

INTERNATIONAL SERIES IN THE EARTH
AND PLANETARY SCIENCES

JON F. CLAERBOUT

DEPARTMENT OF GEOPHISICS
STANFORD UNIVERSITY

FUNDAMENTALS
OF GEOPHISICAL
DATA
PROCESSING

WITH APPLICATIONS
TO PETROLEUM
PROSPECTING

Mc Graw-Hill
Book Company

New York
St. Louis
San Francisco
Auckland
Düsseldorf
Johannesburg

Kuala Lumpur
London
Mexico
Montreal
New Delhi
Panama

Paris
São Paulo
Singapore
Sydney
Tokyo
Toronto

МЕЖДУНАРОДНАЯ СЕРИЯ ПО ЗЕМНЫМ
И ПЛАНЕТАРНЫМ НАУКАМ

ДЖОН Ф. КЛАЕРБОУТ

ГЕОФИЗИЧЕСКОЕ ОТДЕЛЕНИЕ
СТЕНФОРДСКОГО УНИВЕРСИТЕТА

ТЕОРЕТИЧЕСКИЕ
ОСНОВЫ
ОБРАБОТКИ
ГЕОФИЗИЧЕСКОЙ
ИНФОРМАЦИИ
С ПРИЛОЖЕНИЕМ
К РАЗВЕДКЕ НЕФТИ

Перевод с английского
д-ра техн. наук Ю. В. Тимошина

Клаербоут Дж. Ф. Теоретические основы обработки геофизической информации с приложением к разведке нефти. Пер. с англ. М., Недра, 1981. — Пер. изд.: США, 1976. 304 с.

В книге известного американского ученого описаны методы обработки сейсмических материалов с помощью ЭВМ. В ней рассмотрены теория прямых и обратных фильтров, применение метода наименьших квадратов при обработке наблюдений, теория продолжения волновых полей. Особый интерес представляет предложенный автором метод учета сейсмического сноса, позволяющий преобразовывать волновые поля быстро и качественно.

Книга предназначена для инженеров-сейсморазведчиков, занимающихся теорией и методикой обработки наблюдений. Она может быть полезна студентам старших курсов вузов и аспирантам по специальностям «Сейсморазведка» и «Прикладная математика».

Табл. 2, ил. 129, список лит. — 38 назв.

К $\frac{20804-155}{043(01)-81}$ 98—81 1904050000

Copyright © 1976 by McGraw-Hill, Inc. All rights reserved.

© Перевод с английского, издательство «Недра», 1981

Развитие вычислительной техники и численных методов анализа открыло широкие перспективы использования их для повышения эффективности решения обратных задач геофизики. Можно без преувеличения сказать, что дальнейший прогресс разведочной геофизики невозможен без широкого внедрения в практику разведочных работ самых современных методов машинной обработки геофизических данных.

С момента выхода первых на русском языке книг по обработке сейсмических записей на ЭВМ¹ прошло около восьми лет. Это было время бурного развития методов обработки геофизических и особенно сейсмических записей на ЭВМ. Новые веяния в теории и практике обработки геофизических данных нашли отражение в ряде отечественных и переводных работ, вышедших в последние годы. Среди них следует отметить монографии П. Рабинера и Б. Гоулда «Теория и применение цифровой обработки сигналов» (Мир, 1978), И. К. Кондратьева «Линейные обрабатывающие системы в сейсморазведке» (Недра, 1976), Ю. В. Тимошина «Импульсная сейсмическая голография» (Недра, 1978) и др. Предлагаемая вниманию читателя книга известного американского ученого Джона Ф. Клаербуота представляет собой достаточно систематизированное изложение современных методов обработки данных геофизических наблюдений на ЭВМ.

В книге последовательно излагаются идеи и методы дискретного анализа и матричного исчисления, а также способы обработки геофизических данных: от различного рода одномерных и двухмерных фильтраций до решения обратной задачи с помощью дифференциальных уравнений в частных производных и их конечно-разностных аналогов. Ограниченный объем книги и большая насыщенность ее разнообразными фактами, естественно, привели к тому, что многие утверждения в ней даны без доказательств. Однако это, по-видимому, единственный способ создать достаточно полное руководство по методам обработки геофизической информации в целом. Кроме того, книга очень полезна для читателя, который может самостоятельно анализировать предлагаемые методы, идеи и доказательства. Этой же цели служат многочисленные упражнения, приводимые в конце почти каждого раздела. Среди них есть и очень простые, и весьма трудные. Для последних обычно даются указания, облегчающие решения. Несмотря на сравнительно небольшой объем и конспективный во многих местах характер изложения, в книге наряду с достаточно

¹ 1. *Цифровая обработка сейсмических данных*. М., Недра, 1973.

2. Ю. В. Тимошин. *Основы дифракционного преобразования сейсмических записей*. М., Недра, 1972.

полным освещением современного состояния теории и практики обработки геофизических данных на ЭВМ большое внимание уделено рассмотрению математических методов: теории дискретного анализа и статистики, теории матриц и метода наименьших квадратов и т. п. Все это подготавливает читателя к пониманию более сложных математических задач обработки геофизических данных на ЭВМ.

Книга имеет довольно ярко выраженную инженерную направленность, в ней приводится 11 машинных программ на Фортране, реализующих основные процедуры обработки данных. В ней удачно сочетаются достоинства учебника и монографии. Это позволяет надеяться, что она привлечет внимание широкого круга специалистов-геофизиков: инженеров, научных работников, преподавателей и студентов вузов.

Эта книга рассчитана на инженеров и научных работников. Наш опыт работы в Стенфордском университете показал, что односеместровый курс по теории инженерных систем легко усваивается и будет полезен каждому, кто интересуется машинным моделированием и анализом данных в физических науках. Естественно, что эта книга отражает научные интересы автора, которые направлены главным образом на сейсмологические исследования. Вместе с тем за восьмилетний период исследовательской и преподавательской деятельности меня интересовали многие задачи обработки геофизических данных. Предполагается, что книга будет скорее учебником, чем научной монографией. Приводимые во всех разделах упражнения доступны лицам, недавно окончившим институты. Большинство упражнений должно быть досконально проверено.

Я в долгу у огромного множества друзей, коллег и прежних учителей за многое, чему я научился. Я имел много плодотворных бесед со Стивом Симпсоном, Эндрю Робинсоном и Джоном Бургом по анализу временных рядов. Тед Мэдден научил меня многому из того, что написано в этой книге о слоистых средах, но важнее всего, что он увлек меня той мыслью, что пришло время заняться слоистыми средами. Джон Шервуд и Френсис Мюр познакомили меня с теорией сейсмической разведки методом отраженных волн и некоторыми неортодоксальными способами мышления о ней. Много студентов оказало мне большую помощь в обнаружении неточностей в тексте и упражнениях. Фил Шульд, Дон К. Рилей и Стив Догерти подготовили рисунки в заключительных главах. Миссис Сюзана Эрлин отпечатаала почти весь труд и приложила большие усилия для окончательного оформления рукописи. Моя жена, Диана, вдохновляла меня все это время на осуществление замысла.

Я благодарен за финансовую поддержку, полученную за последние восемь лет, главным образом Стенфордскому университету и Компании по разведке нефтеносных районов (Chevron Oil Field Research Company), а также Фонду нефтяных исследований Американского химического общества (Petroleum Research Fund of the American Chemical Society), Национальному фонду

по науке и ведомству ВВС по научным исследованиям (Air Force Office of Scientific Research). Недавняя помощь от фирм, субсидирующих стенфордскую программу исследований (Stanford Exploration Project — SEP), позволила быстро развить обработку сейсмических данных на основе волнового уравнения; этому вопросу посвящена последняя глава. Эти фирмы следующие: Amoco, Arco, Chevron, Continental, Digicon, Dutch Shell, ELF (France), Exxon, GSI, INA (Yugoslavia), Mobil, Petrofina (Belgium), Petty Pay, Preussag (Germany), Seiscom Delta, Seismograph Service, Shell, Sun, Teledine, Texaco, Total (France), Union, U. S. Geological Survey, United Geophysical, Western Geophysical.

Обработка геофизических данных — это применение ЭВМ при анализе результатов геофизических наблюдений. Основной задачей геофизики является более точное определение внутреннего строения Земли. Там, где бурение нерационально или невозможно, проводятся сейсмические, электромагнитные и гравитационные измерения с целью получения необходимых сведений на основании данных этих измерений. Систематически применяя законы физики и статистические методы, можно решить некоторые задачи интерпретации данных геофизических наблюдений с помощью ЭВМ. Когда количество наблюдений невелико, то их можно считать совпадающими с переменными параметрами в известных аналитических решениях уравнений классической физики. Однако в настоящее время получают огромные массивы данных наблюдений, которые содержат намного больше сведений о Земле, чем те, которые могут моделироваться посредством аналитических решений. Так, обычный морской корабль, ведущий исследования методом отраженных волн, может собирать около триллиона (10^{12}) бит информации в месяц.

Такие большие объемы данных требуют как статистической обработки, так и возможности получения теоретических решений для многопараметрических моделей земной коры. Использование цифровых ЭВМ для статистического анализа геофизических данных началось группой геофизического анализа (GAG) в 1953—1957 гг. при Массачусетском технологическом институте [1]. Теоретические расчеты в геофизике начали широко применяться с 1954 г., когда Нормен Хэскелл [2] опубликовал известную статью, в которой изложил возможность вычисления поля сейсмических волн для земной коры, моделируемой произвольным числом плоских параллельных слоев, каждый из которых обладает произвольными упругими параметрами. Это позволило использовать полную форму сигнала сейсмической волны при исследовании модели земной коры с произвольной слоистостью (под «слоистостью» понимается, что свойства породы являются функцией только одной координаты, обычно глубины или радиуса). Методы Хэскелла интенсивно развивались более 20 последних лет, так что теперь можно легко определить результат распространения сейсмических и электромагнитных импульсов от произвольного источника в любой слоистой модели земной коры. Вместе с тем оказалось, что слоистые среды легко заменяются однородными средами, являющимися наиболее простыми моделями в геофизике.

Сейсмограммы часто содержат записи сотен колебаний, происхождение большинства из которых невозможно объяснить. Были

развиты и детально разработаны методики расчета сейсмограмм для моделей слоистых сред со случайными изменениями параметров пластов. Удивительно, но при повторном возбуждении колебаний сейсмограммы полностью повторяются со всеми своими сложными, необъяснимыми деталями. Даже сейсмограммы от землетрясений воспроизводятся в том случае, когда область источника мала. Таким образом, ввод случайных переменных в анализ данных часто служит главным образом для усиления соответствия данных реальным слоистым моделям.

В противоположность нашим хорошо разработанным методам изучения слоистых сред наиболее интересны в настоящее время исследования отклонений строения Земли от слоистой модели. Основными вопросами при этом являются: контроль за дрейфом континентов, понимание причин землетрясений и поиск залежей нефти и минералов. Таким образом, задачи обработки геофизических данных лежат в области согласования полевых данных с двух- и трехмерными неоднородными моделями земной коры.

Обработку геофизических данных можно начать с изучения временного представления сигналов определенной формы с позиций теории фильтров и спектрального анализа. Здесь очень важны математические ограничения, накладываемые принципом причинности. В дальнейшем модели земной коры различной сложности строятся с учетом этого принципа. Устойчивость вычислений часто определяется строгостью их выполнения. Затем рассматриваются основные понятия разрешающей способности, статистики и теории матриц со множеством примеров. Это подготавливает читателя к пониманию общей теории метода наименьших квадратов, играющего, безусловно, принципиальную роль при согласовании экспериментальных данных с теоретическими моделями. Пока он играет еще важную роль, но ему противостоят методы с высоким разрешением (максимума энтропии) и более точные методы (L_1 -норма и линейное программирование).

Согласно такому развитию идей фундаментальной обработки данных последующая часть книги посвящена моделям земной коры с постепенно возрастающей сложностью. Рассматриваются прежде всего многократно-отраженные плоские волны в слоистых средах на основе принципов непрерывности, причинности и сохранения энергии (без привлечения других принципов). Волновые поля могут быть рассчитаны на основе знания структуры и параметров сред, а строение среды определено на основе волнового поля. Затем рассматривается более общая теория физико-математических методов решения задач для слоистых сред и исследуются основные характерные черты конечно-разностного моделирования дифференциальных уравнений в частных производных. Заключительные главы посвящены распространению волн и обработке данных на основе дифференциальных уравнений в частных производных. Я разработывал эти вопросы со своими помощниками в течение последних шести лет в Стенфордском университете [3—8, 36, 37]. В настоящее время этот материал в сущности

неизвестен ученым и инженерам, не работающим в области геофизики. Основная задача схожа с задачей голографии. Волновое поле наблюдается на плоскости (поверхности Земли) с целью создания двух- или трехмерной модели рассеивающих объектов, расположенных по одну сторону (снизу) плоскости. Основные вопросы и используемая техника совершенно отличаются от применяемых в оптической голографии. Скоростная неоднородность среды, дифракция, интерференция и многократное отражение — характерные черты процесса распространения сейсмических волн — практически не встречаются в повседневном опыте. Например, глаз легко обманывается в зеркальной комнате или при взгляде в аквариум.

Материал по обработке данных на основе волнового уравнения быстро устаревает и его необходимо переписывать. Программа стэнфордских исследований в области разведки ныне поддерживается десятью нефтяными компаниями США, восемью разведочными организациями, четырьмя нефтяными компаниями в Европе и их филиалами. Многие из этих групп сами проводят исследования. Поэтому в последних разделах я придавал особое значение основным вопросам, пренебрегая деталями.

Надеюсь, что когда-нибудь мы будем в состоянии использовать волны, возникающие при землетрясениях, для получения сейсмических изображений. Тогда мы сможем увидеть первые четкие картины земной коры на глубинах 10 км и больше.

1. ПРЕОБРАЗОВАНИЯ

Первый шаг анализа данных — изучение способов представления волновых сигналов и обработка их цифровой вычислительной машиной. Время и пространство обычно рассматриваются как непрерывные величины, но для анализа явлений и процессов с помощью ЭВМ необходимо дискретизировать их. Эта операция называется также преобразованием их в цифровую форму, или квантованием. Дискретизация непрерывной функции вначале считалась вынужденной мерой, ухудшающей результаты и необходимой лишь постольку, поскольку данные наблюдений не всегда являются известными аналитическими функциями. Однако после накопления некоторого опыта работы с дискретными функциями стало ясно, что многие математические понятия легче воспринимаются при описании их в дискретном времени, чем в непрерывном. Например, в этой главе рассматривается Z -преобразование и показывается его эквивалентность преобразованию Фурье. Z -преобразование легко понять на основе элементарной алгебры, тогда как преобразование Фурье требует значительного опыта в вычислениях.

1.1. ДИСКРЕТНЫЕ ДАННЫЕ И Z -ПРЕОБРАЗОВАНИЕ

Рассмотрим функцию времени, изображенную на рис. 1.1. Для анализа эту наблюдаемую функцию с помощью ЭВМ необходимо аппроксимировать каким-либо образом таблицей чисел. Обычный путь для этого — выразить в цифрах или измерить значения функции $b(t)$ в равноудаленных друг от друга точках на оси времени. Для нашего примера цифровое представление непрерывной функции можно записать в виде вектора

$$b_t = (\dots, 0, 0, 1, 2, 0, -1, -1, 0, 0, \dots).$$

Естественно, что при взятии отсчетов в более близких друг к другу точках, мы получили бы более точную аппроксимацию функции. Помимо вектора функцию времени можно представить также в виде полинома, коэффициентами которого являются значения $b(t)$ в последовательных точках по оси времени. Для рассматриваемого примера имеем

$$B(Z) = 1 + 2Z + 0Z^2 - Z^3 - Z^4. \quad (1.1.1)$$

Этот полином реализует Z -преобразование функции $b(t)$. Какой смысл имеет величина Z в этом полиноме? Величина Z не принимает численных значений; смысл величины Z состоит в том, что она является оператором единичной задержки. Для примера рас-

смотрим полином

$$ZB(Z) = Z + 2Z^2 - Z^4 - Z^5,$$

коэффициенты которого изображены графически на рис. 1.2. Это тот же сигнал, что и на рис. 1.1, но задержанный на единицу времени. Как видно, временная функция b_t задерживается на n единиц времени, если полином $B(Z)$ умножается на величину

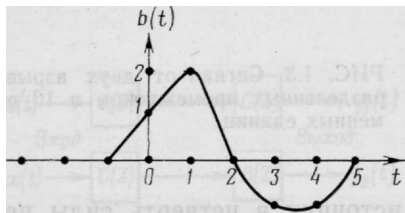


РИС. 1.1. Дискретное представление непрерывной функции времени через одинаковые временные интервалы

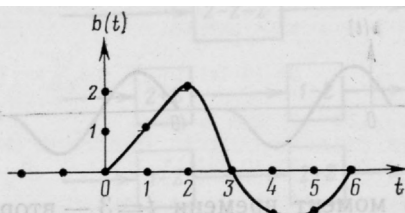


РИС. 1.2. Коэффициенты полинома $ZB(Z)$ сдвинуты на один отсчет коэффициентами полинома $B(Z)$

Z^n . Оператор задержки Z имеет большое значение в анализе волновых процессов уже потому, что волны затрачивают некоторое время на прохождение от одной точки пространства к другой.

Оператор задержки позволяет также строить более сложные функции времени из более простых. Предположим, что $b(t)$ выражает функцию акустического давления или сейсмограмму, наблюдаемую после отдаленного взрыва. Тогда $b(t)$ называется *импульсной реакцией*. Если через $t=10$ временных единиц после первого происходит второй взрыв, то следует ожидать функцию давления $y(t)$, изображенную на рис. 1.3.

С помощью Z -преобразования эту функцию можно записать в виде

$$Y(Z) = B(Z) + Z^{10}B(Z).$$

Если бы за первым источником включился бы сток половинной силы, то

$$Y(Z) = B(Z) - \frac{1}{2} Z^{10}B(Z).$$

Если сигналы источников перекрывают друг друга во времени [что было бы при степени $B(Z)$, большей 10], то сигналы накладываются друг на друга в области перекрытия. Предположение, что они суммируются без какого-либо взаимодействия, называется *допущением линейности*¹. Это допущение очень часто подтверждается практикой. В сейсмологии, несмотря на большую неод-

¹ В отечественной литературе это называется принципом линейной суперпозиции или просто принципом суперпозиции (*Прим. перев.*).

нородность земной коры и различие слагающих ее пород, сейсмические волны, проходя через Землю, не взаимодействуют друг с другом, т. е. они удовлетворяют принципу линейной суперпозиции. Нелинейности проявляются лишь при возмущениях большой амплитуды. Нелинейность не возникает при усложнении геометрии среды.

Теперь предположим, что в момент времени $t=0$ включился источник, а в момент $t=1$ — сток половинной силы, а затем

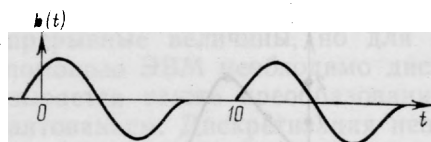


РИС. 1.3. Сигнал от двух взрывов, разделенных промежутком в 10 временных единиц

в момент времени $t=3$ — второй источник в четверть силы первого. Эта последовательность событий определяет «источник» импульсов

$$x_t = \left(1, -\frac{1}{2}, 0, \frac{1}{4}\right).$$

Z -преобразование для такого источника имеет вид

$$X(Z) = 1 - \frac{1}{2}Z + \frac{1}{4}Z^2.$$

Наблюдаемый сигнал y_t для этой последовательности источников и стоков, принятый сейсмоприемником, имеет Z -преобразование $Y(Z)$, заданное следующим образом:

$$\begin{aligned} Y(Z) &= B(Z) - \frac{Z}{2}B(Z) + \frac{Z^2}{4}B(Z) = \\ &= \left(1 - \frac{Z}{2} + \frac{Z^2}{4}\right)B(Z) = X(Z)B(Z). \end{aligned} \quad (1.1.2)$$

Последнее уравнение иллюстрирует принцип, лежащий в основе теории линейных систем, который гласит, что выход $Y(Z)$ может быть выражен как вход $X(Z)$, умноженный на импульсную реакцию $B(Z)$ среды.

Существует много примеров линейных систем. К линейным системам относится, в частности, широкий класс электронных цепей. Сложные линейные системы составляются подключением выхода одной системы на вход другой. Предположим, что имеются две линейные системы, характеризующиеся полиномами $B(Z)$ и $C(Z)$ соответственно. Возникает вопрос, являются ли две комбинированные системы, приведенные на рис. 1.4, эквивалентными. Применяя Z -преобразование, легко показать, что эти две системы эквивалентны, так как произведение полиномов коммутативно, т. е.

$$Y_1(Z) = [X(Z)B(Z)]C(Z) = XBC; \quad (1.1.3)$$

$$Y_2(Z) = [X(Z)C(Z)]B(Z) = XCB = XBC. \quad (1.1.4)$$

Рассмотрим систему с импульсной реакцией $B(Z) = 2 - Z - Z^2$. Этот полином можно разложить на множители $2 - Z - Z^2 = (2 + Z)(1 - Z)$, ввиду чего возможны три эквивалентные системы, показанные на рис. 1.5. Так как любой полином можно разложить на множители, то любую импульсную реакцию можно

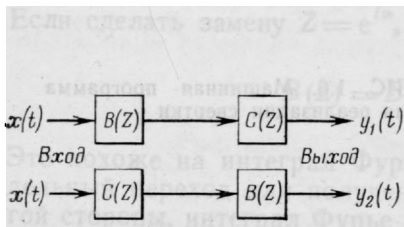


РИС. 1.4. Две эквивалентные фильтрующие системы

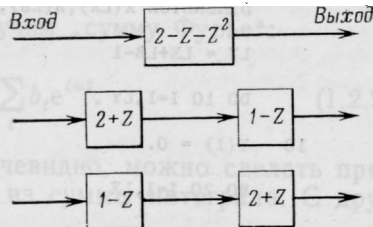


РИС. 1.5. Три эквивалентные фильтрующие системы

смоделировать последовательностью двухчленных фильтров, линейных относительно Z .

Что должна делать ЭВМ, чтобы перемножить два Z -преобразования? Фильтр $2 + Z$ мог бы быть задан в машине путем запоминания коэффициентов $(2, 1)$. Подобным образом для фильтра $1 - Z$ запоминаются числа $(1, -1)$. Программа умножения полиномов должна преобразовать эти исходные данные в последовательность $(2, -1, -1)$. Посмотрим, как происходит процесс вычисления в общем случае, например, для выражения

$$X(Z)B(Z) = Y(Z); \tag{1.1.5}$$

$$\begin{aligned} (x_0 + x_1Z + x_2Z^2 + \dots)(b_0 + b_1Z + b_2Z^2 + \dots) = \\ = (y_0 + y_1Z + y_2Z^2 + \dots). \end{aligned} \tag{1.1.6}$$

Приравнивая коэффициенты при последовательных степенях Z получим:

$$\left. \begin{aligned} y_0 &= x_0b_0; \\ y_1 &= x_1b_0 + x_0b_1; \\ y_2 &= x_2b_0 + x_1b_1 + x_0b_2; \\ y_3 &= x_3b_0 + x_2b_1 + x_1b_2; \\ y_4 &= x_4b_0 + x_3b_1 + x_2b_2; \\ &\vdots \end{aligned} \right\} \tag{1.1.7}$$

$$y_k = \sum_{i=0}^k x_{k-i}b_i. \tag{1.1.8}$$

Выражение (1.1.8) называется *уравнением свертки*. Таким образом, произведение двух полиномов выражается также полиномом, коэффициенты которого находятся с помощью свертки. Простая машинная программа на Фортране, которая делает свертку, включая краевые эффекты на обоих концах, приведена на рис. 1.6.

```

DIMENSION X(LX),B(LB),Y(LY)

LY = LX+LB-1

DO 10 I=1,LY
10  Y(I) = 0.

DO 20 I=1,LX

```

РИС. 1.6. Машинная программа для реализации свертки

Следует отметить, что $X(Z)$ и $Y(Z)$ не обязательно должны быть строго полиномами; они могут содержать как положительные, так и отрицательные степени Z , т. е.

$$X(Z) = \dots \frac{x_{-2}}{Z^2} + \frac{x_{-1}}{Z} + x_0 + x_1 Z + \dots; \quad (1.1.9)$$

$$Y(Z) = \dots \frac{y_{-2}}{Z^2} + \frac{y_{-1}}{Z} + y_0 + y_1 Z + \dots$$

Использование отрицательных степеней Z в $X(Z)$ и $Y(Z)$ позволяет определять сигналы до момента $t=0$ (для отрицательных времен). Эффект использования отрицательных степеней Z в фильтре совершенно отличен. Выражение (1.1.8) показывает, что выходной сигнал y_k , который появляется в момент времени k , — это линейная комбинация текущих и предыдущих входных сигналов, т. е. $(x_i, i \leq k)$. Если бы фильтр $B(Z)$ включал член типа b_{-1}/Z , то выход y_k в момент времени k был бы линейной комбинацией текущих и предыдущих входов, а также x_{k+1} входа, который в действительности еще не получен в момент k . Такой фильтр называется физически *нереализуемым* фильтром, так как в реальных условиях в данный момент он не может реагировать на возбуждение, которое еще не возникло. Однако нереализуемые фильтры изредка полезны в машинном моделировании, где все данные предварительно записаны.

Упражнения

1. Пусть задан полином $B(Z) = 1 + Z + Z^2 + Z^3 + Z^4$. Изобразите графически коэффициенты $B(Z)$ в функции степеней Z , а также коэффициенты полинома $B^2(Z)$.

2. Если $x_t = \cos \omega_0 t$, где t определено в общем виде величиной $b_t = (b_0; b_1)$ и $Y(Z) = X(Z)B(Z)$, каковы A и B в формуле $y_t = A \cos \omega_0 t + B \sin \omega_0 t$?

3. Доказать, что если $x_t = \cos \omega_0 t$ и $b_t = (b_0, b_1, \dots, b_n)$, то y_t всегда имеет формулу $A \cos \omega_0 t + B \sin \omega_0 t$.

1.2. СВЯЗЬ Z-ПРЕОБРАЗОВАНИЯ С ФУРЬЕ-ПРЕОБРАЗОВАНИЕМ

Мы определили Z-преобразование как

$$B(Z) = \sum_t b_t Z^t. \quad (1.2.1)$$

Если сделать замену $Z = e^{i\omega}$, то получим „сумму Фурье“:

$$B(Z) = B(e^{i\omega}) = \sum_t b_t e^{i\omega t}. \quad (1.2.2)$$

Это похоже на интеграл Фурье, и, очевидно, можно сделать предельный переход для получения его из суммы интеграла. С другой стороны, интеграл Фурье

$$B(\omega) = \int_{-\infty}^{+\infty} b(t) e^{i\omega t} dt \quad (1.2.3)$$

приводится к сумме (1.2.2), если $b(t)$ не является непрерывной функцией времени, а определяется выражением

$$b(t) = \sum_k b_k \delta(t - k), \quad (1.2.4)$$

где δ — дельта-функция Дирака.

В предыдущем разделе мы видели, что для перемножения двух полиномов необходимо свернуть их коэффициенты. Такой же процесс на языке Фурье-преобразования означает, что произведение в частотной области соответствует свертке во временной области. Хотя о преобразовании Фурье обычно думают как об интеграле, который трудно или невозможно вычислить, Z-преобразование всегда проще, практически оно тривиально. Чтобы получить Z-преобразование, достаточно присвоить последовательным точкам данные степени Z. Имея полином $B(Z)$, можно сослаться на него как на временную или частотную функцию, в зависимости от того, построены ли коэффициенты полинома либо вычислены и построены значения $B(Z = e^{i\omega})$ для различных частот ω . Следует отметить, что, поскольку ω меняется от нуля до 2π , вектор $Z = e^{i\omega} = \cos \omega + i \sin \omega$ делает один оборот по единичной окружности в направлении против часовой стрелки.

Если осуществление Z-преобразования равнозначно присвоению степеней Z последовательным точкам временной функции, то обратное Z-преобразование должно быть просто идентификацией коэффициентов при различных степенях Z с различными точками во времени. Как может эта простая идентификация коэффициентов давать такой же результат, как, очевидно, более сложная

операция обратного преобразования Фурье? Обратное преобразование Фурье имеет вид

$$b(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} B(\omega) e^{-i\omega t} d\omega. \quad (1.2.5)$$

Прежде всего отметим, что интегрирование Z^n вдоль единичной окружности или $e^{in\omega}$ в интервале $-\pi \leq \omega < +\pi$ дает нуль, если $n \neq 0$, потому что косинусоида и синусоида являются колебательными функциями, т. е.

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{in\omega} d\omega &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos n\omega + i \sin n\omega) d\omega = \\ &= \begin{cases} 1 & n=0; \\ 0 & n - \text{целое число, неравное нулю.} \end{cases} \end{aligned} \quad (1.2.6)$$

В терминах нашей дискретной временной функции обратное преобразование Фурье (1.2.5) имеет вид

$$b_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\dots + b_1 e^{-i\omega} + b_0 + b_1 e^{+i\omega} + \dots) e^{-i\omega t} d\omega. \quad (1.2.7)$$

Как видно, из всех членов интеграла (1.2.7) при учете формулы (1.2.6) остается только один член с коэффициентом b_t , а все остальные осциллируют и уничтожаются. Другими словами, имеется только один коэффициент при члене с нулевой степенью, который дает вклад в интеграл, приводя выражение (1.2.7) к виду

$$b_t = \frac{1}{2\pi} b_t \int_{-\pi}^{+\pi} d\omega = b_t. \quad (1.2.8)$$

Это показывает, что обратное преобразование Фурье совершенно идентично выражается коэффициентами при степенях Z .

