмерного ДПФ, которое может быть записано как набор одномерных ДПФ [71, с. 45]. Например, двухмерное НПФ (6.45) может быть записано в виде:

$$Y(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y(t_1, t_2) e^{i\omega_1 t_1} dt_1 \right] e^{i\omega_2 t_2} dt_2,$$

т. е. в виде одномерного НПФ от одномерного НПФ. А двухмерное ДПФ можно записать в виде (см. (6.52)):

$$\begin{aligned} Y_{lj} &= h_1 h_2 \sum_{k=0}^{N-1} \sum_{m=0}^{M-1} y_{km} e^{i2\pi(lk/N + jm/M)} = \\ &= h_2 \sum_{m=0}^{M-1} \left(h_1 \sum_{k=0}^{N-1} e^{i2\pi lk/N} \right) e^{i2\pi jm/M}, \end{aligned} \quad (6.55)$$

т. е. в виде одномерного ДПФ от одномерного ДПФ. В [71, с. 190] есть СП FTFTC для вычисления двухмерного БПФ.

Использование регуляризации. Задача вычисления ПФ согласно (6.24), (6.38), (6.41)–(6.46), (6.52), (6.53) и т. д. является, строго говоря, некорректной (неустойчивой), так как связана с решением интегрального уравнения Фредгольма I рода, хотя «степень некорректности» в значительной степени снижается благодаря тому, что это уравнение решается аналитически.

В работе [60] для повышения устойчивости использован метод регуляризации n -го порядка Тихонова ($n \in \mathbb{N}$, т. е. $n = 1, 2, \dots$). В результате, например, вместо формулы (6.24) нужно использовать формулу

$$Y_\alpha(\omega) = \int_{-\infty}^{\infty} \frac{y(t)}{1 + \alpha t^{2n}} e^{i\omega t} dt, \quad (6.56)$$

где $\alpha > 0$ — параметр регуляризации.

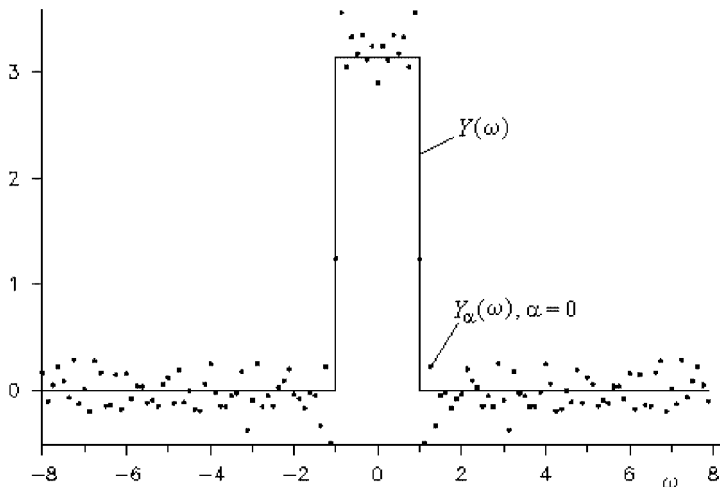


Рис. 6.14

Приведем пример [60], иллюстрирующий эффект использования метода регуляризации n -го порядка Тихонова для повышения устойчивости вычисления ПФ. В данном примере исходная функция $y(t) = \sin(\omega_z t)/(\omega_z t)$. В этом случае $Y(\omega)$, согласно (6.24), равно (см. рис. 6.14)

$$Y(\omega) = \operatorname{Re} Y(\omega) = \begin{cases} \gamma, & \omega \in (-\omega_z, \omega_z), \\ 0, & \omega \notin [-\omega_z, \omega_z], \\ \gamma/2, & \omega = \pm\omega_z, \end{cases}$$

где $\gamma = \pi/\omega_z$. Положено: $\omega_z = 1$, $t_z = \pi/\omega_z = \pi$ — первый ноль функции $y(t)$, число отсчетов $N = 128$, $h = \Delta t = t_z/8$, $t_{\max} = hN = 50.26549$, $a = -b$, $b = t_{\max}/2$, $\omega_{\max} = 2\pi/h = 16$, $\omega_z/\Delta\omega = 8$, $\omega_{\max}/2\omega_z = 8$ (т. е. эффект наложения устранен с восьмикратным запасом).

К дискретным отсчетам y_k были добавлены поточечные погрешности $\delta_1 \in \mathcal{N}(0, 0.05)$, т. е. 5% от y_{\max} . На рис. 6.14 точки — результат расчета $Y_\alpha(\omega_l)$ при $\alpha = 0$, т. е. без регуляризации, а на рис. 6.15 — $Y_\alpha(\omega_l)$ при $n = 1$, $\alpha = 1.6 \cdot 10^{-2}$, причем α выбрано по способу невязки. Расчеты выполнены с помощью программ пакета FFTREG [61].

Из рис. 6.14 видно, что вместо $Y(\omega)$ в виде прямоугольного окна получается окно с осцилляциями (эффект Гиббса), а из рис. 6.15 видно, что регуляризация уменьшает погрешность вычисления $Y(\omega)$ (а значит, увеличивает отношение сигнал/помеха) в 2–3 раза.

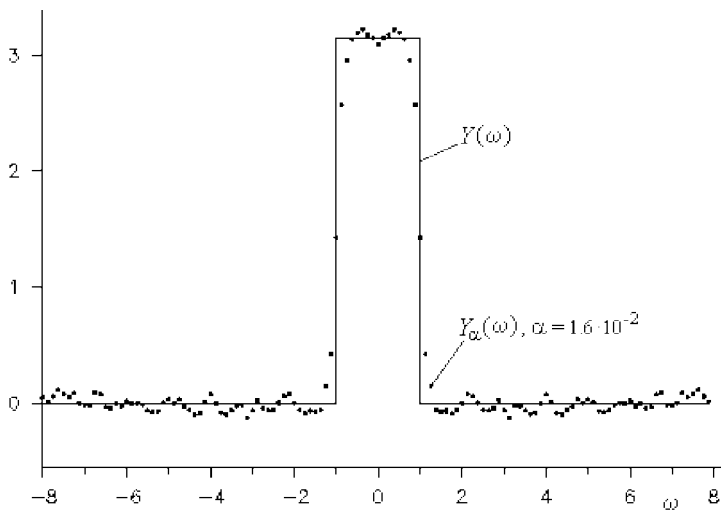


Рис. 6.15

Причина этого состоит в том, что слагаемое αt^{2n} в (6.56) подавляет (умеренно) дальние отсчеты в $y(t)$, чьи погрешности вносят наибольший вклад в погрешности $Y_\alpha(\omega)$.

Преобразование Хартли. Недостатком преобразования Фурье (6.24) является комплексность формул даже в случае вещественности оригинала $y(t)$. Это ведет к тому, что ПФ $Y(\omega)$ получается комплексным, а значит, излишне загружает память компьютера. Еще один нежелательный эффект — это комплексность решений интегральных уравнений Фредгольма I рода типа свертки одно- и двумерных (1.8), (1.16), (1.22), (1.59), (2.15), (2.30), (4.19), (4.21), (4.23)–(4.25) методом ПФ (см. (1.10), (1.17), (1.23), (1.60), (2.18), (2.34)) и др., хотя решения в большинстве случаев должны быть вещественны.

Чтобы избежать комплексности, можно использовать косинус- и синус-преобразования Фурье (6.41)–(6.44) или же преобразование Хартли.

Одномерное непрерывное преобразование Хартли — ПХ (1942 г.) определяется формулой [9, 16] (ср. (6.24), (6.41), (6.43)):

$$Y_H(\omega) = \int_{-\infty}^{\infty} y(t) \operatorname{cas}(\omega t) dt, \quad -\infty < \omega < \infty, \quad (6.57)$$

где (вещественная) функция cas определяется следующим образом:

$$\operatorname{cas} x = \cos x + \sin x \quad (6.58)$$

или

$$\operatorname{cas}(\omega t) = \cos(\omega t) + \sin(\omega t). \quad (6.59)$$

Обратное преобразование Хартли (ОПХ) имеет вид (ср. (6.38), (6.42), (6.44)):

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_H(\omega) \operatorname{cas}(\omega t) d\omega. \quad (6.60)$$

При этом формулы (6.57), (6.60) могут быть записаны в стиле формул (6.39), (6.40) [9, с. 16]:

$$Y_H(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(t) \operatorname{cas}(\omega t) dt, \quad (6.61)$$

$$y(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Y_H(\omega) \operatorname{cas}(\omega t) d\omega. \quad (6.62)$$

Формулы (6.61), (6.62) имеют совершенно одинаковые записи в отличие от формул (6.39), (6.40) преобразования Фурье.

Пример. В качестве примера, иллюстрирующего сходства и различия ПФ и ПХ, рассмотрим смещенную прямоугольную функцию (прямоугольный импульс, стробирующую функцию) [9, с. 20–21] (см. (2.17)):

$$y(t) = \begin{cases} 1/\Delta, & -\Delta \leq t \leq 0, \\ 0, & \text{иначе.} \end{cases} \quad (6.63)$$

ПФ равно (см. (2.22))

$$Y(\omega) = \frac{\sin(\omega\Delta)}{\omega\Delta} + \frac{\cos(\omega\Delta) - 1}{\omega\Delta} i, \quad (6.64)$$

а ПХ согласно (6.57) равно

$$\begin{aligned} Y_H(\omega) &= \frac{1}{\Delta} \int_{-\Delta}^0 \cos(\omega t) dt + \frac{1}{\Delta} \int_{-\Delta}^0 \sin(\omega t) dt = \\ &= \frac{\sin(\omega\Delta)}{\omega\Delta} + \frac{\cos(\omega\Delta) - 1}{\omega\Delta}. \end{aligned} \quad (6.65)$$

На рис. 6.16а представлен оригинал $y(t)$, на рис. 6.16б — вещественная и мнимая части преобразования Фурье $\operatorname{Re} Y(\omega)$ и $\operatorname{Im} Y(\omega)$, а на рис. 6.16в — преобразование Хартли $Y_H(\omega)$.

Рис. 6.16 демонстрирует сходство между ПХ и ПФ:

$$Y_H(\omega) = \operatorname{Re} Y(\omega) + \operatorname{Im} Y(\omega), \quad (6.66)$$

т. е. ПХ есть сумма вещественной и мнимой частей ПФ (без множителя i), но в этом заключено и их различие (это показывают рис. 6.16б и 6.16в).

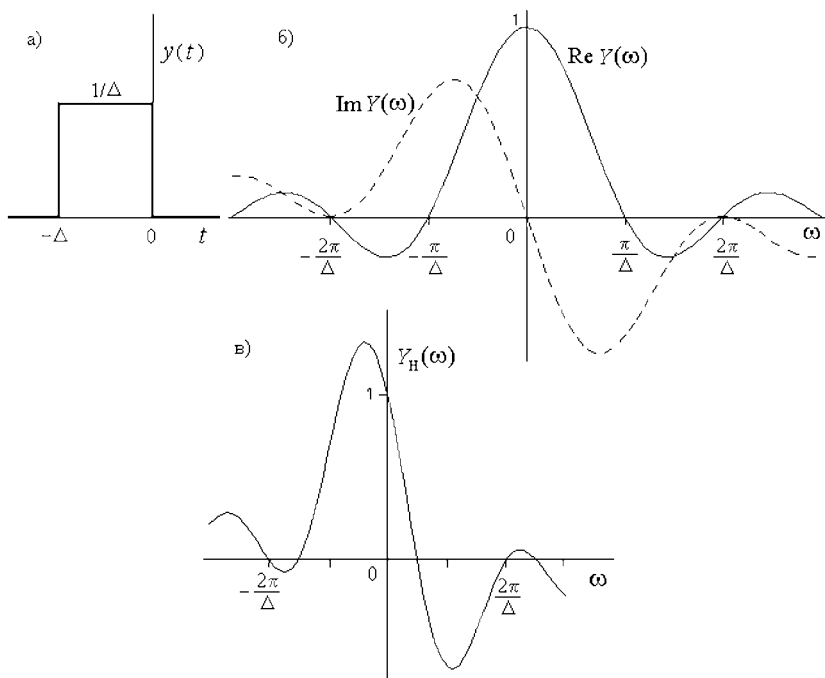


Рис. 6.16

Заметим, что если для ПФ используется формула (ср. (6.39))

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{-i\omega t} dt, \quad (6.67)$$

то вместо (6.66) будем иметь формулу:

$$Y_H(\omega) = \text{Re } Y(\omega) - \text{Im } Y(\omega). \quad (6.68)$$

Непрерывное ПХ расписано до дискретного преобразования Хартли (ДПХ) [9, с. 34]. Для ДПХ, как и для ДПФ, имеют место следующие свойства: периодичность, эффект наложения и др.

Для повышения скорости выполнения ДПХ разработан алгоритм *быстрого преобразования Хартли* (БПХ) [9, с. 91–120], а также разработаны стандартные программы: FHTBAS и FHTSUB на Бейсике [9, с. 132–139] и FHTBAS и FHTFOR на Фортране [9, с. 156–162].

Рассмотрено также двухмерное ПХ [9, с. 65–73].

Преобразование Лапласа. Лежащее в основе операторного метода *одностороннее прямое преобразование Лапласа* есть интегральное преобразование вида [19, с. 94]:

$$\Psi(p) = \int_0^{\infty} \varphi(x) e^{-px} dx, \quad (6.69)$$

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 $\Psi(p) \rightarrow \varphi(x)$ или $\Psi(p) = L[\varphi(x)]$.

Двустороннее преобразование Лапласа отличается от (6.69) нижним пределом интегрирования, равным $-\infty$. Обычно под термином «преобразование Лапласа» подразумевается одностороннее преобразование (6.69).

Оригинал $\varphi(x)$ должен удовлетворять следующим условиям:

а) $\varphi(x)$ — кусочно-непрерывная функция,

б) $\varphi(x) = 0$ при $x < 0$,

в) $|\varphi(x)| < Me^{cx}$ при $x > 0$, где $M > 0$ и $c \geq 0$ — некоторые константы, причем если $|\varphi(x)| \leq |\varphi(0)|$, то $c = 0$.

Тогда справедливо *обратное преобразование Лапласа*, дающее выражение для оригинала:

$$\varphi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Psi(p) e^{px} dp. \quad (6.70)$$

На рис. 6.17 пунктиром отображена прямая, по которой идет интегрирование в (6.70), если $\Psi(p)$ не имеет особых точек, иначе путь интегрирования в (6.70) лежит справа от всех особых точек $\Psi(p)$.

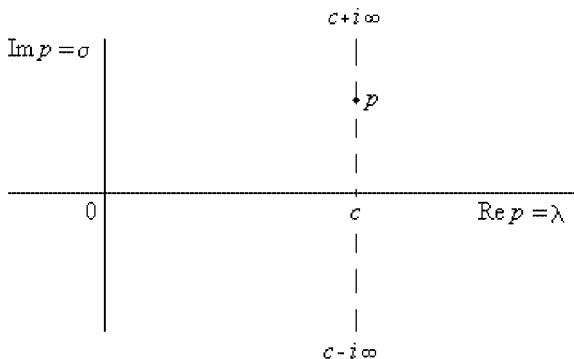


Рис. 6.17

Пример. Пусть задано изображение $L[\varphi(x)] = \Psi(p) = 1/p$, а в отношении оригинала $\varphi(x)$ известно, что $c = 0$. Тогда $p = i\sigma$, $\lambda = 0$ и, в соответствии с (6.70),

$$\begin{aligned} \varphi(x) &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{1}{p} e^{px} dp = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{\sigma} e^{i\sigma x} d\sigma = \\ &= \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\cos \sigma x}{\sigma} d\sigma + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin \sigma x}{\sigma} d\sigma. \end{aligned} \quad (6.71)$$

Первый интеграл в (6.71) равен нулю, так как подынтегральная функция нечетна, а второй (табличный) интеграл равен 2π . В результате

$$\varphi(x) = 1. \quad (6.72)$$

Метод, основанный на применении ПЛ (метод ПЛ, операционный метод) широко используется (наряду с методом ПФ) для решения интегральных уравнений типа свертки, линейных дифференциальных уравнений с постоянными коэффициентами и интегро-дифференциальных уравнений [19, с. 94–105, 137–138].

Контрольные задания и вопросы

1. Запишите одномерные прямые непрерывные преобразования: преобразование Фурье, косинус-преобразование Фурье, преобразование Хартли, преобразование Лапласа и дайте им сравнительную характеристику.
2. Аналогичную запись и характеристику сделайте для обратных преобразований.
3. Решить пример 5: Найти ПФ $Y(\omega)$ для $y(t) = \sin ct$. Получить формулу типа (6.31) и построить график типа рис. 6.4.
4. Решить пример 6: Найти ПФ $Y(\omega)$ для

$$y(t) = e^{-a^2 t^2} + \sum_l b_l \cos(c_l t).$$

Получить формулу типа (6.32) и построить график типа рис. 6.7.

5. В примере 4 дать более подробный вывод формулы (6.33).
6. Исходя из формулы (6.38) для ОПФ, вывести формулу (6.24) для прямого ПФ (аналогично выводу формулы (6.38) из (6.34)).
7. Исходя из формулы (6.39), вывести формулу (6.40), из формулы (6.41) вывести формулу (6.42), из формулы (6.43) вывести (6.44), а из формулы (6.57) вывести (6.60).

8. Из формулы (6.45) вывести формулу (6.46) и наоборот (аналогично выводу формулы (6.38) из (6.34), но используя вместо одномерной формулы (6.36) двухмерную формулу типа (6.23)).

9. Дайте физическую трактовку формул (6.48) и (6.49).

10. Если НПФ $Y(f)$ имеет вид :

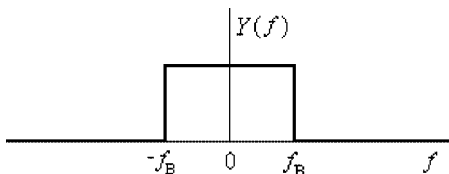


Рис. 6.18

то как будет выглядеть ДПФ при $f_g > 2f_B$?

11. Напишите, какие значения может принимать N в алгоритмах БПФ Кули и Тьюки и Винограда.

12. Более подробно вывести формулу (6.65).

Глава 7

ПРЕДЫСТОРИЯ РЕГУЛЯРНЫХ МЕТОДОВ

В данной главе изложены методы и понятия, предшествующие регулярным (устойчивым) методам решения тех уравнений (в первую очередь, интегрального уравнения Фредгольма I рода), задача решения которых некорректна.

7.1. Корректность и некорректность по Адамару

Определения корректности и некорректности. Ж. Адамар в 1902 г. ввел понятия *корректности* и *некорректности* [19, с. 224], [48, с. 5], [67, с. 15–18]. Рассмотрим операторное уравнение:

$$Ay = f, \quad y \in Y, \quad f \in F, \quad (7.1)$$

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 и нелинейный, интегральный, дифференциальный или алгебраический и т. д.).

Определение. Задача решения уравнения (7.1) называется *корректной* или *корректно поставленной* (*well-posed*), если:

- 1) решение существует,
- 2) решение единственно,
- 3) решение устойчиво.

Если хотя бы одно из этих условий не выполняется, то задача называется *некорректной* или *некорректно поставленной* (*ill-posed*).

Более того, Адамар выдвинул утверждение, что некорректные задачи не имеют физического смысла, другими словами, если уравнение, описывающее некоторую прикладную (физическую, техническую и т. д.) задачу, является некорректным, то или эта задача является искусственной (нереальной), или она описана математически неадекватно, например, описана интегральным уравнением Фредгольма I рода, задача решения которого некорректна, а нужно бы добавить еще ряд ограничений на решение и тогда задача станет корректной. А поскольку, как выяснилось в последние десятилетия, значительная часть прикладных задач является некорректными, то утверждение Адамара привело к замедлению развития многих разделов чистой и прикладной математики.

Примеры. Приведем примеры уравнений или их систем, задача решения которых некорректна.

Пример 1. Переопределенная СЛАУ:

$$\left. \begin{aligned} 2y_1 - 3y_2 &= -4, \\ -y_1 + 2y_2 &= 3, \\ y_1 + 4y_2 &= 15. \end{aligned} \right\} \quad (7.2)$$

Такая СЛАУ не имеет решения y_1, y_2 . Действительно, если рассматривать лишь первые два уравнения, то получим решение: $y_1 = 1, y_2 = 2$, если рассматривать второе и третье уравнения, то получим: $y_1 = y_2 = 3$, а если рассматривать первое и третье уравнения, то $y_1 = 2.635, y_2 = 3.09$, т.е. (единое) решение не существует — нарушен 1-й пункт корректности по Адамару.