В этой книге, как и во многих других, принято, что время, проходящее между отсчетами данных $\Delta t = 1$, является единицей. Для применения данного выражения к другим значениям Δt необходимо заменить ω на $\omega \Delta t$, т. е.

$$\omega_h = \omega_r \Delta t_h = \omega_g \Delta t_g, \quad (1.2.9)$$

С помощью Z -преобразования мы получаем спектр, заданный в диапазоне от 2π до ω_r . В пределе величина Δt_g стремится к нулю, а ω_g имеет, как и интеграл Фурье, те же бесконечные пределы.

При аппроксимации непрерывной дискретной функции необходимо выбрать интервал выборки (дискретизации) Δt_g доста-

точно малым. Известный результат заключается в том, что если спектральные составляющие сигнала на частотах, превышающих некоторую частоту ω_{\max} , не несут заметной энергии, то не будет ощутимой потеря информации, если интервал выборки (шаг квантования по времени) равен $\Delta t = \pi/\omega_{\max}$. Другими словами, косиноида должна быть квантована по крайней мере в двух точках на период (длину волны). Рис. 1.7,а показывает, что недостаточно частое квантование синусоиды часто бывает причиной того, что она кажется синусоидой с более низкой частотой. Поэтому мак-

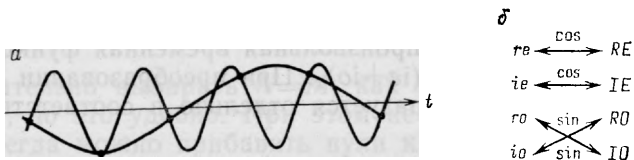


РИС. 1.7. Пример недостаточно частого квантования синусоиды по времени (а) и мнемоническая таблица, иллюстрирующая влияние Фурье-преобразования на свойства четности (нечетности) и действительности (мнимости) функций (б)

симальная частота квантования $\omega_{\max} = \pi/\Delta t$, позволяющая восстановить сигнал истинной частоты, называется частотой свертки¹. Воспроизведение более высоких частот осуществляется с понижением частоты. На практике квазисинусоидальные сигналы всегда квантуются более чем в двух точках на период. Соответствующее теоретическое обоснование для квантования в восьми и более точках на период приводится в разделе 2.8.

В дальнейшем желательно рассмотреть нечетно-четные симметрии с целью оценки их влияния на преобразование Фурье. Четную часть e_t функции времени b_t определим выражением

$$e_t = \frac{b_t + b_{-t}}{2}, \quad (1.2.10)$$

а нечетную часть

$$o_t = \frac{b_t - b_{-t}}{2}. \quad (1.2.11)$$

Функция является суммой этих четной и нечетной частей. Складывая выражения (1.2.10) и (1.2.11), получим

$$b_t = e_t + o_t. \quad (1.2.12)$$

Рассмотрим простую, действительную, четную функцию времени вида $(b_{-1}, b_0, b_1) = (1, 0, 1)$. Ее преобразование $Z + 1/Z = 2 \cos \omega$ является четной функцией ω , так как $\cos \omega = \cos(-\omega)$. Рассмотрим действительную, нечетную функцию времени $(b_{-1}, b_0, b_1) = (-1, 0, 1)$. Ее преобразование $Z - 1/Z = 2(\sin \omega)/i$ является мнимым и нечетным, так как $\sin \omega = -\sin(-\omega)$. Аналогично пре-

¹ Частотой Найквиста (Прим. перев.).

Преобразование мнимой четной функции $(i; 0; i)$ является мнимой четной функцией $i \cos \omega$, а преобразование мнимой нечетной функции $(-i, 0, i)$ является действительной и нечетной функцией. Пусть g и i относятся к действительным и мнимым величинам (рис. 1.7,б), e и o отвечают четным и нечетным величинам, а нижняя и верхняя строки соответствуют временной и частотной функциям.

Более сложные временные функции можно получить сложением двух рассмотренных нами ранее дискретных функций. Так как суммы четных функций являются четными и т. д., то таблица, представленная на рис. 1.7,б, применима ко всем временным функциям. Отметим, что произвольная временная функция принимает вид $b_t = (ge + go)_t + i(ie + io)_t$. При преобразовании b_t каждая из четырех частей преобразуется отдельно в соответствии с таблицей.

Упражнение

Обычно функция определяется только во временной либо только в частотной области. Если одно известно, то другое определяется с помощью преобразования. Пусть теперь половина информации определена во временной области при условии, что $b_t = 0$ при $t < 0$, а половина — в частотной области путем задания в ней действительной части $RE + RO$. Как можно определить остальную часть функции?

1.3. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Записав выражение

$$B(Z) = b_0 + b_1 Z + \dots, \quad (1.3.1)$$

получим как временную функцию, так и ее преобразование Фурье. Если графически изобразить коэффициенты (b_0, b_1, \dots) , то получим временную функцию. Если вычислить и построить график выражения (1.3.1) на множестве действительных ω , то получим результат преобразования. Отметим, что для действительных частот ω величина Z имеет единичную амплитуду, т. е. находится на единичной окружности. Так как ω непрерывная переменная, а в ЭВМ она всегда конечна, то возникает вопрос: как выбрать конечное число значений частоты ω_k для построения графика? Обычно выбирают равномерно распределенные частоты. Самая нижняя частота может быть равна нулю. [Отметим, что $Z(\omega=0) = e^{i0} = 1$]. Высокая частота $\omega = 2\pi [Z(\omega=2\pi) = e^{i2\pi} = 1]$ может не рассматриваться, так как выражение (1.3.1) дает для нее то же значение, что и для нулевой частоты. Выбирая равномерно распределенные в этом интервале частоты, имеем:

$$\omega_k = \frac{(0, 1, 2, \dots, M-1) 2\pi}{M}, \quad (1.3.2)$$

где M — некоторое целое число. Обозначим $B(Z(\omega_k))$ через B_k .

Для частного случая N -точечной функции времени, когда $N=4$, выражение (1.3.1) можно представить в виде произведения матриц:

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W & W^2 & W^3 \\ 1 & W^2 & W^4 & W^6 \\ 1 & W^3 & W^6 & W^9 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad (1.3.3)$$

где

$$W = e^{2\pi i/N}. \quad (1.3.4)$$

Необязательно выбирать $N=M$, как это делалось в выражении (1.3.3), но это удобно. При этом не будет потери общности, так как всегда можно прибавить нули к функции времени перед подстановкой ее в формулу (1.3.3). Удобство выбора $N=M$ заключается в том, что матрица в выражении (1.3.3) будет квадратной и точно обратимой. В самом деле, легко показать, что выражение, обратное формуле (1.3.3), будет иметь вид

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1/W & 1/W^2 & 1/W^3 \\ 1 & 1/W^2 & 1/W^4 & 1/W^6 \\ 1 & 1/W^3 & 1/W^6 & 1/W^9 \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix}. \quad (1.3.5)$$

Так как величины $1/W$ и W являются комплексно-сопряженными, то и матрицы в выражениях (1.3.3) и (1.3.5) — комплексно-сопряженные. Практически нет существенных математических различий между временными и частотными функциями. Эта «двойственность» была бы более выраженной, если бы использовался масштабный коэффициент $N^{-1/2}$, в каждом из выражений (1.3.3) и (1.3.5), а не 1 в формуле (1.3.3) и N^{-1} в формуле (1.3.5). Отметим, что временные и частотные функции могут заменять друг друга в мнемонической таблице, описывающей симметрии. Действительно, приведенное ранее утверждение о том, что произведение двух функций частоты равно свертке соответствующих им двух функций времени, имеет двойственную формулировку, которая гласит, что произведение двух функций времени равно свертке соответствующих им двух функций частоты. Мы не приведем здесь доказательства этой двойственности, поскольку она составляет предмет книг по математике и теории систем. Однако предлагаем читателю по возможности убедиться в том, что в любой теореме понятия «время» и «частота» являются взаимозаменяемыми.

При построении графика преобразования B_k для ($k=0, 1, \dots, \dots, M-1$) значения ω_k на оси частот изменяются в диапазоне $0 \leq \omega_k < 2\pi$. Часто более естественным оказывается представление его в интервале $-\pi \leq \omega < \pi$. Так как преобразование периодически

с периодом 2π , то значения B_k в интервале $\pi \leq \omega < 2\pi$ можно просто переместить в интервал $-\pi \leq \omega < 0$.

Таким образом, при $N=8$ можно нанести последовательно

$$B_4 B_5 B_6 B_7 B_0 B_1 B_2 B_3$$

соответственно для значений ω , равных

$$-\pi, -\frac{3\pi}{4}, -\frac{\pi}{2}, -\frac{\pi}{4}, 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}.$$

Преимущество рассматриваемого интервала состоит в том, что для непрерывных во времени сигналов, квантованных по времени достаточно часто, значения B_k получаются маленькими на обоих концах. Если сигнал является действительным, то действительная часть B_k имеет четную симметрию относительно B_0 , а мнимая часть — нечетную. При этом достаточно вывести лишь половину значений. Выбор нечетного значения N позволил бы поместить $\omega=0$ строго в середину интервала. Однако читатель вскоре увидит, почему мы используем четное число отсчетов данных.

Операция умножения матрицы на временной вектор в выражении (1.3.3) требует N^2 перемножений и сложений. Далее в этом разделе описывается остроумный метод, названный быстрым преобразованием Фурье, позволяющий выполнять перемножение матриц за $N \log_2 N$ операций умножения и суммирования. Так как, например, $\log_2 1024$ равен 10, то это дает огромную экономию машинного времени.

Основной стандартный блок в быстром преобразовании Фурье называется удвоением. Задав ряд $(x_0, x_1, \dots, x_{N-1})$ и его дискретное преобразование Фурье $(X_0, X_1, \dots, X_{N-1})$, а также другой ряд $(y_0, y_1, \dots, y_{N-1})$ и его преобразование Фурье $(Y_0, Y_1, \dots, Y_{N-1})$, найдем преобразование ряда удвоенной длины, полученного чередованием членов этих двух рядов:

$$z_t = (x_0, y_0, x_1, y_1, \dots, x_{N-1}, y_{N-1}).$$

Процесс удвоения используется много раз во время процесса вычисления быстрого преобразования Фурье. Как предполагается из слова *удвоение*, будет удобно считать, что N является целым числом, которое образуется возведением 2 в некоторую целую степень. Пусть $N=8=2^3$. Начнем с разбиения нашего 8-точечного ряда x_0, x_1, \dots, x_7 на восемь различных одноточечных рядов. Применение преобразования Фурье к каждому одноточечному ряду дает точку. На следующем шаге удвоение применяется 4 раза для получения преобразований четырех различных рядов из двух точек (x_0, x_4) , (x_1, x_5) , (x_2, x_6) и (x_3, x_7) . Затем применим удвоение еще дважды для получения преобразований двух различных рядов из четырех точек (x_0, x_2, x_4, x_6) и (x_1, x_3, x_5, x_7) . Наконец, применив удвоение еще раз, получим преобразование исходного ряда из восьми точек $(x_0, x_1, x_2, \dots, x_7)$.

Остается рассмотреть детали процесса удвоения. Пусть

$$V = e^{i2\pi/2N} = W^{1/2}; \quad (1.3.6)$$

$$V^N = e^{i\pi} = -1. \quad (1.3.7)$$

Преобразованиями двух N -точечных рядов по определению являются

$$X_k = \sum_{j=0}^{N-1} x_j V^{2jk} \quad (k=0, 1, \dots, N-1);$$

$$Y_k = \sum_{j=0}^{N-1} y_j V^{2jk} \quad (k=0, 1, \dots, N-1).$$

Преобразованием ряда с чередующимися членами $z_k = (x_0, y_0, x_1, y_1, \dots, x_{N-1}, y_{N-1})$ по определению, получим

$$Z_k = \sum_{l=0}^{2N-1} z_l V^{lk} \quad (k=0, 1, \dots, 2N-1).$$

Для получения Z_k из X_k и Y_k потребуются две формулы: одна для $k=0, 1, \dots, N-1$ и другая для $k=N, N+1, \dots, 2N-1$.

Во-первых,

$$Z_k = \sum_{l=0}^{2N-1} z_l V^{lk} \quad (k=0, 1, \dots, N-1).$$

Разобьем сумму на две части, указывая, что x_j умножается на четные степени V , а y_j — на нечетные:

$$Z_k = \sum_{j=0}^{N-1} x_j V^{2jk} + V^k \sum_{j=0}^{N-1} y_j V^{2jk} = X_k + V^k Y_k. \quad (1.3.8)$$

Оставшуюся половину членов Z_k получим следующим образом:

$$Z_k = \sum_{l=0}^{2N-1} z_l V^{lk} = \sum_{l=0}^{2N-1} z_l V^{l(m+N)}$$

$$(k=N, N+1, \dots, 2N-1; k-N=m=0, 1, \dots, N-1);$$

$$Z_k = \sum_{l=0}^{2N-1} z_l V^{lm} (V^N)^l = \sum_{l=0}^{2N-1} z_l V^{lm} (-1)^l = \sum_{j=0}^{N-1} x_j V^{2jm} -$$

$$- V^m \sum_{j=0}^{N-1} y_j V^{2jm} = X_m - V^m Y_m;$$

$$Z_k = X_{k-N} - V^{k-N} Y_{k-N} \quad (k=N, N+1, \dots, 2N-1). \quad (1.3.9)$$

Первое вычисление на машине по этому алгоритму произведено Верном Гербертом, который широко применял его при интерпретации данных сейсмического метода отраженных волн. Он запрограммировал его на ЭВМ IBM 1401 в 1962 г. В. Герберт никогда не публиковал своего метода. Он был заново найден и опубликован Кули и Тьюки в 1965 г. Таким образом, он стал известен как

```

SUBROUTINE FORK(LX,CX,SIGNI)
C FAST FOURIER
C
C LX
C CX(K) = SQRT(1/LX) SUM (CX(J)*EXP(2*PI*SIGNI*I*(J-1)*(K-1)/LX))
C J=1 FOR K=1,2,...,(LX=2**INTEGER)
COMPLEX CX(LX),CARG,CEXP,CW,CTEMP
J=1
SC=SQRT(1./LX)
DO 30 I=1,LX
IF(I.GT.J) GO TO 10
CTEMP=CX(J)*SC
CX(J)=CX(I)*SC
CX(I)=CTEMP
10 M=LX/2
20 IF(J.LE.M) GO TO 30
J=J-M
M=M/2
IF(M.GE.1) GO TO 20
30 J=J+M
L=1
40 ISTEP=2*L
DO 50 M=1,L
CARG=(0.,1.)*(3.14159265*SIGNI*(M-1))/L
CW=CEXP(CARG)
DO 50 I=M,LX,ISTEP
CTEMP=CW*CX(I+L)
CX(I+L)=CX(I)-CTEMP
50 CX(I)=CX(I)+CTEMP
L=ISTEP
IF(L.LT.LX) GO TO 40
RETURN
END

```

РИС. 1.8. Программа быстрого преобразования Фурье

алгоритм Кули — Тьюки [9]. Программа, выполняющая быстрое преобразование Фурье (в модификации Бреннера), приведена на рис. 1.8. Повторный вызов этой программы приводит к восстановлению исходных данных. При первом обращении оператор SIGNI должен быть равен +1, а при втором обращении —1. Оператор LX должен быть степенью двойки.

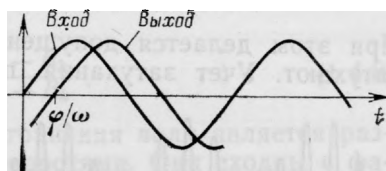
Упражнение

Проверьте, является ли в случае произвольной размерности $N \times N$ матрица в выражении (1.3.5) в самом деле обратной по отношению к матрице в выражении (1.3.3).

1.4. ФАЗОВАЯ И ГРУППОВАЯ ЗАДЕРЖКИ

Многие фильтры сильно изменяют сигналы, проходящие через них. Другие же фильтры незначительно изменяют сигналы или вообще их не изменяют. Ко второй из упомянутых категорий принадлежат преобразователи (приемники) и регистраторы. В таких случаях

РИС. 1.9. Синусоида $\sin \omega t$ на входе фильтра и задержанная синусоида $\sin(\omega t - \varphi)$ на его выходе



основным видом изменения сигнала является задержка его во времени. Один из способов определения задержки фильтра заключается в подаче на вход фильтра синусоидального сигнала и сравнения его фазы с фазой выходного сигнала (рис. 1.9). Если сигнал на входе $\sin \omega t$, а на выходе $\sin(\omega t - \varphi)$, то так называемая фазовая задержка t_p находится следующим образом:

$$\left. \begin{aligned} \sin(\omega t - \varphi) &= \sin \omega(t - t_p); \\ \omega t - \varphi &= \omega t - \omega t_p; \\ t_p &= \frac{\varphi}{\omega}. \end{aligned} \right\} \quad (1.4.1)$$

Интереснее так называемая групповая задержка. Она является аналогом групповой скорости в теории распространения волн. В самом деле, при моделировании волновых процессов на ЭВМ распространение волн, скажем, от точки A до точки B , может быть имитировано с помощью фильтра. Когда форма сигнала, наблюдаемого в точке A , отличается от формы сигнала в точке B , но огибающая энергии в точке A похожа на задержанную огибающую энергии в точке B , тогда возникает ситуация, для которой понятие групповой скорости, характеризующей скорость распространения огибающей энергии, может быть очень полезным.

При суммировании двух косинусоидальных сигналов с близкими частотами возникают биения (пульсации) (рис. 1.10). Когда сигнал такой формы проходит через фильтр, то каждая из частот может приобрести различную задержку, в результате чего огибающая или колебания будут иметь задержку, которая отличается от фазовых задержек обеих частот. Задержка огибающей или групповая задержка не может равняться среднему значению фазовых задержек этих двух частот. Это можно понять из следующего. Предположим, что форма входного импульса x_t имеет вид

$$x_t = \cos \omega_1 t + \cos \omega_2 t. \quad (1.4.2)$$

Используя тригонометрическое тождество, запишем

$$x_t = 2 \cos \frac{\omega_1 + \omega_2}{2} t \cos \frac{\omega_1 - \omega_2}{2} t. \quad (1.4.3)$$

Видим, что сумма двух косинусов представляет собой произведение косинуса средней частоты на косинус половины разности этих частот. Так как частоты взяты близкими друг к другу, то множитель разностной частоты медленно изменяет амплитуду средней частоты. Теперь потребуем, чтобы выходной сигнал фильтра

$$y_t = \cos(\omega_1 t - \varphi_1) + \cos(\omega_2 t - \varphi_2). \quad (1.4.4)$$

При этом делается допущение, что колебания обеих частот не затухают. Учет затухания по дифференциальному закону для

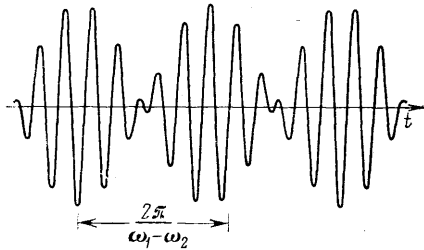


РИС. 1.10. График амплитудно-модулированных колебаний вида $\cos \omega_1 t + \cos \omega_2 t$

обеих частотных компонент значительно усложнил бы анализ. Используя тригонометрическое тождество в формуле (1.4.4), получим

$$y_t = 2 \cos\left(\frac{\omega_1 + \omega_2}{2} t - \frac{\varphi_1 + \varphi_2}{2}\right) \cos\left(\frac{\omega_1 - \omega_2}{2} t - \frac{\varphi_1 - \varphi_2}{2}\right). \quad (1.4.5)$$

Переписав множитель, характеризующий пульсацию амплитуды колебаний, в обозначениях временной задержки t_g , имеем

$$\cos\left(\frac{\omega_1 - \omega_2}{2} t - \frac{\varphi_1 - \varphi_2}{2}\right) = \cos\left[\frac{\omega_1 - \omega_2}{2} (t - t_g)\right]$$

или

$$(\omega_1 - \omega_2) t_g = \varphi_1 - \varphi_2.$$

Тогда групповая задержка определится выражением

$$t_g = \frac{\varphi_1 - \varphi_2}{\omega_1 - \omega_2} = \frac{\Delta\varphi}{\Delta\omega}. \quad (1.4.6)$$

На практике обычно наблюдаются не две чистые косинусоиды, а колебания, занимающие некоторую полосу частот. Поэтому групповая задержка является частотно-зависимой функцией и задается выражением $t_g = d\varphi/d\omega$. Фазовый угол φ можно определить как арктангенс отношения мнимой и действительной частей преобразования Фурье: $\varphi(\omega) = \arctg[\text{Im } B(\omega)/\text{Re } B(\omega)]$. Иногда удобно использовать определение комплексного логарифма.

Тогда

$$B = r e^{i\varphi}; \quad \ln B = \ln |r| + i \varphi = \ln |r| + i \varphi.$$

Итак,

$$\varphi = \text{Im} \ln B;$$

$$t_g = \frac{d\varphi}{d\omega} = \text{Im} \frac{d}{d\omega} [\ln B(\omega)] = \text{Im} \frac{1}{B} \frac{dB}{d\omega}. \quad (1.4.7)$$

Удобным приближением в вычислениях для дискретных значений B является формула

$$t_g \approx \frac{2}{\Delta\omega} \text{Im} \frac{B_{k+1} - B_k}{B_{k+1} + B_k}. \quad (1.4.8)$$

Важным аспектом теории распространения волн является различие между фазовой и групповой скоростями. Они сходны с фазовой и групповой задержками. Например, если волны распространяются вдоль двухмерной поверхности, то фазовая функция может быть задана как

$$\varphi(x, y) = k_x(x - x_0) + k_y(y - y_0). \quad (1.4.9)$$

Здесь координаты (x_0, y_0) определяют местоположение входного сигнала фильтра, а (x, y) — место наблюдения фазы (подобно выходу фильтра). Величины k_x и k_y означают «пространственные частоты», т. е. k_x есть 2π , деленное на длину волны, измеренную по оси x . Соотношение между ω , k_x и k_y устанавливается методами теоретической физики. Обычно оно может быть задано в форме

$$\omega = \omega(k_x, k_y). \quad (1.4.10)$$

Поскольку скорость есть расстояние, деленное на время, то фазовую скорость по направлению x можно определить как

$$(V_{\text{фаз}})_x = \frac{x - x_0}{\varphi/\omega} = \frac{\omega}{k_x}.$$

Групповая скорость для x -компоненты

$$(V_{\text{гр}})_x = \frac{x - x_0}{d\varphi/d\omega} = (x - x_0) \frac{d\omega}{d\varphi}. \quad (1.4.11)$$

Положим, что $y = y_0$. Тогда формула (1.4.9) приводится к виду

$$\varphi = k_x(x - x_0).$$

Откуда (в дифференциальном виде)

$$\frac{\partial k_x}{\partial \varphi} = \frac{1}{x - x_0}.$$

С учетом формулы (1.4.11) это дает

$$(V_{\text{гр}})_x = (x - x_0) \frac{\partial \omega}{\partial k_x} \frac{\partial k_x}{\partial \varphi} = \frac{\partial \omega}{\partial k_x}. \quad (1.4.12)$$

Таким образом, векторная групповая скорость есть $(\partial\omega/\partial k_x, \partial\omega/\partial k_y)$. Иногда физическая теория так сложна, что соотношение, подобное выражению (1.4.10), нельзя найти в явной форме и

вместо него используется неявная зависимость $F(\omega, k_x, k_y)$. В таком случае в соответствии с известным соотношением из теории частных производных имеем

$$\frac{\partial \omega}{\partial k_x} = - \frac{\partial F / \partial k_x}{\partial F / \partial \omega}.$$

В экспериментальной геофизике понятие скорости почти всегда связывается с групповой скоростью, с которой перемещается

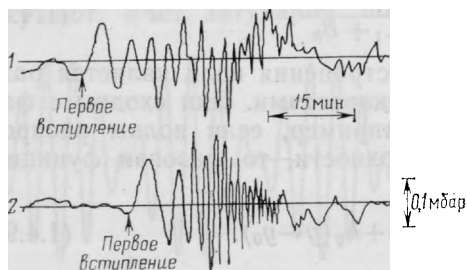


РИС. 1.11. Пример двух волн 1 и 2 звукового давления, которые считаются результатом ядерного взрыва в Азии (они были зарегистрированы в Калифорнии одним из микробарографов автора)

энергия. На рис. 1.11 присутствует излишний «шум» (обычный в экспериментальной геофизике), однако можно видеть, что возмущение проявляется сначала в виде колебаний с большим периодом, а затем уже в виде колебаний с коротким периодом. Групповая скорость находится путем деления расстояния на время распространения волны. Фазовые скорости можно изучать, имея две станции наблюдения недалеко друг от друга и измеряя временную задержку некоторых специфических пересечений нулевой линии. Причина, по которой станции располагают недалеко друг от друга, заключается в том, что форма волны непрерывно изменяется и если станции отстоят слишком далеко друг от друга, то не будет возможности определения тех пересечений нулевой линии, которые должны сравниваться.

1.5. КОРРЕЛЯЦИЯ И СПЕКТРЫ

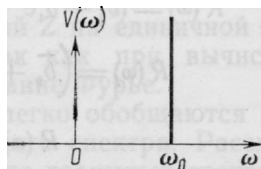
Спектр временной функции есть возведенное в квадрат преобразование Фурье этой функции. В случае действительной функции преобразование Фурье имеет четную действительную часть RE и нечетную мнимую часть IO. Взяв квадраты модулей, имеем:

$$(RE + i IO)(RE - i IO) = (RE)^2 + (IO)^2.$$

Квадрат нечетной функции так же как и квадрат четной функции является, очевидно, четной функцией. Таким образом, спектр действительной функции времени есть четная функция, так что его значения в области положительных частот совпадают с его значениями в области отрицательных частот. Поэтому нет особого смысла выделять отрицательные частоты. Хотя большинство встречающихся на практике функций времени являются действительными, рассмотрение корреляции и спектров не будет математически полным без функций времени с комплексными значе-

ниями. Кроме того, комплексные функции времени весьма полезны во многих физических задачах. В качестве примера рассмотрим векторный двухкомпонентный индикатор скорости ветра: одна компонента, направленная на север, регистрирует n_t , а другая, направленная на запад, записывает ω_t . Теперь, если составить временной ряд с комплексными значениями $v_t = n_t + i\omega_t$, то модули

РИС. 1.12. Спектр комплексного временного ряда $e^{-i(\omega_0 t + \varphi)}$



и фазовые углы комплексных чисел имеют очевидную физическую интерпретацию. Первая часть преобразования $(RE + iIO)$ относится к n_t , а вторая $(RO + iIE)$ — к ω_t . Однако спектр $(RE + RO)^2 + (IE + IO)^2$ не является ни четным, ни нечетным, а неравенство $V(+\omega) \neq V(-\omega)$ должно иметь какое-то объяснение. Действительно, оно имеется, и смысл его состоит в том, что $+\omega$ соответствует вращению по часовой стрелке, а $-\omega$ — вращению в противоположном направлении. Чтобы увидеть это, положим, что $n_t = \cos(\omega_0 t + \varphi)$ и

$\omega_t = -\sin(\omega_0 t + \varphi)$. Тогда $v_t = e^{-i(\omega_0 t + \varphi)}$, а преобразование Фурье имеет вид

$$V(\omega) = \int_{-\infty}^{+\infty} e^{-i(\omega_0 t + \varphi)} e^{i\omega t} dt; \quad (1.5.1)$$

$$V(\omega) = \delta(\omega - \omega_0) e^{-i\varphi}. \quad (1.5.2)$$

Спектр $\delta^2(\omega - \omega_0)$ показан на рис. 1.12. И, наоборот, если бы ω_t имело вид $\sin(\omega_0 t + \varphi)$, то спектр был бы дельта-функцией при $-\omega_0$, означающей, что вектор скорости ветра направлен в другую сторону. В геофизике имеются и другие примеры комплексных временных рядов.

1. Влияние гравитационного притяжения Луны на упругую Землю вызывает локальные наклоны поверхности Земли. Наклон в направлении север — юг мог бы определять временной ряд x_t , а наклон в направлении запад — восток — временной ряд y_t . Тогда ряд $x_t + iy_t$ стремился бы иметь одно значение вращения в северном полушарии и противоположное значение — в южном полушарии.

2. Вертикальные и горизонтальные движения сейсмографа также составляют комплексный временной ряд.

3. Нутации оси земного шара вокруг угловой мгновенной оси (Чендлер Уоббли).

4. Круговая поляризация электромагнитной волны.

Посмотрим, как выражается спектр через Z -преобразование. Пусть спектр

$$R(\omega) = |B(\omega)|^2 = \overline{B(\omega)} B(\omega). \quad (1.5.3)$$

Выразим это в понятиях Z -преобразования по трем точкам:

$$R(\omega) = (\overline{b_0} + \overline{b_1}e^{-i\omega} + \overline{b_2}e^{-i2\omega})(b_0 + b_1e^{i\omega} + b_2e^{i2\omega}); \quad (1.5.4)$$

$$R(\omega) = \left(\overline{b_0} + \frac{\overline{b_1}}{Z} + \frac{\overline{b_2}}{Z^2}\right)(b_0 + b_1Z + b_2Z^2); \quad (1.5.5)$$

$$R(\omega) = \overline{B}\left(\frac{1}{Z}\right)B(Z). \quad (1.5.6)$$

Для определения коэффициентов $R(Z)$ интересно перемножить полиномы $\overline{B}(1/Z)$ и $B(Z)$:

$$R(Z) = \frac{\overline{b_2}b_0}{Z^2} + \frac{(\overline{b_1}b_0 + \overline{b_2}b_1)}{Z} + (\overline{b_0}b_0 + \overline{b_1}b_1 + \overline{b_2}b_2) + (\overline{b_0}b_1 + \overline{b_1}b_2)Z + \overline{b_0}b_2Z \quad (1.5.7)$$

$$R(Z) = \frac{r_{-2}}{Z^2} + \frac{r_{-1}}{Z} + r_0 + r_1Z + r_2Z^2. \quad (1.5.8)$$

Коэффициент r_k при Z^k определяется выражением

$$r_k = \sum_i \overline{b_i}b_{i+k}. \quad (1.5.9)$$

Уравнение (1.5.9) известно как *автокорреляционная* формула. Значение функции автокорреляции r_k при задержке, равной 10, есть r_{10} . Она является мерой подобия величины b_t самой себе при сдвиге на 10 единиц времени. В большинстве случаев величина b_t действительная. При этом с помощью вычислений по формулам (1.5.7) или (1.5.9) можно убедиться, что автокорреляционные коэффициенты также являются действительными величинами, а $r_k = r_{-k}$. Тогда, ограничиваясь действительными временными рядами, имеем:

$$R(Z) = r_0 + r_1\left(Z + \frac{1}{Z}\right) + r_2\left(Z^2 + \frac{1}{Z^2}\right); \quad (1.5.10)$$

$$R(Z) = r_0 + r_1(e^{i\omega} + e^{-i\omega}) + r_2(e^{i2\omega} + e^{-i2\omega}); \quad (1.5.11)$$

$$R(Z) = r_0 + 2r_1 \cos \omega + 2r_2 \cos 2\omega; \quad (1.5.12)$$

$$R(Z) = \sum_k r_k \cos k\omega. \quad (1.5.13)$$

$$R(Z) = \text{косинус-преобразованию Фурье от } r_k. \quad (1.5.14)$$

Таким образом, разложение функции автокорреляции по косинусам равно квадрату модуля Фурье-преобразования. Решение

этой задачи для непрерывного времени относится к наиболее трудным теоремам. Известны два различных вычислительных метода расчета спектра: 1) сначала вычисляют коэффициенты r_k с помощью выражения (1.5.9), а затем формируют сумму косинусов в соответствии с формулой (1.5.13); 2) вычисляют величины $B(Z)$ для некоторых значений Z на единичной окружности и умножают полученные числа на комплексно-сопряженные. Вычисления повторяют для множества значений Z на единичной окружности. Второй метод более выгоден, так как при вычислениях можно использовать быстрое преобразование Фурье.

Понятия автокорреляции и спектра легко обобщаются как на понятия взаимной корреляции и взаимного спектра. Рассмотрим два Z -преобразования $A(Z)$ и $B(Z)$. Тогда взаимный спектр $C(Z)$ определяется выражением

$$C(Z) = \bar{A}\left(\frac{1}{Z}\right)B(Z). \quad (1.5.15)$$

Если какой-либо из коэффициентов c_k в $C(Z)$ превышает все остальные, то говорят, что формы волн a_t и b_t достаточно похожи, если одна из них имеет задержку на k единиц времени по отношению к другой.

Упражнения

1. Предположим, что сейсмический импульс образован комплексными числами. Справедливо ли при этих условиях автокорреляционное соотношение $r_k = r_{-k}$? Являются ли при этом величины r_k и $R(\omega)$ действительными или комплексными?

2. Пусть x_t — некоторая действительная функция времени. Пусть $y_t = x_{t+3}$ — другая функция времени. Постройте фазы взаимного спектра произведения $X(1/Z)Y(Z)$ в функции частоты, вычисленного на ЭВМ, в предположении, что арктангенсы лежат в основных квадрантах $-\pi/2 < \arctg < +\pi/2$. Нанесите масштабы по осям.

3. Если взаимно поменять понятия времени и частоты, то каким станет смысл спектра?

1.6. ПРЕОБРАЗОВАНИЕ ГИЛЬБЕРТА

Фильтр, преобразовывающий синусы в косинусы, называется фильтром со сдвигом фазы на 90° или квадратурным фильтром. Более конкретно, если входной импульс представляет собой $\cos(\omega t + \varphi_1)$, то выходной сигнал должен иметь вид $\cos(\omega t + \varphi_1 + \pi/2)$. Такой фильтр полезен при построении огибающей временной функции. Пусть $X(Z)$ означает Z -преобразование ряда действительных значений, а $Q(Z)$ — квадратурный фильтр, и пусть $Y(Z) = Q(Z)X(Z)$ описывает сигнал на выходе квадратурного фильтра. Тогда огибающую функции времени можно определить выражением $e_t = (x_t^2 + y_t^2)^{1/2}$. С другой стороны, можно построить комплексную функцию времени $u_t = x_t + iy_t$, Z -преобразование которой имеет вид

$$U(Z) = [1 + iQ(Z)]X(Z).$$

Тогда $u_t \bar{u}_t$ представляет собой квадрат огибающей функции. Аналогично фазу φ_t , являющуюся функцией времени, можно опреде-

литель выражением $\varphi_t = \text{arctg}(y_t/x_t)$. Мгновенная частота есть $d\varphi/dt$. Она может быть получена следующим способом:

$$\varphi_t = \text{Im} \ln u_t;$$

$$\frac{d\varphi}{dt} = \text{Im} \frac{1}{u} \frac{du}{dt} \approx \text{Im} \frac{2}{\Delta t} \frac{u_t - u_{t-1}}{u_t + u_{t-1}}.$$

Теперь, имея некоторое представление об использовании фильтра с фазовым сдвигом на 90° , найдем числовые значения q_t . Опера-

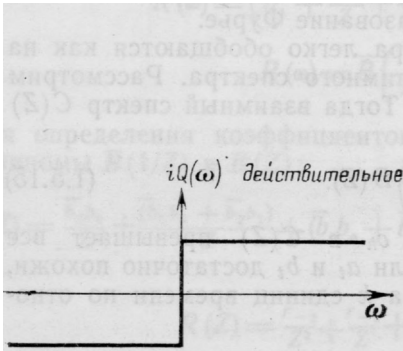


РИС. 1.13. Частотная характеристика фильтра, смещающего фазу на 90° .

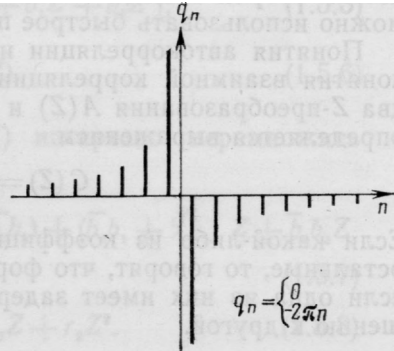


РИС. 1.14. Коэффициенты квадратурного фильтра

ция дифференцирования по времени обладает требуемым свойством сдвига фазы на 90° , который мы ищем. Погрешность дифференцирования заключается в том, что более высокие частоты усиливаются по сравнению с более низкими. В частности,

$$\dot{f}(t) = \int F(\omega) e^{-i\omega t} d\omega;$$

$$\frac{df}{dt} = \int -i\omega F(\omega) e^{-i\omega t} d\omega.$$

Таким образом, дифференцирование по времени соответствует весовому множителю $-i\omega$ в частотной области. Множитель $-i\omega$ имеет нужную нам фазу, но искаженную амплитуду. Требуемый весовой множитель равен $Q(\omega) = -i\omega/|\omega|$. Он является ступенчатой функцией, показанной на рис. 1.13.

Преобразуем $Q(\omega)$ во временную область

$$q_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q(\omega) e^{-i\omega n} d\omega = \frac{i}{2\pi} \int_{-\pi}^0 e^{-i\omega n} d\omega - \frac{i}{2\pi} \int_0^{\pi} e^{-i\omega n} d\omega =$$

$$= \frac{i}{2\pi} \left(\frac{e^{-i\omega n}}{-in} \Big|_{-\pi}^0 - \frac{e^{-i\omega n}}{-in} \Big|_0^{\pi} \right) = \frac{1}{2\pi n} (-1 + e^{in\pi} + e^{-in\pi} - 1) =$$

$$= \begin{cases} 0 & \text{для } n \text{ четных} \\ -\frac{2}{\pi n} & \text{для } n \text{ нечетных} \end{cases}$$

Результат показан на рис. 1.14.

Так как свойства фильтра сохраняются для отрицательных n , то он, очевидно, является физически нереализуемым фильтром (фильтром, который требует будущих входных сигналов для формирования текущего выходного сигнала). Если бы исследование проводилось для непрерывного времени, а не для дискретного, то фильтр мог бы иметь форму $1/t$, т. е. функцию, которая имеет

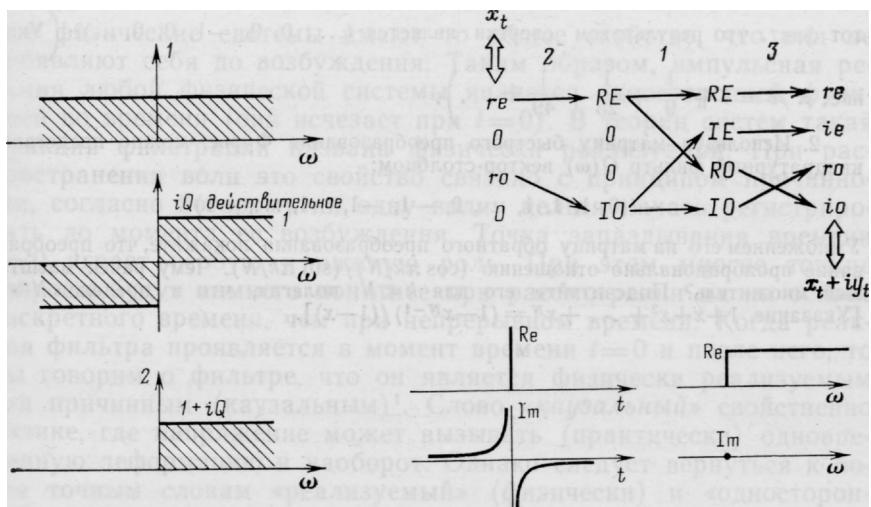


РИС. 1.15. Фильтр $1+iQ(Z)$ действительный и односторонний в частотной области, но комплексный и двухсторонний во временной области

РИС. 1.16. Преобразование Гильберта или квадратурная фильтрация со ступенчатой весовой функцией в частотной области.

1 — весовая функция; 2 — Фурье-преобразование (ФП); 3 — обратное Фурье-преобразование (ФП⁻¹)

РИС. 1.17. Импульс плюс i , прошедший через фильтр, сдвигающий его на 90° , является действительной ступенькой в частотной области

особенность при $t=0$ и интеграл которой по $+t$ расходится. Свертка с коэффициентами фильтра q_n поэтому очень неудобна, так как бесконечная последовательность сходится очень медленно. Свертка с фильтром q называется *преобразованием Гильберта*.

Возвратимся к фильтру $1+iQ(Z)$. Как показано на рис. 1.15, этот фильтр является ступенчатой функцией в частотной области. Операцию фазового сдвига на 90° проще всего осуществить в частотной области. Для этого сигнал записывается в виде $x_t + i \cdot 0$ и преобразуется в частотную область, затем умножается на ступеньку, показанную на рис. 1.15. Наконец, обратное преобразование дает $x_t + iy_t$. Преобразование четной, нечетной, действительной и мнимой частей показано на рис. 1.16.

Функция $1+iQ$ играет особую роль в анализе теоретических временных рядов, который, как будет показано в последующих разделах, связывается с принципом причинности. На рис. 1.17 показаны основные свойства этой функции.

Упражнения

1. Сверните сигнал вида $(2/\pi) \left(\dots, -\frac{1}{5}, 0, -\frac{1}{3}, 0, -1, 0, 1, 0, \frac{1}{3}, 0, \frac{1}{5}, \dots \right)$ с самим собой с помощью частичных дробей. Как объяснить тот факт, что результатом свертки является $(\dots, 0, 0, -1, 0, 0, \dots)$? (Указание. $\pi^2/8 = 1 + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \dots$).

2. Используя матрицу быстрого преобразования Фурье, можно представить квадратурный фильтр $Q(\omega)$ вектор-столбцом:

$$-i(0, 1, 1, 1, \dots, 0, -1, -1, -1, \dots, -1)^T$$

Умножением его на матрицу обратного преобразования покажите, что преобразование пропорционально отношению $(\cos \pi k/N)/(\sin \pi k/N)$. Чему равен масштабный множитель? Подсчитайте его для $k \ll N$, полагая, что в пределе $N \rightarrow \infty$. [Указание. $1+x+x^2+\dots+x^N = (1-x^{N+1})/(1-x)$].

2. ОДНОСТОРОННИЕ ФУНКЦИИ

Все физические системы имеют то общее свойство, что они не проявляют себя до возбуждения. Таким образом, импульсная реакция любой физической системы является односторонней функцией во времени (она исчезает при $t=0$). В теории систем такая функция фильтрации названа физически реализуемой. При распространении волн это свойство связано с принципом причинности, согласно которому ни одну волну нельзя начать регистрировать до момента ее возбуждения. Точка запаздывания времени $t=0$ играет особую и важную роль. При этом многие тонкие вопросы будут намного понятнее при рассмотрении их на основе дискретного времени, чем при непрерывном времени. Когда реакция фильтра проявляется в момент времени $t=0$ и после него, то мы говорим о фильтре, что он является физически реализуемым или причинным (каузальным)¹. Слово «каузальный» свойственно физике, где напряжение может вызывать (практически) одновременную деформацию и наоборот. Однако следует вернуться к более точным словам «реализуемый» (физически) и «односторонний», когда теория фильтров применяется для описания экономических и социальных систем, в которых одновременность достаточно отличается от причины и следствия.

2.1. ИНВЕРСНЫЕ ФИЛЬТРЫ

Чтобы лучше понять причинные фильтры, рассмотрим задачу определения функции такого фильтра. Допустим, что выходной сигнал y_t фильтра b_t известен, а входной сигнал x_t неизвестен (рис. 2.1).

Эта задача часто возникает в приемно-регистрирующей системе. Например, выходной сигнал сейсмоприемника является нанесенной на лист бумаги линией, по которой сейсмолог может определить смещение, скорость или ускорение почвы. Для того чтобы уничтожить фильтрующее действие фильтра $B(Z)$, попытаемся найти другой фильтр² $A(Z)$, который восстановит на выходе исходный сигнал, как это показано на рис. 2.2.

Для нахождения коэффициентов фильтра $A(Z)$ достаточно определить коэффициенты при степенях Z в выражении $E(Z)A(Z)=1$. Если $B(Z)$ трехчленный фильтр, то будем иметь

$$(a_0 + a_1Z + a_2Z^2 + a_3Z^3 + \dots)(b_0 + b_1Z + b_2Z^2) = 1. \quad (2.1.1)$$

¹ Вернее, удовлетворяющим принципу причинности. (Прим. перев.).

² Такой фильтр обычно называют обратным или инверсным (по отношению к предыдущему). (Прим. перев.)

Коэффициенты при Z^0, Z^1, Z^2, \dots в выражении (2.1.1) имеют вид

$$a_0 b_0 = 1; \quad (2.1.2)$$

$$a_1 b_0 + a_0 b_1 = 0; \quad (2.1.3)$$

$$a_2 b_0 + a_1 b_1 + a_0 b_2 = 0; \quad (2.1.4)$$

$$a_3 b_0 + a_2 b_1 + a_1 b_2 = 0; \quad (2.1.5)$$

$$a_4 b_0 + a_3 b_1 + a_2 b_2 = 0; \quad (2.1.6)$$

$$\dots \dots \dots$$

$$a_k b_0 + a_{k-1} b_1 + a_{k-2} b_2 = 0. \quad (2.1.7)$$

Из формулы (2.1.2) можно выразить a_0 через b_0 . Затем из (2.1.3) можно выразить a_1 через a_0 и b_1 , а из (2.1.4) найти a_2 через a_1, a_0

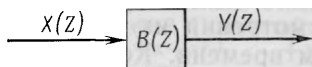


РИС. 2.1. Фильтр с неизвестным входным сигналом. $X(Z)$ — неизвестный сигнал; $Y(Z)$ — известный сигнал

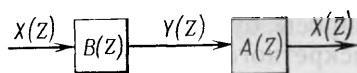


РИС. 2.2. Инверсный (обратный) фильтр $A(Z)$ по отношению к фильтру $B(Z)$

и b_k . Аналогично в общем случае a_k можно найти через a_{k-1}, a_{k-2} и b_k . В частности, коэффициент a_k можно определить из выражения (2.1.7) по рекурсивной формуле:

$$a_k = \frac{-\sum_{i=1}^k a_{k-i} b_i}{b_0} \quad (2.1.8)$$

Рассмотрим пример, когда $B(Z) = 1 - Z/2$. Используя уравнения вида (2.1.2) — (2.1.7), биномиальную теорему, деление полиномов или формулу Тейлора разложения функций в степенные ряды, получим:

$$A(Z) = \frac{1}{1 - Z/2} = 1 + \frac{Z}{2} + \frac{Z^2}{4} + \frac{Z^3}{8} + \dots \quad (2.1.9)$$

Как видно, число коэффициентов фильтра бесконечно, но оно быстро убывает, ввиду чего аппроксимация этих рядов на ЭВМ не представляет трудностей. Однако ситуация не так проста для фильтра вида $B(Z) = 1 - 2Z$:

$$A(Z) = \frac{1}{1 - 2Z} = 1 + 2Z + 4Z^2 + 8Z^3 + 16Z^4 + 32Z^5 + \dots \quad (2.1.10)$$

Коэффициенты этого ряда безгранично возрастают. Выходные сигналы фильтра $A(Z)$ неограниченно сильно зависят от входных сигналов, бесконечно отстающих во времени. [Напомним, что

текущий выходной сигнал фильтра $A(Z)$ есть a_0 , умноженное на текущий входной сигнал x_t , плюс a_1 , умноженное на предыдущий входной сигнал x_{t-1} , и т. д., так что a_n сохраняет память о сигналах, действовавших в предыдущих n единиц времени]. Смысл этого заключается в том, что некоторые фильтры $B(Z)$ не имеют пригодных для их аппроксимации конечных обратных фильтров

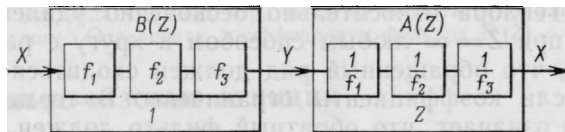


РИС. 2.3. Разложение полиномиального представления фильтра $B(Z)$ на ряд двухчленных фильтров, каждый из которых имеет связанный с ним обратный фильтр.

1 — коэффициенты полинома $B(Z)$; 2 — коэффициенты, обратные коэффициентам полинома $B(Z)$

$A(Z)$, определяемых из соотношений (2.1.2)—(2.1.8). Мы ищем пути получения хороших фильтров из плохих. Для двухимпульсного (двучленного) фильтра критерий заключается в том, что первый член в $B(Z)$ больше второго. Более строгое математическое описание ситуации определяется корнями $B(Z)$, т. е. значениями Z_0 , при которых $B(Z)=0$. Для случая $1-Z/2$ находим, что $Z_0=2$. Соответственно для $1-2Z$ имеем $Z_0=1/2$. Если корни Z_0 уравнения $B(Z)=0$ в общем случае элементарного сейсмического импульса с комплексными коэффициентами лежат внутри единичного круга на комплексной плоскости, то функция $1/B(Z)$ будет расходящейся. Если корень лежит вне единичного круга, то обращенная функция $1/B(Z)$ будет ограничена.

Напомним, что полином (многочлен) $B(Z)$ степени N можно разложить на множители в N подсистемах и что взаимное расположение этих подсистем не имеет значения (рис. 2.3). Вероятно, если при этом любой из N корней функции $B(Z)$ окажется внутри единичного круга, то можно предположить, что возникнут трудности с реализацией $A(Z)$. Это предположение строго доказывается теоремой об абсолютной сходимости рядов в теории комплексного переменного. Эта теорема гласит, что произведение абсолютно сходящихся рядов сходится и, наоборот, произведение любых сходящегося и расходящегося рядов расходится. Другое доказательство основывается на том, что степенной ряд для $1/B(Z)$ сходится внутри окружности вокруг начальной точки с радиусом, равным расстоянию от начальной точки до первого (ближайшего) полюса [нуль функции $B(Z)$ с наименьшим модулем]. Сходимость функции $A(Z)$ в единичном круге означает в терминах теории фильтров, что коэффициенты функции $A(Z)$ уменьшаются. Таким образом, если все нули $B(Z)$ лежат вне единичного круга, то из соотношения (2.1.8) получим сходящийся реализуемый обратный фильтр.

Можно ли что-нибудь вообще сделать, если один или несколько корней $B(Z)$ находится внутри единичного круга? Ответ на этот вопрос подсказывает следующий пример:

$$\frac{1}{1-2Z} = -\frac{1}{2Z} \frac{1}{1-1/2Z} = -\frac{1}{2Z} \left[1 + \frac{1}{2Z} + \frac{1}{(2Z)^2} + \dots \right]. \quad (2.1.11)$$

Уравнение (2.1.11) является разложением в ряд по степеням $1/Z$, т. е. рядом Тейлора относительно бесконечно удаленной точки. Он сходится при $Z \rightarrow \infty$ любым способом к кругу с радиусом $1/2$. Это означает, что обращенный ряд должен сходиться в единичном круге, если коэффициенты ограничены. В терминах теории фильтров это означает, что обратный фильтр должен быть таким фильтром, который реагирует на будущее по времени входные сигналы и, следовательно, является физически нереализуемым, но может быть использован при машинном моделировании.

Поэтому в общем случае надо разложить $B(Z)$ на множители, соответствующие двум областям: $B(Z) = B_{\text{вне}}(Z)B_{\text{вн}}(Z)$, где $B_{\text{вне}}$ содержит корни, лежащие за пределами единичного круга, а $B_{\text{вн}}$ содержит корни, расположенные внутри него. Тогда обращение части фильтра $B_{\text{вне}}$ выражается как ряд Тейлора в окрестности начала отсчета, а обращение части фильтра $B_{\text{вн}}$ выражается как ряд Тейлора на бесконечности. Полное выражение для $1/B(Z)$ называется разложением Лорана для $1/B(Z)$; оно сходится на кольце, окружающем единичный круг. Случаи с нулями на самой единичной окружности представляют определенную проблему. В некоторых случаях трудностей не возникает, но иногда расположение корней на окружности или даже вблизи нее означает, что определенная вычислительная схема на практике не будет хорошо работать.

Наконец, рассмотрим пример из механики. Напряжение (давление) в материале может быть представлено через x_t , а деформация (изменение объема) — через y_t . Следующие два утверждения эквивалентны, т. е. в некоторых ситуациях они оба справедливы, а в других — оба неверны.

Утверждение 1. Напряжение в материале можно представить линейной комбинацией текущих и прошлых деформаций. Аналогично деформацию можно определить на основе текущих и прошлых напряжений.

Утверждение 2. Фильтр, который связывает напряжения с деформациями и наоборот, имеет все полюсы и нули вне единичного круга.

Упражнения

1. Найдите фильтр, обратный по отношению к фильтру $(2-5Z+2Z^2)$. Можно сразу уменьшить количество степеней Z высокого порядка, но точное выражение для коэффициентов любой степени Z предпочтительнее (разбиение на части — полезный, хотя и не обязательный прием). Сделайте набросок импульсной реакции фильтра.

2. Покажите, что умножение на $(1-Z)$ в дискретном времени аналогично дифференцированию по времени при непрерывном времени. Покажите, что деле-

ние на $(1-Z)$ аналогично интегрированию. Какие пределы должен иметь интеграл?

3. Опишите общий метод определения $A(Z)$ и $B(Z)$ по разложению в ряд Тейлора $B(Z)/A(Z) = C_0 + C_1Z + C_2Z^2 + \dots + C_\infty Z^\infty$, где $B(Z)$ и $A(Z)$ являются многочленами неизвестных степеней n и m соответственно. Примените этот метод к случаю $C(Z) = \frac{1}{2} - \frac{3}{4}Z - \frac{3}{8}Z^2 - \frac{3}{16}Z^3 - \frac{3}{32}Z^4 - \dots$. Не пытайтесь решить эту задачу до тех пор, пока вы недостаточно знакомы с определителями. [Указание. Определите коэффициенты фильтра $B(Z) = A(Z)C(Z)$].

2.2. МИНИМАЛЬНО-ФАЗОВЫЕ ФУНКЦИИ

В разделе 2.1. показано, что сходимость ряда Тейлора для функции $1/B(Z)$ в круге $|Z|=1$ эквивалентна отсутствию у функции $B(Z)$ корней внутри единичного круга. Покажем, что эти условия эквивалентны также определенному поведению фазы функции $B(Z)$ в единичном круге. Рассмотрим фазовый сдвиг, вносимый двучленным фильтром:

$$B = 1 - \frac{Z}{Z_0} = 1 - \rho^{-1}e^{i(\omega - \omega_0)} = 1 - \rho^{-1} \cos(\omega - \omega_0) - i\rho^{-1} \sin(\omega - \omega_0)$$

$$(Z_0 = \rho e^{i\omega_0})$$

По определению, фазой является арктангенс отношения мнимой части к действительной.

График фазы как функции частоты выглядит совершенно различно для $\rho < 1$ и для $\rho > 1$. На рис. 2.4. дан случай $\rho > 1$.

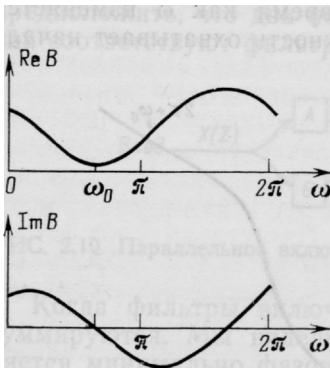


Рис. 2.4. Действительная и мнимая части Z -преобразования функции $1 - Z/(1,25e^{i2\pi/3})$

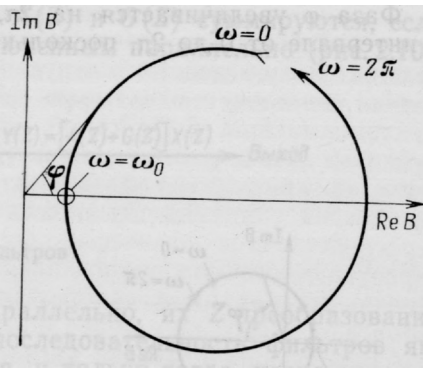


Рис. 2.5. Поведение фазы двучленного фильтра, представленного на рис. 2.4, в комплексной области

Фаза является арктангенсом величины $\text{Im } B / \text{Re } B$. Проще всего следить за поведением фазы на комплексной плоскости B . Это показано на рис. 2.5.

Фаза как функция частоты изображена на рис. 2.6. Отметим, что фаза φ при $\omega=0$ имеет такое же значение, как и при $\omega=2\pi$.

Это следует из того, что действительная и мнимая части функции имеют период 2π . Ситуация изменится, если внутри единичного круга появится нуль, т. е. когда $\rho < 1$. Действительная и мнимая

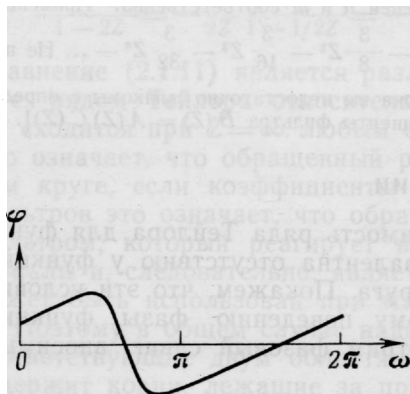


РИС. 2.6. Фазовая характеристика двухчленного минимально фазового фильтра [$\varphi = \arctg(\text{Im}/\text{Re})$]

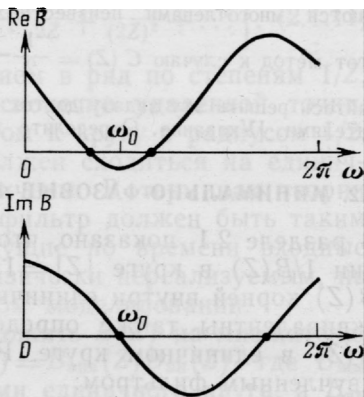


Рис. 2.7. Действительная и мнимая части частотной характеристики двухчленного неминимально фазового фильтра $1 - 1,25Z e^{-i2\pi/3}$

части функции для этого случая изображены на рис. 2.7, а комплексная плоскость на рис. 2.8.

Фаза φ увеличивается на 2π , в то время как ω изменяется в интервале от 0 до 2π , поскольку окружность охватывает начало

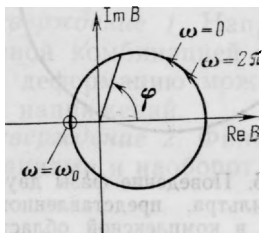


РИС. 2.8. Поведение фазы в комплексной плоскости

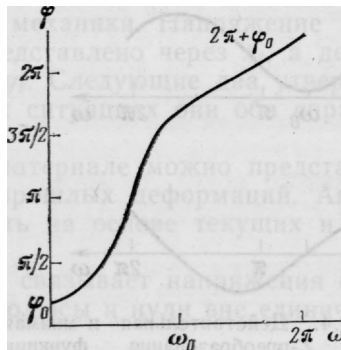


РИС. 2.9. Фазовая характеристика двухчленного неминимально фазового фильтра [$\varphi = \arctg(\text{Im}/\text{Re})$]

координат. Фазовая характеристика рассматриваемой функции показана на рис. 2.9. Случай $\rho > 1$, когда $\varphi(\omega) = \varphi(\omega + 2\pi)$, называют *минимально фазовым* (он соответствует *минимуму задержки*).