Пример 2. Недоопределенная СЛАУ:

$$2y_1 - 3y_2 = -4. \quad (7.3)$$

Она имеет множество решений, 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. СЛАУ:

$$\left. \begin{aligned} 2y_1 - 3y_2 &= 3, \\ -1.33y_1 + 2y_2 &= -1.99. \end{aligned} \right\} \quad (7.4)$$

Решение СЛАУ (7.4) существует и единственно: $y_1 = 3, y_2 = 1$. Однако если правые части немного изменить, а именно:

$$\left. \begin{aligned} 2y_1 - 3y_2 &= 3.01, \\ -1.33y_1 + 2y_2 &= -2, \end{aligned} \right\} \quad (7.4')$$

т.е. внести относительные погрешности $\|\delta f\|/\|f\| < 0.5\%$, то получим новое, заметно отличное, решение: $y_1 = 2$ (относительная погрешность $> 30\%$), $y_2 = 0.33$ (относительная погрешность $> 60\%$), т.е. относительная погрешность решения на два порядка превысит относительную погрешность правой части. Это можно оценить и через число обусловленности $\text{cond}(A)$ (см. п. 5.2). Для этого запишем характеристическое уравнение:

$$\begin{vmatrix} 2 - \lambda & -3 \\ -1.33 & 2 - \lambda \end{vmatrix} = 0, \quad (7.5)$$

откуда $\lambda_1 = 3.997, \lambda_2 = 0.003$, т.е. собственные значения вещественны и положительны (а значит, матрица СЛАУ (7.4) является положительно определенной). Тогда:

$$\text{cond}(A) = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{\lambda_1}{\lambda_2} = 1.332 \cdot 10^3 \quad (7.6)$$

и

$$\frac{\|\delta y\|}{\|y\|} \leq \text{cond}(A) \frac{\|\delta f\|}{\|f\|}. \quad (7.7)$$

Мы видим, что решение СЛАУ (7.4) несколько неустойчиво. А часто СЛАУ имеют гораздо большее число обусловленности, т.е. гораздо более неустойчивы. Но еще более неустойчивы интегральные уравнения Фредгольма I рода, так как у них $\lambda_{\min} = 0$.

Далее будет показано, что если решение не существует, то используют метод типа метода наименьших квадратов Гаусса (получают псевдорешение — см. п. 7.3), если решение неединственно, то используют метод типа метода псевдообратной матрицы Мура-Пенроуза (получают нормальное решение — см. п. 7.4), а если решение неустойчиво, то используют устойчивые (регулярные) методы (регуляризации, фильтрации и др. — см. гл. 8).

Но прежде чем перейти к этим методам, мы остановимся на классических методах (имея в виду, что они обычно дают решения, некорректные по Адамару).

7.2. Классические методы решения интегральных уравнений Фредгольма I рода

Метод квадратур. Рассмотрим интегральное уравнение Фредгольма I рода:

$$\int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (7.8)$$

где $K(x, s)$ — ядро, $y(s)$ — искомая функция, $f(x)$ — правая часть (имеется в виду зашумленная правая часть $\tilde{f}(x)$), $[a, b]$ — область изменения s , а $[c, d]$ — область изменения x . Метод квадратур заключается в следующем:

1) Область $[a, b]$ разбиваем через шаг $\Delta s = h_1 = \text{const}$, а область $[c, d]$ через шаг $\Delta x = h_2 = \text{const}$ (рассмотрим случай постоянства шагов дискретизации h_1 и h_2 , хотя метод можно обобщить и на случай непостоянства h_1 и h_2). Получим число узлов $n = (b - a)/h_1 + 1$ (по s) и $m = (d - c)/h_2 + 1$ (по x).

2) Интеграл в (7.8) заменяем конечной суммой, расписывая его по некоторой квадратурной формуле, например, по формуле трапеций:

$$\int_a^b K(x, s) y(s) ds \approx \sum_{j=1}^n p_j K(x, s_j) y(s_j), \quad (7.9)$$

где

$$p_j = \begin{cases} 0.5h_1, & j = 1 \text{ или } j = n, \\ h_1, & \text{иначе,} \end{cases} \quad (7.10)$$

$$s_j = a + (j - 1)h_1. \quad (7.11)$$

3) Вводя дискретизацию по x :

$$x_i = c + (i - 1)h_2, \quad (7.12)$$

окончательно получим:

$$\sum_{j=1}^n A_{ij} y_j = f_i, \quad i = \overline{1, m}, \quad (7.13)$$

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, получили систему m линейных алгебраических уравнений (7.13) относительно n неизвестных y_j . Решая ее, можно получить решение интегрального уравнения (7.8) в дискретном виде.

Матрица A системы уравнений (7.13), вообще говоря, прямоугольна. Если $m = n$, то матрица A — квадратная и СЛАУ (7.13) можно решать по правилу Крамера, гауссовскими методами и др. Если $m > n$, то СЛАУ (7.13) нужно решать методом наименьших квадратов Гаусса (см. п. 7.3) — получим псевдорешение, а если $m < n$, то нужно использовать метод псевдообратной матрицы Мура-Пенроуза (см. п. 7.4) — получим нормальное решение. Таким образом, первые два пункта корректности по A дамарубудут выполнены.

Однако все эти решения очень неустойчивы, т. е. нарушается 3-й пункт корректности по A дамару Эта неустойчивость обусловлена тем, что минимальное сингулярное число μ_{\min} интегрального оператора уравнения (7.8) равно нулю и число обусловности $\text{cond} = \infty$. Если же m и n конечны, то μ_{\min} может стать несколько отличным от нуля, но решение СЛАУ (7.13) будет по-прежнему очень неустойчивым. На рис. 7.1 приведены результаты решения примера (пример 1 из [59, ч. III]; см. также пример 2, редукция протяженных сигналов, п. 3.2, рис. 3.15) методом квадратур согласно (7.13) при $m = n = 137$.

Непрерывной линией отображено точное решение $y(s)$, а пунктиром — приближенное (дискретное) решение y_j , $j = \overline{1, n}$, СЛАУ (7.13). Видим, что решение y_j получилось в виде так называемой знакопеременной «пилы» большой амплитуды, ничего общего не имеющей с точным решением. Между тем, если для проверки подставить «пилу» в (7.13), то получим совпадение левой и правой частей (7.13) с точностью до 3–5 цифр, если вычисления выполнять с простой точностью (до 7 цифр), или до 6–10 цифр, если вычисления выполнять с двойной точностью (до 14 цифр). Заметим еще, что вид «пилы» зависит от метода решения СЛАУ, от программы и т. д.

Из изложенного можно сделать следующие выводы:

1) метод квадратур решения интегрального уравнения Фредгольма I рода (а также в меньшей степени уравнения Вольтерры

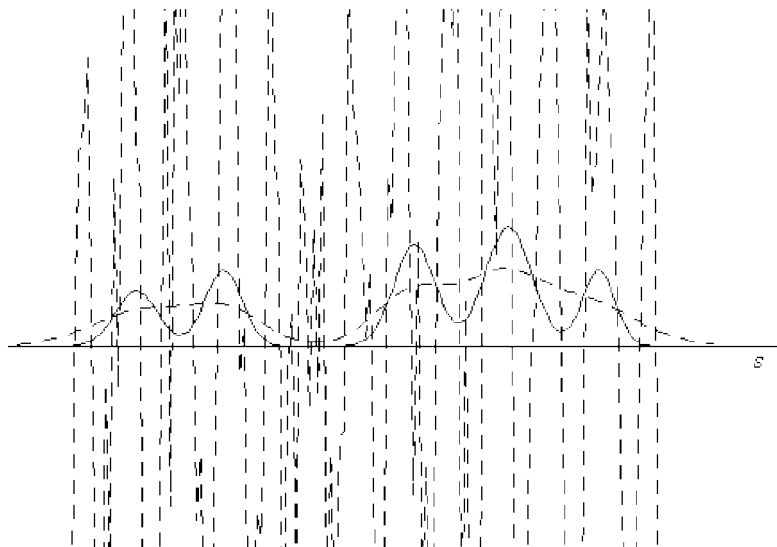


Рис. 7.1

I рода) является крайне неустойчивым (нарушается 3-й пункт корректности по А дамару);

2) классическое определение точного решения \bar{y} как решения, при котором

$$\|A\bar{y} - f\| = 0, \quad (7.14)$$

вообще говоря, не подходит для случая некорректных задач, так как в случае, когда решение не существует (см. пример 1 в п. 7.1), отсутствует такое \bar{y} , для которого выполняется (7.14), а в случае нее единственности решения (см. пример 2 в п. 7.1) существует множество \bar{y} , для которых выполняется (7.14), и наконец, в случае неустойчивости (см. пример 3 в п. 7.1 и рис. 7.1) критерий (7.14) дает неустойчивое решение.

Метод ПФ. Рассмотрим интегральное уравнение Фредгольма I рода типа свертки

$$\int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (7.15)$$

Запишем $y(s)$ в виде ОПФ (ср. (6.38)):

$$y(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega') e^{-i\omega's} d\omega'. \quad (7.16)$$

Умножим (7.15) на $e^{i\omega x}$ и проинтегрируем по x от $-\infty$ до ∞ . Получим (заменив в левой части (7.15) x на x'):

$$\frac{1}{2\pi} \iiint_{-\infty}^{\infty} K(x' - s) Y(\omega') e^{i(\omega x' - \omega' s)} d\omega' ds dx' = F(\omega), \quad (7.17)$$

где

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx. \quad (7.18)$$

Сделаем замену переменной: $x' - s = x$ в (7.17), получим, изменив порядок интегрирования:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} K(x) e^{i\omega x} dx \cdot \int_{-\infty}^{\infty} Y(\omega') \left[\int_{-\infty}^{\infty} e^{i(\omega - \omega')s} ds \right] d\omega' &= \\ &= \int_{-\infty}^{\infty} K(x) e^{i\omega x} dx \cdot \int_{-\infty}^{\infty} Y(\omega') \delta(\omega - \omega') ds = \\ &= \int_{-\infty}^{\infty} K(x) e^{i\omega x} dx \cdot Y(\omega) = F(\omega). \end{aligned}$$

Обозначив

$$\lambda(\omega) = \int_{-\infty}^{\infty} K(x) e^{i\omega x} dx, \quad (7.19)$$

получим

$$\lambda(\omega) Y(\omega) = F(\omega), \quad (7.20)$$

т. е. для уравнения (7.15) справедливо утверждение: произведение преобразований Фурье ядра и искомой функции равно преобразованию Фурье правой части. Из (7.20) имеем окончательно:

$$Y(\omega) = \frac{F(\omega)}{\lambda(\omega)}. \quad (7.21)$$

Итак, уравнение (7.15) имеет аналитическое решение и оно записывается в виде ОПФ:

$$y(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{-i\omega s} d\omega, \quad (7.22)$$

где преобразования Фурье $Y(\omega)$, $F(\omega)$ и $\lambda(\omega)$ от искомой функции, правой части и ядра записываются в виде (7.21), (7.18) и (7.19).

Формулу (7.22) можно записать в ином виде (изменив порядок интегрирования):

$$y(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x) e^{i\omega x} dx \right] \frac{e^{-i\omega s}}{\lambda(\omega)} d\omega = \\ = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\omega(s-x)}}{\lambda(\omega)} d\omega \right] f(x) dx$$

или

$$y(s) = \int_{-\infty}^{\infty} R(s-x) f(x) dx, \quad (7.23)$$

где

$$R(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\omega s}}{\lambda(\omega)} d\omega. \quad (7.24)$$

Формула (7.23) удобна тем, что функция $R(s)$ может быть вычислена заранее и единожды, а затем для ряда функций $f(x)$ могут быть весьма быстро найдены соответствующие решения $y(s)$.

Однако решение (7.22) (или (7.23)), как и решение методом квадратур, также крайне неустойчиво и это можно объяснить следующим образом. Ядро $K(x)$ обычно задается в виде гладкой функции (аналитической формулы типа (3.17) или (8.33)), поэтому его спектр $\lambda(\omega)$ быстро ниспадает с ростом $|\omega|$ и $\lim_{\omega \rightarrow \infty} \lambda(\omega) = 0$.

Функция же $f(x)$ обычно задается в виде таблицы зашумленных значений, т.е. вместо $f(x)$ имеем $\tilde{f}(x) = f(x) + \delta f(x)$, где $\delta f(x)$ — погрешности, обладающие тем свойством, что их спектр при $|\omega| \rightarrow \infty$ обычно стремится к некоторой (пусть даже очень малой) константе (см. рис. 6.8) — уровню «белого шума». Поэтому $\lim_{\omega \rightarrow \infty} F(\omega)/\lambda(\omega) = \lim_{\omega \rightarrow \infty} Y(\omega) = \infty$ и интеграл (7.22) расходится.

Другими словами, неустойчивость метода ПФ обусловлена очень сильной реакцией высоких гармоник в преобразовании Фурье на даже очень малые погрешности измерений $f(x)$. Если же вычисления выполняются по конечным квадратурным формулам, т.е. вместо НПФ $\lambda(\omega)$, $Y(\omega)$ и $F(\omega)$ используются их ДПФ с конечным числом отсчетов N , то эта неустойчивость несколько уменьшается (становится конечной), но тем не менее остается большой.

Заметим, что как показывает решение большого числа примеров [19, 43], метод ПФ дает менее неустойчивое решение, чем метод квадратур, и это обусловлено тем, что во-первых, метод ПФ дает аналитическое решение (7.22) или (7.23) и во-вторых, при численной реализации метода ПФ автоматически срабатывает усечение спектра частот (см. (6.48)).

Метод ПФ для двухмерного уравнения. Для двухмерного интегрального уравнения Фредгольма I рода типа свертки

$$\iint_{-\infty}^{\infty} K(x_1 - s_1, x_2 - s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2), \quad (7.25)$$

$$-\infty < x_1, x_2 < \infty,$$

метод двухмерного ПФ дает следующее решение (ср. (7.22)) [19, с. 268], [71, с. 44]:

$$y(s_1, s_2) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} Y(\omega_1, \omega_2) e^{-i(\omega_1 s_1 + \omega_2 s_2)} d\omega_1 d\omega_2, \quad (7.26)$$

где

$$Y(\omega_1, \omega_2) = \frac{F(\omega_1, \omega_2)}{\lambda(\omega_1, \omega_2)}, \quad (7.27)$$

$$F(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} f(x_1, x_2) e^{i(\omega_1 x_1 + \omega_2 x_2)} dx_1 dx_2, \quad (7.28)$$

$$\lambda(\omega_1, \omega_2) = \iint_{-\infty}^{\infty} K(x_1, x_2) e^{i(\omega_1 x_1 + \omega_2 x_2)} dx_1 dx_2. \quad (7.29)$$

Отметим, что в работе [63] рассмотрен вопрос о решении уравнения (7.15) методом преобразования Хартли, а в работе [62] — вопрос о решении двухмерного уравнения (7.25) методом двухмерного преобразования Хартли.

В заключение добавим, что помимо метода квадратур и метода ПФ для решения уравнений Фредгольма I рода развиты: метод разложения по собственным функциям [19, 36, 67], метод итераций (последовательных приближений) [4, 19, 67] и др. Однако и данные методы дают устойчивые решения лишь при использовании регуляризирующих модификаций, специальных стабилизирующих приемов и т. д.

Контрольные задания и вопросы

1. Сформулируйте определение корректности и некорректности по Адамару.
2. Является ли переопределенной следующая СЛАУ

$$\left. \begin{aligned} 2y_1 - 3y_2 &= -4, \\ -y_1 + 2y_2 &= 3, \\ 4y_1 - 6y_2 &= -8 \end{aligned} \right\}$$

или (подсказка) одно из уравнений есть линейная комбинация другого (или других)?

3. Придумайте примеры СЛАУ типа (7.2)–(7.4).

4. Запишите формулы для метода квадратур (используя формулу трапеций) решения уравнения (7.8) в случае неравномерных сеток узлов: $s_1 < s_2 < \dots < s_j < \dots < s_n$ и $x_1 < x_2 < \dots < x_i < \dots < x_m$ (задание повышенной трудности).

5. В методе квадратур с увеличением m и n амплитуда «пилы» будет увеличиваться или уменьшаться?

6. В методе ПФ неустойчивость решения будет повышаться или уменьшаться с увеличением числа отсчетов N в ДПФ?

7. В методе двумерного ПФ выведите подробно формулу (7.27) аналогично выводу формулы (7.21) (задание повышенной трудности).

7.3. Метод наименьших квадратов Гаусса

Изложим на примере решения СЛАУ метод наименьших квадратов (МНК) Гаусса.

Переопределенная СЛАУ. Рассмотрим систему m линейных алгебраических уравнений (СЛАУ) относительно n неизвестных, причем $m > n$ и $\text{rang}(A|f) > \text{rang}(A)$, т. е. переопределенную СЛАУ (например, (7.2)):

$$Ay = f, \quad (7.30)$$

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, другими словами, нет такого \bar{y} , для которого справедливо

$$\|A\bar{y} - f\| = 0, \quad (7.31)$$

т. е. невязка равна нулю. В МНК Гаусса вместо (7.31) вводится условие

$$\|Ay - f\| = \min_y. \quad (7.32)$$

Определение. Псевдорешением СЛАУ (7.30) называется решение y , удовлетворяющее условию (7.32), 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 . Заметим, что если $\|Ay - f\| = 0$, то псевдорешение y совпадает с точным решением \bar{y} , т. е. псевдорешение обобщает понятие точного решения.

Вывод нормальной СЛАУ. Запишем условие (7.32) в виде:

$$\|Ay - f\|^2 = \min_y, \quad (7.33)$$

а нормы будем определять согласно (5.42). Выведем (нестрого, но наглядно) новую СЛАУ из условия (7.33). Минимизация (7.33) означает равенство нулю вариации (или производной) по y :

$$2(Ay - f)A = 0 \quad (7.34)$$

или, учитывая правила умножения матриц и векторов (5.45), (5.47),

$$A^*(Ay - f) = 0. \quad (7.35)$$

В результате

$$A^*Ay = A^*f. \quad (7.36)$$

Итак, вместо переопределенной СЛАУ (7.30) получена новая СЛАУ (7.36), называемая *нормальной СЛАУ*.

Запишем (7.36) в виде:

$$By = u, \quad (7.37)$$

где

$$B = A^*A, \quad (7.38)$$

$$u = A^*f, \quad (7.39)$$

или в случае вещественности A

$$B = A^T A, \quad (7.40)$$

$$u = A^T f. \quad (7.41)$$

Используя правила (5.46) и (5.48), запишем подробно выражения для элементов новой матрицы B и новой правой части u (в случае вещественности A):

$$B_{ij} = \sum_{k=1}^m A_{ik}^T A_{kj} = \sum_{k=1}^m A_{ki} A_{kj}, \quad (7.42)$$

$$u_i = \sum_{k=1}^m A_{ik}^T f_k = \sum_{k=1}^m A_{ki} f_k. \quad (7.43)$$

СЛАУ (7.36) или (7.37) можно решать по формуле

$$y = (A^*A)^{-1}A^*f \quad (7.44)$$

или

$$y = B^{-1}u \quad (7.45)$$

или же по правилу Крамера, методами Гаусса, Краута-Холецкого и др.

МНК применительно к интегральному уравнению. Если применить МНК Гаусса к интегральному уравнению (7.8), то получим следующее новое интегральное уравнение (ср. (7.37)):

$$\int_a^b B(t, s) y(s) ds = u(t), \quad a \leq t \leq b, \quad (7.46)$$

где (ср. (7.42), (7.43))

$$B(t, s) = B(s, t) = \int_c^d K(x, t) K(x, s) dx, \quad (7.47)$$

$$u(t) = \int_c^d K(x, t) f(x) dx. \quad (7.48)$$

Основные особенности МНК Гаусса:

1) Матрица B — квадратная $n \times n$, т.е. решается система n уравнений относительно n неизвестных и в случае $\det(B) \neq 0$ решение СЛАУ (7.37) существует и является единственным.

2) Матрица B и новое ядро $B(t, s)$ являются симметричными и положительно определенными.

3) Решения уравнений (7.37) и (7.46) неустойчивы.

Пример. Рассмотрим пример 1 из п. 7.1 (см. (7.2)). Матрица СЛАУ (7.2) равна (см. (5.57))

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \\ 1 & 4 \end{pmatrix}. \quad (7.49)$$

Используя формулы (7.42), (7.43), найдем

$$B = \begin{pmatrix} 6 & -4 \\ -4 & 29 \end{pmatrix}, \quad (7.50)$$

$$u = \begin{pmatrix} 4 \\ 78 \end{pmatrix}, \quad (7.51)$$

т.е. приходим к новой СЛАУ

$$\left. \begin{aligned} 6y_1 - 4y_2 &= 4, \\ -4y_1 + 29y_2 &= 78. \end{aligned} \right\} \quad (7.52)$$

Ее решение: $y_1 = 2.71$, $y_2 = 3.06$. Этому соответствует невязка (минимально возможная) $\|Ay - f\| = 0.3993 \approx 0.4$ (согласно (5.42)). Собственные значения матрицы $B = A^*A$ равны (см. пример 2 в п. 5.2): $\lambda_1(A^*A) = 29.675$, $\lambda_2(A^*A) = 5.325$, т.е. λ_1 и λ_2 вещественны и неотрицательны, что подтверждает положительную определенность матрицы $B = A^*A$.

Контрольные задания и вопросы

1. Сформулируйте основную идею МНК Гаусса.
2. Проставьте размерности у матриц и векторов в (7.34) и найдите нарушение правил (5.45), (5.47) умножения вектора на матрицу. Проставьте аналогичные размерности в (7.35) и подтвердите, что нарушений правил умножения нет.
3. Как называются матрицы A^* , A^T и B^{-1} ?
4. Почему МНК не решает проблему устойчивости решения?
5. Найдите сингулярные числа $\mu_1(A)$ и $\mu_2(A)$ и число обусловленности $\text{cond}(A)$ матрицы A (7.49).

7.4. Метод псевдообратной матрицы Мура-Пенроуза

Как и предыдущий метод (МНК Гаусса), изложим метод псевдообратной матрицы (МПОМ) Мура-Пенроуза сначала на примере решения СЛАУ.

Недоопределенная СЛАУ. Рассмотрим СЛАУ

$$Ay = f, \quad (7.53)$$

где A — $m \times n$ -матрица, y — искомый n -вектор, f — заданная правая часть — m -вектор, причем $m < n$. Такая недоопределенная СЛАУ имеет множество решений \bar{y} , т. е. нарушается 2-й пункт корректности по Адамару. Например, СЛАУ (7.3) имеет множество решений: 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 всех них выполняется равенство (7.31).

Нормальное решение и псевдообратная матрица. Недоопределенная СЛАУ может быть решена *методом псевдообратной матрицы* (МПОМ) Мура-Пенроуза (1930 г.) [19, с. 508], [21], [26].

Определение. *Нормальное решение* — это решение с минимальной нормой среди множества решений, т. е. удовлетворяющее условию

$$\|y\| = \min_y \quad (7.54)$$

или

$$\|y\|^2 = \min_y. \quad (7.55)$$

Нормальное решение — это наиболее гладкое из решений.

Согласно МПОМ, среди множества решений недоопределенной СЛАУ выбирается нормальное решение. Доказывается [26], что нормальное решение существует и является единственным и находится по формуле:

$$y = A^+ f, \quad (7.56)$$

где A^+ — *псевдообратная* $n \times m$ -матрица Мура-Пенроуза. Матрица A^+ определяется соотношением:

$$AA^+A = A \quad (7.57)$$

или (теоретическая асимптотическая формула)

$$A^+ = \lim_{\alpha \rightarrow 0} (\alpha E + A^* A)^{-1} A^*. \quad (7.58)$$

Однако соотношения (7.57) и (7.58) неудобны для практического нахождения A^+ . Практически удобный алгоритм отыскания A^+ приведен в [26], [19, с. 508].

В случае квадратной невырожденной матрицы A имеем: $A^+ = A^{-1}$ (ср. (5.37)), а в случае переопределенной СЛАУ $A^+ = (A^* A)^{-1} A^*$ (ср. (7.44)), т.е. запись (7.56) является общей для недоопределенной, определенной и переопределенной СЛАУ. Кроме того, решение (7.56), которое уместно записать в виде $y^+ = A^+ f$, дает нулевую невязку $\|A y^+ - f\| = 0$, т.е. оно является псевдорешением (ср. (7.32)) и среди всех псевдорешений (которых в случае недоопределенной СЛАУ множество) имеет, как нормальное решение, минимальную норму. Другими словами, нормальное решение является и псевдорешением. Однако заметим, что пользоваться формулой (7.44) для отыскания любого решения, в том числе и нормального, при $m < n$ нельзя, так как в этом случае матрица $A^* A$ является вырожденной.

Пример. В качестве примера рассмотрим недоопределенную СЛАУ (7.3). Мы уже указали (см. 7.1) несколько ее решений: 1) $y_1 = 1, y_2 = 2$, норма решения (в соответствии с (5.42)) равна $\|y\| = \sqrt{1^2 + 2^2} = 2.24$; 2) $y_1 = 2, y_2 = 8/3, \|y\| = 3.34$; 3) $y_1 = 0, y_2 = 4/3, \|y\| = 1.33$. Нормальным же решением является следующее: $y_1 = -1/2, y_2 = 1, \|y\| = 1.12$. Далее

$$A = \begin{pmatrix} 2 & -3 \end{pmatrix}, \quad A^* = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \quad A^* A = \begin{pmatrix} 4 & -6 \\ -6 & 9 \end{pmatrix},$$

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

$$\text{cond}(A) = \mu(A)_{\max} / \mu(A)_{\min} = \infty, \quad \det(A^* A) = 0,$$

т.е. матрица $A^* A$ является вырожденной и обратная матрица $(A^* A)^{-1}$ не существует (ее норма $\|(A^* A)^{-1}\| = 1/\mu(A)_{\min} = \infty$).

МПОМ применительно к другим уравнениям. Если под записью (7.53) подразумевать иные уравнения, в частности, интегральное уравнение Фредгольма I рода (7.8) или операторное уравнение (5.34), то изложенный метод также применим, а именно, в качестве решения интегрального или операторного уравнения выбирается нормальное решение, удовлетворяющее условию (7.54) и найденное по формуле (7.56), где A^+ — *псевдообратный оператор*, определяемый посредством (7.58).

Общий вывод. По материалам двух последних пунктов можно сделать следующий вывод: если в качестве решения брать так называемое *нормальное псевдорешение* (нормальное или псевдорешение), то будут выполнены два первых пункта корректности по Адамару. Однако МНК и МПОМ не решают проблему неустойчивости решения, т.е. третий пункт корректности по Адамару для нормального псевдорешения, вообще говоря, не выполняется.

Контрольные задания и вопросы

1. СЛАУ (ср. (7.3))

$$\left. \begin{aligned} 2y_1 - 3y_2 &= -4, \\ 4y_1 - 6y_2 &= -8 \end{aligned} \right\}$$

является определенной (ранг $r = n = 2$) или ее следует отнести к типу недоопределенных СЛАУ ($r < n$)?

2. В чем заключается основная идея метода псевдообратной матрицы (оператора)?

3. Поставьте размерность векторов и матриц в (7.56)–(7.58).

4. Исследуйте СЛАУ (второе уравнение из (7.2)):

$$-y_1 + 2y_2 = 3.$$

Найдите ряд решений и нормы $\|y\|$ для них. Запишите A , A^* , A^*A , найдите $\lambda(A^*A)_{\max}$, $\lambda(A^*A)_{\min}$, $\mu(A)_{\max}$, $\mu(A)_{\min}$, $\text{cond}(A)$, $\det(A^*A)$, $\|(A^*A)^{-1}\|$.

5. Почему МПОМ не решает проблему устойчивости решения?

Глава 8

МЕТОДЫ РЕГУЛЯРИЗАЦИИ, ФИЛЬТРАЦИИ И АППРОКСИМАЦИИ

В данной главе изложены устойчивые методы регуляризации Тихонова, фильтрации Калмана и Винера, сплайн-аппроксимации и т. д. решения тех уравнений (интегральное уравнение Фредгольма I рода и др.), задача решения которых некорректна.

8.1. Метод регуляризации Тихонова

Метод регуляризации Тихонова (1963 г.) [4, 15, 19, 23, 36, 45, 48, 67, 71] 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 применительно к операторному уравнению:

$$Ay = f, \quad y \in L_2, \quad f \in L_2, \quad (8.1)$$

где A — линейный вполне непрерывный оператор, f — заданная правая часть, а y — искомое решение, причем вместо точных f и A известны их приближения \tilde{f} и \tilde{A} такие, что

$$\|\tilde{f} - f\|_{L_2} \leq \delta, \quad (8.2)$$

$$\|\tilde{A} - A\| \leq \xi, \quad (8.3)$$

где $\delta \geq 0$ и $\xi \geq 0$ — погрешности правой части и оператора (точнее, их верхние оценки, поэтому в (8.2) и (8.3) стоят знаки \leq , а не знаки $=$), т. е. решается уравнение

$$\tilde{A}\tilde{y} = \tilde{f}, \quad \tilde{y} \in L_2, \quad \tilde{f} \in L_2. \quad (8.4)$$

Однако для упрощения записи мы далее будем использовать запись (8.1), подразумевая, что в действительности рассматривается уравнение (8.4).

В методе регуляризации Тихонова ставятся два условия: условие минимизации невязки типа (7.33), как в МНК Гаусса, и условие минимизации нормы решения типа (7.55), как в МПОМ Мура-Пенроуза. Это — задача условной минимизации и она решается методом неопределенных множителей Лагранжа, а именно,

$$\|Ay - f\|_{L_2}^2 + \alpha \|y\|_{L_2}^2 = \min_y, \quad (8.5)$$

где $\alpha > 0$ — параметр регуляризации, играющий роль неопределенного множителя Лагранжа. Из условия (8.5) вытекает *уравнение Тихонова* (ср. (7.36)):

$$(\alpha E + A^*A)y_\alpha = A^*f, \quad (8.6)$$

где E — единичный оператор ($Ey = y$). Итак, вместо уравнения I рода получено уравнение II рода (8.6).

Анализ метода. Проанализируем условие (8.5) и уравнение (8.6).

Если $\alpha = 0$, то метод регуляризации Тихонова переходит в МНК Гаусса с крайне неустойчивым решением, но минимальной невязкой $\|Ay - f\|^2$. С увеличением же α решение 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 будем иметь и умеренную гладкость, и умеренную невязку. Некоторые способы выбора α изложены ниже.

Если $\delta, \xi \rightarrow 0$, то $\alpha \rightarrow 0$ и

$$y_\alpha = \lim_{\alpha \rightarrow 0} (\alpha E + A^*A)^{-1} A^*f \equiv A^+f \quad (8.7)$$

(см. (7.58)), 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 эта устойчивость обусловлена следующими обстоятельствами. Оператор A^*A в (8.6) является 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 (8.6) увеличивает все $\lambda_i(A^*A)$ на α , поэтому $\alpha E + \lambda(A^*A)_{\min} = \alpha$. 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 (8.6) есть

$$y_\alpha = (\alpha E + A^*A)^{-1} A^*f. \quad (8.8)$$

В методе регуляризации Тихонова рассматривается также более общий подход, когда [19, с. 238]

$$\|Ay - f\|^2 + \alpha \|y - \psi\|^2 = \min_y, \quad (8.9)$$

где ψ — начальное приближение (матожидание, прогноз) решения y . В этом случае решение равно

$$y_\alpha = \psi + (\alpha E + A^*A)^{-1} A^*(f - A\psi). \quad (8.10)$$

Однако соотношения (8.9), (8.10) используются редко из-за неопределенности ψ . Мы их приводим лишь для сопоставления метода регуляризации Тихонова и метода оптимальной фильтрации Калмана (см. п. 8.2).

Регуляризованное интегральное уравнение. Применительно к интегральному уравнению Фредгольма I рода

$$Ay \equiv \int_a^b K(x, s) y(s) ds = f(x), \quad c \leq x \leq d, \quad (8.11)$$

соотношение (8.6) приобретает вид интегрального уравнения Фредгольма II рода с положительно определенным ядром (ср. (7.46)–(7.48)) [19, с. 241]:

$$\alpha y_\alpha(t) + \int_a^b R(t, s) y_\alpha(s) ds = F(t), \quad a \leq t \leq b, \quad (8.12)$$

где

$$R(t, s) = R(s, t) = \int_c^d K(x, t) K(x, s) dx, \quad (8.13)$$

$$F(t) = \int_c^d K(x, t) f(x) dx. \quad (8.14)$$

Способы выбора параметра регуляризации α . Разработан ряд способов выбора α в методе регуляризации Тихонова [19, 48, 67]. Приведем три из них.

1-й способ — *способ невязки*, согласно которому α выбирается из условия [19, с. 244], [23, с. 156], [36, с. 75–77], [67, с. 71–80] при $\xi = 0$:

$$\|Ay_\alpha - f\|_{L_2} = \delta. \quad (8.15)$$

Если $\|f\|_{L_2} \geq \delta$, то решение уравнения (8.15) относительно α существует и является единственным. При $\xi \neq 0$ способ невязки переходит в *обобщенный принцип невязки* [19, с. 242], [36, с. 83], [48, с. 63], [71, с. 13].

2-й способ — *способ подбора* [62, 63, 104]. Согласно нему, находятся решения y_α для ряда «разумных» значений α и окончательный выбор α делается на основе дополнительной информации о решении, в основном, визуальной. Способ весьма напоминает выбор контраста телеизображения. Действительно, уменьшение α соответствует повышению неустойчивости решения y_α , т. е. увеличению контраста изображения, если под изображением подразумевать y_α , и наоборот, увеличение α соответствует уменьшению контраста (см. рис. 2.3, 2.7). Несмотря на простоту способа, он

может быть весьма эффективен в случае, когда имеется немалая информация о решении (степень гладкости, число экстремумов и т. д.), а также выполнена обработка предшествующих «близких» примеров, позволяющая выделить область возможных значений α (см. дополнительно п. 2.1, 2.2).

3-й способ — *асимптотический способ*, основанный на следующей зависимости при $\xi = 0$ и $\delta \rightarrow 0$ [19, с. 240, 245], [36, с. 134]:

$$\alpha = C\delta^2, \quad (8.16)$$

где $C > 0$ — некоторая константа. Этот способ можно использовать при малых значениях δ .

Разработаны также другие способы выбора α : способ квазиоптимального (квазинаилучшего) α , способ отношения, способ независимых реализаций, способ перекрестной значимости, способ моделирования и т. д. [17, с. 135–137], [19, с. 245–249], [23, с. 156–165].

Численный алгоритм. Рассмотрим вопрос о численном решении интегрального уравнения (8.12). Остановимся на одном из наиболее эффективных алгоритмов — *методе квадратур*.

Пусть правая часть $f(x)$ задана таблично на следующей, вообще говоря, неравномерной x -сетке узлов:

$$c = x_1 < x_2 < x_3 < \dots < x_l = d, \quad (8.17)$$

а решение $y_\alpha(s)$ ищется на другой неравномерной s -сетке узлов, совпадающей с t -сеткой узлов:

$$a = s_1 = t_1 < s_2 = t_2 < s_3 = t_3 < \dots < s_n = t_n = b, \quad (8.18)$$

причем $l \leq n$. Распишем интеграл в (8.12) по некоторой квадратурной формуле, лучше всего по формуле трапеций. Получим:

$$\alpha y_k + \sum_{j=1}^n r_j R_{kj} y_j = F_k, \quad k = \overline{1, n}, \quad (8.19)$$

где $y_k = y_\alpha(t_k)$, $y_j = y_\alpha(s_j)$, $R_{kj} = R(t_k, s_j)$, $F_k = F(t_k)$. Аналогично интегралы в (8.13) и (8.14) аппроксимируем конечными суммами по квадратурной формуле. Получим:

$$R_{kj} = R_{jk} = \sum_{i=1}^l p_i K_{ik} K_{ij}, \quad k, j = \overline{1, n}, \quad (8.20)$$

$$F_k = \sum_{i=1}^l p_i K_{ik} f_i, \quad k = \overline{1, n}, \quad (8.21)$$

где $K_{ik} = K(x_i, t_k)$, $K_{ij} = K(x_i, s_j)$, $f_i = f(x_i)$, а r_j и p_i — коэффициенты квадратурных формул.

Запись (8.19) есть СЛАО относительно y_j , $j = \overline{1, n}$. Подробности — в [19, с. 249–251].

Программы. В [19, с. 371–379] приведены программы ТІКН1, ТІКН2, ТІКН3, ТІКН4, ТІКН5, а в [71, с. 104–117, 157–174] — программы РТІМР, РТІЗР на Фортране, реализующие метод регуляризации Тихонова согласно формул типа (8.17)–(8.21) с различными способами выбора параметра регуляризации α .

Численные примеры. Приведем несколько примеров решения интегрального уравнения (8.11) методом регуляризации Тихонова.

Пример 1. Это — пример 4.2 из [19, с. 280]. В данном примере точное решение равно

$$y(s) = \left\{ \left[1 - \left(\frac{s}{0.85} \right)^2 \right]^2 + 0.5 \sin^4 \left(\frac{\pi s}{0.85} \right) \cos \left(\frac{6.5\pi s}{0.85} \right) \right\} \left(1 - \frac{s}{4} \right), \quad (8.22)$$

ядро

$$K(x, s) = \sqrt{\frac{Q}{\pi}} e^{-Q(x-s)^2/(1+x^2)}, \quad (8.23)$$

$a = c = -0.85$, $b = d = 0.85$, шаги дискретизации $\Delta x = \text{const} = 0.05$ ($l = 35$), $\Delta s = \Delta t = \text{const} = 0.025$ ($n = 69$), $Q = 59.9$. В примере значения f_i , $i = \overline{1, l}$, вычислялись численно по квадратурной формуле типа (7.9):

$$f_i = \sum_{j=1}^N p_j K(x_i, s_j) y(s_j)$$

с шагом, значительно меньшим, чем $\Delta s = 0.025$, т. е. $N \gg n$. Затем с помощью датчика случайных чисел RNDAN к значениям f_i добавлялись погрешности δf_i , распределенные по нормальному закону с нулевым математическим ожиданием и среднеквадратическим отклонением $\delta = 0.5 \cdot 10^{-2}$ и $\delta = 0.5 \cdot 10^{-3}$ (этому соответствует $\delta_{\text{отн}} \approx 1\%$ и $\delta_{\text{отн}} \approx 0.1\%$). Кроме того, вместо точного $Q = 59.9$ полагалось $\tilde{Q} = 60$ (этому соответствует $\xi_{\text{отн}} \approx 1\%$) и $\tilde{Q} = 59.91$ (этому соответствует $\xi_{\text{отн}} \approx 0.1\%$). На рис. 8.1 представлены точное решение $y(s)$ и точная правая часть $f(x)$.

Особенности примера: искомая функция $y(s)$ имеет значительные флуктуации, но ядро $K(x, s)$ неузкое настолько, что флуктуации в $f(x)$ отсутствуют. Данный пример характерен для задачи восстановления сигнала на входе измерительного устройства — антенны, спектрометра, системы управления и т. д. и задача состоит в восстановлении микроструктуры входного сигнала $y(s)$ по измеренному (с погрешностями) выходному сигналу $f(x)$ и аппаратной функции $K(x, s)$.

Пример решался по программам ТІКН1 и ТІКН2. На рис. 8.2 приведена относительная погрешность $\|y_\alpha - y\|_{L_2} / \|y\|_{L_2}$ численного решения в функции α (по программе ТІКН2) при $\delta_{\text{отн}} \approx \xi_{\text{отн}} \approx 1\%$ (кривая 1) и $\delta_{\text{отн}} \approx \xi_{\text{отн}} \approx 0.1\%$ (кривая 2).

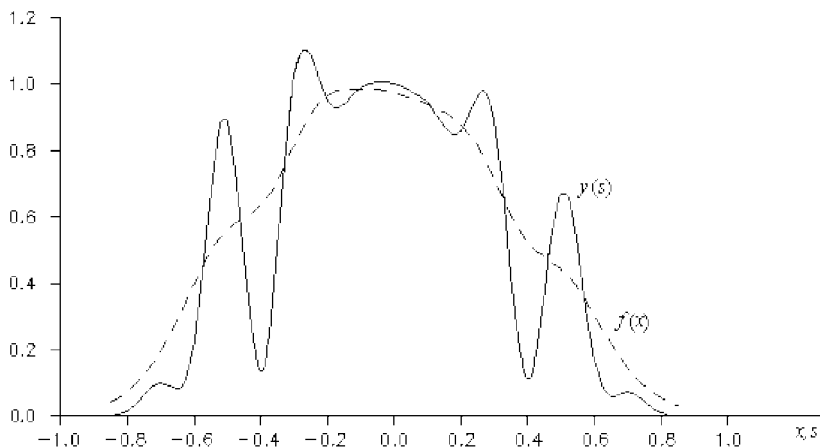


Рис. 8.1

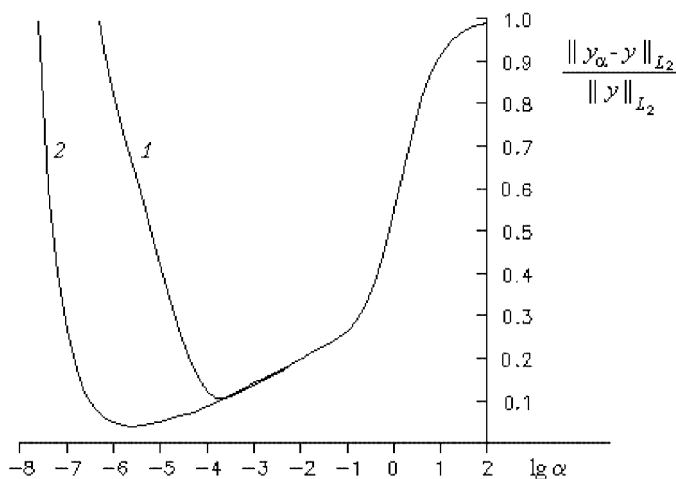


Рис. 8.2

Рассмотрим случай $\delta_{\text{отн}} \approx \xi_{\text{отн}} \approx 1\%$. Видим, что

$$\|y_{\alpha} - y\|_{L_2} / \|y\|_{L_2} = 0.105 = \min$$

при $\alpha = \alpha_{\text{opt}} = 10^{-3.6}$. По программе же ТИКН1, использующей принцип обобщенной невязки выбора α , получается $\alpha = \alpha_d = 10^{-1.9}$ и $\|y_{\alpha_d} - y\|_{L_2} / \|y\|_{L_2} = 0.204$. На рис. 8.3 представлены решения $y_{\alpha_{\text{opt}}}(s)$ (\bullet) и $y_{\alpha_d}(s)$ ($+$) при $\delta_{\text{отн}} \approx \xi_{\text{отн}} \approx 1\%$.