Теперь рассмотрим сложный фильтр

$$B(Z) = \frac{(Z - c_1)(Z - c_2) \dots}{(Z - a_1)(Z - a_2) \dots}. \quad (2.2.1)$$

По правилам перемножения комплексных чисел фазовая характеристика функции $B(Z)$ равна сумме фазовых характеристик сомножителей числителя минус сумму фазовых характеристик сомножителей знаменателя. Поскольку мы рассматриваем физически реализуемые фильтры, то все сомножители знаменателя должны быть минимально фазовыми и, следовательно, фазовая характеристика знаменателя будет суммой кривых, подобных изображенным на рис. 2.6. Сомножители числителя могут быть или не быть минимально фазовыми. Поэтому фазовая характеристика числителя будет суммой кривых, подобных изображенным или на рис. 2.6, или на рис. 2.9. Если какие-либо сомножители имеют фазовые характеристики, подобные изображенным на рис. 2.9, то суммарная фазовая характеристика будет похожа на характеристику, представленную на рис. 2.9, тем, что фаза в точке $\omega = 2\pi$ больше фазы в точке $\omega = 0$. Тогда фильтр будет неминимально фазовым.

2.3. ПАРАЛЛЕЛЬНОЕ ВКЛЮЧЕНИЕ ФИЛЬТРОВ

Известно, что при последовательном включении фильтров полиномы, описывающие их характеристики, перемножаются. Можно предположить, что два полинома $A(Z)$ и $G(Z)$ суммируются, если они соответствуют фильтрам, включенным параллельно (рис. 2.10).

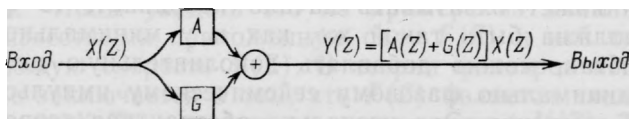


РИС. 2.10. Параллельное включение фильтров

Когда фильтры включены параллельно, их Z -преобразования суммируются. Мы видели, что последовательность фильтров является минимально фазовой тогда, и только тогда, когда каждый член произведения минимально фазовый. Получим достаточное (но не необходимое) условие, при котором сумма $A(Z) + G(Z)$ будет минимально фазовой. Прежде всего предположим, что фильтр $A(Z)$ — минимально фазовый. Тогда можно записать

$$A(Z) + G(Z) = A(Z) \left[1 + \frac{G(Z)}{A(Z)} \right].$$

Следовательно, вопрос о том, является ли сумма $A(Z) + G(Z)$ минимально фазовой, превратился в вопрос о том, являются ли оба фильтра, $A(Z)$ и $1 + G(Z)/A(Z)$, минимально фазовыми. Ранее

мы предполагали, что $A(Z)$ минимально фазовый. Перед тем, как задаваться вопросом о том, является ли фильтр $1+G(Z)/A(Z)$ минимально фазовым, нужно убедиться в том, что он реализуемый. Поскольку фильтр $1/A(Z)$ представляется (по условию) только положительными степенями Z , то фильтр $G(Z)/A(Z)$ также является реализуемым. В дальнейшем будет показано достаточное

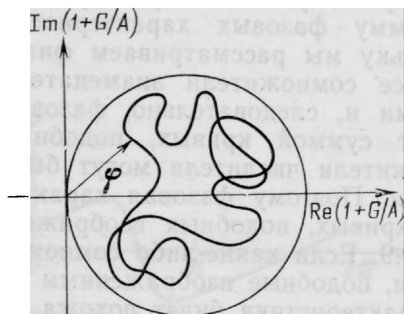


РИС. 2.11. Поведение фазы положительной действительной функции в интервале $\pm\pi/2$

условие минимальной фазовости фильтра $1+G(Z)/A(Z)$, заключающееся в том, что спектр фильтра A превосходит спектр фильтра G на всех частотах¹. Другими словами, для любой действительной частоты $\omega|A| > |G|$. Таким образом, если изобразить кривую $G(Z)/A(Z)$ на комплексной плоскости для действительных ω в диапазоне $0 \leq \omega \leq 2\pi$, то она будет лежать всюду внутри единичного круга. Если теперь прибавить единицу, чтобы перейти к $1+G(Z)/A(Z)$, то соответствующая кривая всегда будет иметь положительную действительную часть (рис. 2.11). Поскольку эта кривая не может охватывать начало координат, то фаза этой функции должна быть такой же, как и у минимально фазовой функции. Итак, можно добавлять дополнительную информацию (шум) к минимально фазовому сейсмическому импульсу, но добавлять ее умеренно. Эта несколько абстрактная теорема имеет непосредственное физическое следствие. Представим себе волну, характеризующуюся минимально фазовым сигналом $A(Z)$, излученную источником и зарегистрированную приемником через некоторое время. В еще более позднее время приходит отраженный от близлежащих объектов сигнал и также регистрируется приемником. Приемник регистрирует $Y(Z) = A(Z) + Z^n \alpha A(Z)$, где n определяет задержку отраженного сигнала относительно прихода фронта прямой волны, а α представляет уменьшение амплитуды отраженного сигнала. Чтобы увидеть, что $Y(Z)$ является минимально фазовым сигналом, заметим, что модуль Z^n равен единице и коэффициент отражения α должен быть меньше единицы (чтобы избежать вечного движения). При этом $Z^n \alpha A(Z)$ играет роль функции $G(Z)$. Таким образом, минимально фазовый сигнал

¹ В отечественной литературе вместо термина «спектр фильтра» обычно употребляется термин «частотная характеристика фильтра». (Прим. перев.)

вместе со своим отраженным сигналом тоже является минимально фазовым. В дальнейшем рассматриваются ситуации распространения волн, в которых учитываются многократно-отраженные сигналы (до бесконечности).

Упражнения

1. Найдите два неминимально фазовых сейсмических импульса, сумма которых будет минимально фазовой.

2. Пусть полином $A(Z)$ степени N является минимально фазовым. Пусть $A'(Z) = Z^N \bar{A}(1/Z)$. Найдите на комплексной Z плоскости корни функции $A'(Z)$, которая называется максимално фазовым сигналом. [Указание. Рассмотрите сначала простой случай $A(Z) = a_0 + a_1 Z_{нач}$].

3. Предположим, что полином $A(Z)$ — максимално фазовый, а степень полинома $G(Z)$ меньше либо равна степени полинома $A(Z)$. Допустим, что $|A| > |G|$. Покажите, что сумма полиномов $A(Z)$ и $G(Z)$ является максимално фазовой функцией.

4. Пусть полином $A(Z)$ — минимально фазовый. Где находятся корни функции $A(Z) + cZ^N \bar{A}(1/Z)$ в трех случаях: $|c| < 1$; $|c| > 1$; $|c| = 1$? [Указание. Корни полинома являются непрерывной функцией коэффициентов полинома.]

2.4. ПОЛОЖИТЕЛЬНЫЕ ДЕЙСТВИТЕЛЬНЫЕ ФУНКЦИИ

Две подобные функции, называемые *функцией полной проводимости* $Y(Z)$ и *функцией полного сопротивления* (импеданса) $I(Z)$, встречаются во многих физических задачах. В электронике они являются отношениями тока к напряжению и напряжения к току; в акустике полным сопротивлением называется отношение давления к скорости. Если в рассматриваемой электрической сети или акустической области не содержится источников энергии, то рассматриваемые величины являются действительными и положительными. Чтобы увидеть это на примере из механики, можно считать известными приложенную силу $F(Z)$ и наблюдаемую результирующую скорость $V(Z)$. В теории фильтров подобное рассмотрение заключается в том, что $F(Z)$ является входным сигналом фильтра $Y(Z)$, дающего на выходе сигнал $V(Z)$. Тогда

$$V(Z) = Y(Z)F(Z). \quad (2.4.1)$$

Фильтр $Y(Z)$ является, очевидно, реализуемым. Если это верно, то можно сделать наоборот, т. е. задать скорость и наблюдать силу, так как должна существовать сходящаяся реализуемая функция $I(Z)$:

$$F(Z) = I(Z)V(Z). \quad (2.4.2)$$

Поскольку функции Y и I предполагаются обратными по отношению друг к другу, а также ограниченными и реализуемыми, то обе они должны быть минимально фазовыми.

Прежде чем рассматривать физику явлений отметим, что если комплексное число $a + ib$ имеет положительную действительную часть a , то действительная часть числа $(a + ib)^{-1}$, т. е. $a / (a^2 + b^2)$ также является положительной. Используя $a + ib$ для представле-

ния величины $Y(Z)$ или $I(Z)$ на единичном круге, видим, что если одна из этих величин (либо Y , либо I) является действительной и положительной, то и другая величина обладает теми же свойствами.

Рассеянная мощность W выражается произведением силы на скорость, т. е.

$$W = \dots + f_0 v_0 + f_1 v_1 + f_2 v_2 + \dots \quad (2.4.3)$$

Это можно записать в терминах Z -преобразования в следующем виде:

$$W = \frac{1}{2} \text{коэффициента при } Z^0 \left[V\left(\frac{1}{Z}\right) F(Z) + F\left(\frac{1}{Z}\right) V(Z) \right];$$

$$W = \frac{1}{2} \cdot \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left[V\left(\frac{1}{Z}\right) F(Z) + F\left(\frac{1}{Z}\right) V(Z) \right] d\omega. \quad (2.4.4)$$

Используя выражение (2.4.1) для замены $V(Z)$, получим

$$W = \frac{1}{2} \cdot \frac{1}{2\pi} \int_{-\pi}^{+\pi} F\left(\frac{1}{Z}\right) \left[Y\left(\frac{1}{Z}\right) + Y(Z) \right] F(Z) d\omega.$$

Выражение в скобках $Y(Z) + Y(1/Z)$ выглядит внешне как спектр, поскольку коэффициент при члене Z^k равен коэффициенту при Z^{-k} , что указывает на симметрию автокорреляционной функции. Введем обозначение

$$R(Z) = Y(Z) + Y\left(\frac{1}{Z}\right) \quad (2.4.5)$$

и преобразуем выражение (2.4.4) к виду

$$W = \frac{1}{2} \cdot \frac{1}{2\pi} \int_{-\pi}^{+\pi} R(Z) F\left(\frac{1}{Z}\right) F(Z) d\omega. \quad (2.4.6)$$

Подынтегральная функция в формуле (2.4.6) является произведением спектра произвольной положительной исходной силы и $R(Z)$. Если рассеяние мощности считается положительным на всех частотах (для всех $\bar{F}F$), то, очевидно, функция $R(Z)$ должна быть положительной на всех частотах; таким образом, $R(Z)$ в самом деле является спектром. Поскольку только что установлено, что сумма $Y(Z) + Y(1/Z)$ должна быть положительной на всех частотах, то оказывается, что величина $Y(Z)$ не является произвольным минимально фазовым фильтром. Действительная часть функций как $Y(Z)$, так и $Y(1/Z)$ имеет вид

$$\text{Re} [Y(Z)] = \text{Re} \left[Y\left(\frac{1}{Z}\right) \right] = y_0 + y_1 \cos \omega + y_2 \cos 2\omega + \dots$$

Поскольку действительная часть суммы должна быть положительной, то, очевидно, действительная часть каждой из равных частей также должна быть положительной.

Если изучаемый предмет или механизм является пассивным (не содержащим источником энергии), то рассеяние энергии должно быть положительным в любом промежутке времени от минус бесконечности до любого момента времени t . Определим выражение для энергии, рассеянной в этом интервале времени. Для простоты будем считать силу и скорость исчезающими при $t=0$. Пусть конец временного интервала включает точку $t=2$ и не включает точку $t=3$.

Определим

$$f'_t = \begin{cases} f_t & \text{при } t \leq 2; \\ 0 & \text{при } t > 2. \end{cases} \quad (2.4.7)$$

Для определения совершенной за все время работы можно проинтегрировать выражение (2.4.6) по всем частотам. Чтобы найти работу в выбранном временном интервале, необходимо заменить F на F' и проинтегрировать выражение (2.4.6) по всем частотам, а именно:

$$W_2 = \frac{1}{2} \cdot \frac{1}{2\pi} \int_{-\pi}^{+\pi} F' \left(\frac{1}{Z} \right) R(Z) F'(Z) d\omega. \quad (2.4.8)$$

Как мы уже видели, в этом интеграле используется только коэффициент при Z^0 в подынтегральном выражении. Выведем это. Прежде всего соберем коэффициенты при степенях Z в произведении $R(Z)F'(Z)$. Тогда

$$Z^0: r_0 f'_0 + r_{-1} f'_1 + r_{-2} f'_2;$$

$$Z^1: r_1 f'_0 + r_0 f'_1 + r_{-1} f'_2;$$

$$Z^2: r_2 f'_0 + r_1 f'_1 + r_0 f'_2.$$

Для получения коэффициента при Z^0 в выражении $F'(1/Z) \times [R(Z)F'(Z)]$ необходимо умножить верхний из представленных выше рядов на f'_0 , второй ряд на f'_1 и третий ряд на f'_2 . Результат можно представить в очень компактном виде:

$$W_2 = \frac{1}{2} [f_0 f_1 f_2] \begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} = \frac{1}{2} [f_0 f_1 f_2] \begin{bmatrix} 2y_0 & y_1 & y_2 \\ y_1 & 2y_0 & y_1 \\ y_2 & y_1 & 2y_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix}. \quad (2.4.9)$$

Квадратичная форма размерности 3×3 в формуле (2.4.9) должна быть положительной (т. е. $W_2 > 0$ для произвольной f_t). Вместе с тем подобные квадратичные формы размерности $t \times t$ в выражениях для W_t также должны быть положительными.

В заключение отметим: свойство положительности и действительности в частотной области означает, что сумма $Y(Z) + Y(1/Z)$ положительна для любой действительной частоты ω , а свойство положительности и действительности во временной области означает, что все матрицы размерности $t \times t$, подобные приведенной

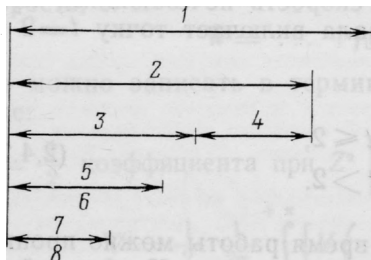


РИС. 2.12. Основные функции времени.

1 — все функции времени; 2 — односторонние функции времени (удовлетворяющие принципу причинности); 3 — функции с ограниченной энергией (физически реализуемые); 4 — функции с бесконечной энергией; 5 — минимально фазовые; 6 — поддающиеся обращению; 7 — полная проводимость или полное сопротивление; 8 — положительные, действительные

в формуле (2.4.9), являются положительно определенными. На рис. 2.12 приведены результаты рассмотрения свойств этих классов функций.

Упражнения

1. В механике сила и скорость свободной единичной массы выражаются зависимостями $dv/dt=f$ или $v = \int_{-\infty}^t f dt$. Вычислите рассеянную мощность как функцию частоты, если интеграл аппроксимируется сверткой из $(0,5, 1, 1, 1, \dots)$. [Указание. Разложите $(1+Z)/2(1-Z)$ по положительным степеням Z .]
2. Приведите пример простой функции, которая была бы минимально фазовой, не будучи положительной и действительной.

2.5. УЗКОПОЛОСНЫЕ ФИЛЬТРЫ

Фильтры часто используются для изменения спектра регистрируемых сигналов. Для входного сигнала $X(Z)$ фильтра $B(Z)$ и выходного сигнала $Y(Z)$ имеем $Y(Z) = B(Z)X(Z)$ и соответственно для сопряженного по Фурье выходного сигнала $\bar{Y}(1/Z) = \bar{B}(1/Z)\bar{X}(1/Z)$. Перемножив эти два выражения, получим соотношение

$$\bar{Y}Y = (\bar{B}B)(\bar{X}X),$$

которое показывает, что спектр входного сигнала, умноженный на спектр фильтра, равен спектру выходного сигнала. Обычно фильтры характеризуются формой их спектра. Примеры спектров некоторых фильтров показаны на рис. 2.13.

В дальнейшем нам часто придется иметь дело с синусоидальными временными функциями, Z -преобразование которых имеет вид

$$\frac{1}{1 - Ze^{i\omega_0}} = 1 + Ze^{i\omega_0} + Z^2e^{i2\omega_0} + \dots \quad (2.5.1)$$

Функция времени, соответствующая этому Z -преобразованию, имеет вид $e^{i\omega_0 t}$, но она „включается“ в момент времени $t=0$. Действительно, левая часть выражения (2.5.1) содержит полюс точно на единичной окружности, так что ряд находится на границе между областями сходимости и расходимости. Это может привести к парадоксальным ситуациям [можно было бы разложить функцию (2.5.1) в ряд так, чтобы синусоида выключалась при $t=0$], кото-

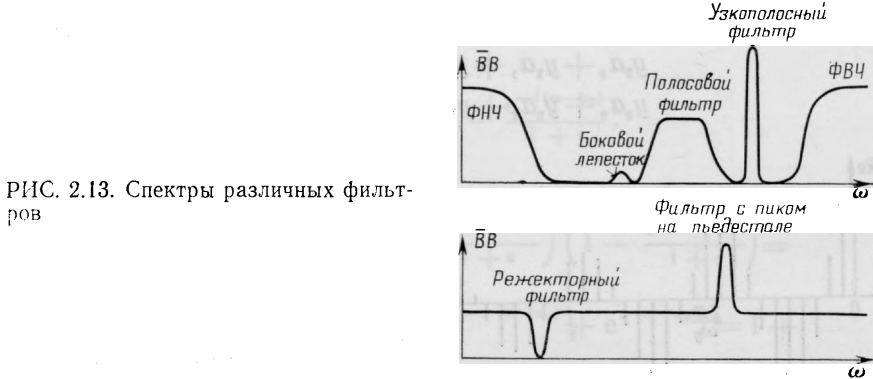


РИС. 2.13. Спектры различных фильтров

рых мы будем избегать путем смещения полюса с единичной окружности на небольшое расстояние ϵ за ее пределы. Пусть $Z^0 = (1 + \epsilon) e^{i\omega_0}$, тогда

$$B(Z) = \frac{1}{A(Z)} = \frac{1}{1 - Z/Z_0} = 1 + \frac{Z}{Z_0} + \left(\frac{Z}{Z_0}\right)^2 + \dots \quad (2.5.2)$$

Функция времени, соответствующая величине $B(Z)$, равна нулю при $t=0$ и равна $e^{-i\omega_0 t} / (1 + \epsilon)^t$ при $t > 0$. Это — затухающая синусоидальная функция, которая постепенно убывает со временем по закону $(1 + \epsilon)^{-t}$. Коэффициенты ее показаны на рис. 2.14. Вероятно, хотя позднее это будет доказано, свертка с коэффициентами выражения (2.5.2), изображенными на рис. 2.14, есть операция узкополосной фильтрации. Если полюс перемещается все ближе к единичной окружности, то полоса пропускания фильтра становится уже, а коэффициенты $B(Z)$ спадают все медленнее. Для выполнения свертки необходимо усекавать ряд (2.5.2), т. е. отбрасывать степени Z выше определенного практического предела. Это приводит к тому, что свертка с помощью коэффициентов фильтра $B(Z)$ оказывается невыгодной. Значительно более экономичным способом узкополосной фильтрации является деление полиномов на функцию $A(Z)$. Для выходного сигнала $Y(Z)$ запишем

$$Y(Z) = B(Z)X(Z); \quad (2.5.3)$$

$$Y(Z) = \frac{X(Z)}{A(Z)}; \quad (2.5.4)$$

Умножая обе части выражения (2.5.4) на функцию $A(Z)$, имеем:

$$Y(Z)A(Z) = X(Z). \quad (2.5.5)$$

Для определенности положим, что x_t и y_t равны нулю при $t=0$. Теперь определим коэффициенты при последовательных степенях Z :

$$\begin{aligned} y_0 a_0 &= x_0; \\ y_1 a_0 + y_0 a_1 &= x_1; \\ y_2 a_0 + y_1 a_1 + y_0 a_2 &= x_2; \\ y_3 a_0 + y_2 a_1 + y_1 a_2 + y_0 a_3 &= x_3. \end{aligned} \quad (2.5.6)$$

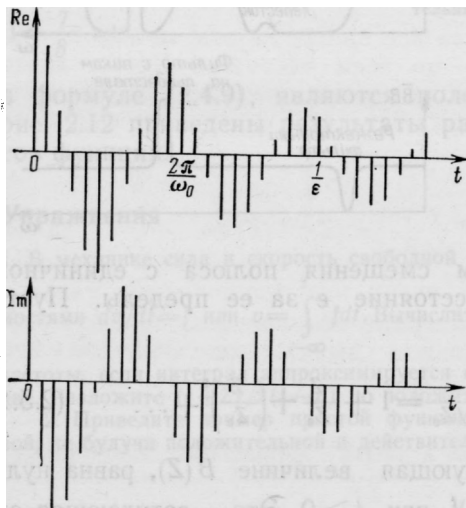


РИС. 2.14. Коэффициенты временной функции, имеющей простой полюс, лежащий вблизи единичной окружности за ее пределами в точке

$$Z_0 = 1,1e^{i\pi/5}$$

Общее выражение имеет вид

$$y_k a_0 + \sum_{i=1}^{\infty} y_{k-i} a_i = x_k. \quad (2.5.7)$$

Решая последнее выражение относительно y_k , получим

$$y_k = \frac{x_k - \sum_{i=1}^{\infty} y_{k-i} a_i}{a_0}. \quad (2.5.8)$$

Уравнение (2.5.8) можно использовать для определения y_k , если известны y_{k-1}, y_{k-2}, \dots . Таким образом, это решение рекурсивно и оно не будет расходиться, если a_i являются коэффициентами минимально фазового полинома. Практически бесконечная верхняя граница суммы усекается, как только истощается запас коэффициентов у $A(Z)$ или у $Y(Z)$. Для примера рассмотрим функцию $B(Z) = 1/A(Z) = 1/(1-Z/Z_0)$, сумма которой имеет только один

член. Фильтрация таким способом называется *фильтрацией с обратной связью*. Для узкополосной фильтрации этот способ экономичнее фильтрации с помощью свертки, так как полином $A(Z)$ имеет намного меньше коэффициентов, чем полином $B(Z) = 1/A(Z)$. Наконец, рассмотрим спектр полинома $B(Z)$. Имеем

$$A(Z) = 1 - \frac{Z}{Z_0} = 1 - \frac{e^{i\omega}}{(1+\varepsilon)e^{i\omega_0}} = 1 - \frac{e^{i(\omega-\omega_0)}}{1+\varepsilon}$$

и

$$\bar{A}\left(\frac{1}{Z}\right) = 1 - \frac{e^{-i(\omega-\omega_0)}}{1+\varepsilon}.$$

Следовательно,

$$\begin{aligned} \bar{A}\left(\frac{1}{Z}\right)A(Z) &= \left(1 - \frac{e^{-i(\omega-\omega_0)}}{1+\varepsilon}\right)\left(1 - \frac{e^{i(\omega-\omega_0)}}{1+\varepsilon}\right) = \\ &= 1 + \frac{1}{(1+\varepsilon)^2} - \frac{1}{1+\varepsilon}(e^{-i(\omega-\omega_0)} + e^{i(\omega-\omega_0)}) = 1 + \\ &+ \frac{1}{(1+\varepsilon)^2} - \frac{2\cos(\omega-\omega_0)}{1+\varepsilon} = 1 + \frac{1}{(1+\varepsilon)^2} - \frac{2}{1+\varepsilon} + \\ &+ \frac{1}{1+\varepsilon} \frac{2}{\varepsilon} [1 - \cos(\omega-\omega_0)] = \left(1 - \frac{1}{1+\varepsilon}\right)^2 + \frac{4}{1+\varepsilon} \sin^2 \frac{\omega-\omega_0}{2}; \\ \bar{B}\left(\frac{1}{Z}\right)B(Z) &= \frac{(1+\varepsilon)^2}{\varepsilon^2 + 4(1+\varepsilon)\sin^2 \frac{\omega-\omega_0}{2}}. \end{aligned} \quad (2.5.9)$$

Эту функцию можно достаточно точно аппроксимировать выражением $1/[\varepsilon^2 + (\omega-\omega_0)^2]$. График функции (2.5.9) показан на рис. 2.15.

Теперь понятно, почему эта функция называется узкополосным фильтром. Этот фильтр пропускает очень узкую полосу частот и ослабляет все другие. Ширина полосы пропускания (частотного окна) этого фильтра равна $\Delta\omega \approx 2\varepsilon$. Временное окно $\Delta t = 1/\varepsilon$, т. е. равно постоянной времени затухания синусоиды b_t .

Практическим недостатком рассматриваемого фильтра является то, что при подаче на его вход действительного сигнала сигнал на выходе будет комплексным. В то же время во многих приложениях фильтр с действительными коэффициентами оказывается более предпочтительным.

Одно из приближений заключается в том, чтобы заменить фильтр $[1, e^{i\omega_0}/(1+\varepsilon)]$, заданный во временной области, комплексно-сопряженным [фильтром $[1, e^{-i\omega_0}/(1+\varepsilon)]$]. Сложный составной оператор во временной области теперь имеет вид $[1, (2\cos \omega_0)/(1+\varepsilon), 1/(1+\varepsilon)^2]$ и является действительным. [Отметим, что ком-

плексно-сопряженная функция этого фильтра в частотной области имеет вид $\bar{B}(1/Z)$, а во временной — $\bar{B}(Z) = \bar{b}_0 + \bar{b}_1 Z + \dots$. Составной фильтр можно записать в виде произведения $B(Z)\bar{B}(Z)$. Спектр этого фильтра равен $[B(Z)\bar{B}(1/Z)] [B(Z)B(1/Z)]$. Можно легко удостовериться в том, что спектр $\bar{B}(Z)$ подобен спектру $B(Z)$, но максимум его находится в точке $-\omega_0$ вместо точки $+\omega_0$. Таким

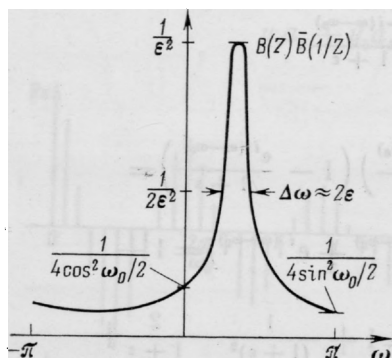


РИС. 2.15. Спектр, соответствующий функции с простым полюсом в точке $Z_0^* = (1 + \epsilon) e^{i\omega_0}$

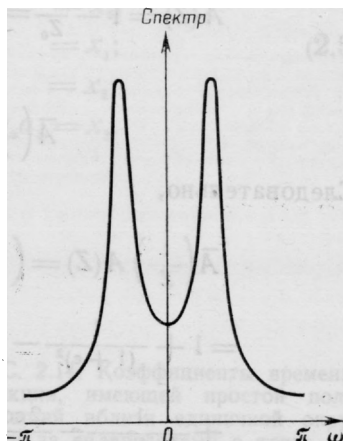


РИС. 2.16. Спектр двухполюсного фильтра, в котором один полюс такой же, как у фильтра, показанного на рис. 2.15, а второй — сопряжен с первым

образом, составной спектр является произведением спектра, показанного на рис. 2.15, на точно такой же спектр, обращенный вдоль оси частот. Это показано на рис. 2.16.

Упражнения

1. Простой операцией с обратной связью является $y_t = (1 - \epsilon)y_{t-1} + x_t$. Эта операция называется интегрированием с утечкой (затуханием). Предложите выражение для определения выходного сигнала y_t в замкнутой форме, если x_t — некоторый импульс. Чему равно время затухания τ в вашем решении (время, которое требуется для того, чтобы выходной сигнал y_t уменьшился до значения $e^{-1}y_0$)? Чему равно τ для малых ϵ , например $\epsilon = 0,1$; $0,001$ или $0,0001$?

2. Каковы удаления полюсов выражения $1/(1 - 0,1Z + 0,9Z^2)$ от единичной окружности? Чему равны значения времени затухания фильтра и его резонансной частоты?

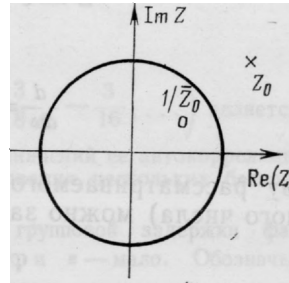
3. Найдите трехчленный действительный фильтр с обратной связью, пропускающий интервал частот 49—51 Гц данных наблюдений, квантуемых в 500 точках/с. Где располагаются полюсы? Каково время затухания этого фильтра?

2.6. ФАЗОВЫЕ ФИЛЬТРЫ

В разделе рассматриваются фильтры с постоянными единичными спектрами, т. е. $B(Z)\bar{B}(1/Z) = 1$. Другими словами, в частотной области фильтр $B(Z)$ принимает форму $e^{i\varphi(\omega)}$, где φ — действи-

тельная функция, которая называется *фазовым сдвигом*. Ясно, что $B\bar{B}=1$ для всех действительных φ . Для построения фильтра с любым требуемым фазовым сдвигом достаточно сделать преобразование Фурье функции $e^{i\varphi(\omega)}$ во временную область. Если $\varphi(\omega)$ произвольно, то получающаяся функция времени является в общем случае двухсторонней. Так как нас интересуют только физически реализуемые процессы, то хотелось бы знать, какой класс функций

РИС. 2.17. Пример характеристики широкополосного фильтра, полюс (X) которого лежит вне единичного круга, а нуль (O) — внутри него. Полюс и нуль лежат на одном радиусе



$\varphi(\omega)$ соответствует односторонним функциям времени. Проще начать с однополюсного фазового фильтра, имеющего один нуль. Тогда более сложные фазовые фильтры можно составить с помощью каскадного включения этих простых фильтров. Рассмотрим фильтр

$$P(Z) = \frac{Z - 1/\bar{Z}_0}{1 - Z/Z_0}. \quad (2.6.1)$$

Этот простой случай функций вида $Z^N \bar{A}(1/z)/A(Z)$, где $A(Z)$ — полином степени, меньшей либо равной N . Отметим, что спектр фильтра p_i в действительности является не зависящей от частоты постоянной и имеет вид

$$\bar{P}\left(\frac{1}{Z}\right) P(Z) = \frac{1/Z - 1/Z_0 Z - 1/\bar{Z}_0}{1 - 1/(Z\bar{Z}_0) Z - 1/Z_0}. \quad (2.6.2)$$

Умножим числитель и знаменатель левой дроби на Z . Тогда

$$\bar{P}\left(\frac{1}{Z}\right) P(Z) = \frac{1 - Z/Z_0}{Z - 1/\bar{Z}_0} \frac{Z - 1/\bar{Z}_0}{1 - Z/Z_0} = 1. \quad (2.6.3)$$

Легко показать, что равенство $\bar{P}(1/Z)P(Z)=1$ характерно для общего случая фильтра $P(Z)=Z^N \bar{A}(1/Z)/A(Z)$. Если Z_0 выбрано вне единичного круга, то знаменатель выражения (2.6.1) можно разложить по положительным степеням Z . Этот ряд будет сходиться на единичном круге. Это означает, что физическая реализуемость эквивалентна расположению Z_0 вне единичного круга. Положив числитель выражения $P(Z)$ равным нулю, видим, что нуль $Z=1/\bar{Z}_0$ находится внутри единичного круга. Эта ситуация показана на рис. 2.17. Чтобы показать, что полюс и нуль располагаются при этом на одном радиусе, выразим Z_0 в полярных координатах, т. е. в виде $Z_0 = r_0 e^{i\varphi_0}$.