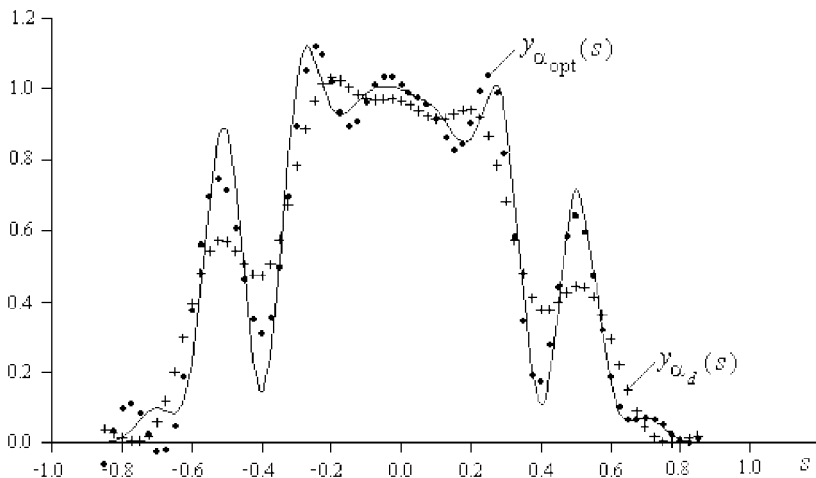


Рис. 8.3

Рис. 8.2 (кривая 1) и 8.3 показывают следующее. Во-первых, при малых α ($\lesssim 10^{-6}$) решение $y_\alpha(s)$ неустойчиво (норма $\|y_\alpha - y\|/\|y\|$ велика) и это обусловлено некорректностью задачи, а при больших α решение $y_\alpha(s)$ слишком заглажено (норма $\|y_\alpha - y\|/\|y\| \approx 1$). Но есть область умеренных значений α (в данном примере от $\approx 10^{-4.5}$ до $\approx 10^{-1}$), при которых решение $y_\alpha(s)$ близко к $y(s)$. Во-вторых, принцип обобщенной невязки дает завышенное значение α и, как следствие, заглаженное решение. В-третьих, метод регуляризации Тихонова позволил (при $\alpha = \alpha_{\text{opt}}$) восстановить все флуктуации в решении $y(s)$ (хотя и с некоторыми смещениями), причем без ложных флуктуаций. Это равносильно тому, что измерительное устройство, сопряженное с вычислительным устройством (с заложенной в него программой, реализующей метод регуляризации Тихонова), приводит к новому измерительному устройству с более высокой разрешающей способностью.

Пример 2. Это — пример 1 из [59, ч. III] (см. также пример 2 из задачи редукции протяженных сигналов, п. 3.2, рис. 3.15). В данном примере рассматривается уравнение (8.11), причем точное решение равно

$$y(s) = 6.5e^{-\left(\frac{s+0.66}{0.085}\right)^2} + 9e^{-\left(\frac{s+0.41}{0.075}\right)^2} + 12e^{-\left(\frac{s-0.14}{0.084}\right)^2} + 14e^{-\left(\frac{s-0.41}{0.095}\right)^2} + 9e^{-\left(\frac{s-0.67}{0.065}\right)^2}, \quad (8.24)$$

$a = -0.85$, $b = 0.85$, $c = -1$, $d = 1$, ядро $K(x, s)$ имеет вид (8.23), где точное $Q = 59.924$. Шаги дискретизации $\Delta x = \Delta s = \text{const} = 0.0125$, число узлов $n = 137$, $m = 161$. На рис. 8.4 представлены точное решение $y(s)$ и правая часть $f(x)$.

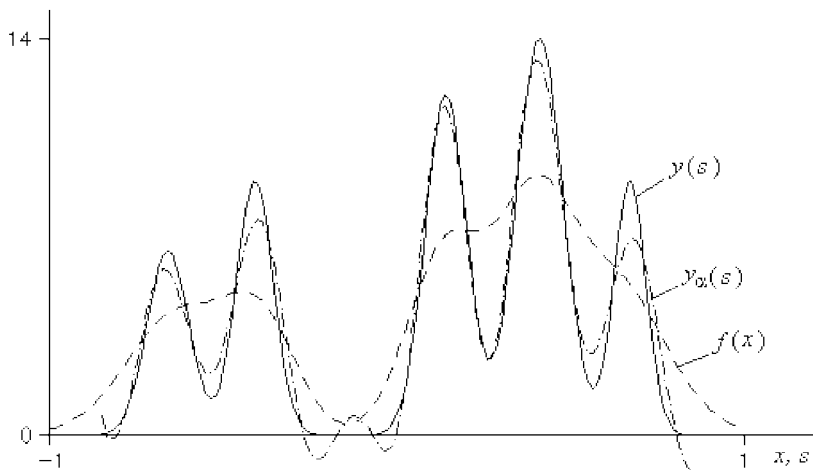


Рис. 8.4

К значениям $f(x)$ добавлялись погрешности $\delta f(x) \in \mathcal{N}(0, 0.0513)$, а точное значение Q заменялось на приближенное $\tilde{Q} = 60$, т.е. $\delta_{\text{отн}} \approx \xi_{\text{отн}} \approx 10^{-2}$ (погрешности правой части и оператора $\approx 1\%$). Значение параметра регуляризации α выбрано способом моделирования (путем решения ряда близких примеров) [59, ч. III] с помощью программы ТИХН2 [19, с. 252, 374–376] ($\alpha = 10^{-3.5}$).

На рис. 8.4 приведено решение $y_\alpha(s)$ методом регуляризации Тихонова по программе ТИХН3 (решение с выбранным α) [19, с. 253, 376–377].

Другие характерные примеры см. в [71, с. 34, 35, 113, 117].

Метод регуляризации для уравнения типа свертки. Рассмотрим частный случай интегрального уравнения Фредгольма I рода — уравнение типа свертки одномерное (см. (5.3), (7.15), (1.16), (2.15), (4.19), (4.24), (4.25)) и двухмерное (см. (5.5), (7.25), (1.8), (1.22), (2.30), (2.51), (4.21)). Если уравнение общего вида (8.11) при его численном решении методом квадратур требует размещения в компьютерной памяти матрицы СЛАУ (см. (8.19)) и это ограничивает возможности метода, то для решения одномерного уравнения типа свертки возможно применение метода преобразования Фурье, оперирующего лишь с векторами, что существенно расширяет возможности метода в отношении памяти и времени решения. Сказанное в еще большей степени характерно для двухмерного уравнения.

Рассмотрим *одномерное интегральное уравнение Фредгольма I рода типа свертки*:

$$Ay \equiv \int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (8.25)$$

Применительно к нему в методе регуляризации Тихонова решение находится из условия *минимума сглаживающего функционала* (ср. (8.5)):

$$\int_{-\infty}^{\infty} [Ay - f(x)]^2 dx + \alpha \int_{-\infty}^{\infty} M(\omega) |Y(\omega)|^2 d\omega = \min_y, \quad (8.26)$$

где

$$M(\omega) = |\omega|^{2q} \quad (8.27)$$

— *регуляризатор q -го порядка*, причем $q \geq 0$ — задаваемый порядок регуляризации, например, $q = 1$.

Из условия (8.26) получается следующее выражение для регуляризованного решения (ср. (7.22)):

$$y_{\alpha}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-\omega)F(\omega)}{L(\omega) + \alpha M(\omega)} e^{-i\omega s} d\omega, \quad (8.28)$$

где

$$L(\omega) = |\lambda(\omega)|^2 = \lambda(\omega)\lambda(-\omega) = \text{Re}^2\lambda(\omega) + \text{Im}^2\lambda(\omega) \quad (8.29)$$

(см. также (7.18), (7.19)), или (ср. (7.23))

$$y_{\alpha}(s) = \int_{-\infty}^{\infty} R_{\alpha}(s-x)f(x)dx, \quad (8.30)$$

где (ср. (7.24))

$$R_{\alpha}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-\omega)}{L(\omega) + \alpha M(\omega)} e^{-i\omega s} d\omega. \quad (8.31)$$

Сравним классическое решение (7.22) и регуляризованное решение (8.28). В (8.28) за счет слагаемого $\alpha M(\omega)$ подынтегральная функция стремится к нулю при $|\omega| \rightarrow \infty$, т.е. слагаемое $\alpha M(\omega)$ подавляет реакцию высоких гармоник на погрешность исходных данных, 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 его увеличением сильнее

подавляются все гармоники. Поэтому в отношении q следует руководствоваться следующим правилом: если искомое решение имеет флуктуации (типа рис. 3.15, 7.1, 8.1, 8.4, 8.5), то q следует взять поменьше (например, $q = 1$), а если решение гладкое (типа рис. 6.5), то q можно повысить, например, $q = 2$. Что же касается α , то способы его выбора те же, что и для уравнения (8.11) (способы невязки, по дбора и др.).

Разработан ряд численных алгоритмов получения решения $y_\alpha(s)$ [71, с. 38–43, 123–124], [19, с. 263–267]. Все они основаны на замене интегралов в (8.28), (7.18), (7.19), (8.30), (8.31) конечными суммами (по формулам прямоугольников или трапеций), переходе, тем самым, от НПФ к ДПФ и использовании алгоритма БПФ (ср. п. 6.3).

Согласно численным алгоритмам разработаны и опубликованы следующие программы для решения уравнения (8.25) методом регуляризации Тихонова: РТИКР [71, с. 124–130, 178–179], CONV1, CONV2, CONV3, CONV4, CONV5 [19, с. 379–388], CONVOL [61] и др.

Приведем результаты решения следующего примера [91] (близкого к примеру [71, с. 43, 127–129], [104], изображенному на рис. 5.1):

$$\int_0^1 k(t-\tau)y(\tau)d\tau = f(t), \quad 0 \leq t \leq 2, \quad (8.32)$$

где ядро

$$k(t) = e^{-80(t-0.5)^2}, \quad (8.33)$$

точное решение

$$y(\tau) = \left[0.45e^{-\left(\frac{\tau-0.29}{0.18}\right)^2} + e^{-\left(\frac{\tau-0.71}{0.18}\right)^2} \right] \sqrt{1 - \left(\frac{\tau-0.5}{0.5}\right)^2}, \quad (8.34)$$

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 (8.32), хотя и имеет конечные пределы изменения τ и t , но является уравнением типа свертки, так как вне пределов переменных 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 [71, с. 39, 123], вместо (8.32) решалось уравнение:

$$\int_{-0.5}^{1.5} k(t-\tau)y(\tau)d\tau = f(t), \quad 0 \leq t \leq 2, \quad (8.35)$$

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 64. Вместо точной $f(t)$ использовалась $\tilde{f}(t) = f(t) + v(t)$ со случайной погрешностью $v(t) \in \mathcal{N}(0, 0.0164)$, т. е. среднеквадратическая погрешность $\approx 10\%$ от $\max_{t \in [0, 2]} f(t)$.

На рис. 8.5 приведены точные $k(t)$, $y(\tau)$, $f(t)$, а также $\tilde{f}(t)$.

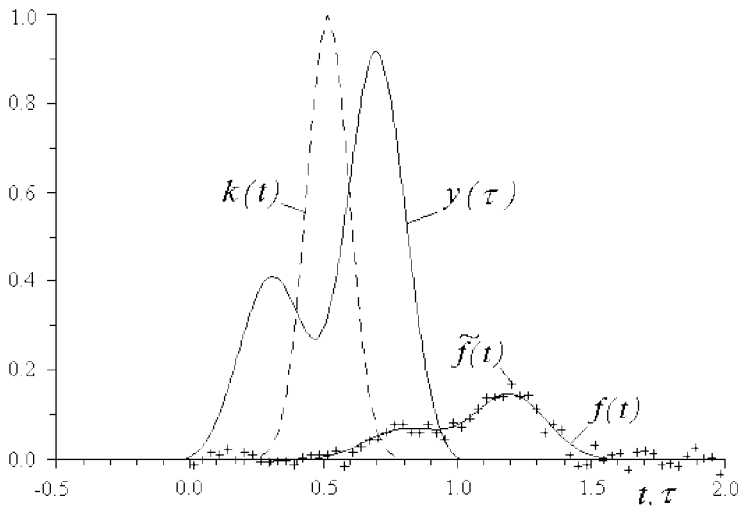


Рис. 8.5

На рис. 8.6 приведены: $y(\tau)$ — точное решение, $y_{\alpha_1}(\tau)$ — решение методом регуляризации Тихонова с α , выбранным по обобщенному принципу невязки ($\alpha = \alpha_1 = 4 \cdot 10^{-4}$) — слишком гладкое решение и $y_{\alpha_2}(\tau)$ — решение методом регуляризации Тихонова при $\alpha = \alpha_2 \ll \alpha_1$, а именно, $\alpha_2 = 10^{-6}$ — решение с гораздо лучшим разрешением.

Этот пример еще раз показывает, что при больших погрешностях исходных данных ($\sim 10\%$) способ невязки или обобщенный принцип невязки дает завышенное значение α и нужно использовать дополнительные способы, например, способ подбора.

Рассмотрим кратко *двухмерное интегральное уравнение Фредгольма I рода типа свертки*:

$$Ay \equiv \iint_{-\infty}^{\infty} K(x_1 - s_1, x_2 - s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2),$$

$$-\infty < x_1, x_2 < \infty. \quad (8.36)$$

Применим для решения уравнения (8.36) метод регуляризации Тихонова, используя аналогию с одномерным уравнением (8.25).

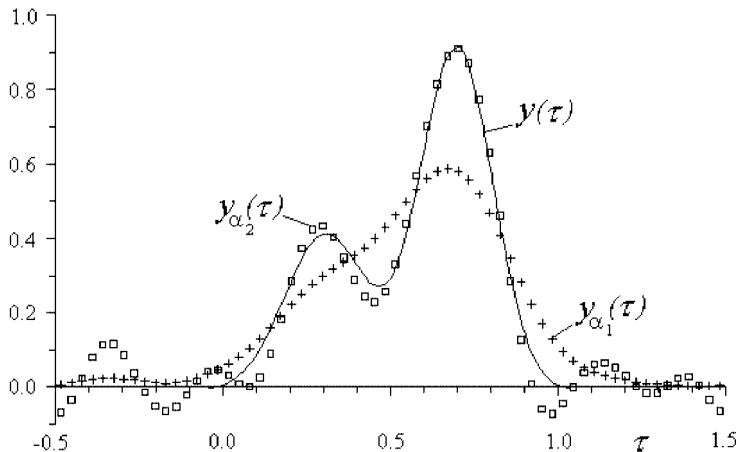


Рис. 8.6

Введем условие минимума функционала (ср. (8.26))

$$\|Ay - f\|^2 + \alpha \int_{-\infty}^{\infty} M(\omega_1, \omega_2) |Y(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2 = \min_y. \quad (8.37)$$

Из условия (8.37) следует регуляризованное решение (ср. (7.26), (8.28))

$$y_\alpha(s_1, s_2) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \frac{\lambda(-\omega_1, -\omega_2) F(\omega_1, \omega_2)}{L(\omega_1, \omega_2) + \alpha M(\omega_1, \omega_2)} e^{-i(\omega_1 s_1 + \omega_2 s_2)} d\omega_1 d\omega_2, \quad (8.38)$$

где (ср. (8.29))

$$L(\omega_1, \omega_2) = |\lambda(\omega_1, \omega_2)|^2 = \lambda(-\omega_1, -\omega_2) \lambda(\omega_1, \omega_2) = \\ = \operatorname{Re}^2 \lambda(\omega_1, \omega_2) + \operatorname{Im}^2 \lambda(\omega_1, \omega_2) \quad (8.39)$$

(см. также (7.28), (7.29)), или (ср. (8.30))

$$y_\alpha(s_1, s_2) = \iint_{-\infty}^{\infty} R_\alpha(s_1 - x_1, s_2 - x_2) f(x_1, x_2) dx_1 dx_2, \quad (8.40)$$

где (ср. (8.31))

$$R_\alpha(s_1, s_2) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \frac{\lambda(-\omega_1, -\omega_2)}{L(\omega_1, \omega_2) + \alpha M(\omega_1, \omega_2)} e^{-i(\omega_1 s_1 + \omega_2 s_2)} d\omega_1 d\omega_2. \quad (8.41)$$

Регуляризатор $M(\omega_1, \omega_2)$ (ср. (8.27)) в работе [71, с. 44] выбран в виде:

$$M(\omega_1, \omega_2) = 1 + (\omega_1^2 + \omega_2^2)^2, \quad (8.42)$$

однако сложение безразмерного слагаемого 1 и слагаемого $(\omega_1^2 + \omega_2^2)^2$, имеющего размерность, строго говоря, недопустимо, так как в зависимости от выбранной системы единиц соотношение слагаемых будет меняться и будет меняться степень подавления высоких гармоник в решении. Поэтому более эффективным является, например, выражение

$$M(\omega_1, \omega_2) = (\omega_1^2 + \omega_2^2)^2, \quad (8.43)$$

хотя и в нем слагаемые ω_1^2 и ω_2^2 имеют, вообще говоря, разную физическую размерность.

При практической реализации, когда x_1, s_1, x_2, s_2 задаются дискретно и в конечных пределах, двумерные НПФ (7.28), (7.29), (8.38), (8.41) заменяются на двумерные ДПФ, которые вычисляются как набор одномерных ДПФ по схеме (6.55). При этом одномерные ДПФ вычисляются по алгоритму БПФ.

Параметр регуляризации α выбирается теми же способами, что и в случае одномерного уравнения (8.25) (способами невязки, обобщенной невязки, подбора и др.).

В [71, с. 130–136, 185–186] приведена программа RTTR на Фортране для решения двумерного интегрального уравнения Фредгольма I рода типа свертки (8.36) методом регуляризации Тихонова, а также программа FTFTC [71, с. 190] для вычисления двумерного БПФ.

В заключение отметим, что в п. 6.3 изложен вопрос об использовании метода регуляризации Тихонова для устойчивого решения еще одной задачи (помимо решения интегрального уравнения), а именно, вычисления преобразования Фурье (НПФ и ДПФ).

Отметим также, что для устойчивого решения уравнений I рода (интегральных, дифференциальных, алгебраических и т. д.), помимо метода регуляризации Тихонова, развиты следующие методы: методы регуляризации Лаврентьева, Денисова, «погружения» Бакушинского, максимальной энтропии Берга, итеративной регуляризации Фридмана, Бакушинского, Морозова, локальной регуляризации Арсенина, генератор РА (регуляризирующих алгоритмов) Бакушинского, квазирешений Иванова, метод поиска решения на компакте, дескриптивной регуляризации Морозова и др. [4, 19, 36, 43, 45, 48, 67–69, 71, 104]. Это всё *методы детерминистской регуляризации*.

Есть еще *методы статистической регуляризации*, использующие статистический (вероятностный) подход. Это (в порядке повышения точности решения и количества дополнительной информации о решении) — метод максимального правдоподобия, статистической регуляризации Турчина, Халфина, Лаврентьева,

Жуковского, субоптимальной фильтрации, управляемой линейной фильтрации, оптимальной фильтрации Калмана-Бьюси, Винера, максимальной апостериорной вероятности и др. [15, 16, 19, 23, 67, 69, 91, 104].

Два из них (методы Калмана-Бьюси и Винера) изложены ниже.

Контрольные задания и вопросы

1. Анализируя соотношение (8.5), сформулируйте основную идею метода регуляризации Тихонова.
2. Опишите соотношение слагаемых в левой части (8.5) при изменении α .
3. Как ведет себя регуляризованное решение y_α при изменении α ?
4. Чем обусловлена устойчивость метода регуляризации Тихонова?
5. В способе подбора α проведите аналогию между значением α и контрастностью телеизображения.
6. Напишите выражения для коэффициентов r_j и p_i квадратурной формулы трапеций при замене интегралов в (8.12)–(8.14) конечными суммами с использованием неравномерных x - и s -сеток узлов (8.17) и (8.18) (задание повышенной трудности).
7. В чем состоят преимущества компьютерного решения уравнения типа свертки (8.25) по сравнению с общим уравнением (8.11)?
8. Сопоставьте спектр $Y(\omega)$ в случае гладкого решения $y(s)$ и решения с флуктуациями. В какую сторону нужно изменять α и q в случае увеличения (уменьшения) гладкости $y(s)$?
9. Опишите роль слагаемого $\alpha M(\omega)$ в (8.28) и (8.31).
10. Что означает запись: $v(t) \in \mathcal{N}(0, 0.0164)$?
11. Почему выражение (8.43) является более правильным, чем (8.42)? Предложите другие выражения для $M(\omega_1, \omega_2)$ типа (8.42) или (8.43), но включающие только безразмерные слагаемые (подсказка: введите $\omega_{1\max}$ и $\omega_{2\max}$).

8.2. Метод оптимальной фильтрации Калмана-Бьюси

В пп. 7.2–7.4, 8.1 были рассмотрены некоторые детерминистские методы решения уравнений. Из них наиболее эффективен (в отношении точности, устойчивости, времени компьютерной реализации, требуемой компьютерной памяти, исходной информации и т. д.) метод регуляризации Тихонова, как показывает решение большого числа прикладных задач [4, 15, 19, 36, 67–71, 99, 100, 104–106], несмотря на то, что он использует минимум априорной дополнительной информации: лишь значения погрешностей δ и ξ (а также иногда прогноз решения ψ).

Еще более точными являются оптимальные методы фильтрации Калмана (Калмана-Бьюси) и Винера — методы, использующие среди устойчивых (регулярных) методов наибольшее количество априорной информации: в методе Калмана — ковариации ошибок и матожидания правой части и решения, а в методе Винера — спектральные плотности мощности шумов правой части и решения. Эти методы относятся к методам статистической регуляризации.

Одношаговый (однократный) фильтр Калмана [8, 16, 19]. Рассматривается СЛАУ

$$Ay + v = f \quad (8.44)$$

или

$$\sum_{j=1}^n a_{ij}y_j + v_i = f_i, \quad i = \overline{1, m}, \quad (8.44')$$

где A — $m \times n$ -матрица, y — искомый n -вектор, f — измеренный m -вектор (замер), v — m -вектор — помеха.

Если рассматривается интегральное уравнение типа (7.8), то вместо него нужно рассматривать его дискретный аналог типа (7.13).

В методе фильтрации Калмана делаются следующие предположения:

1) Матожидание случайного вектора v равно нулю:

$$E[v] = 0, \quad (8.45)$$

где запись $E[v]$ означает матожидание (или среднеарифметическое значение) по ансамблю реализаций, т.е.

$$E[v] \equiv \langle v \rangle = \lim_{Q \rightarrow \infty} \frac{1}{Q} \sum_{q=1}^Q v_{i_q} = 0, \quad i = \overline{1, m}, \quad (8.45')$$

где q — номер реализации, или эксперимента, а Q — число реализаций случайного процесса v .

2) Задана симметричная положительно определенная $m \times m$ -матрица — *ковариация ошибок правой части*:

$$R = E[vv^T] \quad (8.46)$$

или по подробнее:

$$R_{il} = \lim_{Q \rightarrow \infty} \frac{1}{Q} \sum_{q=1}^Q v_{i_q} v_{l_q}, \quad i, l = \overline{1, m}. \quad (8.46')$$

Каждый диагональный элемент матрицы R есть квадрат среднеквадратической погрешности измерения f_i , т.е. $R_{ii} = \delta_i^2 = \sigma_i$,

а внедиагональный элемент R_{il} , $i \neq l$, определяет корреляцию погрешностей v_i и v_l .

3) Задан n -вектор

$$\psi = E[y] \quad (8.47)$$

— *матожидание* (начальное приближение, априорная оценка, прогноз) *вектора* y или подробнее:

$$\psi_j = \lim_{Q \rightarrow \infty} \frac{1}{Q} \sum_{q=1}^Q y_{jq}, \quad j = \overline{1, n}. \quad (8.48)$$

4) Задана симметричная положительно определенная $n \times n$ -матрица — *априорная ковариация ошибок решения*:

$$M = E[(y - \psi)(y - \psi)^T]. \quad (8.49)$$

Далее искомое решение y находится из условия минимума квадратичной формы (ср. (8.9)):

$$(Ay - f)^T R^{-1}(Ay - f) + (y - \psi)^T M^{-1}(y - \psi) = \min_y. \quad (8.50)$$

Из условия (8.50) получается (ср. (8.10)) *решение* (апостериорная оценка y , свертка замера с прогнозом):

$$\hat{y} = \psi + (M^{-1} + A^T R^{-1} A)^{-1} A^T R^{-1} (f - A\psi), \quad (8.51)$$

причем *апостериорная $n \times n$ -матрица ковариаций ошибок решения* \hat{y} равна

$$P \equiv E[(\hat{y} - y)(\hat{y} - y)^T] = (M^{-1} + A^T R^{-1} A)^{-1}. \quad (8.52)$$

Итак, если помимо f и A известны дополнительно R , ψ и M , то уточненное решение уравнения (8.44) согласно метода фильтрации Калмана выразится формулой (8.51), а уточненная матрица ковариаций ошибок решения — формулой (8.52).

Сравнение одношагового фильтра Калмана с методом регуляризации Тихонова. Если A , y и f вещественны, то условие минимизации (8.9) в методе регуляризации Тихонова можно записать в виде:

$$(Ay - f)^T (Ay - f) + \alpha (y - \psi)^T (y - \psi) = \min_y. \quad (8.53)$$

Сравнение (8.50) и (8.53) показывает, что роль α играет (символически) R/M . Наиболее же отчетливо сравнение методов Калмана и Тихонова выполняется в случае, когда

$$M = \varepsilon^2 E, \quad R = \delta^2 E, \quad (8.54)$$

где ε и δ — априорные среднеквадратические ошибки решения и правой части, а E — единичная матрица. В этом случае решение методом Калмана имеет вид (см. (8.51)):

$$\hat{y} = \psi + \left(\frac{\delta^2}{\varepsilon^2} E + A^T A \right)^{-1} A^T (f - A\psi), \quad (8.55)$$

а апостериорная матрица ковариаций ошибок решения равна (см. (8.52)):

$$P = \delta^2 \left(\frac{\delta^2}{\varepsilon^2} E + A^T A \right)^{-1}. \quad (8.56)$$

Решение же методом регуляризации Тихонова имеет вид (см. (8.10)):

$$y_\alpha = \psi + (\alpha E + A^T A)^{-1} A^T (f - A\psi). \quad (8.57)$$

Сравнение (8.55) и (8.57) показывает, что при

$$\alpha = \frac{\delta^2}{\varepsilon^2} \quad (8.58)$$

и справедливости (8.54) методы Калмана и Тихонова дают одинаковые решения. При этом, как следует из (8.56) и (8.58), апостериорная матрица ковариаций ошибок решения равна

$$P_\alpha = \delta^2 (\alpha E + A^T A)^{-1}, \quad (8.59)$$

откуда

$$\|\Delta y_\alpha\| = \sqrt{\|P_\alpha\|} \leq \frac{\delta}{\sqrt{\alpha}} \quad (8.60)$$

— оценка ошибки решения по норме.

Многошаговый (многократный) фильтр Калмана. Для одношагового процесса, когда имеется лишь одна реализация вектора f , требование об априорном знании ψ и M , содержащееся в методе Калмана, трудно выполнимо. Поэтому фильтр Калмана обычно применяется для многошаговых процессов, когда в функции времени поступают новые реализации f , а ψ и M итеративно уточняются. Схема многошагового фильтра Калмана выглядит следующим образом [8].

Из априорных соображений выбираются *начальные приближения* для решения $y_0 \equiv \psi$ и матрицы ковариаций ошибок решения $P_0 \equiv M$. Для выбора начальных приближений можно использовать метод регуляризации Тихонова и положить (см. (8.8), (8.59)):

$$y_0 = (\alpha E + A^T A)^{-1} A^T f, \quad (8.61)$$

$$P_0 = \delta^2 (\alpha E + A^T A)^{-1}. \quad (8.62)$$

Последующие приближения будут найдены согласно следующей *итерационной схеме* (ср. (8.51), (8.52)):

$$y_k = y_{k-1} + (P_{k-1}^{-1} + A^T R_k^{-1} A)^{-1} A^T R_k^{-1} (f_k - A y_{k-1}), \quad (8.63)$$

$$P_k = (P_{k-1}^{-1} + A^T R_k^{-1} A)^{-1}, \quad k = 1, 2, 3, \dots \quad (8.64)$$

На рис. 8.7 отображен (качественно) процесс итераций согласно (8.63). Он показывает, что в то время как значения $f_k \equiv f_{i_k}$

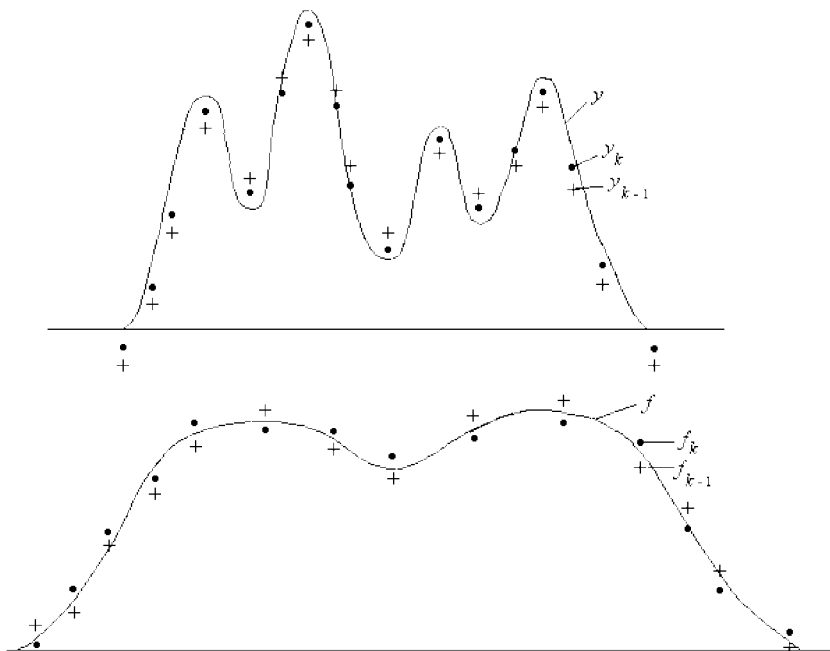


Рис. 8.7

будут флуктуировать вокруг точных f , значения $y_k \equiv y_{j_k}$ от итерации к итерации будут приближаться к точным значениям y , демонстрируя сходимость процесса итераций (8.63) к точному решению.

При этом может иметь место *стационарный случай*, когда R_k от n не зависит, т.е. $R_k = R$, а индекс k у y и P означает номер уточняющей итерации искомым y и P . Может иметь место и *нестационарный случай*, когда индекс k означает не только номер уточняющей итерации, но и переменность (зависимость от времени $t = t_k$) значений y , P , f , R .

Видим, что многошаговый фильтр Калмана обладает большими возможностями обработки результатов измерений, но требует большого объема данных.

Контрольные задания и вопросы

1. Сформулируйте одношаговый фильтр Калмана.
2. Анализируя (8.50), скажите, когда метод Калмана приближается к МНК Гаусса и когда к МПОМ Мура-Пенроуза (с ростом или убыванием элементов матриц R и M).

3. Проставьте размерности всех векторов и матриц в (8.50)–(8.53), (8.55)–(8.57), (8.59), (8.61)–(8.64).
4. Что означают записи (8.54)?
5. В чем различие q (см. (8.45')) и k (см. (8.63), (8.64))?
6. В методе регуляризации Тихонова требуется одна реализация погрешностей v или несколько (т. е. $Q = 1$ или $Q > 1$)?
7. Запишите подробно выражение для M_{il} (см. (8.49)) типа (8.46').

8.3. Метод оптимальной линейной фильтрации Винера

Существо метода. Рассматривается интегральное уравнение Фредгольма I рода типа свертки (ср. (5.3), (8.25)):

$$\int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (8.65)$$

В методе фильтрации Винера делаются следующие предположения [19, с. 315]:

1) Искомая функция $y(s)$ и погрешность правой части $v(x)$ являются реализациями стационарных, некоррелированных между собой случайных процессов.

2) Полагаются известными статистические характеристики этих процессов: *спектральная плотность мощности (СПМ) искомого решения*

$$R_y(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} \mathbb{E} \left[\left| \int_{-T}^T y(s) e^{i\omega s} ds \right|^2 \right] \quad (8.66)$$

и *СПМ помехи*

$$R_v(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} \mathbb{E} \left[\left| \int_{-T}^T v(x) e^{i\omega x} dx \right|^2 \right]. \quad (8.67)$$

Заметим, что согласно теореме Винера-Хинчина [19, с. 315]

$$R_y(\omega) = \int_{-\infty}^{\infty} r_y(x) e^{i\omega x} dx, \quad (8.68)$$

$$R_v(\omega) = \int_{-\infty}^{\infty} r_v(x) e^{i\omega x} dx, \quad (8.69)$$

где $r_y(x) = \mathbb{E}[y(s+x)y(s)]$, $r_v(x) = \mathbb{E}[v(x'+x)v(x')]$ — *автокорреляционные функции искомого решения и помехи*.

В методе Винера решение $y_R(s)$ ищется исходя из условия минимума величины $E[y_R(s) - y(s)]^2$, где $y(s)$ — точное решение. В результате [19, с. 316]

$$y_R(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-\omega) F(\omega)}{L(\omega) + R_v(\omega)/R_y(\omega)} e^{-i\omega s} d\omega, \quad (8.70)$$

$$E[y_R(s) - y(s)]^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{R_v(\omega)}{L(\omega) + R_v(\omega)/R_y(\omega)} d\omega. \quad (8.71)$$

Формула (8.70) дает решение уравнения (8.65) методом фильтрации Винера, а формула (8.71) — значение среднеквадратического отклонения этого решения от точного, являющееся минимально возможным — поэтому фильтр называется оптимальным (см. также (7.18), (7.19), (8.29)).

Сравнение методов Винера и Тихонова. Сопоставление формул (8.28) и (8.70) показывает, что регуляризованные решения, даваемые методами Тихонова и Винера, переходят одно в другое при

$$\alpha M(\omega) = R_v(\omega)/R_y(\omega). \quad (8.72)$$

Заметим следующее. Если функцию $R_v(\omega)$ (СПМ помехи) можно получить в результате спектральной обработки ряда реализаций чисто шумового процесса (без полезного сигнала), то функцию $R_y(\omega)$ (СПМ решения) аналогичным образом получить невозможно. Поэтому метод фильтрации Винера следует рассматривать, строго говоря, не как рабочий метод, а как теоретический метод, являющийся предельно точным методом среди всевозможных методов решения уравнения (8.65).

Отметим также, что как следует из (8.70), устойчивость метода Винера тем выше, чем больше отношение $R_v(\omega)/R_y(\omega)$, играющее роль отношения помеха/сигнал. Другими словами, помеха стабилизирует решение. Однако, как следует из (8.71), с ростом помехи $R_v(\omega)$ растет погрешность решения. Вывод: в методе Винера имеет место компромисс между устойчивостью и точностью решения. Впрочем, такого рода компромисс имеет место и в методе регуляризации Тихонова, но он в нем не является оптимальным.

Контрольные задания и вопросы

1. Сформулируйте метод фильтрации Винера (исходные данные, решение и т.д.).
2. Когда методы Винера и Тихонова переходят друг в друга?

3. Опишите, как в зависимости от $R_v(\omega)$ и $R_y(\omega)$ изменяются устойчивость и точность метода Винера.

4. Почему метод фильтрации Винера называется оптимальным?

8.4. Интерполяция, экстраполяция, сглаживание и аппроксимация

Во многих задачах, в том числе прикладных, используется прием замены функции $y(x)$ (заданной аналитически или таблично) некоторой другой функцией, удобной для обработки и устойчивой по отношению к погрешностям. Мы рассмотрим использование для этих целей полиномов.

Линейная интерполяция и экстраполяция. Пусть на сетке узлов

$$a = x_0 < x_1 < \dots < x_{j-1} < x_j < \dots < x_n = b, \quad (8.73)$$

вообще говоря, неравномерной заданы значения функции $y(x)$:

$$y_0, y_1, \dots, y_{j-1}, y_j, \dots, y_n, \quad (8.74)$$

где $y_j = y(x_j)$ (см. рис. 8.8).

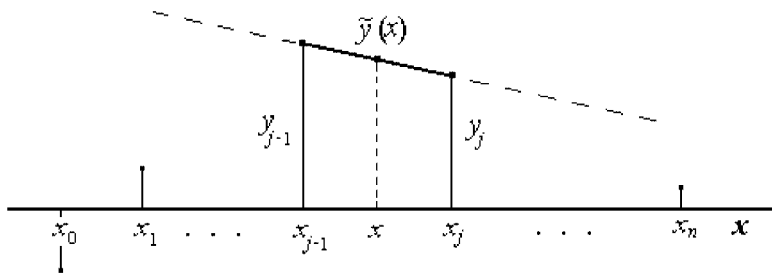


Рис. 8.8

Пусть далее требуется определить значение функции $y(x)$ в точке x , промежуточной между x_{j-1} и x_j , т. е. при $x \in [x_{j-1}, x_j]$. Воспользуемся для этого приемом *линейной интерполяции*. Он заключается в следующем. Через точки y_{j-1} и y_j проводим прямую линию, т. е. полином 1-го порядка (см. рис. 8.8):

$$\tilde{y}(x) = ux + v, \quad (8.75)$$

где u и v — некоторые коэффициенты, которые могут быть определены через систему двух уравнений:

$$\left. \begin{aligned} ux_{j-1} + v &= y_{j-1}, \\ ux_j + v &= y_j, \end{aligned} \right\} \quad (8.76)$$

решение которой:

$$u = \frac{y_j - y_{j-1}}{x_j - x_{j-1}}, \quad v = \frac{x_j y_{j-1} - x_{j-1} y_j}{x_j - x_{j-1}}. \quad (8.77)$$

Следовательно,

$$\tilde{y}(x) = \frac{y_j - y_{j-1}}{h_j} x + \frac{x_j y_{j-1} - x_{j-1} y_j}{h_j}, \quad (8.78)$$

где $h_j = x_j - x_{j-1}$, или

$$\tilde{y}(x) = \left(\frac{x - x_j}{x_{j-1} - x_j} \right) y_{j-1} + \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right) y_j \quad (8.79)$$

или

$$\tilde{y}(x) = y_{j-1} + \frac{x - x_{j-1}}{h_j} \Delta y_j, \quad (8.80)$$

где $\Delta y_j = y_j - y_{j-1}$.

Любая из формул (8.78), (8.79) или (8.80), используемых на практике [11, с.12], позволяет выполнить линейную интерполяцию. Слагаемое $\frac{x - x_{j-1}}{h_j} \Delta y_j$ в (8.80) называется *интерполяционной поправкой*.

Пример 1. Требуется определить $y(x) = e^x$ при $x = 0.853$, а в таблице [11, с.47] даны значения $y_{j-1} = e^{0.85} = 2.3396$ и $y_j = e^{0.86} = 2.3632$. По формуле (8.80) имеем:

$$\tilde{y}(0.853) = 2.3396 + 0.0071 = 2.3467.$$

Если с помощью полинома 1-го порядка нужно определить $y(x)$ в точке x вне интервала $[x_{j-1}, x_j]$, т.е. при $x \notin [x_{j-1}, x_j]$, то такая операция называется *линейной экстраполяцией* (пунктирная часть прямой на рис. 8.8). Для линейной экстраполяции подходит также любая из формул (8.78), (8.79) или (8.80).

Пример 2. Определим $y(x) = e^x$ при $x = 0.853$ из примера 1, но используем другие значения y_{j-1} и y_j , а именно, $y_{j-1} = e^{0.84} = 2.3164$ и $y_j = e^{0.85} = 2.3396$. По формуле (8.80) получим: $\tilde{y}(0.853) = 2.3164 + 0.0302 = 2.3466$, т.е. практически то же значение, что и в примере 1.

Однако, если x отстоит далеко от интервала $[x_{j-1}, x_j]$, то линейная экстраполяция может оказаться весьма неточной. Это демонстрирует

Пример 3. Нужно определить $y(x) = e^x$ при $x = 0.853$, используя $y_{j-1} = e^{0.1} = 1.1052$ и $y_j = e^{0.11} = 1.1163$. По формуле (8.80) получим: $\tilde{y}(0.853) = 1.1052 + 0.8358 = 1.9410$, т.е. со значительной погрешностью.

Квадратичная интерполяция и экстраполяция. Более точной является квадратичная интерполяция и экстраполяция. Пусть требуется определить значение функции $y(x)$. Используем значения $y_{j-1} = y(x_{j-1})$, $y_j = y(x_j)$ и $y_{j+1} = y(x_{j+1})$ такие, что $x \in [x_{j-1}, x_{j+1}]$, причем $x_{j-1} < x_j < x_{j+1}$. Проводим через точки $y(x_{j-1})$, $y(x_j)$ и $y(x_{j+1})$ полином 2-го порядка (см. рис. 8.9).

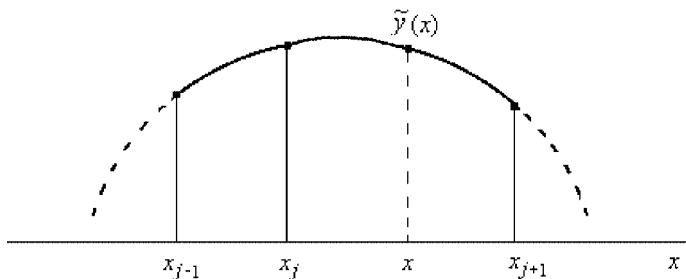


Рис. 8.9

Уравнение полинома будет (обобщение формулы (8.79)):

$$\tilde{y}(x) = \frac{(x-x_j)(x-x_{j+1})}{(x_{j-1}-x_j)(x_{j-1}-x_{j+1})} y_{j-1} + \frac{(x-x_{j-1})(x-x_{j+1})}{(x_j-x_{j-1})(x_j-x_{j+1})} y_j + \frac{(x-x_{j-1})(x-x_j)}{(x_{j+1}-x_{j-1})(x_{j+1}-x_j)} y_{j+1}. \quad (8.81)$$

Вычисление $\tilde{y}(x)$ по формуле (8.81) при $x \in [x_{j-1}, x_{j+1}]$ называется *квадратичной интерполяцией*, а при $x \notin [x_{j-1}, x_{j+1}]$ — *квадратичной экстраполяцией*. Формула (8.81) может быть преобразована к виду, дающему квадратичную интерполяцию по Бесселю [11, с. 12].

Пример 4. Нужно определить $y(x) = e^x$ при $x = 0.853$, используя $y_{j-1} = e^{0.1} = 1.1052$, $y_j = e^{0.11} = 1.1163$ и $y_{j+1} = e^{0.12} = 1.1275$. По формуле (8.81) получим

$$\tilde{y}(0.853) = 3009.565 - 6161.407 + 3154.063 = 2.2209,$$

т. е. результат, значительно более точный, чем в примере 3, однако требующий вычислений с повышенной точностью (этого не требует квадратичная интерполяция по Бесселю).