Из раздела 2.2 (о минимально фазовых функциях) следует, что числитель в P не является минимально фазовым и его фаза увеличивается на 2π при изменении ω от 0 до 2π . Таким образом, средняя групповая задержка $d\varphi/d\omega$ положительна. Фактически положительной оказывается не только средняя задержка, но и групповая на всех частотах. Чтобы понять это заметим, что

$$Z = e^{i\omega}; \quad \frac{dZ}{d\omega} = ie^{i\omega} = iZ; \quad (2.6.4)$$

$$\frac{d}{d\omega} = \frac{dZ}{d\omega} \frac{d}{dZ} = iZ \frac{d}{dZ}.$$

Фазу рассматриваемого фазового фильтра (или любого комплексного числа) можно записать в виде

$$\varphi = \text{Im} \ln P(Z). \quad (2.6.5)$$

Так как $|P|=1$, то действительная часть логарифма исчезает, и, следовательно, для фазового фильтра (только) можно преобразовать выражение (2.6.5) к виду

$$\varphi = \frac{1}{i} \ln P(Z) = \frac{1}{i} \ln \frac{Z - 1/\bar{Z}_0}{1 - Z/Z_0} = \frac{1}{i} \left[\ln \left(Z - \frac{1}{Z_0} \right) - \ln \left(1 - \frac{Z}{Z_0} \right) \right]. \quad (2.6.6)$$

Используя формулу (2.6.4), найдем групповую задержку:

$$\begin{aligned} \tau_g &= \frac{d\varphi}{d\omega} = iZ \frac{d\varphi}{dZ} = Z \left(\frac{1}{Z - 1/\bar{Z}_0} + \frac{1/Z_0}{1 - Z/Z_0} \right) = \frac{1}{1 - 1/\bar{Z}_0 Z} + \frac{Z/Z_0}{1 - Z/Z_0} = \\ &= \frac{1 - Z/Z_0 + (1 - 1/\bar{Z}_0 Z) (Z/Z_0)}{(1 - 1/\bar{Z}_0 Z) (1 - Z/Z_0)} = \frac{1 - 1/Z_0 \bar{Z}_0}{(1 - 1/\bar{Z}_0 Z) (1 - Z/Z_0)}. \end{aligned} \quad (2.6.7)$$

Числитель выражения (2.6.7) является положительным действительным числом (так как $|Z_0| > 1$), а знаменатель имеет форму $\bar{A}(1/Z)A(Z)$, т. е. является спектром, и также положителен. Таким образом, мы показали, что групповая задержка этого физически реализуемого фазового фильтра всегда положительна.

Если теперь последовательно с некоторым фильтром включить фазовый фильтр, то сумма фаз и групповая задержка составного фильтра, естественно, должны быть больше групповой задержки исходного фильтра. По тем же соображениям минимально фазовый фильтр должен иметь меньшую групповую задержку, чем любой другой фильтр с таким же спектром.

В заключение отметим, что фазовый фильтр с одним полюсом и одним нулем пропускает все частотные компоненты с постоянным коэффициентом усиления и фазовым сдвигом, который может регулироваться положением полюса. Выбор Z_0 вблизи единичной

окружности приводит к тому, что наибольшие фазовые сдвиги концентрируются около частоты, на которой располагается полюс. Расположение полюса на большом расстоянии от единичного круга вызывает увеличение диапазона задержек для более высоких частот. Большие фазовые сдвиги или групповые задержки можно реализовать путем последовательного включения нескольких однополюсных фильтров.

Упражнения

1. Временная функция $p_t = \left(\frac{1}{2}, -\frac{3}{4}, -\frac{3}{8}, -\frac{3}{16}, \dots \right)$ является примером фазового фильтра. Рассчитайте несколько значений ее автокорреляционной функции (для разных задержек) путем суммирования нескольких бесконечных рядов.
2. Нарисуйте графики амплитуды, фазы и групповой задержки фазового фильтра $(1 - \bar{Z}_0 Z) / (Z_0 - Z)$, где $Z_0 = (1 + \epsilon) e^{i\omega_0}$ и ϵ — мало. Обозначьте на кривой наиболее важные параметры.
3. Покажите, что коэффициенты фазового фильтра, образованного последовательным включением фильтров $(1 - \bar{Z}_0 Z) / (Z_0 - Z)$ и $(1 - Z_0 \bar{Z}) / (\bar{Z}_0 - Z)$, являются действительными.
4. Некоторая непрерывная функция времени является непрерывной во времени импульсной реакцией фазового фильтра. Опишите эту функцию во временной и в частотной областях. Поменяйте местами слова «время» и «частота» в описании этой функции. Каков физический смысл такой функции? Что произойдет с утверждением: «групповая задержка фазового фильтра положительна».
5. График групповой задержки $\tau_g(\omega)$ для уравнения (2.6.7) свидетельствует о том, что величина τ_g положительна для всех частот. Какую область занимает τ_g в интервале $0 < \omega < 2\pi$. [Указание. Этот каверзный вопрос вы должны решить в уме.]

2.7. РЕЖЕКТОРНЫЙ ФИЛЬТР И ФИЛЬТР С ПИКОМ НА ПЬЕДЕСТАЛЕ

В некоторых случаях требуется подавить очень узкую полосу частот, оставляя остальную часть спектра практически неизменной. Это необходимо, например, для исключения наводок в 50 Гц от сети переменного тока. Такой фильтр называется режекторным (заграждающим) и может быть легко реализован путем незначительного изменения фазового фильтра. В фазовом фильтре полюс и нуль имеют равное (логарифмически) относительное расстояние от единичного круга. Нам необходимо поместить нуль ближе к кругу. Практически нет причин, мешающих поместить нуль прямо на окружности. В этом случае частота, на которой располагается нуль, в спектре входных данных подавляется. Если нет необходимости полностью подавить нежелательные частоты, то нуль можно поместить в непосредственной близости от окружности (внутри или вне ее). По мере удаления нуля от окружности впадина в спектре режекторного фильтра становится менее глубокой до тех пор, пока нуль не окажется дальше от круга, чем полюс. Тогда впадина становится выбросом (пиком): получившийся таким

образом фильтр будем называть *фильтром с пиком на пьедестале*. Он во многом похож на узкополосный фильтр. Характеристики некоторых из этих фильтров приведены на рис. 2.18 и 2.19. Различие между фильтром с пиком на пьедестале и узкополосным фильтром заключается в асимптотическом поведении вдали от частоты ω_0 .

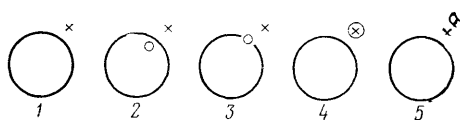


РИС. 2.18. Положение полюсов и нулей для некоторых простых фильтров. Окружности являются единичными окружностями на Z -плоскости.

Фильтры: 1 — узкополосный, 2 — фазовый, 3 — режекторный, 4 — постоянный, 5 — с пиком на пьедестале

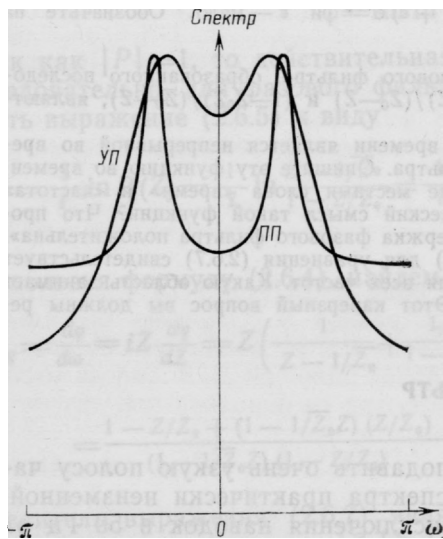


Рис. 2.19. Зависимости амплитуды от частоты для действительных во временной области узкополосного фильтра (УП) и фильтра с пиком на пьедестале (ПП). Каждый фильтр имеет один полюс в точке $Z_0 = 1,2e^{i\pi/3}$. Второй полюс в точке $Z_0 = 1,2e^{-i\pi/3}$

Первый из них имеет в этой области плоскую частотную характеристику, а второй — спадающую с увеличением модуля разности $(\omega - \omega_0)$. Это делает фильтр с пиком на пьедестале более удобным для создания фильтров со сложной формой спектров путем последовательного включения однополюсных фильтров.

Узкополосные фильтры и фильтры с крутым срезом должны применяться с осторожностью. Возможное осложнение заключается в том, что выходные сигналы этих фильтров медленно затухают во времени. Если для некоторых случаев это не имеет значения, то для других — это важно. Очевидно, если продолжительность сбора данных меньше или сравнима с длительностью импульсной реакции узкополосного фильтра, то переходные эффекты в начале эксперимента не будут успевать затухать. Точно так же впадина характеристики фильтра, устраняющего колебания с частотой 50 Гц, не должна быть слишком узкой. Сигналы на выходе полосового фильтра (который легче реализовать с помощью быстрого преобразования Фурье, чем с помощью нескольких полюсов) также имеют определенную скорость затухания во времени, которая

в некоторых экспериментах может быть слишком малой. В радиолокации и сейсмологии методом отраженных волн ценность (значимость) сигнала не связана с его силой. Отраженные сигналы, приходящие с большим временем запаздывания, могут быть очень

слабыми, но они содержат информацию, которая отсутствует в более ранних отражениях. При использовании фильтра со слишком крутым срезом частотной характеристики переходный процесс фильтра, вызванный ранее прошедшими сильными волнами, может продолжаться достаточно долго и занять время, в течение которого приходят более поздние слабые отраженные сигналы.

Упражнения

1. Рассмотрите симметричный (нереализуемый) фильтр, который пропускает все частоты меньше ω_0 с единичным коэффициентом усиления. Частоты, превышающие ω_0 , полностью затухают. Чему равна скорость затухания амплитуды во времени для этого фильтра?

2. Энергия волн, распространяющихся от точечного источника, затухает обратно пропорционально квадрату расстояния (сферическое расхождение). Амплитуда спадает как корень квадратный из энергии. Это подразумевает определенную скорость затухания энергии во времени. Скорость затухания во времени такая же, как и при отражении волн от плоских границ раздела. Какой степени времени t пропорционально затухание амплитуды сигналов? Энергия волн, рассеиваемых точечными отражателями в обратную сторону, т. е. в сторону источника, затухает обратно пропорционально расстоянию в четвертой степени. Чему равно затухание во времени?

3. Обсудите возможность использования фильтра из упражнения 1 для данных из упражнения 2.

4. Сконструируйте однополюсный режекторный фильтр с одним нулем для подавления частот в интервале 49—51 Гц данных, квантованных в 500 точках/с.

2.8. БИЛИНЕЙНОЕ ПРЕОБРАЗОВАНИЕ

Z -преобразование и преобразование Фурье связаны соотношениями $Z = e^{i\omega}$ и $i\omega = \ln Z$. Особенность этих соотношений состоит в том, что простые отношения полиномов степеней Z не переводятся в отношении полиномов степеней ω и наоборот. Приближенное выражение

$$-i\hat{\omega} = 2 \frac{1-Z}{1+Z} \quad (2.8.1)$$

легко разрешается относительно Z

$$Z = \frac{1 + i\hat{\omega}/2}{1 - i\hat{\omega}/2}. \quad (2.8.2)$$

Выражения (2.8.1) и (2.8.2) часто оказываются полезными. Они являются усечениями точных разложений в степенные ряды

$$-i\omega = -\ln e^{i\omega} = -\ln Z = 2 \left[\frac{1-Z}{1+Z} + \frac{1}{3} \frac{(1-Z)^3}{(1+Z)^3} + \frac{1}{5} \dots \right]; \quad (2.8.3)$$

$$Z = e^{i\omega} = \frac{e^{i\omega/2}}{e^{-i\omega/2}} = \frac{1 + i\omega/2 + (i\omega/2)^2/2! + \dots}{1 - i\omega/2 + (i\omega/2)^2/2! - \dots}. \quad (2.8.4)$$

Для того чтобы Z -преобразование функции $B(Z)$ было минимально фазовым, любой из корней Z_0 уравнения $B(Z)=0$ должен находиться вне единичного круга. Так как

$$Z_0 = \exp \{i [\operatorname{Re}(\omega_0) + i \operatorname{Im}(\omega_0)]\}, \text{ а } |Z_0| = e^{-\operatorname{Im}(\omega_0)},$$

то для получения минимально фазовой функции мнимая часть $\operatorname{Im}(\omega_0)$ должна быть отрицательной. Другими словами, ω_0 должна располагаться в нижней полуплоскости. Следовательно, зависимость $Z = e^{i\omega}$ отображает внешнюю часть единичного круга на нижнюю полуплоскость. Рассматривая рис. 2.20 и 2.21, видим, что

Z	$\omega = 2\pi n - i \ln Z$	$\hat{\omega} = 2i \frac{1-Z}{1+Z}$
A 1	$2\pi n + 0$	0
B i	$2\pi n + \pi/2$	2
C -1	$2\pi n + \pi$	$\pm \infty$
D $-i$	$2\pi n - \pi/2$	-2
E $\frac{1}{2}$	$2\pi n + 0,693i$	$i\frac{2}{3}$
F 2	$2\pi n - 0,693i$	$-i\frac{2}{3}$

РИС. 2.20. Некоторые типичные точки на Z -, ω - и $\hat{\omega}$ -плоскостях

билинейное приближение (2.8.1) или (2.8.2) также отображает внешнюю часть единичного круга на нижнюю полуплоскость.

Таким образом, хотя билинейная аппроксимация является некоторым приближением, она обеспечивает точное сохранение минимально фазовых свойств функции. Это очень важно, поскольку преобразование устойчивого дифференциального уравнения в разностное уравнение с помощью выражения (2.8.1) приводит к устойчивому разностному уравнению. Можно привести много примеров, когда аппроксимация производной по времени умножением на $1-Z$ преобразует устойчивое дифференциальное уравнение в неустойчивое разностное уравнение.

Соотношение (2.8.1) удобно запомнить, учитывая, что $-i\omega$ соответствует дифференцированию по времени преобразования Фурье, а $1-Z$ является оператором первой разности, величина $1+Z$ в знаменателе «центрирует» выражение относительно $Z^{1/2}$.

Чтобы показать, что билинейная аппроксимация является низкочастотным приближением, умножим числитель и знаменатель выражения (2.8.1) на $Z^{-1/2}$:

$$-i\hat{\omega} = 2 \frac{Z^{-1/2} - Z^{1/2}}{Z^{-1/2} + Z^{1/2}} = -2i \frac{\sin \omega/2}{\cos \omega/2}; \quad \hat{\omega} = 2 \operatorname{tg} \omega/2. \quad (2.8.5)$$

Уравнение (2.8.5) подразумевает скорость квантования с частотой в одну выборку на 1 с. При выборе произвольной частоты кванто-

вания Δt приближение (2.8.5) приобретает вид

$$\hat{\omega}\Delta t \approx 2 \operatorname{tg} \omega\Delta t/2. \quad (2.8.6)$$

График этого приближения показан на рис. 2.22. Очевидно, погрешность можно сделать сколь угодно малой путем достаточного

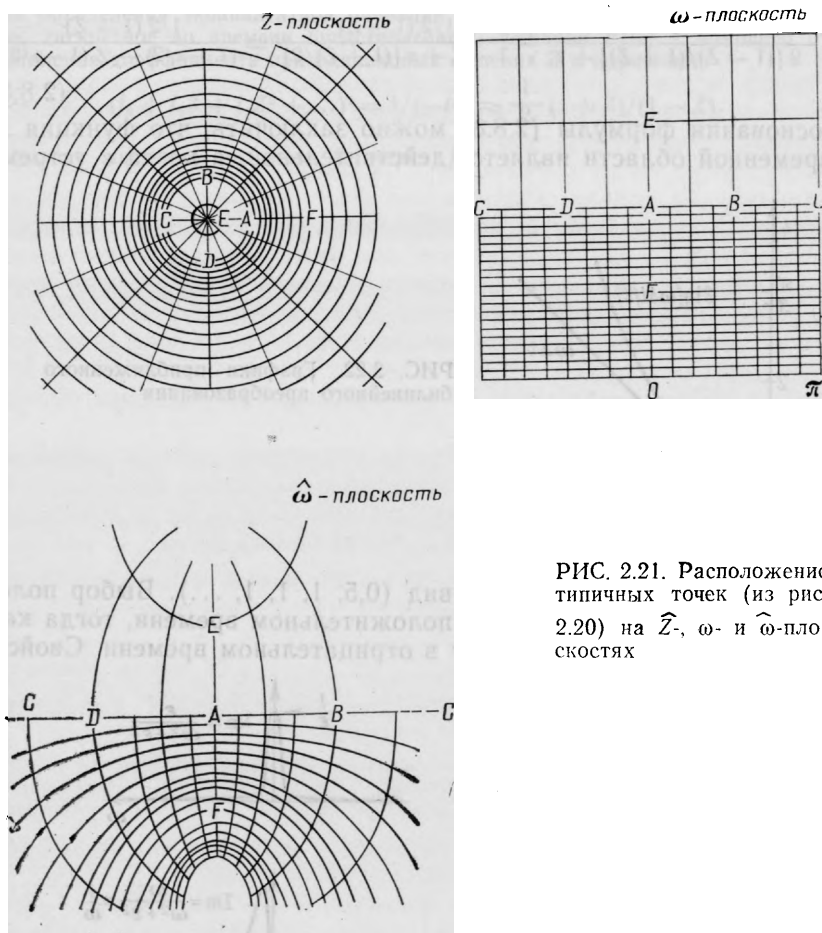


РИС. 2.21. Расположение типичных точек (из рис. 2.20) на \hat{Z} -, ω - и $\hat{\omega}$ -плоскостях

увеличения частоты квантования, т. е. выбирая Δt достаточно малым. Из рис. 2.22 видно, что погрешность будет составлять лишь несколько процентов, если выбрать Δt достаточно малым так, чтобы $\omega_{\max}\Delta t \leq 1$. Читатели, хорошо знакомые с теоремой свертки, вспомнят, что она дает менее строгое ограничение $\omega_{\max}\Delta t < \pi$. Очевидно, что теорема свертки слишком слаба для случаев, включающих билинейное преобразование.

Теперь рассмотрим в качестве примера функцию с полюсом $1/(-i\omega)$ на нулевой частоте. Это соответствует интегрированию. По причинам, которые вскоре станут ясными, рассмотрим функцию (с незначительно отличающимся от ранее указанного полюсом)

$$P = \frac{1}{-i\omega + \epsilon}, \quad (2.8.7)$$

где ϵ — малая величина. Применяя билинейное преобразование, получим:

$$P = \frac{1}{2[(1-Z)/(1+Z)] + \epsilon} = \frac{1/2(1+Z)}{1-Z + \epsilon[(1+Z)/2]} = \frac{1/2(1+Z)}{(1+\epsilon/2) - Z(1-\epsilon/2)}. \quad (2.8.8)$$

На основании формулы (2.8.8) можно заключить, что функция P во временной области является действительной, а когда ϵ устрем-

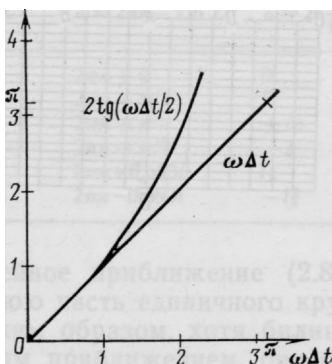


РИС. 2.22. Графики приближенного билинейного преобразования

ляется к нулю, она принимает вид $(0,5, 1, 1, 1, \dots)$. Выбор положительного ϵ дает ступеньку в положительном времени, тогда как отрицательное ϵ дает ступеньку в отрицательном времени. Свойст-

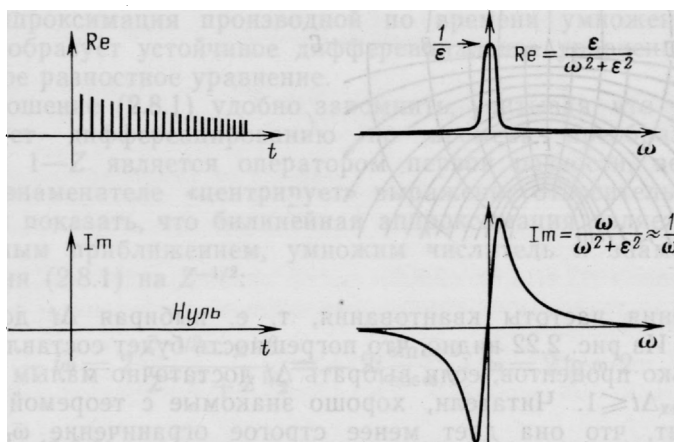


РИС. 2.23. Свойства оператора интегрирования

ва этой функции показаны на рис. 2.23. Интересно отметить, что если временную и частотную области всюду поменять местами, то мы получим интегрирующий (квадратурный) фильтр, характеристики которого показаны на рис. 1.17.

Упражнение

В решении диффузионных задач часто возникает множитель $F(\omega) = 1/(-i\omega)^{1/2}$. Для определения эквивалентной операции свертки найдите физически реализуемое, дискретное во времени представление f_t функции $F(\omega)$ с помощью отождествления коэффициента при одинаковых степенях Z в выражении

$$(f_0 + f_1 Z + f_2 Z^2 + \dots)^2 = 1/(-i\omega) \approx \frac{1}{2} (1 + Z)/(1 - Z).$$

Выразите численно значение f_0 через f_1 .

3. ФАКТОРИЗАЦИЯ СПЕКТРА

Как мы увидим позже, существует бесконечное число временных функций, соответствующих любому заданному спектру¹. Факторизацией спектра будем называть методы нахождения единственной функции времени, которая обладает заданным спектром и в то же время является минимально фазовой. Минимально фазовые функции находят широкое применение. Только такие функции могут быть использованы для фильтрации с обратной связью. Эти функции будут часто появляться в задачах распространения волн в последующих главах. Минимально фазовые функции возникают в теории предсказания и регулирования по заданному спектру. В дальнейшем будет показано, что для минимально фазовой функции характерна насколько возможно плотная концентрация энергии около момента $t=0$. Эта функция определяет наименьшее значение дисперсии при распространении волн в среде с поглощением. Она находит применение также в теории двухмерного потенциала, когда по наблюдениям модуля вектора поля необходимо определить его компоненты.

В главе рассматриваются четыре различных метода определения минимально фазового сейсмического импульса по заданному спектру. Эти методы представляют разные возможности понимания сущности факторизации спектра и минимальной фазы.

3.1. МЕТОД КОРНЕЙ

Функция времени (2, 1) имеет такой же спектр, как и функция времени (1,2). Их автокорреляционная функция имеет вид (2, 5.2). Это обстоятельство используется для исследования множества всех функций времени с такими автокорреляционными функциями и спектрами. Можно считать, что любая действительная функция, обращенная во времени, будет иметь такую же автокорреляционную функцию, что и первоначальная функция. На самом деле, в некоторых областях могут встречаться и комплексные временные ряды; поэтому необходимо сформулировать более точное утверждение, что любой сейсмический импульс и его комплексно-сопряженное обращение во времени обладают одинаковыми автокорреляционными функциями и спектрами. Проверим это для простых двухточечных функций времени. Спектр функции (b_0, b_1) имеет вид

$$\bar{B}\left(\frac{1}{Z}\right)B(Z) = \left(\bar{b}_0 + \frac{\bar{b}_1}{Z}\right)(b_0 + b_1 Z) = \frac{\bar{b}_1 b_0}{Z} + (\bar{b}_0 b_0 + \bar{b}_1 b_1) + b_0 \bar{b}_1 Z. \quad (3.1.1)$$

¹ Амплитудному спектру. (Прим. перев.)

Сопряженно-обращенная временная функция (\bar{b}_1, b_0) с Z -преобразованием $b_z(Z) = \bar{b}_1 + \bar{b}_0 Z$ имеет спектр

$$\bar{B}_r \left(\frac{1}{Z} \right) B_r(Z) = \left(b_1 + \frac{b_0}{Z} \right) (\bar{b}_1 + \bar{b}_0 Z) = \frac{b_0 b_1}{Z} + (b_0 \bar{b}_0 + b_1 \bar{b}_1) + b_1 \bar{b}_0 Z. \quad (3.1.2)$$

Как видно, спектры (3.1.1) и (3.1.2) действительно являются тождественными. Теперь желательно распространить это свойство на функции времени с тремя и более точками. Можно осуществить обобщение трехточечных функций, например, вида $B(Z) = b_0 + b_1 Z + b_2 Z^2$. Применяв основную теорему алгебры (которая утверждает, что многочлен степени n имеет ровно n корней), разложим $B(Z)$ на множители:

$$B(Z) = b_2 (Z_1 - Z) (Z_2 - Z). \quad (3.1.3)$$

Спектр этого выражения имеет вид

$$R(Z) = \bar{B} \left(\frac{1}{Z} \right) B(Z) = \bar{b}_2 b_2 \left(\bar{Z}_1 - \frac{1}{Z} \right) (Z_1 - Z) \left(\bar{Z}_2 - \frac{1}{Z} \right) (Z_2 - Z). \quad (3.1.4)$$

Очевидно, что b_2 можно умножить на любое комплексное число с единичным модулем. При этом сейсмический сигнал (3.1.3) изменится, а его спектр (3.1.4) останется прежним. Остальные члены $(Z_i - Z)$ в формуле (3.1.4) можно представить как произведение сомножителей в форме $(\bar{Z}_i - 1/Z) (Z_i - Z)$. Но этот множитель похож на множитель (3.1.1). Функцией времени для $(Z_i - Z)$ является $(Z_i - 1)$, а ее комплексно-сопряженная обращенная функция времени имеет вид $(-1, \bar{Z}_i)$. Таким образом, любой множитель $(Z_i - Z)$ в формуле (3.1.3) можно заменить множителем $(-1 + \bar{Z}_i Z)$. В общем случае выражение (3.1.3) может иметь N множителей $[(Z_i - Z), i=1, 2, \dots, N]$. Любая их комбинация может быть обращена. Следовательно, существует 2^N различных сейсмических сигналов, которые могут формироваться путем перестановок, и все эти сигналы имеют один и тот же спектр. Рассмотрим единичный круг на комплексной плоскости. Сомножитель $(Z_i - Z)$ означает, что Z_i является корнем как функции $B(Z)$, так и ее спектра $R(Z)$. Если заменить $(Z_i - Z)$ на $(-1 + \bar{Z}_i Z)$ в функции $B(Z)$, то тем самым корень Z_i из функции $B(Z)$ устраняется и заменяется $Z = 1/\bar{Z}_i$. Корни спектра $R(Z)$ при этом не меняются, потому что величины Z_i и $1/\bar{Z}_i$ являются первоначальными корнями спектра $R(Z)$ и при замене просто переходят друг в друга. Таким образом, если функция $B(Z)$ имеет корни Z_i в комплексной плоскости, которые располагаются где угодно, то спектр $R(Z)$ должен иметь все корни Z_i и, кроме того, корни $1/\bar{Z}_i$. Замена любого отдельного корня Z_i на $1/\bar{Z}_i$ изменяет функцию $B(Z)$, но не оказывает влияния

на спектр $R(Z)$. Операцию замены корня Z_i на корень $1/\bar{Z}_i$ можно записать в виде

$$B'(Z) = \frac{Z - 1/\bar{Z}_i}{1 - Z/\bar{Z}_i} B(Z). \quad (3.1.5)$$

Множитель перед $B(Z)$ является фазовым фильтром, рассмотренным выше. Принимая это во внимание, легко установить, что функция $B'(Z)$ имеет такой же спектр, как и функция $B(Z)$. В самом деле, нет никаких реальных причин, мешающих Z_i быть корнем функции $B(Z)$. Если Z_i является корнем $B(Z)$, то $B'(Z)$ будет многочленом; в противном случае это будет бесконечный ряд.