Полином Лагранжа. Логическим продолжением линейной и квадратичной интерполяции и экстраполяции является *полином Лагранжа* [11, с. 502]. Пусть в узлах (8.73) заданы значения функции (8.74). Тогда через них можно провести (единый) полином n -й степени (обобщение формулы (8.81)):

$$\tilde{y}(x) = \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} y_0 + \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} y_1 + \dots + \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})} y_n \quad (8.82)$$

ИЛИ

$$\tilde{y}(x) = \sum_{j=0}^n \left[\prod_{\substack{i=0 \\ i \neq j}}^n \left(\frac{x-x_i}{x_j-x_i} \right) \right] y_j. \quad (8.83)$$

Если $x \in [x_0, x_n]$, то полином Лагранжа будет *интерполяционным*, а если $x \notin [x_0, x_n]$, то *экстраполяционным*.

Если $x = x_0$, то $\tilde{y}(x) = y_0$; если $x = x_1$, то $\tilde{y}(x) = y_1$; ... если $x = x_n$, то $\tilde{y}(x) = y_n$, т.е. полином Лагранжа проходит точно через заданные значения функции: $\tilde{y}(x_j) = y_j$, $j = \overline{0, n}$, однако между узлами интерполяции x_j полином, как правило, неустойчив, причем степень неустойчивости растет с увеличением n (см. рис. 8.10).

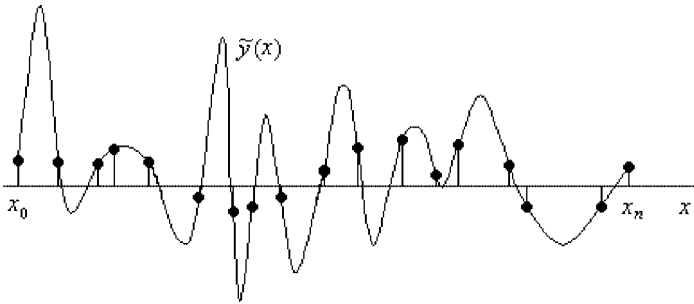


Рис. 8.10

Кроме того, даже небольшие погрешности значений y_j могут сильно изменить ход полинома между узлами. Отмеченные эффекты неустойчивости проявляются, когда $n \gtrsim 15$, а в практических задачах, когда $n \sim 100$, неустойчивость полинома Лагранжа очень велика.

Отмеченные недостатки полинома Лагранжа породили идею создания вместо единого полинома кусочно-полиномиальной функции, или сплайна.

Сплайны [2, 11, 19, 22, 23, 29, 48, 66, 90].

Определение. *Сплайн* — это кусочно-полиномиальная функция такая, что:

- 1) Вся область $[a, b]$ разбита на подобласти (участки, отрезки), вообще говоря, неодинаковой длины, в каждой из которых функция есть полином (многочлен) степени $m \geq 0$.
- 2) На границах подобластей полиномы стыкуются до производных $d = (m - p)$ -го порядка, где $p \in [1, m + 1]$ — дефект сплайна, а $d \in [-1, m - 1]$ — максимальный порядок непрерывной производной.

Другими словами, сплайн есть совокупность полиномов степени m , различных для каждой подобласти, но стыкующихся на границах подобластей до производных $(m - p)$ -го порядка включительно. Стыковка производных позволяет делать плавным переход между

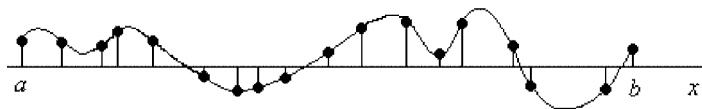


Рис. 8.11

подобластями (тем более плавный, чем меньше p). На рис. 8.11 изображен один из возможных сплайнов по данным рис. 8.10.

Мы видим, что кривая на рис. 8.11 значительно более плавная, чем на рис. 8.10.

Особенность сплайна состоит в том, что он является гораздо более устойчивой и гладкой функцией, чем полином Лагранжа, причем гладкость можно регулировать значением p (чем меньше p , тем более гладким будет сплайн), а также степенью m .

Если $d = -1$, то сплайны называются *разрывными* (возможен разрыв значений полиномов на границах подобластей), или *B-сплайнами* (хронологически первыми сплайнами) — см. рис. 8.12.

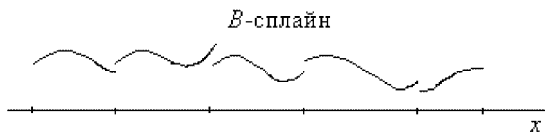


Рис. 8.12

Если $d = 0$, то на границах подобластей стыкуются лишь значения полиномов (*непрерывные сплайны*) — см. рис. 8.13.

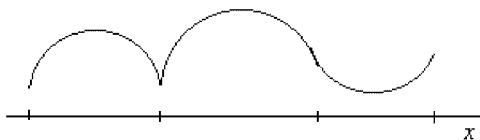


Рис. 8.13

Частный случай непрерывных сплайнов — это случай, когда $m = 1$ (при этом $d = 0$, $p = 1$). В этом случае сплайн дает *линейную интерполяцию* (см. рис. 8.14).

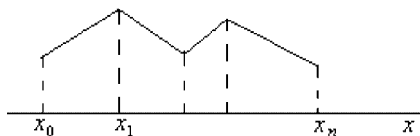


Рис. 8.14

А если $d = m - 1$, то на границах подобластей стыкуются производные от полиномов до $(m - 1)$ -го порядка (*классические сплайны*). Пример: нижеследующие кубические сплайны, получившие наибольшие приложения в физике и технике.

Кубические интерполирующие сплайны. Рассматривается задача интерполяции функции $y(x)$ с использованием значений функции (8.74) в узлах (8.73).

Определение. *Кубическим сплайном дефекта $p = 1$* называется функция $S(x)$, удовлетворяющая следующим требованиям [11, с. 504], [19, с. 84], [22], [23, с. 35], [48, с. 140]:

1) $S(x)$ непрерывна вместе со своими производными до второго порядка включительно: $S(x_j - 0) = S(x_j + 0)$, $S'(x_j - 0) = S'(x_j + 0)$, $S''(x_j - 0) = S''(x_j + 0)$;

2) на каждом отрезке $[x_{j-1}, x_j]$ она является кубическим полиномом (т. е. $m = 3$):

$$S(x) = S_j(x) = \sum_{l=0}^3 a_l^{(j)} (x_j - x)^l, \quad j = \overline{1, n}; \quad (8.84)$$

3) в узлах сетки (8.73) выполняются равенства $S(x_j) = y_j$, $j = \overline{0, n}$;

4) для $S''(x)$ выполняются граничные условия:

$$S''(a) = S''(b) = 0. \quad (8.85)$$

Выполнение приведенных требований приводит к тому, что выражение (8.84) для сплайна принимает конкретный вид:

$$S(x) = S_j(x) = \mu_{j-1} \frac{(x_j - x)^3}{6h_j} + \mu_j \frac{(x - x_{j-1})^3}{6h_j} + \left(y_{j-1} - \frac{\mu_{j-1} h_j^2}{6} \right) \frac{x_j - x}{h_j} + \left(y_j - \frac{\mu_j h_j^2}{6} \right) \frac{x - x_{j-1}}{h_j}, \quad (8.86)$$

где $h_j = x_j - x_{j-1}$ (шаг сплайна), а величины $\mu_j = S''(x_j)$, $j = \overline{0, n}$, подлежат определению (см. рис. 8.15).

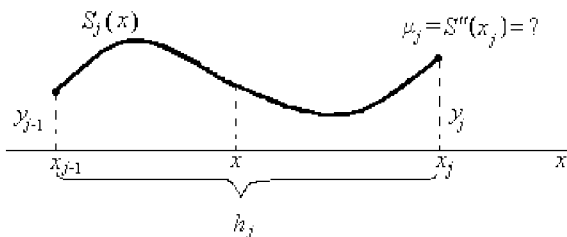


Рис. 8.15

Из (8.85) следует: $\mu_0 = \mu_n = 0$, а $\mu_1, \mu_2, \dots, \mu_{n-1}$ определяются из СЛАУ:

$$A\mu = Dy, \quad (8.87)$$

где квадратная матрица A размера $(n-1) \times (n-1)$ имеет вид (3-ленточная симметричная матрица):

$$A = \begin{pmatrix} \frac{h_1 + h_2}{3} & \frac{h_2}{6} & 0 & \dots & 0 & 0 \\ \frac{h_2}{6} & \frac{h_2 + h_3}{3} & \frac{h_3}{6} & \dots & 0 & 0 \\ 0 & \frac{h_3}{6} & \frac{h_3 + h_4}{3} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{h_{n-1}}{6} & \frac{h_{n-1} + h_n}{3} \end{pmatrix}, \quad (8.88)$$

прямоугольная матрица D размера $(n-1) \times (n+1)$ имеет вид (3-ленточная матрица):

$$D = \begin{pmatrix} \frac{1}{h_1} \left(-\frac{1}{h_1} - \frac{1}{h_2} \right) & \frac{1}{h_2} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & \frac{1}{h_2} & \left(-\frac{1}{h_2} - \frac{1}{h_3} \right) & \frac{1}{h_3} & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \frac{1}{h_{n-1}} \left(-\frac{1}{h_{n-1}} - \frac{1}{h_n} \right) & \frac{1}{h_n} \end{pmatrix}, \quad (8.89)$$

вектор $y = (y_0, y_1, \dots, y_n)^T$ задан, а вектор $\mu = (\mu_1, \mu_2, \dots, \mu_{n-1})^T$ является искомым.

Матрица A является положительно определенной и неособенной, поскольку имеет преобладающие диагональные элементы, значит, система (8.87) однозначно разрешима, а вычисления $\mu_1, \mu_2, \dots, \mu_{n-1}$ будут выполняться с высокой точностью.

Если шаг сплайна $h_j = h = (b-a)/n = \text{const}$, то

$$A = \begin{pmatrix} \frac{2}{3}h & \frac{1}{6}h & 0 & \dots & 0 & 0 \\ \frac{1}{6}h & \frac{2}{3}h & \frac{1}{6}h & \dots & 0 & 0 \\ 0 & \frac{1}{6}h & \frac{2}{3}h & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{6}h & \frac{2}{3}h \end{pmatrix}, \quad (8.90)$$

$$D = \begin{pmatrix} \frac{1}{h} & -\frac{2}{h} & \frac{1}{h} & \dots & 0 & 0 & 0 \\ 0 & \frac{1}{h} & -\frac{2}{h} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{h} & -\frac{2}{h} & \frac{1}{h} \end{pmatrix}. \quad (8.91)$$

Например, $n = 2$ (см. рис. 8.16). Тогда $\mu_0 = \mu_2 = 0$,

$$A = \frac{2}{3}h, \quad D = \left(\frac{1}{h}, -\frac{2}{h}, \frac{1}{h}\right), \quad y = (y_0, y_1, y_2)^T, \quad Dy = \frac{y_0 - 2y_1 + y_2}{h}, \quad (8.92)$$

$$\mu_1 = \frac{Dy}{A} = \frac{3}{2} \frac{y_0 - 2y_1 + y_2}{h^2}. \quad (8.93)$$

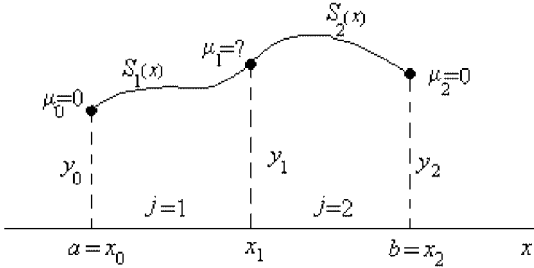


Рис. 8.16

В результате (см. (8.86))

$$\begin{aligned} S_1(x) &= \frac{\mu_1}{6h} (x - x_0)^3 + y_0 \cdot \frac{x_1 - x}{h} + \left(y_1 - \frac{\mu_1 h^2}{6}\right) \cdot \frac{x - x_0}{h} = \\ &= \frac{y_0 - 2y_1 + y_2}{4h^3} (x - x_0)^3 + y_0 \cdot \frac{x_1 - x}{h} - \frac{y_0 - 6y_1 + y_2}{4h} \cdot \frac{x - x_0}{h}, \end{aligned} \quad (8.94)$$

$$S_2(x) = \frac{\mu_1}{6h} (x_2 - x)^3 + \left(y_1 - \frac{\mu_1 h^2}{6}\right) \cdot \frac{x_2 - x}{h} + y_2 \cdot \frac{x - x_1}{h}. \quad (8.95)$$

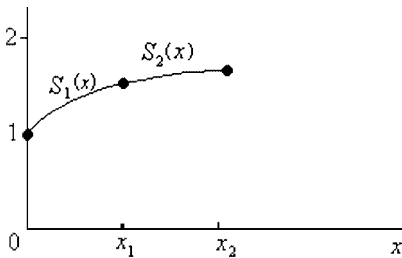


Рис. 8.17

Пример. Пусть $a = x_0 = 0$, $x_1 = 1$, $b = x_2 = x_n = 2$, $h = 1$, $y_0 = 1$, $y_1 = 1.5$, $y_2 = 1.6$ (см. рис. 8.17).

Тогда $\mu_1 = S''(1) = -0.6$ (см. (8.93)), а с помощью (8.94), (8.95) найдем:

$$S_1(x) = -0.1x^3 + 0.6x + 1, \quad (8.96)$$

$$S_2(x) = -0.1(2 - x)^3 + 1.6. \quad (8.97)$$

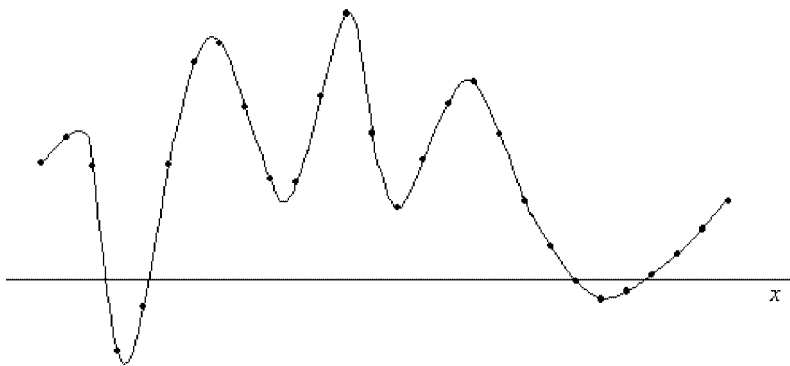


Рис. 8.18

Если n велико (например, 50), то будет «сшито» много полиномов в единый сплайн (см. рис. 8.18).

В [58, с. 134] рассмотрен случай, когда для кубических сплайнов заданы не только значения функции y_j , но и ее производных y'_j в узлах (8.73), причем для дефекта $p = 2$.

Отметим следующие положительные особенности сплайнов:

— они хорошо строятся графически (в ряде компьютерных редакторов, например, Графор, Grapher и др., есть операция сплайн-интерполяции),

— сплайны удобно дифференцируются (например, можно дважды непрерывно дифференцировать выражения (8.86), (8.94)–(8.97)), поэтому они используются для дифференцирования таблично заданных функций,

— сплайны удобно интегрируются (как интегралы от полиномов), поэтому они используются для интегрирования таблично заданных функций,

— сплайны применяются для решения интегральных уравнений [19, с. 86].

Отметим еще, что сплайн-интерполяция используется также для функции двух переменных $f(x, y)$ и большего числа переменных [2, 66].

Изложенные выше линейная, квадратичная и сплайн-интерполяция (а также экстраполяция) применимы, главным образом, в случае, когда значения функции (8.74) являются незашумленными, например, табличные значения функции e^x или $\sin x$ и т. д. Однако картина может резко измениться, когда в качестве табличных значений функции (8.74) выступают зашумленные (в первую очередь, экспериментальные) значения. На рис. 8.19а изображены точками точные значения некоторой функции ($h = \text{const}$) и через них проведен кубический сплайн (дефекта $p = 1$). Мы видим, что сплайн дал хорошую интерполяционную кривую и даже хорошие кривые для первой и второй производных (не приведенные на рисунке). На рис. 8.19б точками отображены зашумленные значения

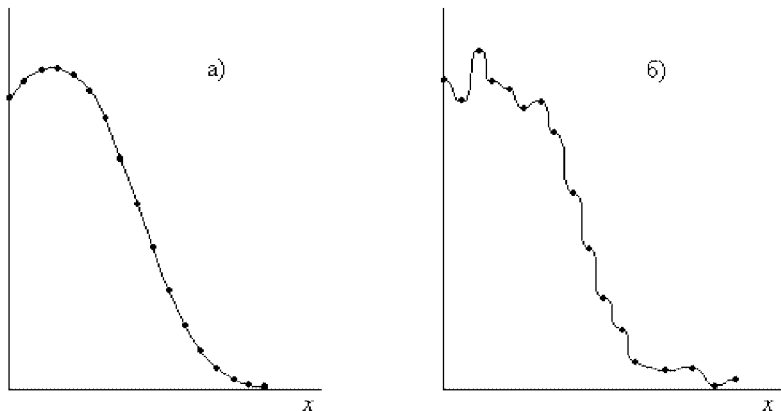


Рис. 8.19

функции и проведенный через них кубический сплайн ($p = 1$). Мы видим, что сплайн хотя и аккуратно соединил точки, но дал флуктуирующую кривую, которую, конечно, нельзя использовать для отыскания первой и второй производных.

Главная причина флуктуаций — в том, что используемый сплайн является интерполяционным. Для данного случая более эффективными являются сглаживающие, или аппроксимирующие сплайны. Поэтому рассмотрим вопрос о сглаживании и аппроксимации полиномами.

Линейная аппроксимация (линейное сглаживание). Ставится следующая задача: используя точки (8.73), (8.74), провести такую (единую) прямую линию

$$\tilde{y}(x) = \beta x + \gamma, \quad (8.98)$$

где β и γ — некоторые коэффициенты, чтобы

$$\sum_{j=0}^n [\tilde{y}(x_j) - y_j]^2 = \min_{\beta, \gamma}, \quad (8.99)$$

т. е. чтобы невязка между прямой $\tilde{y}(x)$ и заданными значениями y_j была минимальна. Из (8.99) видно, что прямая (8.98) находится методом наименьших квадратов.

Запишем (8.99) иначе:

$$\sum_{j=0}^n (\beta x_j + \gamma - y_j)^2 = \min_{\beta, \gamma}. \quad (8.100)$$

Приравнявая производные от (8.100) по β и γ нулю, получим

$$\left. \begin{aligned} \sum_{j=0}^n x_j (\beta x_j + \gamma - y_j) &= 0, \\ \sum_{j=0}^n (\beta x_j + \gamma - y_j) &= 0. \end{aligned} \right\} \quad (8.101)$$

Запишем (8.101) в виде системы двух алгебраических уравнений относительно β и γ :

$$\left. \begin{aligned} \left(\sum_{j=0}^n x_j^2 \right) \beta + \left(\sum_{j=0}^n x_j \right) \gamma &= \sum_{j=0}^n x_j y_j, \\ \left(\sum_{j=0}^n x_j \right) \beta + (n+1) \gamma &= \sum_{j=0}^n y_j. \end{aligned} \right\} \quad (8.102)$$

Прямая линия (8.98) осуществляет линейную аппроксимацию (или линейное сглаживание). Эта прямая может использоваться как для интерполяции (при $x \in [a, b]$), так и для экстраполяции (при $x \notin [a, b]$) — см. рис. 8.20.

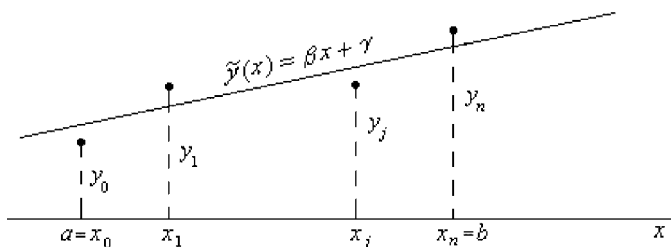


Рис. 8.20

Пример: $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $y_0 = 1$, $y_1 = 2.5$, $y_2 = 2$ ($n = 2$). СЛАУ (8.102) получается в виде:

$$\left. \begin{aligned} 5\beta + 3\gamma &= 6.5, \\ 3\beta + 3\gamma &= 5.5, \end{aligned} \right\} \quad (8.103)$$

откуда $\beta = 0.5$, $\gamma = 4/3$. Следовательно,

$$\tilde{y}(x) = 0.5x + 4/3. \quad (8.104)$$

Квадратичная аппроксимация (квадратичное сглаживание). Более точной является квадратичная аппроксимация, согласно которой нужно найти такой квадратичный полином:

$$\tilde{y}(x) = \beta x^2 + \gamma x + \eta, \quad (8.105)$$

где β, γ, η — некоторые коэффициенты, что

$$\sum_{j=0}^n [\tilde{y}(x_j) - y_j]^2 = \min_{\beta, \gamma, \eta}. \quad (8.106)$$

Сглаживающие (аппроксимирующие) кубические сплайны [22], [23, с. 43], [29], [48, с. 27]. Логическим продолжением линейной и квадратичной аппроксимации являются кубические сплайны (дефекта $p = 1$), не проходящие точно через узлы (8.73)–(8.74), а проходящие «в среднем» через них. Такие сплайны используются в случае зашумленности значений y_j (см. рис. 8.196), когда вместо значений (8.74) заданы значения

$$\tilde{y}_0, \tilde{y}_1, \dots, \tilde{y}_{j-1}, \tilde{y}_j, \dots, \tilde{y}_n. \quad (8.107)$$

Такие сплайны называются *сглаживающими*, или *аппроксимирующими*.