Теперь определим функцию $B(Z)$ по заданному ее спектру $R(Z)$. Отметим, что корни спектра $R(Z)$ являются по определению

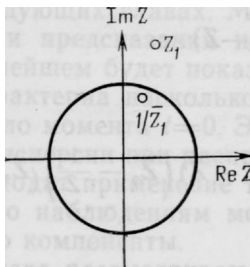


РИС. 3.1. Корни функции $\bar{B}(1/Z)B(Z)$

решениями уравнения $R(Z)=0$. Если умножить $R(Z)$ на Z^N [где $R(Z)$ может иметь заданную степень вплоть до N], то $Z^N R(Z)$ является полиномом и решения Z_i уравнения $Z^N R(Z)=0$ будут такими же, как и решения уравнения $R(Z)=0$. Нахождение всех корней этого полинома является хотя и трудной, но стандартной процедурой. Предполагая ее выполненной, можно затем проверить, являются ли все корни Z_i и $1/\bar{Z}_i$ парными. Если такие пары не охватывают все корни, то функция $R(Z)$ не является реальным спектром. В случае парных корней каждому нулю внутри единичного круга должен соответствовать нуль вне его (рис. 3.1). Таким образом, чтобы сделать функцию $B(Z)$ минимально фазовым сейсмическим сигналом со спектром $R(Z)$, следует собрать все корни, лежащие вне единичного круга. Тогда функция $B(Z)$ приобретает вид

$$B(Z) = b_N (Z - Z_1) (Z - Z_2) \dots (Z - Z_N). \quad (3.1.6)$$

Этим завершается нахождение минимально фазового сейсмического сигнала по заданному спектру. Когда N велико, то это неудобно для вычислений и является недостатком по сравнению с другими методами, которые пока не обсуждались. Значение метода корней состоит в том, что он демонстрирует некоторые основные принципы.

1. Каждому спектру соответствует минимально фазовый сейсмический сигнал, который является единственным с погрешностью до комплексного масштабного коэффициента с единичным модулем

2. Существует бесконечное множество функций времени, соответствующих любому заданному спектру.

3. Не все функции являются потенциальными автокорреляционными функциями.

Метод корней в факторизации спектров был развит, по-видимому, экономистами в 1920-х и 1930-х гг. Ряд ссылок на эти ранние работы можно найти во всемирно известной книге Г. Уолда «Стационарные временные ряды» [10].

Упражнения

1. Как можно найти масштабный множитель b_N в выражении (3.1.6)?

2. Вычислите автокорреляционные функции для каждого из следующих четырех сейсмических импульсов $(4, 0, -1)$, $(2, 3, -2)$, $(-2, 3, 2)$, $(1, 0, -4)$.

3. Наблюденный энергетический спектр соответствует формуле $P(\omega) = 38 + 10 \cos \omega - 12 \cos 2\omega$. Какие сейсмические импульсы отвечают этому спектру? Какой из них является минимально фазовым? [Указание. $\cos 2\omega = 2 \cos^2 \omega - 1$; $2 \cos \omega = Z + 1/Z$; используйте квадратичную формулу.]

4. Покажите, что если сейсмический импульс $b_t = (b_0, b_1, \dots, b_n)$ является действительным, то корни его спектра R составляют четверки (квадруполь) $Z_0, 1/Z_0, \bar{Z}_0$ и $1/\bar{Z}_0$. Рассмотрите случай корней, расположенных точно на единичном круге и на действительной оси. Каково наименьшее число таких корней?

3.2. ТЕОРЕМА РОБИНСОНА О ЗАДЕРЖКЕ ЭНЕРГИИ [11]

Покажем, что минимально фазовый сейсмический сигнал имеет меньшую задержку энергии, чем любой другой, односторонний сейсмический сигнал с таким же спектром. Это означает, что энергия,

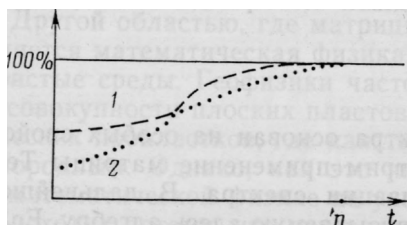


РИС. 3.2. Полная энергия сигнала в фильтре в интервале времени от 0 до t .

1 — все нули вне окружности; 2 — некоторые нули внутри окружности

суммируемая от нуля до любого момента времени, для минимально фазового сейсмического сигнала больше или равна энергии в том же интервале любого другого сейсмического сигнала с таким же спектром (рис. 3.2).

Сравним два сейсмических сигнала $P_{\text{вн}}$ и $P_{\text{вне}}$, которые идентичны во всем, за исключением одного нуля, который лежит вне ($P_{\text{вне}}$) и внутри ($P_{\text{вн}}$) единичного круга. Это можно записать как

$$P_{\text{вне}}(Z) = (b + sZ)P(Z);$$

$$P_{\text{вн}}(Z) = (s + bZ)P(Z),$$

где b больше s , а P является произвольной функцией степени n .

Теперь составим таблицу членов исследуемого выражения (табл. 1). Разность, которая приводится в правой графе таблицы, очевидно, всегда положительна.

Т а б л и ц а 1

t	$P_{\text{вне}}$	$P_{\text{вн}}$	$P_{\text{вне}}^2 - P_{\text{вн}}^2$	$\sum_{k=0}^t (P_{\text{вне}}^2 - P_{\text{вн}}^2)$
0	bp_0	sp_0	$(b^2 - s^2) p_0^2$	$(b^2 - s^2) p_0^2$
1	$bp_1 + sp_0$	$sp_1 + bp_0$	$(b^2 - s^2) (p_1^2 - p_0^2)$	$(b^2 - s^2) p_1^2$
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
k	$bp_k + sp_{k-1}$	$sp_k + bp_{k-1}$	$(b^2 - s^2) (p_k^2 - p_{k-1}^2)$	$(b^2 - s^2) p_k^2$
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
$n + 1$	sp_n	bp_n	$(b^2 - s^2) (-p_n^2)$	0

Чтобы доказать, что минимально фазовый сейсмический сигнал дает наименьшую задержку энергии, вычисления главного аргумента повторяются для каждого корня до тех пор, пока все они не выйдут за пределы единичного круга.

Упражнение

Докажите предшествующую теорему (о минимуме задержки энергии) для комплексных значений b , s и P . [Предупреждение. Что справедливо: $P_{\text{вн}}(s + bZ)P$, или $P_{\text{вн}} = (\bar{s} + \bar{b}Z)P$?

3.3. МЕТОД ТЕПЛИЦА

Метод Теплица факторизации спектра основан на особых свойствах матриц Теплица [12]. Рассмотрим применение матрицы Теплица для осуществления факторизации спектра. В дальнейшем будем многократно ссылаться на описываемую здесь алгебру. Если требуется предсказывать временные ряды (сигналы), то это можно осуществить с помощью так называемого *предсказывающего фильтра*, который находится как решение системы уравнений Теплица. Норман Левинсон в своем приложении к книге «*Временные ряды*» Норберта Винера впервые ознакомил инженеров с матрицей Теплица. До этого она была широко известна и применялась только в экономике. Появление ее сначала в экономике совершенно естественно, так как в ней рассматриваются данные, заданные в дискретных точках во времени, тогда как в инженерном деле мысль о дискретном времени была в некоторой степени искусственной до появления цифровых ЭВМ. Необходимость предсказания в экономике очевидна. В сейсмологии больший интерес представляет не предсказание волнового процесса само по себе, а ошибка

предсказания. При использовании отраженных волн рассматриваемая ситуация схожа на радиолокационную, где время задержки прямо пропорционально физическому расстоянию. Это справедливо для *однократно (первично)*-отраженных волн. Серьезное осложнение возникает на практике при разведке в мелких морях, где интенсивные звуковые волны распространяются вперед и назад, попеременно отражаясь от поверхности и дна моря. Эти волны называются *многократно*-отраженными¹. Разделение первичных и многократно-отраженных волн возможно на основе предсказания. Действительно, момент появления многократно-отраженной волны можно предсказать на основе ранее принятых отраженных сигналов. В то же время первичный отраженный сигнал таким способом непредсказуем. Таким образом, полезную информацию несет в себе та часть сейсмограммы, которая является *непредсказуемой*². ЭВМ, предназначенные для интерпретации данных сейсмической разведки, обычно решают за день около 100 000 систем совместных уравнений Теплица.

Другой важной областью применения алгебры, связанной с матрицами Теплица, является спектральный анализ с высоким разрешением, необходимый, например, для определения энергетического спектра сигнала на основании короткой выборки данных (во времени или в пространстве). Общепринятые статистические и инженерные познания по этому предмету основываются на допущениях, которые часто оказываются неприемлемыми геофизически. Эту ситуацию хорошо изучил Джон Бург, используя некоторые специфические свойства матриц Теплица для разработки своей процедуры максимально-энтропийного спектрального анализа, описываемой в одном из последующих разделов.

Другой областью, где матрицы Теплица играют основную роль, является математическая физика, с помощью которой описываются слоистые среды. Геофизики часто используют модель Земли в виде совокупности плоских пластов или в виде концентрических сферических пропластков, где каждый пропласток или пласт является однородным. Однако, как это ни удивительно, во многих книгах по математической физике не упоминаются матрицы Теплица. Это происходит потому, что они посвящены *прямым* задачам, т. е. расчетам волн (или потенциалов), наблюдаемых в средах с известным строением. В геофизике же нас интересуют не только *прямые*, но и *обратные* задачи, в которых волны наблюдаются на поверхности Земли, а мы хотим определить строение среды внутри Земли. В последующих разделах рассматривается решение обратной задачи с помощью матрицы Теплица.

Зададим некоторую функцию времени x_1 , которая может быть минимально фазовой или не быть ею. Определим ее спектр выра-

¹ Многократно-отраженные волны, возникающие в мелкой воде, обычно называют реверберационными волнами. (Прим. перев.)

² Следует отметить, что кратно-отраженные волны также дают некоторую полезную информацию, например, о поглощении и рассеянии волн в среде и т. п. (Прим. перев.)

женем $R(Z) = \bar{X}(1/Z)X(Z)$. Как было показано ранее, если задан только спектр $R(Z)$, то нет возможности определить, был ли он вычислен для минимально фазовой функции или для неминимально фазовой. Можно предположить, что существует минимально фазовая функция $B(Z)$, соответствующая данному спектру, т. е. $R(Z) = \bar{B}(1/Z)B(Z)$. Так как $B(Z)$ по предположению является минимально фазовой функцией, то она имеет некоторую обратную функцию $A(Z) = 1/B(Z)$, которую можно найти следующим образом:

$$R(Z) = \bar{B} \left(\frac{1}{Z} \right) B(Z) = \frac{\bar{B}(1/Z)}{A(Z)}; \quad (3.3.1)$$

$$R(Z)A(Z) = \bar{B} \left(\frac{1}{Z} \right) = \bar{b}_0 + \frac{\bar{b}_1}{Z} + \dots \quad (3.3.2)$$

Для определения функции $A(Z)$ необходимо приравнять коэффициенты при степенях Z . Например, для случая, когда функция $A(Z)$ является квадратичной ($a_0 + a_1Z + a_2Z^2$), коэффициент при Z^0 в выражении (3.3.2) имеет вид

$$r_0a_0 + r_{-1}a_1 + r_{-2}a_2 = \bar{b}_0. \quad (3.3.3a)$$

Коэффициент при Z^1 соответственно равен

$$r_1a_0 + r_0a_1 + r_{-1}a_2 = 0, \quad (3.3.3б)$$

и, наконец, коэффициент при Z^2 равен

$$r_2a_0 + r_1a_1 + r_0a_2 = 0. \quad (3.3.3в)$$

Объединяя их, получим совместные уравнения

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \bar{b}_0 \\ 0 \\ 0 \end{bmatrix}. \quad (3.3.4)$$

Из последнего выражения ясно, как обобщить его для системы совместных уравнений произвольной размерности. Главная диагональ матрицы содержит r_0 во всех позициях. Диагональ, расположенная под главной диагональю, всюду содержит r_1 . Вся матрица заполнена аналогично. Такая матрица называется матрицей Теллица. Пусть $a'_k = a_k/a_0$. Напомним, что согласно алгоритму деления многочленов $\bar{b}_0 = 1/\bar{a}_0$. Зададим положительное число $v = 1/a_0\bar{a}_0$. Теперь, разделив векторы с обеих сторон выражения (3.3.4) на a_0 , получим хорошо известную форму этих уравнений:

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a'_1 \\ a'_2 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix} \quad (3.3.5)$$

Это выражение дает три уравнения для трех неизвестных a'_1 , a'_2 и v . Для приведения выражения (3.3.5) к виду, удобному для при-

менения стандартных программ решения систем уравнений, нужно разделить векторы с обеих сторон последнего равенства на u . После решения уравнений получим коэффициент a_0 , учитывая, что он имеет модуль, равный $1/\sqrt{v}$, а его фаза произвольна, так же как и при факторизации спектров методом корней.

Здесь пессимист мог бы воскликнуть, что полином $A(Z) = a_0 + a_1Z + a_2Z^2$, найденный из решения системы совместных уравнений, не обязательно должен быть минимально фазовым, так что мы не могли бы непосредственно вычислить $B(Z)$ из зависимости $B(Z) = 1/A(Z)$. Он мог бы утверждать, что при решении, вероятно, возникнут трудности, особенно, если размерность системы (3.3.5) не была взята достаточно большой. В действительности же экспериментаторы давно знают, что пессимисты не правы. Доказать это можно достаточно легко совместно с описанием машинного алгоритма, который используется для решения системы уравнений (3.3.5).

Стандартные машинные алгоритмы для решения систем уравнений требуют затрат времени пропорционально величине n^3 и объема машинной памяти пропорционально величине n^2 . Машинный алгоритм Левинсона [13] для матриц Теплица требует затрат времени пропорционально величине n^2 и объема памяти пропорционально величине n . Заметим, что матрица Теплица содержит много одинаковых элементов. Левинсон использовал эту специфическую симметрию матриц Теплица для разработки своего «быстрого» метода (подобного БПФ).

Этот метод основан на приближении, называемом рекурсией. Сущность такого подхода заключается в том, что, задав решение системы уравнений размерности $k \times k$, мы указываем как найти решение системы размерности $(k+1) \times (k+1)$. Сначала необходимо получить решение для $k=1$; затем следует повторно (рекурсивно) применить систему формул, увеличивая k на единицу на каждом этапе. Покажем, как осуществляется рекурсия для функций действительного времени ($r_k = r_{-k}$), идя от системы уравнений размерности 3×3 к системе размерности 4×4 . Общий случай предлагаем для разработки читателю.

Зададим систему уравнений размерностью 3×3 и тогда ее решение

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix}. \quad (3.3.6)$$

Затем определим величину e при заданном r_3 (или r_3 , если задано e):

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (3.3.7)$$

Первые три строки в выражении (3.3.7) такие же, как и в системе (3.3.6); последняя строка представляет собой новое определение величины e . Рекурсия, по Левинсону, показывает, как получить решение a' для системы уравнений размерности 4×4 , которая похожа на систему (3.3.6), но имеет большую размерность:

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (3.3.8)$$

Важен следующий прием: на основании выражения (3.3.7) записывают «обращенную» систему уравнений (если действия над матрицами затруднительны), надо записать выражение (3.3.7) как систему уравнений, а затем проделать обращение степеней неизвестных и порядка уравнений:

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix} = \begin{bmatrix} e \\ 0 \\ 0 \\ v \end{bmatrix}. \quad (3.3.9)$$

Рекурсия по Левинсону заключается в вычитании еще неизвестной части c_3 выражения (3.3.9) из формулы (3.3.7) так, чтобы получить в результате системы (3.3.8). Таким образом,

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \left\{ \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix} - c_3 \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} v \\ 0 \\ 0 \\ e \end{bmatrix} - c_3 \begin{bmatrix} e \\ 0 \\ 0 \\ v \end{bmatrix} \right\}. \quad (3.3.10)$$

Чтобы сделать правую часть выражения (3.3.10) похожей на правую часть системы (3.3.8), необходимо убрать нижний элемент. Для этого достаточно положить $c_3 = e/v$. Это предполагает, что $v' = v - c_3 e = v - e^2/v = v[1 - (e/v)^2]$. Таким образом, решение системы размерности 4×4 получено из решения системы размерности 3×3 следующим путем:

$$e \leftarrow \sum_{i=0}^2 a_i r_{3-i}; \quad (3.3.11)$$

$$\begin{bmatrix} 1 \\ a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} \leftarrow \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix} - \frac{e}{v} \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix}; \quad (3.3.12)$$

$$v' \leftarrow v[1 - (e/v)^2]. \quad (3.3.13)$$

Мы показали, как получить решение системы уравнений Тейлица размерности 4×4 , зная решение системы уравнений Тейлица размерности 3×3 . Рекурсия по Левинсону предполагает выполне-

ние таких шагов, начиная с размерности 1×1 до размерности $n \times n$.

Рассмотрим еще раз вычисления, чтобы увидеть, почему функция $A(Z)$ оказывается минимально фазовой. Для начала отметим, что величины $v=1/a_0\bar{a}_0$ и $v'=1/\bar{a}'_0a'_0$ всегда положительны. Из формулы (3.3.13) следует, что $-1 < e/v < +1$. (То, что величина $c=e/v$ ограничена единицей, будет показано позже в связи с тем, что коэффициенты отражения волн также ограничены). Далее выражение (3.3.12) можно записать в виде полинома

$$A'(Z) = A(Z) - (e/v) Z^3 A(1/\bar{z}). \quad (3.3.14)$$

Известно, что величина Z^3 имеет единичный модуль на единичной окружности. Аналогично (для действительных временных рядов) спектр функции $A(Z)$ равен спектру функции $A(1/Z)$. Тогда (следуя теореме аддитивного сложения помех с минимально фазовым сейсмическим сигналом), если функция $A(Z)$ является минимально фазовой, то и функция $A'(Z)$ также будет минимально фазовой. Следовательно, следующие три утверждения эквивалентны.

1. $R(Z)$ имеет форму $\bar{X} \left(\frac{1}{2} \right) X(Z)$.

2. $|c_k| < 1$.

3. Функция $A(Z)$ является минимально фазовой.

Если хотя бы одно из этих утверждений несправедливо, то все они несправедливы. Программа для вычисления величин a_k и c_k

РИС. 3.3. Машинная программа, выполняющая рекурсию, по Левинсону. Предполагается, что входное значение r нормализовано делением на r_0 . Комплексная арифметика необязательна

```

COMPLEX R,A,C,E,BOT,CONJG
C(1)=-1.; R(1)=1.; A(1)=1.; V(1)=1.
200 DO 220 J=2,N
A(J)=0.
E=0.
DO 210 I=2,J
210 E=E+R(I)*A(J-I+1)
C(J)=E/V(J-1)
V(J)=V(J-1)-E*CONJG(C(J))
JH=(J+1)/2
DO 220 I=1,JH
BOT=A(J-I+1)-C(J)*CONJG(A(I))
A(I)=A(I)-C(J)*CONJG(A(J-I+1))
220 A(J-I+1)=BOT

```

по r_k приведена на рис. 3.3. В главе 8, посвященной вопросам распространения волн в слоистых средах, приводятся программы для вычисления r_k по заданным a_k или c_k .

Упражнения

1. Пусть верхняя строка в системе совместных уравнений Тейлица размерности 4×4 , подобной строке (3.3.8), имеет вид

$\left(1, -\frac{1}{4}, -\frac{1}{16}, -\frac{1}{4} \right)$. Как вычислить a_k ?

2. Как изменить способ Левинсона для случая комплексных временных функций? В частности, где встречаются комплексно-сопряженные величины в формулах (3.3.11), (3.3.12) и (3.3.13)?

3. Пусть величина $A_m(Z)$ обозначает некоторый полином, коэффициенты которого являются решениями системы уравнений Теллица размерности $m \times m$. Покажите, что если $B_h(Z) = Z^h A_h(Z^{-1})$, то соотношение

$$v_n \delta_{nm} = \frac{1}{2\pi} \int_0^{2\pi} R(Z) B_m(Z) Z^{-n} d\omega \quad \text{при } n \leq m$$

означает ортогональность полиномов $B_m(Z)$ и Z^n на единичной окружности при положительно взвешенной функции R . Используя этот результат, объясните, почему полиномы B_m и B_n ортогональны, т. е.

$$v_n \delta_{n:n} = \frac{1}{2\pi} \int_0^{2\pi} R(Z) B_m(Z) \overline{B_n} \left(\frac{1}{Z} \right) d\omega.$$

[Указание. Сначала рассмотрите случай $n \leq m$, а затем все n]. Матрицы Теллица рассматриваются в математической литературе на тему об ортогональности полиномов на единичной окружности. Автор особенно рекомендует по этому вопросу книгу Аткинсона [14].

3.4. ЭКСПОНЕНЦИАЛЬНО-ЛОГАРИФМИЧЕСКИЙ МЕТОД ВИТЛЯ [15]

В основе метода факторизации спектра лежит замена степенных рядов одного вида на ряды другого вида. Как и в методе корней, это хорошо для изучения, но неудобно для вычислений. Начнем с некоторой заданной автокорреляционной функции r_t , где

$$R(Z) = \dots + r_{-1}Z^{-1} + r_0 + r_1Z + r_2Z^2 + \dots$$

Если $|R| > 2$ на единичной окружности, то необходимо разделить функцию на этот масштабный множитель. Подставим этот степенной ряд в степенной ряд для логарифма

$$\begin{aligned} U(Z) = \ln R(Z) &= (R-1) - \frac{(R-1)^2}{2} + \frac{(R-1)^3}{3} - \dots = \\ &= \dots + u_{-1}Z^{-1} + u_0 + u_1Z + u_2Z^2 + \dots \quad (\text{при } 0 < R \leq 2) \end{aligned}$$

На практике это требует много усилий, но может быть проделано автоматически на ЭВМ. Теперь введем временной ряд u^+_t , полученный путем исключения отрицательных степеней Z из $U(Z)$:

$$U^+(Z) = \frac{u_0}{2} + u_1Z + u_2Z^2 + \dots$$

Подставим его в степенной ряд для экспоненты:

$$B(Z) = e^{U^+(Z)} = 1 + U^+ + \frac{(U^+)^2}{2!} + \frac{(U^+)^3}{3!} + \dots$$

Функция $B(Z)$ является требуемым минимально фазовым сейсмическим сигналом со спектром $R(Z)$. Для наглядности рассмотрим

следующие тождества:

$$\begin{aligned}
 R(Z) &= \exp [\ln R(Z)] = \exp \left(\frac{u_0}{2} + \sum_{-\infty}^{-1} u_k Z^k + \frac{u_0}{2} + \sum_{+1}^{\infty} u_k Z^k \right) = \\
 &= \exp \left(\frac{u_0}{2} + \sum_{-\infty}^{-1} u_k Z^k \right) \exp \left(\frac{u_0}{2} + \sum_{+1}^{\infty} u_k Z^k \right) = \\
 &= \exp \left[\bar{U}^+ \left(\frac{1}{Z} \right) \right] \exp [U^+(Z)] = \bar{B} \left(\frac{1}{Z} \right) B(Z).
 \end{aligned}$$

Таким образом, получаем разложение $R(Z)$ на требуемые сопряженные части. Наконец, посмотрим, почему функция $B(Z) = \exp [U^+(Z)]$ является минимально фазовой. Для этого достаточно знать о ряде $U^+(Z)$, что он ограничен и не содержит отрицательных степеней Z . Существует два доказательства.

Первое доказательство следует из того, что мнимая часть в $U(Z)$ на единичной окружности характеризует фазовый угол величины $B(Z)$. Для того чтобы функция $B(Z)$ была минимально фазовой, ее фаза не должна увеличиваться при вращении Z по единичному кругу на кратные 2π значения; другими словами, фаза должна быть периодичной с периодом 2π . Очевидно, что фаза функции $u_1 \sin \omega + u_2 \sin 2\omega + \dots$ удовлетворяет этому условию. Второе доказательство более абстрактно. Заметим, что существует единственная возможность для функции $B(Z) = \exp [U^+(Z_0)]$ обратиться в нуль при некотором значении Z_0 , когда ряд $U^+(Z_0)$ равен $-\infty$; другими словами, это возможно, если ряд $U^+(Z)$ расходящийся. Однако это невозможно внутри единичного круга, так как мы использовали для функции $U(Z)$ логарифмический ряд, который абсолютно сходится на единичной окружности; это означает, что ряд $U^+(Z)$ сходится (ограничен) внутри круга. Поскольку функция $B(Z)$ не может иметь нулей внутри единичной окружности, то она должна быть минимально фазовой.

Упражнения

1. Как можно получить функцию $B(Z)$, если $|R| > 2$ на единичной окружности?

2. Положим, что $U^+(Z) = aZ$. Для значений $a = -1, 0, 1, 100$ и i изобразите функции времени, связанные с $B(Z)$. Нарисуйте также их спектры. Найдите среди них один, который похож на гауссовую кривую.

3. Сейсмограммы могут быть *спектрально уравнены* путем деления Фурье-преобразования каждой сейсмограммы на ее модуль и последующего умножения на средний модуль. Это равносильно свертке с помощью симметричного фильтра. Как можно определить спектральное уравнивание сейсмограммы с помощью физически реализуемых фильтров?

4. Рассмотрите коэффициенты производной по Z функции $\ln B(Z) = U(Z)$ и найдите рекуррентные соотношения для определения u_k по заданному b_k и для определения b_k по заданному u_k .

3.5. МЕТОД КОЛМОГорова [16]

Если в ЭВМ введены коэффициенты (x_0, x_1, x_2, \dots) полинома $X(Z)$, то говорят, что мы «работаем во временной области». Если мы изобразим полином $X(Z)$ некоторым числом точек на единичной окружности, то имеем числа вида $X(e^{i\omega_0}), X(e^{i\omega_1}), \dots, X(e^{i\omega_n})$, которые называются представлением полинома $X(Z)$ в частотной области. В то же время быстрое преобразование Фурье является очень экономным средством перевода функций из временной области в спектральную (частотную) и обратно. Поэтому применение метода факторизации Виттля в частотной области предпочтительно. Кроме того, поясним факторизацию спектра уже с другой точки зрения.

Начнем с некоторой функции времени или Z -преобразования $X(Z) = x_0 + x_1 Z + \dots$. Обозначим через X_k преобразование функции времени, т. е. $X(Z)$, вычисленной для некоторого числа $(k=0, 1, \dots, n)$ точек на единичной окружности. Рассмотрим тождества

$$R_k = \bar{X}_k X_k = e^{\ln(R_k)} = e^{U_k}.$$

Теперь прибавим к U_k и вычтем гладкую произвольную функцию Φ_k :

$$R_k = \exp\left[\frac{1}{2}(U_k - i\Phi_k)\right] \exp\left[\frac{1}{2}(U_k + i\Phi_k)\right] = \bar{B}_k B_k.$$

Теперь главный вопрос заключается в том, какой вид должна иметь функция Φ_k , чтобы гарантировать преобразование функции B_k в минимально фазовую одностороннюю функцию времени. При рассмотрении метода Виттля мы отметили, что важнейшими свойствами ряда $U^+(Z)$ являются его ограниченность и обращение

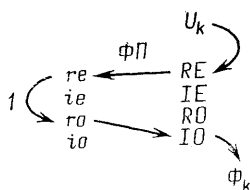


РИС. 3.4. Определение фазовой функции.

ФП — Фурье-преобразование; l — функция, равная и противоположная по знаку для отрицательных значений t

временной функции u_t в нуль при $t=0$. Тем самым мы предполагаем, что функция Φ_k должна быть выбрана таким образом, чтобы сумма $U_k + i\Phi_k$, преобразованная во временную область, давала в результате временную функцию u^{+t} , равную нулю при отрицательных значениях времени. Это может быть сделано способом, показанным на рис. 3.4.

Чтобы увидеть, как легко на самом деле найти мнимую нечетную часть IO , выберем интегрирующий фильтр из раздела 2.8 (билинейное преобразование) и изобразим его на рис. 3.5.

Для определения функции Φ_k возьмем функцию U_k во временной области, т. е. получим u_t , а затем умножим ее на действительную ступенчатую временную функцию, показанную на рис. 3.5. Тогда получим $u^{+}_t = u_0, u_1, \dots$, тем самым предполагается, что в частотной области произведена свертка функций U_k и $\delta_{k=0} + iX$

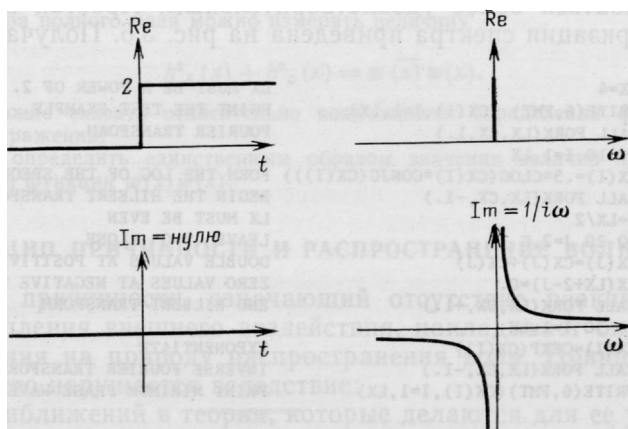


РИС. 3.5. Пары трансформаций, используемых в преобразовании Гильберта

\times (фильтр с фазовым сдвигом на 90°). Таким образом, функция Φ найдена.

Рассмотрим вновь операцию исключения всех отрицательных степеней Z в функции $U(Z)$, как это было сделано в предыдущем разделе при построении функции $U^+(Z)$. Для простоты рассмотрим случай действительных r_i ; тогда и функция u_t действительна:

$$U(Z) = \dots + u_1 Z^{-1} + u_0 + u_1 Z + u_2 Z^2 + \dots;$$

$$U = u_0 \cos 0 + 2u_1 \cos \omega + 2u_2 \cos 2\omega + \dots$$

Теперь составим новую функцию Φ путем замены косинусов на синусы в предыдущем выражении:

$$\Phi = 2u_1 \sin \omega + 2u_2 \sin 2\omega + \dots$$

Объединяя функции U и $i\Phi$, получим функцию U^+ :

$$\frac{1}{2}(U + i\Phi) = \frac{1}{2}u_0 + u_1 Z + u_2 Z^2 + \dots = U^+(Z).$$

Заметим, что операция замены $\cos t$ на $\sin t$ может рассматриваться как фильтрация с фазовым сдвигом на 90° . В рассматриваемом случае замена $\cos \omega$ на $\sin \omega$ дает результат, заключающийся в том, что $U^+(Z)$ имеет только положительные коэффициенты при Z .