Для данных кубических сплайнов (дефекта $p = 1$) используется определение кубических интерполирующих сплайнов (см. выше), однако из определения исключается требование 3 и добавляется условие минимума функционала [22], [23, с. 43], [48, с. 27]:

$$\sum_{j=0}^n [S_\alpha(x_j) - \tilde{y}_j]^2 + \alpha \int_a^b S_\alpha''^2(x) dx = \min_{S_\alpha(x)}, \quad (8.108)$$

где $S_\alpha(x)$ — искомый сглаживающий сплайн, а $\alpha > 0$ — параметр сглаживания. Условие (8.108) напоминает условие (8.5) в методе регуляризации Тихонова, а α — параметр регуляризации.

При $\alpha \rightarrow 0$ сплайн $S_\alpha(x)$ переходит в интерполирующий, т. е. будет выполняться требование 3: $S_0(x_j) = \tilde{y}_j, j = \overline{0, n}$, а при завышенных α функция $S_\alpha(x)$ будет слишком гладкой, т. е. необходимо выбрать умеренное значение α .

В работе [22] приведены тексты программ SMF1V1 и др. для минимизации функционала (8.108), выбора параметра α , вычисления значений сплайна $S_\alpha(x)$ и его первой и второй производных в узлах сетки, вообще говоря, отличной от (8.73). На рис. 8.21 приведен пример (типа рис. 8.196) расчета сглаживающего кубического сплайна $S_\alpha(x)$ по зашумленным данным $\tilde{y}_j, j = \overline{0, 39}$.

Контрольные задания и вопросы

1. Вычислить $y(x) = \cos x$ при $x = 1.022$ рад по формуле (8.80) линейной интерполяции, используя значения: $y_{j-1} = \cos 1.02 = = 0.5234$ и $y_j = \cos 1.03 = 0.5148$.

2. Вычислить $y(x) = \cos x$ при $x = 1.022$ по формуле (8.80) линейной экстраполяции, используя «далекие» отсчеты: $y_{j-1} = = \cos 0.3 = 0.9553$ и $y_j = \cos 0.31 = 0.9523$.

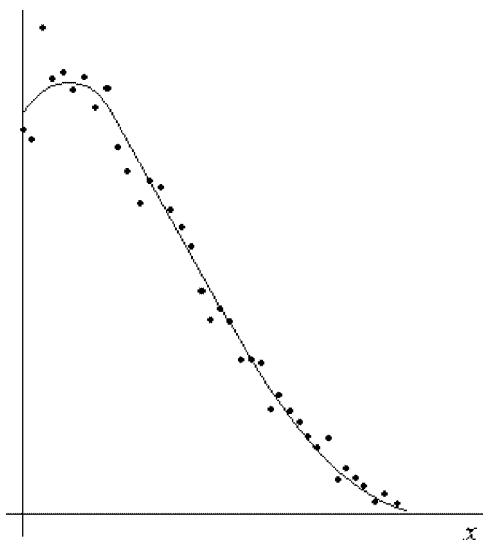


Рис. 8.21

3. Вычислить $y(x) = \cos x$ при $x = 1.022$ по формуле (8.81) квадратичной экстраполяции, используя также «далекие» отсчеты: $y_{j-1} = \cos 0.3 = 0.9553$, $y_j = \cos 0.31 = 0.9523$, $y_{j+1} = \cos 0.32 = 0.9492$. Сравнить полученные результаты.

4. Отметить достоинства и недостатки линейной и квадратичной интерполяции (и экстраполяции) и полинома Лагранжа.

5. Какие значения абсцисс соответствуют границам подобластей сплайна?

6. Нужно ли задавать значения функции $y(x)$ внутри подобластей или только на их границах для построения сплайна?

7. Используя (8.86), проверить, выполняются ли равенства: $S(x_j) = y_j$, $S(x_{j-1}) = y_{j-1}$, $S'''(x_j) = \mu_j$, $S'''(x_{j-1}) = \mu_{j-1}$.

8. При $n = 2$ (см. (8.92)–(8.95)) проставить размерности A , D , y и более подробно вывести (8.92), (8.93).

9. Используя (8.96), (8.97), определить значения $S_1(0)$, $S_1(1)$, $S_2(1)$, $S_2(2)$, $S'_1(0)$, $S'_1(1)$, $S'_2(1)$, $S'_2(2)$, $S''_1(0)$, $S''_1(1)$, $S''_2(1)$, $S''_2(2)$.

10. Для квадратичной аппроксимации, исходя из (8.105), (8.106), вывести систему трех алгебраических уравнений относительно β , γ , η , аналогичную системе (8.102) для линейной аппроксимации.

11. Анализируя рис.8.21, ответить, каким будет сплайн при $\alpha = 0$.

СПИСОК ЛИТЕРАТУРЫ

1. *Абрагам А.* Ядерный магнитный резонанс. — М.: Изд-во иностр. лит., 1963.
2. *Алберг Дж., Нильсон Э., Уолш Дж.* Теория сплайнов и ее приложения. — М.: Мир, 1972.
3. *Арсенин В. Я., Рубашов И. Б.* О решении некоторых математических задач компьютерной томографии // Вестн. МГУ. Сер. 15, 1986, № 3, с. 52–59.
4. *Бакушинский А. Б., Гончарский А. В.* Некорректные задачи. Численные методы и приложения. — М.: Изд-во МГУ, 1989.
5. *Березин И. С., Жидков Н. П.* Методы вычислений. В 2-х т. Т. 1. — М.: Наука, 1966.
6. *Бойков И. В.* Пассивные и адаптивные алгоритмы приближенного вычисления сингулярных интегралов. Ч. 1, 2. — Пенза: Изд-во ПГТУ, 1995.
7. *Бородин П. М.* (ред.). Ядерный магнитный резонанс. — Л.: Изд-во ЛГУ, 1982.
8. *Брайсон А., Хо Ю-ши.* Прикладная теория оптимального управления. — М.: Мир, 1972.
9. *Брейсуэлл Р.* Преобразование Хартли. — М.: Мир, 1990.
10. *Брижман М. С.* Интегральные модели в современной теории управления. — Рига.: Зинатне, 1979.
11. *Бронштейн И. Н., Семендяев К. А.* Справочник по математике для инженеров и учащихся втузов. Изд-е 13-е. — М.: Наука, 1986.
12. *Бурдун Г. Д., Марков Б. Н.* Основы метрологии. — М.: Изд-во стандартов, 1972.
13. *Бьёрне Л.* (ред.). Подводная акустика и обработка сигналов. — М.: Мир, 1985.
14. *Вапник В. Н.* (ред.). Алгоритмы и программы восстановления зависимостей. — М.: Наука, 1984.
15. *Василенко Г. И.* Теория восстановления сигналов: О редукции к идеальному прибору в физике и технике. — М.: Сов. радио, 1979.
16. *Васильев В. Н., Гуров И. П.* Компьютерная обработка сигналов в приложении к интерферометрическим системам. — СПб.: БХВ — Санкт-Петербург, 1998.
17. *Верлань А. Ф., Абдусатаров Б. Б., Игнатченко А. А., Максимович Н. А.* Методы и устройства интерпретации экспериментальных зависимостей при исследовании и контроле энергетических процессов. — Киев: Наук. думка, 1993.

18. *Верлань А. Ф., Москалюк С. С.* Математическое моделирование непрерывных динамических систем. — Киев: Наук. думка, 1988.
19. *Верлань А. Ф., Сизиков В. С.* Интегральные уравнения: методы, алгоритмы, программы. — Киев: Наук. думка, 1986.
20. *Воеводин В. В.* Численные методы алгебры: Теория и алгоритмы. — М.: Наука, 1966.
21. *Воеводин В. В.* Линейная алгебра. — М.: Наука, 1980.
22. *Воскобойников Ю. Е., Ицкович Е. И.* Пакет подпрограмм для построения сглаживающих кубических сплайнов. Препринт 46–79. — Новосибирск: Ин-т теплофизики АН СССР, 1979.
23. *Воскобойников Ю. Е., Преображенский Н. Г., Седельников А. И.* Математическая обработка эксперимента в молекулярной газодинамике. — Новосибирск: Наука, 1984.
24. *Галайдин П. А.* Исследование и разработка элементов и устройств компьютерных магниторезонансных томографов. Дис. ... докт. техн. наук. — СПб: ИТМО, 1996.
25. *Галайдин П. А., Замятин А. И., Иванов В. А.* Основы магниторезонансной томографии. — СПб: Изд-во ИТМО, 1998.
26. *Гантмахер Ф. Р.* Теория матриц. — М.: Наука, 1966.
27. *Глазов М. В., Болохова Т. А.* Решение редукционной проблемы Рэлея с использованием различных модификаций метода регуляризации // Оптика и спектроскопия, 1989, т. 67, вып. 3, с. 533–537.
28. *Горшков А. В.* Улучшение разрешения изображений при обработке данных физического эксперимента и нахождение неизвестной аппаратной функции по программам пакета REIMAGE // Приборы и техника эксперимента, 1995, № 2, с. 68–78.
29. *Гребенников А. И.* Метод сплайнов и решение некорректных задач теории приближений. — М.: Изд-во МГУ, 1983.
30. *Днестровский Ю. Н., Костомаров Д. П.* Математические задачи диагностики плазмы // Некорректные задачи естествознания / Под ред. А. Н. Тихонова и А. В. Гончарского. — М.: Изд-во МГУ, 1987, с. 103–134.
31. *Жуков В. Б.* Расчет гидроакустических антенн по диаграмме направленности. — Л.: Судостроение, 1977.
32. *Журавлев А. К., Лукашкин А. П., Поддубный С. С.* Обработка сигналов в адаптивных антенных решетках. — Л.: Изд-во ЛГУ, 1983.
33. *Задирака В. К.* Теория вычисления преобразования Фурье. — Киев: Наук. думка, 1983.
34. *Иванов В. А.* Способ определения внутреннего строения материальных объектов. А. с. № 1112266 // Открытия, изобретения, 1984, № 33 (Приоритет от 21.03.60).
35. *Иванов В. А.* Внутривидение (ЯМР-томография). — Л.: Знание, 1989.

36. *Иванов В. К., Васин В. В., Танана В. П.* Теория линейных некорректных задач и ее приложения. — М.: Наука, 1978.
37. *Ишлинский А. Ю., Черный Г. Г.* (ред.). Метод граничных интегральных уравнений (Новое в зарубежной науке. Механика. Сер. 15). — М.: Мир, 1978.
38. *Кей С. М., Марпл С. Л.* Современные методы спектрального анализа (обзор) // Тр. Ин-та инж. по электротехнике и радиоэлектрон., 1981, т. 69, № 11, с. 5–51.
39. *Клепиков Н. П., Соколов С. Н.* Анализ и планирование экспериментов методом максимума правдоподобия. — М.: Наука, 1964.
40. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. — М.: Наука, 1968.
41. *Котельников В. А.* Теория потенциальной помехоустойчивости. — М.-Л.: Госэнергоиздат, 1956.
42. *Красильников В. А.* Звуковые и ультразвуковые волны в воздухе, воде и твердых телах. — М.: Физматгиз, 1960.
43. *Краснов М. Л., Киселев А. И., Макаренко Г. И.* Интегральные уравнения. — М.: Наука, 1976.
44. *Крауляня Э. К., Лиена С. Я., Пикалов В. В., Скудра А. Я.* К проблеме исследования атомной сенсibilизированной флуоресценции по контурам спектральных линий // Некорректные обратные задачи атомной физики / Под ред. Н. Г. Преображенского. — Новосибирск: Изд-во ИТПМ, 1976, с. 61 — 72.
45. *Лаврентьев М. М., Романов В. Г., Шишатский С. П.* Некорректные задачи математической физики и анализа. — М.: Наука, 1980.
46. *Ландсберг Г. С.* Оптика (общий курс физики). — М.: Наука, 1976.
47. *Монзинго Р. А., Миллер Т. У.* Адаптивные антенные решетки: Введение в теорию. — М.: Радио и связь, 1986.
48. *Морозов В. А.* Регулярные методы решения некорректно поставленных задач. — М.: Наука, 1987.
49. *Наттерер Ф.* Математические аспекты компьютерной томографии. — М.: Мир, 1990.
50. *Пиблз, Беркович.* Многолучевой моноимпульсный радиолокатор // Зарубеж. радиоэлектрон., 1969, № 10.
51. *Потеев М. И., Сизиков В. С.* Повышение разрешающей способности измерительных устройств путем компьютерной обработки результатов измерений. // СПб: Изд-во ИТМО, 1992.
52. *Преображенский Н. Г., Пикалов В. В.* Неустойчивые задачи диагностики плазмы. — Новосибирск: Наука, 1982.
53. *Проخورов А. М.* (ред.) Физический энциклопедический словарь. — М.: Сов. Энциклопедия, 1984.
54. *Пытьев Ю. П.* Математические методы интерпретации эксперимента. Учеб. пособие для ВУЗов. — М.: Высш. шк., 1989.

55. *Пытьев Ю. П., Чуличков А. И.* Прибор + ЭВМ = новые возможности. — М.: Знание, 1983.
56. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов. — М.: Мир, 1978.
57. *Розенвассер Е. Н.* Периодически нестационарные системы управления. — М.: Наука, 1973.
58. *Самохин А. Б., Самохина А. С.* Численные методы и программирование на Фортране для персонального компьютера. — М.: Радио и связь, 1996.
59. *Сизиков В. С.* Обобщенный метод редукции измерений. I, II, III // Электрон. моделирование, 1991, т. 13, № 4, с. 7–14, № 5, с. 9–14, № 6, с. 3–9.
60. *Сизиков В. С.* Использование регуляризации для устойчивого вычисления преобразования Фурье // Ж. вычисл. матем. и матем. физики, 1998, т. 38, № 3, с. 376–386.
61. *Сизиков В. С.* Пакеты программ CONF, FFTREG, INEQF, SLNE на Fortran'e MS 5 и Fortran'e 90. — СПб: СПбГИТМО (ТУ), 1998.
62. *Сизиков В. С., Кузьмин А. В., Козаченко А. В.* Обработка дефокусированных изображений методами двухмерного преобразования Хартли и регуляризации Тихонова // Изв. вузов. Приборостроение, 1999, т. 42, № 8, с. 12–16.
63. *Сизиков В. С., Российская М. В., Козаченко А. В.* Обработка смазанного изображения методами дифференцирования, преобразования Хартли и регуляризации Тихонова // Изв. вузов. Приборостроение, 1999, т. 42, № 7, с. 11–15.
64. *Слепян Л. И., Яковлев Ю. С.* Интегральные преобразования в нестационарных задачах механики. — Л.: Судостроение, 1980.
65. *Смарышев М. Д.* Направленность гидроакустических антенн. — Л.: Судостроение, 1973.
66. *Стечкин С. Б., Субботин Ю. Н.* Сплаины в вычислительной математике. — М.: Наука, 1976.
67. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. — М.: Наука, 1986.
68. *Тихонов А. Н., Арсенин В. Я., Рубашов И. Б., Тимонов А. А.* Первый советский компьютерный томограф // Природа, 1984, № 4, с. 11–21.
69. *Тихонов А. Н., Арсенин В. Я., Тимонов А. А.* Математические задачи компьютерной томографии. — М.: Наука, 1987.
70. *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.

71. *Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г.* Численные методы решения некорректных задач. — М.: Наука, 1990.
72. *Тюрин А. М.* Введение в теорию статистических методов в гидроакустике. — Л.: Изд-во ВМОЛА, 1963.
73. *Тюрин А. М., Сташкевич А. П., Таранов Э. С.* Основы гидроакустики. — Л.: Судостроение, 1966.
74. *Уэбб С.* (ред.). Физика визуализации изображений в медицине. В 2-х т. — М.: Мир, 1991.
75. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры. — М.: Физматгиз, 1963.
76. *Фалькович С. Е., Коновалов Л. Н.* Разрешение неизвестного числа сигналов // Радиотехника и электрон., 1982, т. 27, № 1, с. 92–97.
77. *Фриш С. Э., Тиморева А. В.* Курс общей физики. В 3-х т. — М.: Физматгиз, 1962.
78. *Фрост О. Л.* Алгоритм линейно-ограниченной обработки сигналов в адаптивной решетке // Тр. Ин-та инж. по электротехнике и радиоэлектрон., 1972, т. 60, № 8, с. 5–16.
79. *Хермен Г.* Восстановление изображений по проекциям. Основы реконструктивной томографии. — М.: Мир, 1983.
80. *Химмельблау Д.* Прикладное нелинейное программирование. — М.: Мир, 1975.
81. *Хуанг Т.* (ред.). Обработка изображений и цифровая фильтрация. — М.: Мир, 1979.
82. *Чейссон Э. Дж.* Первые результаты с космического телескопа «Хаббл» // В мире науки, 1992, № 8, с. 6–14.
83. *Шендеров Е. Л.* Волновые задачи гидроакустики. — Л.: Судостроение, 1972.
84. *Эндрюс Г.* Применение вычислительных машин для обработки изображений. — М.: Энергия, 1977.
85. *Эрнст Р., Боденхаузен Дж., Вокаун А.* ЯМР в одном и двух измерениях. — М.: Мир, 1990.
86. *Яноши Л.* Теория и практика обработки результатов измерений. — М.: Мир, 1968.
87. *Яценко Ю. П.* Интегральные модели систем с управляемой памятью. — Киев: Наук. думка, 1991.
88. *Adamiak K.* Method of the magnetic field synthesis on the axis of cylinder solenoid // Appl. Phys., 1978, v. 16, p. 417–423.
89. *Anderson V. C.* DICANNE, a realizable adaptive process // J. Acoust. Soc. Amer., 1969, v. 45, № 2, p. 398–405.
90. *Brunner H., Houwen P. J. van der.* The numerical solution of Volterra equations. — Amsterdam: North-Holland, 1986.
91. *Brunner H., Sizikov V.* On a suboptimal filtration method for solving convolution-type integral equations of the first kind // J. Math. Analysis and Appl., 1998.

92. *Butler J. P., Mohler J. G.* Estimating a distribution's central moments: a specific tidal ventilation application // *J. Appl. Physiol.: Respirat. Environ. Exercise Physiol.*, 1979, v. 46, № 1, p. 47–52.
93. *Cho Z. H., Jones J. P., Singh M.* Foundations of medical imaging. — New York: Wiley, 1993.
94. *Eveson S. P.* An integral equation arising from a problem in mathematical biology // *Bull. London Math. Soc.*, 1991, v. 23, № 3, p. 293–299.
95. *Holt J. H., Bracken A. J.* First kind Fredholm integral equation of liver kinetics: numerical solutions by constrained least squares // *Math. Biosci.*, 1980, v. 51, № 1/2, p. 11–24.
96. *Kawanaka A., Takagi M.* Estimation of static magnetic field and gradient fields from NMR image // *J. Phys. E: Sci. Instrum.*, 1986, v. 19, p. 871–875.
97. *Reddi S. S.* Multiple source location — A digital approach // *IEEE Trans. Aerosp. Electron. Syst.*, 1979, AES-15, № 1, p. 95–105.
98. *Silvia M. T.* Deconvolution // In: *Handbook of digital signal processing. Engineering applications* / D. F. Elliot (ed.). — San Diego: Acad. Press, 1987, p. 741–788.
99. *Sizikov V. S.* Integral equations in NMR-tomography: magnetic field synthesis on a coil axis // *Proc. of 5th Int. Conf. IMSE98* / B. S. Bertram (ed.). — Houghton, USA, 1998, p. 76–77.
100. *Sizikov V. S.* Integral equations in NMR-tomography: reconstruction of NMR images with a regularization // *Proc. of 5th Int. Conf. IMSE98* / B. S. Bertram (ed.). — Houghton, USA, 1998, p. 74–75.
- * * *
101. *Бейтс Р., Мак-Доннелл М.* Восстановление и реконструкция изображений. — М.: Мир, 1989.
102. *Белов И. А.* Пакет программ IMAGE на Visual C++. — СПб: СПбГИТМО(ТУ), 1999.
103. *Гельфанд И. М., Шилов Г. Е.* Обобщенные функции и действия над ними. — М.: Физматгиз, 1958.
104. *Сизиков В. С.* Анализ методов локальной регуляризации и формулировка метода субоптимальной фильтрации решения уравнений I рода // *Ж. вычисл. матем. и матем. физики*, 1999, т. 39, № 5, с. 718–733.
105. *Сизиков В. С., Белов И. А.* Реконструкция смазанных и дефокусированных изображений методом регуляризации // *Оптич. ж.*, 2000, т. 67, № 3.
106. *Sizikov V. S., Belov I. A.* Modelling of problem of distorted image reconstruction by regularization method // *Proc. of 2nd Intern. Confer. "Tools for Mathem. Modelling"* / G. S. Osipenko, Yu. G. Ivanov (eds.). — St-Petersburg, 2000.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Автокорреляционные функции**
решения и помехи 210
- Алгоритм интегральной аппроксимации** 88, 113
- Ансамбль протонов** 36
– реализаций 108, 206
- Антенна** 106
- Апостериорная матрица ковариаций**
ошибок решения 207
- Аппаратная функция (АФ)** 10, 82, 101, 115, 134
- Аппаратурные искажения** 26, 135
- Аппроксимация** 221
- Априорная информация о решении** 77, 195
– ковариация ошибок решения 207
- Белый шум** 22, 29
- Быстрое преобразование Фурье (БПФ)** 32, 71, 170
- V-сплайны (разрывные сплайны)** 216
- Вектор-столбец** 142
– ядерной намагниченности **М** 36
- Весовые коэффициенты преобразователей** 106
- Визуализация результатов** 30, 136
- Восстановление дефокусированных изображений** 72, 136
– искаженных изображений 63
– биологических микрообъектов 124, 135
– сигнала в динамической системе 119, 137
– смазанных изображений 63, 137
– спектра 83, 135
- Время поперечной релаксации**
 T_2, T_2^* 39
– продольной релаксации T_1 39
- Вход** 10, 100
- Выход** 10, 100
- Гармоника Фурье** 161
- Гидроакустика** 9, 100, 110
- Гиромангнитное отношение γ** 34
- Гладкость решения** 195
- Градиентные поля G_x, G_y, G_z** 44
- Граничные условия** 132
- Датчик случайных чисел RNDAN** 150
- Детектор** 18
- Дефект сплайна** 215
- Дефокусированное изображение** 72
- Дефокусировка** 74
- Диагностика плазмы** 91
- Дискретизация** 57, 87, 166, 180
- Дисперсия** 148
- Дифференциальное уравнение** 34, 128, 132
– Гельмгольца 132
 δ -функция Дирака 156
- Задача компьютерной томографии** 17
– корректная (well-posed) 178
– некорректная (ill-posed) 178
– неустойчивая 178
– обратная 11, 100
– прикладная 11
– спектроскопии 79
– редукции к идеальному спектральному прибору 83
– удельной приливной вентиляции в легких 127
– устойчивая 178
- Закон Бера** 18
– распределения ошибок 148
- Замер** 206
- Идеальное измерительное устройство** 111
- Измерительно-вычислительный комплекс (ИВК)** 118
- Измерительное устройство** 9, 110
- Изображение** 63, 140
– по Лапласу 175
– по Фурье 160
- Импульс высокочастотный (ВЧ)** 41
- Импульсная переходная функция** 121
- Импульсы $\pi/2$ и π Карра-Парселла** 41
- Индикаторный процесс (ИП)** 101, 110, 113, 134
- Индукция магнитного поля** 34, 45
- Интеграл Фурье** 158
- Интегральное (непрерывное) преобразование** 138
– Лапласа 175
– Фурье (НПФ) 31, 129, 139, 160
– Хартли (НПХ) 172