Факторизация спектра методом Колмогорова очень быстро реализуется на ЭВМ, так как в нем могут применяться быстрые

преобразования Фурье. Принципиальным недостатком этого метода является то, что результат суммирования вдоль всей единичной окружности всегда незначительно отличается от результата интегрирования по окружности. Когда спектр прост, а полюсы расположены очень близко к единичной окружности, метод Теплица может оказаться более удовлетворительным. Простая программа для факторизации спектра приведена на рис. 3.6. Получаемый с ее

<pre> LX=4 WRITE(6,FMT) (CX(I),I=1,LX) CALL FORK(LX,CX,1.) DO 10 I=1,LX 10 CX(I)=.5*CLOG(CX(I)*CONJG(CX(I))) CALL FORK(LX,CX,-1.) K=LX/2 DO 20 J=2,K CX(J)=CX(J)+CX(J) 20 CX(LX+2-J)=0. CALL FORK(LX,CX,+1.) DO 30 I=1,LX 30 CX(I)=CEXP(CX(I)) CALL FORK(LX,CX,-1.) WRITE(6,FMT) (CX(I),I=1,LX) </pre>	<pre> LX MUST BE A POWER OF 2. PRINT THE TEST EXAMPLE. FOURIER TRANSFORM. FORM THE LOG OF THE SPECTRUM BEGIN THE HILBERT TRANSFORM LX MUST BE EVEN LEAVE T=0 ALONE DOUBLE VALUES AT POSITIVE TIME ZERO VALUES AT NEGATIVE TIME END HILBERT TRANSFORM EXPONENTIATE INVERSE FOURIER TRANSFORM PRINT MINIMUM PHASE WAVELET. </pre>
<pre> IN: (0.0, 0.0), (0.0, 8.0), (16.0, 0.0), (0.0, 0.0), OUT: (15.6,-0.0), (0.0,-8.7), (-0.4,-0.0), (0.0,-0.7), </pre>	

РИС. 3.6. Программа, выполняющая факторизацию спектра с помощью быстрого преобразования Фурье и преобразования Гильберта. Комплексный расчет обязателен

помощью результат является приближенным, так как интегрирование по единичной окружности аппроксимируется суммированием по четырем точкам.

Упражнения

1. Нанесите на рис. 3.4 дополнительные стрелки, которые потребуются при работе с комплексными функциями времени.

2. Какой смысл будет иметь понятие минимально фазовой формы сигнала, если временную и частотную области взаимно поменять?

3. Покажите, как произвести обратное преобразование Гильберта, отыскивая u по заданному φ . Как объяснить тот факт, что нельзя получить величину u_0 ?

4. Рассмотрите модель части Земли, в которой x является координатой в направлении на север, $+z$ представляет высоту над поверхностью Земли, а магнитные тела распределены в Земле таким образом, что они не создают компоненты магнитного поля в направлении с востока на запад. Показать, что в рассматриваемом случае магнитное поле h над поверхностью Земли можно представить выражением

$$\begin{bmatrix} h_x(x, z) \\ h_z(x, z) \end{bmatrix} = \int_{-\infty}^{+\infty} F(k) \begin{bmatrix} -ik \\ |k| \end{bmatrix} e^{ikx - |k|z} dk,$$

где $F(k)$ — некоторый спектр пространственных частот.

а. Как вычислить компоненту $h_x(x, 0)$ по заданной компоненте $h_z(x, 0)$ и наоборот, используя преобразование Фурье?

- б. Задана компонента $h_z(x, 0)$; как вычислить величину $h_z(x, z)$?
 в. Заметим, что при $z=0$

$$w(x) = h_z(x) + ih_x(x) = \int_{-\infty}^{+\infty} e^{ikx} F(k) (|k| + k) dk,$$

а также, что $F(k)(|k|+k)$ является односторонней функцией от k . С помощью магнитометра полного поля можно измерять величину

$$h_x^2(x) + h_z^2(x) = \overline{w(x)} w(x).$$

Что можно сказать относительно возможности определения функции $F(k)$ из этого выражения?

г. Как определить единственным образом значения величин $h_x(x)$ и $h_z(x)$ по заданной функции $w(x)\overline{w(x)}$.

3.6. ПРИНЦИП ПРИЧИННОСТИ И РАСПРОСТРАНЕНИЕ ВОЛН

Принцип причинности, означающий отсутствие реакции системы до наступления внешнего воздействия, накладывает определенные ограничения на природу распространения волн. Принцип причинности часто нарушается вследствие:

- 1) приближений в теории, которые делаются для ее упрощения;
- 2) неправильного подбора кривых при описании экспериментальных данных;
- 3) приближений, вводимых в процессе преобразования дифференциальных уравнений в разностные уравнения.

Нарушения принципа причинности, по-видимому, часто являются результатом нелогичного приближения спектральных частот, находящихся за границами диапазона, представляющего практический интерес. По-видимому, такие приближения часто приобретают значение вследствие очень слабой сходимости преобразования Гильберта. Что касается пунктов 1 и 2, то указанные в них нарушения принципа причинности могут оказаться оправданными благодаря низким затратам на расчеты в сравнении с затратами на более точный анализ. В отношении пункта 3 и вычислений вообще следует сказать, что вытекающее из него нарушение принципа причинности обычно имеет катастрофические последствия. Например, даже если разностные уравнения дают правильное решение в области интересующих нас частот, то на других частотах почти всегда наблюдается огромный рост экспоненты, полностью разрушающий требуемое решение.

Большинство теорий распространения волн разработано в частотной области для однородных сред. Если они описаны в декартовой системе координат, то конечный результат состоит в том, что распространение волны вдоль оси x описывается выражением

$$Y(x) = Y(0)e^{ikx}. \quad (3.6.1)$$

Волновая теория предусматривает выражение величины k через частоту и другие физические параметры, такие как скорость, затухание и т. д. Рассмотрим вид функции $k(\omega)$ после возбуждения

и посмотрим, удовлетворяет ли она принципу причинности. Другими словами, соотношение (3.6.1) подобно системе фильтров, изображенной на рис. 3.7, где $Y(0)$ представляет входной сигнал фильтра, $Y(x)$ — выходной сигнал, а e^{ikx} — фильтр, который моделирует эффект распространения волны. Вопрос заключается в том,

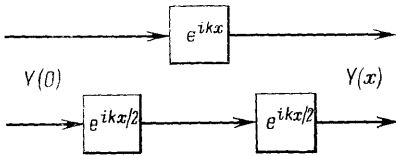


РИС. 3.7. Моделирование распространения волны с помощью фильтров

являются ли фильтры односторонними. Это зависит от конкретного выражения величины $k(\omega)$.

Сначала рассмотрим простой случай, когда функция $k(\omega)$ справедлива для всех действительных частот ω . Тогда функция e^{ikx} соответствует фазовому фильтру. Он физически реализуем, как было показано ранее (см. стр. 52), тогда и только тогда, когда $k(\omega)x$ является монотонно возрастающей функцией частоты ω .

Затем предположим, что функция $k(\omega)$ комплексная. Это значит, что будет затухание. Очевидно, что принцип сохранения энергии требует, чтобы действительная часть выражения ikx была отрицательной.

Определим функцию $P(Z)$ как физически реализуемый фазовый фильтр, а функцию $B(Z)$ как физически реализуемый минимально фазовый сейсмический сигнал. Любую физически реализуемую функцию можно представить в виде $B(Z)P(Z)$. Потребуем такого представления функции e^{ikx} :

$$B(Z)P(Z) = e^{ikx} = \exp [i(k_2 + ik_i)x]. \quad (3.6.2)$$

Производя операцию логарифмирования и используя φ для обозначения фазы, имеем:

$$\ln |B| + i\varphi_B + i\varphi_P = -k_i x + ik_r x. \quad (3.6.3)$$

Разбивая это выражение на действительную и мнимую части, получим:

$$\ln |B| = -k_i x;$$

$$\varphi_P = k_r x - \varphi_B.$$

Прежде всего, на основе затухания $-k_i x = \ln |B|$ можно с помощью преобразования Гильберта вычислить фазу φ_B . Если ее вычесть из $k_r x$, то остается фазовый сдвиг φ_P фазового фильтра. Вспомним из раздела 2.6 (о фазовых фильтрах), что такой фильтр будет физически реализуемым тогда и только тогда, когда фаза φ_P монотонно увеличивается с частотой. В заключение отметим, что проверка принципа причинности для функции $k(\omega)$ состоит в вычислении фазы $\varphi_P(\omega)$ и в решении вопроса о том, является ли она монотонно возрастающей для всех частот.

Большая путаница возникает при учете затухания волн, которые, вероятно, имеют вид $k_i \approx \omega^2$ (гауссово затухание) или $k_i \approx |\omega|$ (иногда называемое затуханием с постоянной добротностью Q). Причина этой путаницы заключается в том, что эти функции не интегрируемы, поэтому они не имеют сходящихся преобразований Гильберта; т. е. фаза φ_V оказывается бесконечной. Эти функции соответствуют временным импульсам, похожим на гауссовы временные функции $\exp(-t^2)$, которые не равны нулю ни в какие моменты времени. Такие функции не могут рассматриваться в строгой связи с любым физически реализуемым процессом; однако они дают превосходные практические приближения. В действительности физическое затухание энергии обычно связывается с явлением релаксации, обладающей конечным временем. Для частот, лежащих ниже частоты релаксации, затухание может казаться неопределенно возрастающим с частотой, однако в действительности затухание уменьшается для частот, превышающих характеристическую частоту релаксации. Затухание также создает искажение фазовой скорости из-за добавления фазы φ_V к фазе φ_P , но в большинстве случаев считается, что искажение слишком мало, чтобы его можно было измерить.

Теперь рассмотрим пример моделирования распространения волн на ЭВМ. Простейшим примером является распространение волн без затухания. Тогда $k = \omega/v$, где v — скорость распространения волн в среде. До построения фильтра, который будет переносить волны на большое расстояние, попробуем перенести их на небольшое расстояние Δx . Такой фильтр можно будет использовать N раз для моделирования распространения волн на большое расстояние, равное $X = N\Delta x$. Используем билинейное усечение степенного ряда для показательной функции:

$$e^u \approx \frac{1 + \frac{u}{2}}{1 - \frac{u}{2}} \quad (3.6.4)$$

и билинейное представление для

$$i\omega\Delta t \approx 2 \left(\frac{Z-1}{Z+1} \right). \quad (3.6.5)$$

С помощью этих выражений с учетом зависимости $k = \omega/v$, а также обозначения $a = \Delta x/v\Delta t$ получим

$$\begin{aligned} e^{ik\Delta x} &\approx e^{i\hat{k}\Delta x} = \frac{1 + a \left(\frac{Z-1}{Z+1} \right)}{1 - a \left(\frac{Z-1}{Z+1} \right)} = \frac{(Z+1) + a(Z-1)}{(Z+1) - a(Z-1)} = \\ &= \frac{(1-a) + (1+a)Z}{(1+a) + (1-a)Z}. \end{aligned} \quad (3.6.6)$$

Последняя формула представляет собой общее выражение для фазового фильтра. В этом легко убедиться, обозначив $Z_0 = -(1+a)/(1-a)$. Тогда формула (3.6.6) приобретает вид

$$e^{i\hat{k}\Delta x} = \frac{-1/Z_0 + Z}{1 - Z/Z_0}. \quad (3.6.7)$$

Если знаменатель выражения (3.6.7) будет минимально фазовой функцией, то фазовый фильтр будет физически реализуемым, если $Z_0 > 1$ или $\Delta x/v\Delta t > 0$. Это означает, что текущие и прошлые значения входных сигналов будут определять выходной сигнал, если волна распространяется в направлении оси $+x$. Если необходимо направить данную волну в противоположную сторону в пространстве (Δx — отрицательное), то и Δt должно быть отрицательным, в результате чего текущий выходной сигнал будет определяться текущими и будущими значениями входного сигнала.

Упражнение

Чему равна погрешность определения скорости в функции частоты ω для фильтра, описываемого выражением (3.6.6), при значениях $a = 0, 1; 1$ и 2 .

Указание. Пусть $\hat{k} = \omega/\hat{v}$ и, следовательно, $v/\hat{v} = \hat{k}v/\omega$. Учтите также, что

$$\begin{aligned} i\hat{k}\Delta x &= \ln e^{i\hat{k}\Delta x} = \ln Z + \ln \left(\frac{-1}{Z_0 Z} + 1 \right) - \ln \left(1 - \frac{Z}{Z_0} \right) = \\ &= +i\omega\Delta t - 2 \operatorname{Im} \ln \left(1 - \frac{Z}{Z_0} \right). \end{aligned}$$

4. РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ

При поисках положения очага землетрясения или залежи нефти обычно некоторая неопределенность в установлении места, скажем $(\Delta x, \Delta y, \Delta z)$, связанная с погрешностями измерений и физическими размерами целевого объекта. При измерении электрического напряжения погрешность измерения равна Δv . Сигнал полезной сейсмической волны занимает полосу частот $\Delta \omega$. Время возникновения землетрясения имеет неопределенность, зависящую от длительности толчков Δt . Телескоп диаметром Δd имеет в лучшем случае разрешающую силу, измеряемую определенным угловым интервалом $\Delta \theta$. При измерениях желательно уменьшить величины $\Delta x, \Delta y, \Delta z, \Delta d, \Delta v, \Delta \omega, \Delta t$ и $\Delta \theta$ до возможно меньших значений. Оценки (меры) разрешающей способности (называемые дисперсиями, допусками, неопределенностями, полосами, продолжительностями, протяженностями, диапазонами и т. д.) иногда взаимодействуют друг с другом таким образом, что любая модификация эксперимента, уменьшающая одну из них, должна с необходимостью увеличивать другую или некоторую комбинацию других. В главе обсуждаются обычно встречающиеся ситуации, когда возникают подобные противоречивые влияния.

В этой главе величина Δt используется для обозначения длительности сигнала во времени, а величина τ — для обозначения интервала времени между точками выборки (квантования). В других главах Δt совпадает с τ , т. е. интервалом выборки (шагом дискретизации по времени).

4.1. ЧАСТОТНО-ВРЕМЕННОЕ РАЗРЕШЕНИЕ

Известный в квантовой механике «принцип неопределенности» является следствием наблюдений, показавших, что элементарные частицы ведут себя подобно волнам с частотами, пропорциональными моментам количества движения частиц. Законы классической механики позволяют предсказывать будущее поведение механической системы путем экстраполяции от известного состояния и момента количества движения. Однако вследствие волновой природы материи, обладающей моментом количества движения, пропорциональным частоте, такое предсказание требует одновременного знания двух величин: положения тела и частоты волны. Синусоидальная волна имеет точно определенную частоту, но она занимает бесконечно длинный интервал на оси времени. С другой стороны, дельта-функция сжимается в точку на оси времени, но содержит все частоты. Математический анализ принципа неопределенности превращается, таким образом, в анализ связи функций со своими преобразованиями Фурье.

Такой анализ начинается с определения продолжительности сигнала и занимаемой им полосы частот. Длительность затухающей экспоненциальной функции бесконечна, если под длительностью понимать интервал ненулевых значений функций. Однако почти для всех практических целей интервал существования этой функции во времени выбирается до момента, когда амплитуда сигнала уменьшается в e раз по сравнению с первоначальным значением. Для многих функций этот интервал определяется расстоянием между точками на оси времени или на оси частот, в которых кривая (или ее огибающая) спадает до половины максимального значения. Выбор такого интервала основан на том, что промежуток времени Δt или диапазон частот $\Delta \omega$ должен заключать в себе большую (основную) часть полной энергии сигнала, но необязательно всю ее. Точное определение величин Δt и $\Delta \omega$ несколько произвольно и с целью упрощения анализа оно может быть различным. Общий принцип состоит в том, что для любой функции интервал времени Δt и диапазон (полоса) частот $\Delta \omega$ связываются соотношением

$$\Delta \omega \Delta t \geq 2\pi. \quad (4.1.1)$$

Выражение (4.1.1) легко проверить для многих частных случаев, но очень трудно доказать как общий принцип. Однако это сделал Д. Габор [17]. Он определил величины Δt и $\Delta \omega$ через вторые моменты.

Подобным, а возможно, более важным, чем произведение временного и частотного интервалов, понятием является зависимость между полосой частот и временем нарастания функции отклика системы. *Время нарастания* отклика системы Δt определяется несколько произвольно, часто как промежуток времени между моментом возбуждения системы и моментом, когда отклик системы достигает половины своего максимального значения. В принципе отклик системы в широкой полосе частот может быть следствием как быстрого во времени затухания, так и быстрого во времени нарастания процесса в системе. *Строгое выполнение* неравенства (4.1.1) характерно для систем, в которых определенное время нарастания процесса быстро сменяется равным временем его затухания. *Слабое выполнение* неравенства (4.1.1) может связываться с увеличением неравенства времен нарастания и затухания процесса в системе, а также может быть следствием других комбинаций нарастания и затухания процесса, в том числе результатом случайных комбинаций. Многие системы реагируют на возбуждение очень быстро по сравнению со скоростью, с которой они впоследствии затухают. Выделив такие системы, можно теперь попытаться получить неравенство (4.1.1) для времени нарастания и ширины полосы частот. Первым шагом является выбор определения времени нарастания. Этот выбор ограничен не только стремлением к ясности и полезности, но и необходимостью обеспечения просто-

гы последующего анализа. Время нарастания

$$\Delta t = \frac{\int_0^{\infty} b^2(t) dt}{\int_0^{\infty} \frac{1}{t} b^2(t) dt}, \quad (4.1.2)$$

где $b(t)$ — отклик рассматриваемой функции. Числитель является нормирующим множителем. Знаменатель показывает, что Δt определяется через первый, отрицательный момент. Например, если $b(t)$ — ступенчатая функция, то интеграл в знаменателе расходится, давая необходимое время нарастания $\Delta t = 0$. Если величина $b^2(t)$ растет линейно от нуля до t_0 , а затем равна нулю, то время нарастания равно $t_0/2$, что также соответствует разумному определению.

Хотя метод Z -преобразования оказывает огромную помощь при изучении ситуаций, где расходимость (вида $1/t$) играет основную роль, он обладает тем недостатком, что нарушает формальное соответствие между областью времени и областью частот. Вероятно, этот недостаток не является принципиальным, так как всегда можно перейти к предельному процессу, в котором область дискретного времени стремится к непрерывности. Чтобы использовать аналитическую простоту Z -преобразования, рассмотрим двойственность задачи о времени нарастания. Вместо функции времени, квадрат которой тождественно равен нулю для отрицательных значений времени, рассмотрим спектр $\bar{B}(1/Z)B(Z)$, который равен нулю для отрицательных частот. Определим, как быстро растут значения этой спектральной функции, начиная с частоты $\omega = 0$. Это необходимо для оценки длительности Δt комплексной функции времени b_t . Точнее, необходимо определить самую нижнюю существенную для данного спектра частотную компоненту $\omega = \Delta\omega$, т. е. найти выражение, аналогичное формуле (4.1.2), в спектральной области

$$\Delta\omega = \frac{\int_{-\infty}^{+\infty} \bar{B}B d\omega}{\int_{-\infty}^{+\infty} \frac{1}{\omega} \bar{B}B d\omega}. \quad (4.1.3)$$

Без потери общности можно предположить, что этот спектр нормализован таким образом, что интеграл в числителе равен единице. Другими словами, нулевая задержка функции автокорреляции для b_t равна $+1$. Тогда

$$\frac{1}{\Delta\omega} = \int_{-\infty}^{+\infty} \frac{1}{\omega} \bar{B}B d\omega. \quad (4.1.4)$$

Теперь вспомним билинейное преобразование, которое дает различные выражения Z -преобразования для функции $(-i\omega)^{-1}$. Одно из них, обычно используемое нами, является интегралом вида $(\dots 0, 0, 0,5, 1, 1, \dots)$. Мы могли бы использовать также выражение $(\dots 1, 1, 1, 0,5, 0, 0, \dots)$. Наличие полюса, расположенного прямо на единичной окружности при $Z=1$, приводит к некоторой неоднозначности. Поскольку $1/i\omega$ является мнимой нечетной частотной функцией, то мы выберем разложение так, чтобы оно было нечетной функцией времени, заданной в виде

$$(-i\omega)^{-1} = \frac{(\dots - Z^{-2} - Z^{-1} + 0 + Z + Z^2 + \dots)}{2}. \quad (4.1.5)$$

Приводя интеграл (4.1.4) к интегралу на единичной окружности, будет иметь (в обозначениях Z -преобразования):

$$\begin{aligned} \frac{1}{\Delta\omega} &= \int_{-\infty}^{+\infty} -i \frac{1}{-i\omega} \bar{B} B d\omega = \\ &= \frac{-i}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{2} (\dots - Z^{-2} - Z^{-1} + Z + Z^2 + \dots) \bar{B} \left(\frac{1}{Z}\right) B(Z) d\omega. \end{aligned} \quad (4.1.6)$$

Поскольку в этом интеграле используется коэффициент при степени Z^0 своего аргумента, то

$$\frac{1}{\Delta\omega} = \frac{-i}{2} [(r_{-1} - r_1) + (r_{-2} - r_2) + \dots], \quad (4.1.7)$$

где r_t — автокорреляционная функция для b_t . Преобразуем выражение (4.1.7) к виду

$$\frac{1}{\Delta\omega} = \sum_{t=1}^{\infty} \text{Im}(-r_t) \leq \sum_{t=1}^{\infty} |r_t|. \quad (4.1.8)$$

Сумма в выражении (4.1.8) похожа на интеграл, определяющий площадь под функцией $|r_t|$. Представьте, что функция $|r_t|$ заменена на прямоугольную функцию равной площади. Это определило бы $\Delta t_{\text{авто}}$ для функции $|r_t|$. Любая автокорреляционная функция удовлетворяет неравенству $r_t < r_0$ (нормированное $r_0=1$). Таким образом, мы расширили неравенство (4.1.8) следующим образом:

$$\frac{1}{\Delta\omega} \leq \sum_{t=1}^{\infty} |r_t| < \Delta t_{\text{авто}}. \quad (4.1.9)$$

Наконец, необходимо связать длительность функции времени Δt с длительностью ее автокорреляционной функции $\Delta t_{\text{авто}}$. Вообще говоря, легко найти длинную функцию времени, имеющую корот-

кую автокорреляционную функцию. Например, можно взять произвольную короткую функцию времени и свернуть ее с длинным фазовым фильтром. Полученная таким способом новая функция времени будет длинной, а ее автокорреляционная функция — короткой. Если функция времени имеет n ненулевых точек, то ее автокорреляционная функция будет иметь $2n-1$ ненулевых точек. Очевидно, невозможно получить длинную автокорреляционную функцию для короткой функции времени. Нельзя даже просто сказать, что автокорреляционная функция вдвое длиннее первоначальной функции времени, поскольку она должна принадлежать к классу сужающихся к концам функций. Для того чтобы построить функцию времени с возможно более длинной автокорреляционной функцией, лучше всего сконцентрировать энергию сигнала в двух максимумах — по одному на каждом конце функции времени. Даже из этого утрированного примера видно, что

$$\Delta t \geq \Delta t_{\text{авто}}. \quad (4.1.10)$$

Подставляя выражение (4.1.10) в формулу (4.1.9), получим соотношение неопределенности:

$$\Delta t \Delta \omega \geq 1. \quad (4.1.11)$$

В более привычной форме принципа неопределенности используются частотная переменная $f=2\omega$ и величина Δt , соответствующая скорее *длительности отклика*, чем его *времени нарастания*. Тогда соотношение неопределенности имеет вид

$$\Delta t \Delta f \geq 1 \quad (\Delta t — \text{длительность}). \quad (4.1.12)$$

Выбор величины 2π в качестве масштабирующего множителя для преобразования *времени нарастания* в *длительность отклика* указывает на приближенный характер приводимых выше неравенств.

Упражнения

1. Рассмотрите функцию $B(Z) = [1 - (Z/Z_0)^n] / [1 - Z/Z_0]$, в пределе устремляя Z_0 к единичной окружности. Изобразите функцию времени и квадрат ее амплитуды. Изобразите частотную функцию частоты и квадрат ее амплитуды. Выберите Δf и Δt .

2. Временной ряд, составленный из двух частот, можно записать в виде $b_t = A \cos \omega_1 t + B \sin \omega_1 t + C \cos \omega_2 t + D \sin \omega_2 t$. Покажите, как рассчитать амплитуды

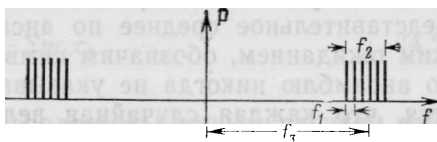


РИС. У4.1.

и фазы двух синусоидальных компонент, если величины A, B, C, D, ω_1 и ω_2 заданы.

3. Рассмотрите функцию частоты, график которой приведен на рис. У4.1. Опишите приближенно соответствующую ей функцию времени, указав моменты времени, отвечающие величинам $1/f_1, 1/f_2$ и $1/f_3$. Постарайтесь избежать алгебраических вычислений. Изобразите приблизительный результат

Задача для исследований

Можете ли Вы найти метод определения $\Delta\omega$ и Δt для односторонних сейсмических сигналов таким образом, чтобы неравенство, соответствующее принципу неопределенности, становилось равенством только для минимально фазовых сигналов?

4.2. СТАТИСТИКО-ВРЕМЕННОЕ РАЗРЕШЕНИЕ

Если бы Вы подбросили монету 100 раз, то возможно, что Вы получили бы точно 50 «орлов» и 50 «решек». Более вероятно, что эти числа были бы между 60—40 и 40—60. Какое типичное отклонение от 50 ожидали бы Вы увидеть? В среднем эта величина должна равняться 50, но почти всегда из случайной выборки получают другое значение, которое называется *выборочным средним*. Мы, вероятно, хотели бы знать, какую разницу следует ожидать между выборочным средним и истинным средним значениями. Средний квадрат этой разности называется *дисперсией выборочного среднего*. Для очень большой выборки выборочное среднее должно быть пропорционально более близким к истинному среднему, чем для меньшей выборки. Это также приводит к соотношению неопределенности между вероятной погрешностью оценки среднего и размером (объемом) выборки.

«Истинное значение» средней величины можно найти, подбрасывая монету n раз и мысленно представляя себе n стремящихся к бесконечности. Более удобным способом определения «истинного значения» является этот воображаемый эксперимент, проведенный раздельно в тождественных условиях неограниченным числом людей (ансамблем). Этот прием позволит нам определить средние значения величин, которые изменяются со временем.

Полезность такого истолкования ансамбля часто подвергается серьезной критике как теоретиками с позиций статистики, так и экспериментаторами, применяющими статистические методы. Тем не менее в большей части геофизической литературы используется этот прием, предполагается существование необозримого ансамбля. Сторонники использования ансамблей (последователи Гиббса) имеют преимущество перед своими противниками (последователями Байеса), заключающееся в том, что их математический аппарат легче усваивается (и понятнее).

Представительное среднее по ансамблю, называемое математическим ожиданием, обозначим символом E . Индекс суммирования по ансамблю никогда не указывается в явном виде; предполагается, что каждая случайная величина обладает им. Таким образом, истинное среднее в момент времени t можно определить

$$m_t = E(x_t). \quad (4.2.1)$$

Если среднее не меняется во времени, то можно написать

$$m = E(x_t) \quad (\text{для всех } t). \quad (4.2.2)$$

Для нас представляет интерес также свойство случайной величины x_t , называемое *дисперсией*, которая является мерой отклонения от среднего:

$$\sigma^2_t = E[(x_t - m_t)^2]. \quad (4.2.3)$$

Случайные величины x_t могут определяться таким образом, что величины σ или m , или обе одновременно, являются либо функциями времени, либо константами. Если обе величины константы, то имеем

$$\sigma^2 = E[(x_t - m)^2]. \quad (4.2.4)$$

При вычислениях конкретных выражений E ведет себя как знак суммирования, а именно:

$$E = (\lim n \rightarrow \infty) \frac{1}{n} \sum_1^n. \quad (4.2.5)$$

Заметим, что индекс суммирования не задан, так как сумма берется по ансамблю, а не по времени.

Пусть теперь случайная величина x_t является временным рядом, составленным из (равномерно распределенных, независимо выбранных) случайных чисел таким образом, что величины m и σ не зависят от времени. Предположим, что имеется выборка из n точек случайной величины (функции) x_t и делается попытка определить ее среднее значение m . Тогда можно оценить \hat{m} среднее значение случайной величины m с помощью формулы

$$\hat{m} = \frac{1}{n} \sum_{t=1}^n x_t. \quad (4.2.6)$$

Несколько более точным методом оценки среднего является определение взвешенного среднего. Пусть w_t определяет набор весов, нормализованных таким образом, что

$$\sum w_t = 1. \quad (4.2.7)$$

Тогда более точная оценка среднего \hat{m} с учетом весов имеет вид

$$\hat{m} = \sum w_t x_t. \quad (4.2.8)$$

Очевидно, что формула (4.2.6) является частным случаем выражения (4.2.8), когда значения всех весов одинаковы: $w_t = 1/n$ ($t = 1, 2, \dots, n$).