- Интегральное уравнение 134
- Абеля 97
 - Вольтерры (I, II, III рода) 137
 - Вольтерры-Урысона 138
 - двухмерное 135
 - Радона 136
 - регуляризованное 194
 - сингулярное 137
 - с параметром 137
 - типа свертки 135, 137
 - двухмерное 135
 - Урысона (нелинейное) 61, 136
 - Фредгольма (I, II, III рода) 134
 - Цейделя 97, 137
- Интенсивность 18
- Интервал (шаг) дискретизации
 $h = \Delta t$ по t и Δf по частоте 166
- Интерполяция (квадратичная, линейная, по Бесселю, по Лагранжу) 212
- Искомая функция 134
- Канал пространственный** 103, 112
- Катушка ЯМР-томографа 55
- Ковариация ошибок правой части 206
- решения 207
- Кодирование пространства (фазовое, частотное) 45
- Компьютерная томография (КТ) 17
- Контраст изображения 30, 55
- Конус прецессии 35
- Корректность и некорректность по Адамару 178
- Косинус-преобразование Фурье 139, 161, 165
- Кусочно-полиномиальная функция 215
- Ларморова прецессия** 35
- Линейная алгебра 142
- Локальный носитель supp 201
- Магнитное поле** (градиентное, переменное, поляризирующее, постоянное) 34, 36, 44
- Магнитный момент протона** 33
- Магожидание ψ (начальное приближение, априорная оценка, прогноз) решения 193, 207
- Матрица 142
- M - априорная ковариация ошибок решения 207
 - P - апостериорная ковариация ошибок решения 207
 - R - ковариация ошибок правой части 206
- Метод граничных интегральных уравнений 133, 136
- дифференцирования 68
 - импульсный 41
 - квадратур 180
 - наименьших квадратов (МНК) Гаусса 186
 - неопределенных множителей Лагранжа 192
 - обобщенный редукции измерений 113
 - оптимальной фильтрации Винера 210
 - Калмана-Бьюси 205
 - преобразования Фурье (ПФ) 21, 27, 51, 69
 - двухмерного 31, 76
 - псевдообратной матрицы (МПОМ) Мура-Пенроуза 189
 - регуляризации Тихонова 192
 - для уравнения типа свертки 199
 - «чувствительной точки» Хиншоу 51
 - эха Ганна 43
- Методы адаптации 105
- обработки сигналов (Андерсона, когерентной компенсации, максимальный, Фроста) 103, 109
 - пеленгования классические 103
 - регулярные 85, 206
 - численные 142, 195
- Минимизация функционала 200
- Многочлен 215
- Напряженность магнитного поля** 34, 45
- Начальное приближение 193, 207
- Невязка 186, 192
- Недоопределенная СЛАУ 189
- Некорректная задача 178
- Неоднородность полей ΔB_0 , ΔG_x , ΔG_y 51, 54
- Неопределенный множитель Лагранжа 193

- Неустойчивость решения 70, 180
 Норма (вектора, матрицы, оператора, решения) 140, 141, 144
 Нормальное псевдорешение 191
 – распределение 148
 – решение 189
Области применения РТ 26
 – спектрального анализа 80
 – ЯМР-томографии 61
 Область измерения 134
 – поиска решения 134
 Обобщенные функции 155
 Обобщенный метод редукции измерений 113
 – принцип невязки выбора α 194
 Обработка измерений 9
 – изображений 63
 – сигналов (signal processing) 100
 Обратная задача 11, 16
 –– диагностики плазмы 91, 136
 –– речевой акустики 126, 135
 – связь 119
 Обратное ПФ 160
 – ПХ 173
 Обратные задачи биофизики 123
 –– механики 119
 –– оптики 63
 –– спектроскопии 79, 134
 –– теории управления 119
 Обращение матрицы 144
 Объект 18, 63, 110, 131
 Ограничения (на решение) 109
 Оператор 10, 101, 112, 141, 178, 192
 Определитель 143
 Оригинал 140, 160, 175
 Ошибка – см. погрешность
Пакет программ 32, 71, 78, 90, 228
 –– IMAGE 71, 78, 230
 Параметр 66, 91, 101, 124, 137, 151
 – регуляризации α 22, 30, 55, 71, 77, 96, 170, 193, 223
 Переопределенная СЛАУ 186
 Периодичность ДПФ 169
 «Пила» 29, 58, 181
 Плазма 91
 Плотность вероятности 148
 – вещества 17
 – распределения 148
 – спектральная 210
 Погрешность 10, 22, 28, 144, 179
 – исходных данных (оператора, правой части, ядра) 10, 134, 144
 – решения 55, 144, 179, 211
 Поле постоянное B_0 38
 – статическое поляризующее 45
 Полином 212
 – Лагранжа 214
 Поля градиентные G_x, G_y, G_z 44
 Помеха 10, 29, 103, 206
 Порядок регуляризации q 200
 Правая часть 134
 Преобразование Лапласа 140, 175
 – Радона 19
 – Фурье (ПФ) 128, 129, 160
 –– быстрое (БПФ) 32, 170
 –– двухмерное 76, 139, 166
 –– дискретное (ДПФ) 32, 166
 –– непрерывное (НПФ) 31, 160
 –– обратное (ОПФ) 50, 76, 160, 164
 –– прямое 160, 164, 166
 – Хартли (ПХ) 139, 172
 –– быстрое (БПХ) 174
 –– дискретное (ДПХ) 174
 –– обратное (ОПХ) 173
 Преобразователи 107
 – аналого-цифровые (АЦП) 118
 – цифро-аналоговые (ЦАП) 118
 Прецессия магнитного момента протона 35
 Приемник 18, 94, 105, 115
 Прикладные задачи 11, 134
 Примеры модельные, численные 52, 57, 88, 171, 181, 196, 201
 – прикладных задач 11
 – реконструкции изображений 23, 65, 74, 75, 78, 79
 Прогноз решения 193, 207
 Программа CONVOL 201
 – RNDAN 150
 – SMF1V1 223
 Программы CONV1, CONV2, CONV3, CONV4, CONV5 201, 228
 – FFT, FTFC, FTFTC 170, 204
 – LOCAL0, LOCALINF, LOCALN, SUBOPT, OPT 228
 Программы РТИКР, РТИМР, РТИР, РТИТР, РТИЗР 196, 201

- Программы ТІКН1, ТІКН2, ТІКН3,
 ТІКН4, ТІКН5 196
 Прозрачность 20
 Прокол в ХН 105, 107
 Пространства L_1, L_2, W_2^1 140
 Пространственный канал 103, 112
 Псевдообратная матрица Мура-
 Пенроуза 189
 Псевдорешение 186
Рабочий объем томографа 53
 Радиопульс 122
 Разрешающая способность измери-
 тельного устройства 11, 83, 111,
 118, 198
 -- томограмм 51
 Ранг матрицы r 143
 Распад клеток 126, 134
 - радиоактивных элементов 126, 134
 Распределение ошибок (нормальное,
 равномерное) 148
 Расстройка по частоте $\Delta\omega = \omega_0 - \omega$
 37
 Расфазирование (потеря резонанса)
 39
 Реализация 206
 Регуляризатор q -го порядка 200
 Регуляризация 192
 - «интуитивная» 23
 - локальная 23
 Регуляризованное решение 200, 203,
 211
 - уравнение 194
 Редукционная проблема Рэлея 111
 Редукция измерений к идеальному
 измерительному устройству 110
 - локальных сигналов 112, 141
 - наблюдений биологических
 микроорганизмов 125
 - протяженных сигналов 114, 134
 Резонанс ($\omega_0 = \omega$) 37
 Результат измерений 10
 - обработки 10
 Реконструкция (формирование)
 изображений 63
 - рентгеновских изображений 20
 - ЯМР-изображений 45
 Релаксация 38
 - спин-решетчатая (продольная) 39
 - спин-спиновая (поперечная) 39
 Рентгеновская томография (РТ) 17
 Рентгенография 17
 Решение: неустойчивое, нормальное,
 регуляризованное, точное,
 устойчивое 178, 189, 193
Свертка 207
 Сглаживание (линейное, квадратич-
 ное, сплайнами) 221
 Сглаживающий функционал 223
 Сетка узлов (по s, t, x, ω) 32, 166,
 195, 212
 Сигнал входной 100, 134
 - выходной 100, 134
 - локальный (дискретный) 100
 - протяженный (распределенный)
 100
 Симметрия (круговая, цилиндриче-
 ская, шаровая) 96, 98
 Сингулярное число μ 144
 Синтез магнитного поля 134
 Синус-преобразование Фурье 139,
 161, 165
 Система линейно-нелинейных
 уравнений (СЛНУ) 86, 113, 141
 - линейных алгебраических уравне-
 ний (СЛАУ) 142
 ---- недоопределенная 189
 ---- переопределенная 186
 - нелинейных уравнений (СНУ) 54,
 87
 Сканирование 24, 51, 82
 Сканирующая функция (СФ) 134
 Слой 46
 Случайный процесс 108
 Смаз 66
 Собственное значение (число) 143
 Совокупность интегральных урав-
 нений 66, 137
 Спектр: дискретный, непрерывный,
 полосатый, экспериментальный
 80
 - излучения плазмы 92
 - Фурье 129, 160
 Спектральная плотность мощности
 (помехи, решения) 210
 - чувствительность (СЧ) 82
 Спектральные задачи механики 131
 Спектральный анализ 79
 Спектроскопия 79
 Спин 33
 Сплаины 215

- Способы выбора α : моделирования, монотонного решения, невязки, обобщенной невязки, подбора 194
- Сравнение методов Тихонова и Винера 211
- Тихонова и Калмана 207
- Существование решения 178, 188
- Схема измерений и обработки 9
- обработки в РТ 32
- пассивной диагностики плазмы 93
- Теорема Винера-Хинчина** 210
- Котельникова 166
- Техническая реализация алгоритмов редукции 118
- Типы интегральных уравнений 134
- спектрального анализа 81
- спектров 80
- Томограф рентгеновский (5 поколений) 24
- ядерно-магнитно-резонансный 55
- Угол компенсации** 106
- прихода сигнала 101
- Узлы (по s, t, x, ω) 166, 195
- Умножение матриц и векторов 145
- Уравнение дифференциальное 128
- Гельмгольца 132
- Лармора 34
- интегральное 134
- Абея 97
- Вольтерра (I, II, III рода) 137
- типа свертки 137
- Вольтерра-Урысона 138
- двухмерное 135
- линейное 134
- нелинейное 61, 136
- одномерное 134
- Радона 136
- регуляризованное 194
- типа свертки 135
- Фредгольма (I, II, III рода) 134
- Цейпеля 97, 137
- операторное 141, 178, 192
- Уравнения Блоха 38
- Устойчивость решения 30, 77, 211
- Устойчивые методы 132
- Фильтр Винера** 211
- инверсный 27, 31
- Калмана 206
- Формула Грина 132
- квадратурная прямоугольников 167
- трапеций 180, 195
- Эйлера 158
- Функционал 200, 203, 223
- Функция автокорреляционная 210
- аппаратная (АФ) 10, 82, 101, 134
- Дирака (δ -функция) 156
- зашумленная 69
- искомая 21, 27, 49, 61, 97, 112, 124, 127, 134, 180, 210
- обобщенная 155
- Хэвисайда 155
- Фурье-образ 129, 160
- Характеристика направленности (ХН) антенны** 106, 134
- чувствительности (ХЧ) пленки 67
- Характеристическое уравнение 143
- число 143
- Частота дискретизации (максимальная частота) $f_g = f_{\max}$** 166
- Частотная характеристика (ЧХ) 82
- Численные примеры 23, 51, 65, 74, 78, 88, 112, 145, 161, 171, 173, 179, 182, 190, 196, 213, 219, 222, 224
- Число обусловленности cond 144
- узлов 180, 198, 202
- характеристическое 143
- Шаг дискретизации** 166
- Шум 10, 29, 103, 108, 206
- Экстрополяция (квадратичная, линейная)** 212
- Элементы теории вероятностей 147
- обобщенных функций 155
- Эффект ЯМР 33
- Эхо-сигнал 41
- Ядро интегрального уравнения** 134
- разностное 134
- симметричное 188
- со слабой сингулярностью 97
- ЯМР-томография 33

V. S. Sizikov

STABLE METHODS

FOR PROCESSING OF MEASUREMENT RESULTS

Textbook

CONTENTS

Abbreviations	7
Introduction	9
General scheme of measurements and processing. Some examples of applied problems. Necessity of using stable methods. Comparison with other origins. Brief contents of the textbook. Control tasks and questions.	
Section I. INVERSE APPLIED PROBLEMS	16
Chapter 1. Problems of computer tomography	17
1.1. X-ray tomography (XT)	17
Shortages of usual X-radiography. Idea of XT. Setting of a problem. The Baire law. The Radon equation. Historical reference. Fredholm integral equation of the first kind. Solution by FT method without regularization and with regularization. Numerical illustrations. Five generations of X-ray tomographs. Application fields of XT. Removal of influence of instrument distortions. Visualization of results (representation of a slice on display). About algorithms and programs. General scheme of processing in XT. Control tasks and questions.	
1.2. Nuclear magnetic resonance (NMR) tomography	33
NMR effect. The Larmor equation. Ensemble of protons. Motion of magnetic moments of isolated protons in fixed and variable magnetic fields. The Bloch equations. Echo-signal, $\pi/2$ - and π -pulses. Gradient fields. Reconstruction of NMR images. Examples of image reconstruction. Influence of field nonhomogeneity on resolving power of tomograms. Mathematical taking account of technical nonhomogeneities of fields. Magnetic field synthesis on a coil axis of NMR-tomograph. Application fields of NMR-tomography. Control tasks and questions.	
Chapter 2. Some inverse problems of optics and spectroscopy	63
2.1. Restoration of blurred images	63
Setting of a problem. Deduction of an integral equation. Taking account of sensitivity characteristic of a film. Methods for solving of equations. On programs. Control tasks and questions.	

2.2. Restoration of defocused images.....	72
Setting of a problem. Deduction of basic relation. Standard shape of an equation. Solution by two-dimensional Fourier transform method. Application of Tikhonov regularization method. On programs. Control tasks and questions.	
2.3. Inverse problems of spectroscopy	79
Spectral analysis and apparatuses. Application fields of spectral analysis. Types of spectrums. Types of spectral analysis. Experimental spectrum. Problem of reduction to ideal spectral device. Continuous spectrum. Discrete spectrum. Solving system of linear-nonlinear equations. Brief statement of integral approximation algorithm. Model example. Other inverse problems of spectroscopy. On programs. Control tasks and questions.	
2.4. Inverse problem of plasma diagnostics	91
Plasma concept. Plasma characteristics. Radiation spectrum of plasma. Plasma diagnostics. A scheme of passive plasma diagnostics. Methods for solving of an equation. Case of cylindrical symmetry. Case of globular symmetry. The solving of equations. Control tasks and questions.	
Chapter 3. Generalized statement of inverse problems	100
3.1. Signal processing.....	100
Setting of a problem. Types of signal processing. Enumeration of methods of secondary signal processing. Classical methods of bearing. Compensating methods of local signals-noises. Adaptive methods. Other methods of signal processing. Control tasks and questions.	
3.2. Reduction of measurements to ideal measuring device.....	110
Introduction into problem. The Rayleigh reduction problem. Example 1 (reduction of local signals). Example 2 (reduction of extended signals). On spread function. Technical realization of reduction algorithms. Control tasks and questions.	
Chapter 4. Some inverse problems of the control theory, biophysics and mechanics	119
4.1. Inverse problems of the control theory	119
Signal restoration in dynamic system. Signal restoration in dynamic system without feed-back. Signal restoration in system not being dynamic. Control tasks and questions.	
4.2. Inverse problems of biophysics.....	123
Restoration of distorted images of biological microobjects. Tomography of biological microobjects. Inverse problem of speech acoustics. Disintegration of cells and radioactive elements. Problem of specific tidal ventilation in the lung. Control tasks and questions.	
4.3. Use of Fourier transform in applied problems of mechanics.....	128
Fourier Transform. Spectral problems of mechanics. Control tasks and questions.	

Section II. STABLE METHODS FOR SOLVING OF EQUATIONS	132
Chapter 5. Basic types of equations and concepts accompanying them	134
5.1. Basic types of equations	134
Integral equations. Systems of linear algebraic equations. Systems of linear-nonlinear equations. Operator equations. Control tasks and questions.	
5.2. Some information from linear algebra	142
System of linear algebraic equations. Characteristic equation and types of matrixes. Norms of vectors and matrixes. Condition number. Multiplication of matrixes and vectors. Examples. Control tasks and questions.	
5.3. Elements of probability theory	147
Basic definitions. Program RNDAN.	
Chapter 6. Generalized functions, the Euler formula and integral transforms	155
6.1. Elements of generalized functions theory	155
Definition of generalized function. The Heaviside function. The Dirac δ -function. Control tasks and questions.	
6.2. The Euler formula	158
Definition of the Euler formula. The Euler formula and δ -function. Control tasks and questions.	
6.3. Integral transforms	160
Continuous Fourier transform (FT). Connection between direct and inverse FT. Two-dimensional FT. Discrete FT. On FFT algorithms. Use of a regularization. The Hartley transform. The Laplace transform. Control tasks and questions.	
Chapter 7. A previous history of regular methods	178
7.1. Well-posedness and ill-posedness by Hadamard	178
Definition of well-posedness and ill-posedness. Examples.	
7.2. Classical methods for solving Fredholm integral equations of the first kind	180
Quadrature method. FT method. FT method for two-dimensional equation. Control tasks and questions.	
7.3. Gauss least-squares method (LSM)	186
Overdetermined system of linear algebraic equations (SLAE). Deduction of normal SLAE. LSM conformably to integral equation. Control tasks and questions.	
7.4. Moore-Penrose pseudoinverse matrix method (PIMM)	189
Undetermined SLAE. Normal solution and pseudoinverse matrix. Example. PIMM conformably to other equations. General conclusion. Control tasks and questions.	

Chapter 8. Methods of regularization, filtration and approximation	192
8.1. The Tikhonov regularization method	192
Essence of a method. Analysis of a method. Regularized integral equation. Principles for choosing of regularization parameter α . Numerical algorithm. Programs. Numerical examples. Regularization method for equation of convolution type. Control tasks and questions.	
8.2. Kalman-Bucy optimal filtration method	205
One-step (single) Kalman filter. Comparison of one-step Kalman filter with the Tikhonov regularization method. Multistage (multiple) Kalman filter. Control tasks and questions.	
8.3. Wiener optimal linear filtration method	210
Essence of a method. Comparison of Wiener and Tikhonov methods. Control tasks and questions.	
8.4. Interpolation, extrapolation, smoothing and approximation	212
Linear interpolation and extrapolation. Square-law interpolation and extrapolation. Lagrange polynomial. Splines. Cubic interpolating splines. Linear approximation (linear smoothing). Square-law approximation (square-law smoothing). Smoothing (approximating) cubic splines. Control tasks and questions.	
Bibliography	225
Subject index	231