Целью этого раздела является определение вероятной близости оценки среднего \hat{m} случайной величины к истинному среднему m для выборки длины n . Одно из возможных выражений для опре-

деления этого отклонения Δm имеет вид

$$(\Delta \widehat{m})^2 = E [(\widehat{m} - m)^2]; \quad (4.2.9)$$

$$(\Delta m)^2 = E \{[(\sum \omega_t x_t) - m]^2\}. \quad (4.2.10)$$

Теперь, используя тот факт, что $m = m \sum \omega_t = \sum \omega_t m$, запишем:

$$(\Delta m)^2 = E \left\{ \left[\sum_t \omega_t (x_t - m) \right]^2 \right\}; \quad (4.2.11)$$

$$(\Delta m)^2 = E \left\{ \left[\sum_t \omega_t (x_t - m) \right] \left[\sum_s \omega_s (x_s - m) \right] \right\}; \quad (4.2.12)$$

$$(\Delta m)^2 = E \left[\sum_t \sum_s \omega_t \omega_s (x_t - m) (x_s - m) \right]. \quad (4.2.13)$$

Здесь символ математического ожидания E означает знак суммирования и его можно внести под знаки сумм по t и по s . Тогда

$$(\Delta m)^2 = \sum_t \sum_s \omega_t \omega_s E [(x_t - m) (x_s - m)]. \quad (4.2.14)$$

Благодаря случайности величин x_t и x_s математическое ожидание в правой части, т. е. сумма по всему ансамблю, дает нуль, если не выполняется условие $s=t$. Если $s=t$, то математическое ожидание является дисперсией, определяемой соотношением (4.2.4). Тогда

$$(\Delta m)^2 = \sum_t \sum_s \omega_t \omega_s \sigma^2 \delta_{ts}; \quad (4.2.15)$$

$$(\Delta m)^2 = \sum_t \omega_t^2 \sigma^2 \quad (4.2.16)$$

или

$$\Delta m = \sigma \left(\sum_t \omega_t^2 \right)^{1/2}. \quad (4.2.17)$$

Теперь рассмотрим этот конечный результат для случая, когда каждый из весов имеет значение $1/n$, т. е.

$$\Delta m = \sigma \left[\sum_1^n \left(\frac{1}{n} \right)^2 \right]^{1/2} = \frac{\sigma}{(n)^{1/2}} = \frac{\sigma}{\sqrt{n}}. \quad (4.2.18)$$

Это наиболее важное свойство случайных чисел не является, вероятно, очевидным. Для случая нулевого среднего оно может быть выражено следующей фразой: « n случайных чисел с единичными модулями при их сложении образуют величину, близкую к квадратному корню из n ».

При попытках вычислить среднее значение ряда случайных величин с меняющимся во времени средним приходится сталкиваться с фундаментальной дилеммой. Если включить в сумму много чисел, чтобы обеспечить малую величину Δm , то за время измерения величина m может заметно измениться. В противоположность этому оценка среднего значения \bar{m} , полученная на основании короткой выборки из этого же ряда, может сильно отклониться от истинного среднего m (определенного по всему ансамблю с бесконечным числом членов в любой момент времени). Это представляет собой основную дилемму, перед которой стоит, например, биржевой маклер, когда клиент говорит ему: «Так как цена участка (земли) на рынке меняется, то я хочу, чтобы вы продали мое имущество в тот момент, когда его цена превысит среднюю продажную цену».

Предположим, что выборка из временного ряда производилась каждые τ секунд и что $\Delta t = n\tau$ обозначает длину выборки. Тогда выражение (4.2.18) можно записать в виде

$$(\Delta m)^2 \Delta t = \sigma^2 \tau. \quad (4.2.19)$$

Очевидно, желательно, чтобы значения Δm и Δt были сколь угодно малы. Если первоначальные случайные числа x_i коррелированы друг с другом, например, если x_i являются приближением непрерывной функции, то сумма этих n чисел не сводится к корню из n . Это выражается следующим неравенством:

$$(\Delta m)^2 \Delta t \geq \sigma^2 \tau. \quad (4.2.20)$$

Неравенство (4.2.20) можно назвать соотношением неопределенности между точностью и разрешением во времени.

Рассматривая другие наборы весов, можно предложить физически более обоснованное определение величины Δt , чем $\Delta t = n\tau$. Например, если веса w_i задаются с помощью дискретизированной гауссовой функции, как показано на рис. 4.1, то можно было бы выбрать в качестве Δt отрезок времени, разделяющий точки с половинным значением амплитуды, точки, в которых амплитуда спадает в e раз, временной интервал, заключающий в себя 95% площади, а также множество других «разумных» определений. Если задан маленький разброс в определении величин Δm и Δt , то неравенство (4.2.20), очевидно, не может строго соблюдаться.

Задав выборку случайного временного ряда x_i с нулевым средним, можно определить другой ряд y_i как $y_i = x_i^2$. Задача опреде-

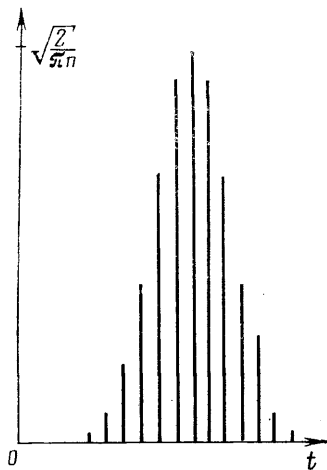


РИС. 4.1. Биноминальные коэффициенты, распределенные по закону гауссовой функции. Представлены коэффициенты при Z^t для функции $(0,5 + 0,5Z)^{20}$

ления дисперсии $\sigma^2 = p$ для ряда x_t тождественна задаче определения среднего m для ряда y_t . Если выборка короткая, то возникает погрешность Δp определения дисперсии. Таким образом, строго говоря, было бы желательно для среднего писать:

$$m = \hat{m} \pm \Delta m; \quad (4.2.21)$$

$$m = \hat{m} \pm \sigma / \sqrt{n}. \quad (4.2.22)$$

Однако в связи с тем, что дисперсия σ^2 часто неизвестна для того или другого ряда, то необходимо использовать оценку дисперсии $\hat{\sigma}$, т. е.

$$m = \hat{m} \pm \hat{\sigma} / \sqrt{n}. \quad (4.2.23)$$

Конечно, соотношение (4.2.23) на самом деле неверно, так как в действительности мы должны прибавить какую-то величину, чтобы показать дополнительную неопределенность, обусловленную погрешностью оценки $\hat{\sigma}$. Эта погрешность оценки будет опять иметь погрешность и так до бесконечности. Чтобы действительно получить правильный результат, необходимо знать функцию распределения плотности вероятности для расчета всех $E(x^n)$, которые требуются. Функцию распределения вероятности можно либо вычислить на основании данных измерений, либо выбрать теоретически. На практике в силу причин, рассматриваемых в следующей главе, часто выбирается гауссова функция. Можно показать, что

$$\Delta p = p \sqrt{\frac{2}{n}}. \quad (4.2.24)$$

Возведя последнее выражение в квадрат с учетом равенства $\Delta t = n\tau$, получим

$$\left(\frac{\Delta p}{p}\right)^2 \frac{\Delta t}{2\tau} \geq 1. \quad (4.2.25)$$

Это неравенство применяется в случае, когда случайные числа x_t не являются полностью непредсказуемыми. Если значения x_t аппроксимируют непрерывную функцию, то они хорошо предсказуемы, а неравенство для них частично ослабляется.

Функция корреляции подобна косинусу. Косинус определяет угол между двумя векторами. Он задается скалярным произведением двух векторов, деленным на произведение их модулей, т. е.

$$c = \frac{(xy)}{[(xx)(yy)]^{1/2}}.$$

Функция корреляции выражается подобным же образом, за исключением того, что x и y — скалярные случайные переменные. Функция корреляции

$$c = \frac{E(xy)}{[E(x^2)E(y^2)]^{1/2}}.$$

На практике почти никогда не встречается ансамблей случайных величин. Поэтому в практических задачах среднее по ансамблю приходится заменять средним по выборке. Эта задача возникает для малых выборок и наиболее ярко иллюстрируется выборкой, включающей только один элемент. Тогда корреляция выборки

$$\hat{c} = \frac{xy}{|x||y|} = \pm 1$$

независимо от того, какое из значений случайных чисел следует взять: x или y . На самом деле оказывается, что значения функции корреляции выборки \hat{c} будут всегда рассеяны достаточно далеко от нуля.

Это, несомненно, объясняет многие ложные «открытия». Проблема определения смещения и дисперсии корреляции является сложной, но, по-видимому, следует ожидать, что смещения и дисперсии \hat{c} имеют порядок $1/\sqrt{n}$ для выборки объема n .

Упражнения

1. Предположим, что оценка среднего по выборке случайных чисел произведена треугольной взвешивающей функцией, т. е.

$$\hat{m} = s \sum_{i=0}^n (n-i) x_i.$$

Найдите такой масштабный множитель s , что $E(\hat{m}) = m$. Вычислите Δm . Определите допустимый интервал Δt . Исследуйте соотношение неопределенности.

2. Среднее значение случайного ряда x_i , имеющего переменное во времени среднее значение, может оцениваться с помощью уравнения с обратной связью

$$\hat{m}_t = (1 - \epsilon) \hat{m}_{t-1} + b x_t.$$

а. Выразите \hat{m}_t как функцию от x_t, x_{t-1}, \dots , не зависящую от \hat{m}_{t-1} .

б. Чему равно эффективное время усреднения Δt ?

в. Найдите масштабный множитель b таким образом, чтобы при $m_t = m$, выполнялось условие $E(\hat{m}_t) = m$.

г. Подсчитайте случайную погрешность $\Delta m = [E(\hat{m} - m)^2]^{1/2}$ (результат стремится к $\sigma(\epsilon/2)^{1/2}$ при стремлении ϵ к нулю).

д. Чему равно в этом случае $(\Delta m)^2 \Delta t$?

3. Покажите, что $(\Delta P)^2 = \frac{1}{n} [E(x^4) - \sigma^4]$.

4. Определите поведение ряда x_i с независимыми значениями и средним по времени, равным нулю, с помощью вероятностей достижения различных амплитуд. Вычислите $E(x_i), E(x_i^2), (\Delta P)^2$. Если вы изучали курс теории вероятностей, то используйте гауссову функцию плотности вероятности для представления ряда x_i .

Указание:

$$P(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/2\sigma^2}$$

и

$$\int_0^{\infty} x^{2n} e^{-ax^2} dx = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^{n+1} a^n} \sqrt{\frac{\pi}{a}}.$$

4.3. ЧАСТОТНО-СТАТИСТИЧЕСКОЕ РАЗРЕШЕНИЕ

Наблюдения за уровнем моря в течение длительного периода времени можно выразить в терминах нескольких статистических средних значений, таких как средняя высота m и дисперсия σ^2 . Другим важным видом статистического среднего используемым в таких геофизических временных рядах является *энергетический спектр*. Некоторые математические модели объясняют только статистические средние исходных данных, а не сами данные. Для того чтобы узнать о возможных ловушках и понять некоторые основные ограничения при использовании энергетических спектров, рассмотрим сначала идеализированный пример.

Пусть x_t будет временным рядом, составленным из независимо выбранных случайных чисел. Предположим, что имеется n таких чисел. Тогда можно определить многочлен $X(Z)$, описывающий эту выборку данных

$$X(Z) = x_0 + x_1 Z + x_2 Z^2 + \dots + x_{n-1} Z^{n-1}. \quad (4.3.1)$$

Теперь найдем спектр энергии¹ $\hat{R}(Z)$ для этой выборки случайных чисел в виде

$$\hat{R}(Z) = \frac{1}{n} \bar{X} \left(\frac{1}{Z} \right) X(Z). \quad (4.3.2)$$

Разница между этим и более ранним определением спектра состоит в том, что *энергетический* спектр имеет делитель n , который компенсирует линейное увеличение ожидаемого результата пропорционально довольно произвольному объему выборки n .

Энергетический спектр определяется математическим ожиданием величины \hat{R} :

$$R(Z) = E[\hat{R}(Z)]. \quad (4.3.3)$$

Может показаться, что при практическом определении величины \hat{R} выражение (4.3.2) допускает устремление n в бесконечность. Однако такое заключение ошибочно. Рассмотрение этого вопроса является основной темой настоящего раздела. В частности, на основании рис. 4.2 можно заключить, что функция $\hat{R}(Z)$ обладает намного большим уровнем шумов, чем функция $R(Z)$, ввиду чего

$$R(Z) \neq \lim_{n \rightarrow \infty} \hat{R}(Z). \quad (4.3.4)$$

Для понимания этого приравняем коэффициенты при одинаковых степенях Z в соотношении (4.3.2). Тогда

$$\hat{r}_k = \frac{1}{n} \sum_{t=0}^{n-k-1} \bar{x}_t x_{t+k} \quad (k = -n+1, \dots, n-1). \quad (4.3.5)$$

¹ Точнее, спектральную плотность. (Прим. перев.)

Это позволяет нам переписать выражение (4.3.2) в виде действительного временного ряда (при $x_t = \bar{x}_t$):

$$\widehat{R} = \widehat{r}_0 + 2 \sum_{k=1}^{n-1} \widehat{r}_k \cos k\omega. \quad (4.3.6)$$

Исследуем (4.3.6) для больших n . Для этого необходимо использовать некоторые статистические свойства случайных чисел. Пусть случайные числа имеют нулевое среднее значение $m = E(x_t)$ и обладают известной постоянной дисперсией $\sigma^2 = E(x_t^2)$. Вспомним наше предположение о независимости случайных величин, которое

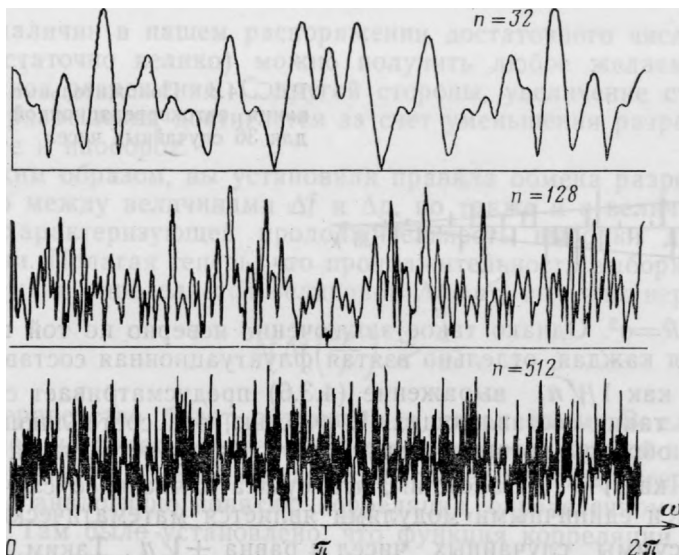


РИС. 4.2. Амплитудные спектры $[\widehat{R}(Z)]^{1/2}$ выборок из n случайных чисел. Функции осциллируют с одинаковой амплитудой как при $n=512$, так и при $n=32$. При $n \rightarrow \infty$ следует ожидать неограниченного увеличения частоты колебаний

означает, что $E(x_t x_{t+s}) = 0$, если $s \neq 0$. Вследствие случайности колебаний мы должны ожидать, что \widehat{r}_0 станет равной дисперсии σ^2 , сложенной с компонентой случайного колебания, которая уменьшается с увеличением объема выборки пропорционально $1/\sqrt{n}$:

$$\widehat{r}_0 = \sigma^2 \pm \frac{\sigma^2}{\sqrt{n}}. \quad (4.3.7a)$$

Аналогично \widehat{r}_1 должно стать равным нулю, но в соответствии с определением (4.3.5) следует ожидать появления флуктуационной компоненты. Тогда

$$\widehat{r}_1 = \pm \frac{n-1}{n} \frac{\sigma^2}{\sqrt{n}}. \quad (4.3.7b)$$

Для k -го значения функции корреляции $k > 1$ ожидается флуктуация:

$$\hat{r}_k = \pm \frac{n-k}{n} \frac{\sigma^2}{\sqrt{n}}. \quad (4.3.7)$$

На рис. 4.3 показан график значений \hat{r}_k , рассчитанный по выражению (4.3.7) для частного набора случайных чисел. Теперь можно было бы предположить, что при $n \rightarrow \infty$ флуктуационные составляющие исчезают и выражение (4.3.2) принимает предельную

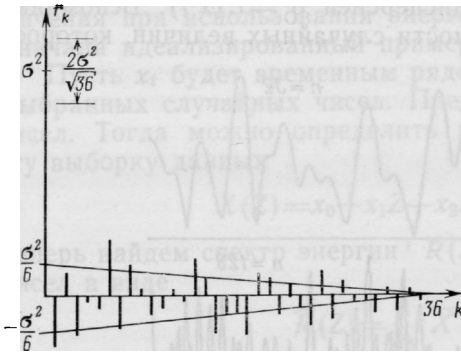


РИС. 4.3. Положительные запаздывания автокорреляционной функции для 36 случайных чисел

форму: $\hat{R} = \sigma^2$. Однако такое заключение неверно по той причине, что, хотя каждая, отдельно взятая флуктуационная составляющая спадает как $1/\sqrt{n}$, выражение (4.3.6) предусматривает суммирование n таких составляющих. К счастью, эти составляющие случайным образом погашают друг друга, ввиду чего сумма не расходуется как \sqrt{n} . Напомним, что сумма n чисел со случайными знаками и единичными модулями является математическим ожиданием суммы случайных чисел и равна $\pm\sqrt{n}$. Таким образом, сумма (4.3.6) приводится к виду

$$\hat{R} \approx \left(1 \pm \frac{\sqrt{n}}{\sqrt{n}}\right) \sigma^2 = (1 \pm 1) \sigma^2. \quad (4.3.8)$$

Это — основной результат, заключающийся в том, что энергетический спектр, вычисленный на основании плотности энергии выборки случайных чисел, флуктуирует от частоты к частоте и от выборки к выборке так же сильно, как и ожидаемый спектр.

Должно быть понятно, что при $n \rightarrow \infty$ мы не приходим к теоретическому результату $\hat{R} = \sigma^2$. Проблема заключается в том, что при увеличении n растет разрешение по частоте, а не статистическое разрешение. Для увеличения статистического разрешения необходимо осуществить усреднение по ансамблю. Существуют два способа усреднения: 1) взять выборку из n точек и разбить ее на k сегментов равной длины по n/k точек в каждом, а затем вычислить $R(\omega)$ для каждого сегмента и просуммировать $R(\omega)$, полученные для всех k ; 2) определить $R(\omega)$ для выборки из n точек независимых амплитуд, заменив каждую из них средней

по k точек в их ближайшей окрестности. При использовании любого из двух способов будет установлено, что $\Delta f = 0,5k/n\tau$ и $(\Delta P/P)^2 =$ обратному числу степеней свободы, в среднем равному $1/k$. Таким образом, имеем

$$\Delta f \left(\frac{\Delta p}{p} \right)^2 = \frac{1}{2n\tau}.$$

Если некоторые данные не использовались или использовались неэффективно, то получим обычное условие:

$$\Delta f \left(\frac{\Delta p}{p} \right)^2 \geq \frac{1}{2n\tau}.$$

При наличии в нашем распоряжении достаточного числа данных (n достаточно велико) можно получить любое желаемое статистическое разрешение. С другой стороны, увеличение статистического разрешения достигается за счет уменьшения разрешения по частоте и наоборот.

Таким образом, вы установили правила обмена разрешения не только между величинами Δf и Δp , но также и с величиной $\Delta t = n\tau$, характеризующей продолжительность выборки данных во времени. Полагая теперь, что продолжительность выборки данных во времени определяется величиной $\Delta t = n\tau$, получим неравенство:

$$\Delta f \Delta t \left(\frac{\Delta p}{p} \right)^2 > \frac{1}{2}. \quad (4.3.9)$$

Это неравенство будет объяснено в дальнейшем. В следующем разделе оно будет получено заново на основе другого подхода.

В анализе временных рядов существует понятие когерентности, которое аналогично понятию корреляции, определенному в разделе 4.2. Там было установлено, что функция корреляции для двух случайных переменных x и y имеет вид

$$c = \frac{E(xy)}{[E(x^2)E(y^2)]^{1/2}}$$

Если x_t и y_t — временные ряды, то они могут иметь между собой соотношение, которое зависит от временной задержки (сдвига между ними), выбора масштаба или даже фильтрации. Например, возможно, что $Y(Z) = F(Z)X(Z) + N(Z)$, где $F(Z)$ — фильтр, а n_t — независимый шум. Обобщение понятия корреляционной функции приводит к следующему определению функции когерентности:

$$C = \frac{E \left[X \left(\frac{1}{Z} \right) Y(Z) \right]}{[E(\bar{X}X)E(\bar{Y}Y)]^{1/2}}.$$

Если функция корреляции является действительной и скалярной, то функция когерентности — комплексная и выражает зависимость функции корреляции от частоты. При формировании оценки когерентности всегда существенное значение имеет усреднение по не-

которому ансамблю. Отметим, что если усреднение по ансамблю будет опущено, то вычисление квадрата функции когерентности даст следующий результат:

$$|C|^2 = \overline{CC} = \frac{(\overline{XY})(\overline{XY})}{(\overline{XX})(\overline{YY})} = +1,$$

который констатирует, что квадрат функции когерентности равен +1 независимо от данных. Поскольку функция корреляции практически всегда имеет ненулевые значения, то, следовательно, и квадрат функции когерентности также смещен от нуля.

4.4. ЧАСТОТНО-ВРЕМЕННОЕ РАЗРЕШЕНИЕ С УЧЕТОМ СТАТИСТИКИ

Многие временные функции не являются совершенно случайными при переходе от точки к точке, но становятся все более случайными, чем длиннее отрезок времени, на котором они рассматриваются. Популярной математической моделью, воплощающей это свойство, является формирование так называемых *стационарных*

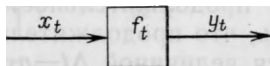


РИС. 4.4. Модель генерирования случайных временных рядов. x_t — случайные числа; f_t — фильтр; y_t — случайные временные ряды

временных рядов путем пропускания случайных чисел через фильтр, как это показано на рис. 4.4. Величины x_t на входе фильтра могут быть *независимыми случайными числами* или «белым светом»¹. Эти два термина означают на практике почти одно и то же, но первый термин точнее; он означает, что x_t никак не связано с x_s , если $t \neq s$, тогда как термин «белый свет» означает, что $E(x_t x_s) = 0$, если $t \neq s$. Если f_t — фильтр нижних частот, то случайный временной ряд y_t на выходе фильтра будет изменяться от точки к точке значительно медленнее, чем исходные случайные числа. Это обычный случай при формировании непрерывных функций времени. Полученный случайный временной ряд может быть назван *стационарным случайным временным рядом*, если этот фильтр и свойства случайных чисел (такие как среднее значение m или дисперсия σ) не изменяются во времени. Стационарность часто предполагается даже тогда, когда она не может быть строго выполнена.

Эта модель будет полезна позднее при рассмотрении задачи предсказания будущей точки для временного ряда на основе знания прошлых значений ряда. Сейчас мы используем ее для полу-

¹ Автор применяет термин «белый свет» (white light) для обозначения случайного набора синусоидальных колебаний, ограниченных во времени. (Прим. перев.)

чения спектра временного ряда y_t , заданного выборкой из n точек ряда y_t . Пусть для начала задано очень точное значение для спектра временного ряда y_t . Тогда

$$Y(Z) = F(Z)X(Z) \quad (4.4.1)$$

и сопряженное с ним

$$\bar{Y}\left(\frac{1}{Z}\right) = \bar{F}\left(\frac{1}{Z}\right)\bar{X}\left(\frac{1}{Z}\right). \quad (4.4.2)$$

Перемножая выражения (4.4.1) и (4.4.2), получим

$$\bar{Y}\left(\frac{1}{Z}\right)Y(Z) = \bar{X}\left(\frac{1}{Z}\right)X(Z)\bar{F}\left(\frac{1}{Z}\right)F(Z). \quad (4.4.3)$$

В предыдущем разделе установлено, что $E(\bar{X}X) = \sigma^2$. Полагая дисперсию σ^2 равной единице, получим, что ожидаемый энергетический спектр выходного сигнала Y является энергетическим спектром фильтра F . Общая ситуация изображена на рис. 4.5. Важный

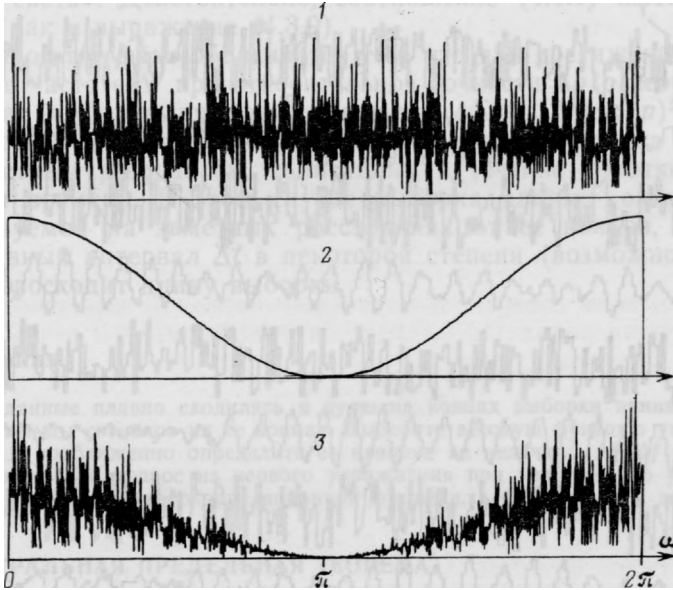


РИС. 4.5. Определение спектра

1 — амплитудный спектр выборки «белого» шума; 2 — амплитудный спектр фильтра; 3 — несглаженный энергетический спектр выборки данных на выходе фильтра

вопрос состоит в том, насколько хорошо можно определить спектр, имея выборку из n точек временного ряда y_t . Рассмотрим три способа с различными вычислительными схемами, обладающие однако одними и теми же основными ограничениями.

Первый способ использует последовательность фильтров, как показано на рис. 4.6. Когда случайные числа возбуждают узко-

полосный фильтр, сигнал на его выходе отчасти подобен синусоиде. Вместе с тем он отличается от нее одним важным свойством. Синусоида имеет постоянную амплитуду, а амплитуда выходного сигнала узкополосного фильтра меняется в широком диапазоне. Это иллюстрируется рис. 4.7. При узкой полосе пропускания

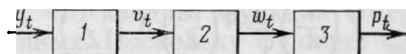


РИС. 4.6. Определение спектра случайного ряда.

y_t — временной ряд; v_t — выход фильтра; w_t — весовые функции; p_t — искомый спектр; 1 — узкополосный фильтр; 2 — квадратор; 3 — интегратор с затуханием

фильтра амплитуда изменяется медленно. Если импульсная реакция фильтра имеет длительность Δt_Φ , то амплитуда выходного сигнала в момент времени t будет случайным образом связана с амплитудой в момент времени $t + \Delta t_\Phi$. Таким образом, при стати-

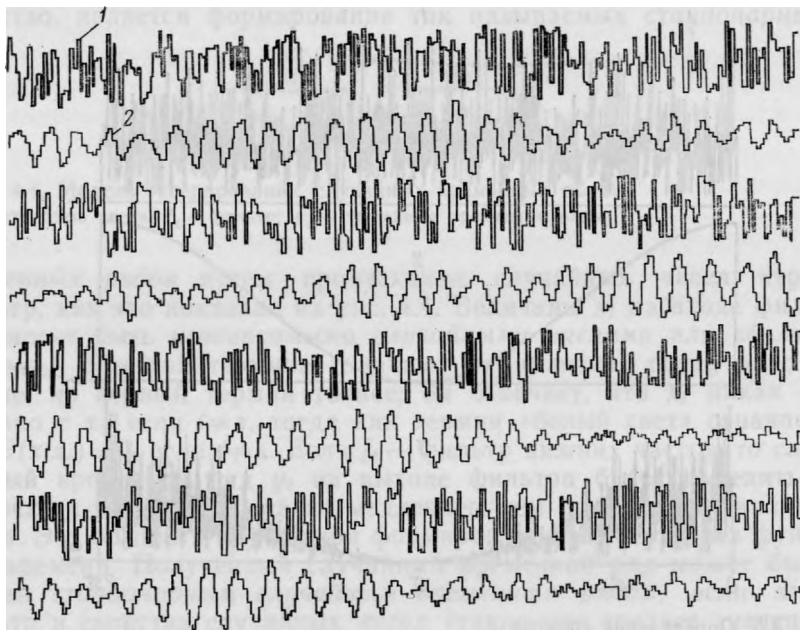


РИС. 4.7. Выборка из 1024 случайных чисел до (1) и после (2) узкополосной фильтрации. Фильтр описывается выражением $(1 - Z)/(1 - Z/Z_0)(1 - \bar{Z}_0)$ при $Z_0 = 1,02e^{i\pi/5}$

стическом усреднении надежность среднего значения увеличивает не количество временных точек, а число интервалов Δt_Φ . Следовательно, время затухания интегратора Δt_Φ следует выбирать большим, чем $\Delta t_\Phi = 1/\Delta f$. Изменчивость Δp выходного сигнала p умень-

шается с увеличением Δt_n . Поскольку величина v_t имеет независимые значения на временных интервалах, равных примерно $\Delta t_{\Phi} = 1/\Delta f$, то «степени свободы», сглаженные по всем диапазонам, можно записать как $\Delta t_n/\Delta t_{\Phi} = \Delta f \Delta t_n$. Относительная изменчивость $\Delta p/p$ обратно пропорциональна квадратному корню из числа степеней свободы и, следовательно, описывается выражением

$$\left(\frac{\Delta p}{p}\right)^2 = \frac{1}{\Delta f \Delta t_n}.$$

Представляя это выражение в форме обычного неравенства, имеем

$$\Delta t \Delta f \left(\frac{\Delta p}{p}\right)^2 > 1. \quad (4.4.4)$$

Неравенство (4.4.4) указывает на неопределенность трех параметров, которая является основой при оценке энергетических спектров случайных функций. Два способа определения энергетического спектра временного ряда y_t по выборке длины n являются такими же, как и описанные в разделе 4.3 способы определения спектра «белого света». Действительно, соотношение (4.4.4) имеет такой же вид, как и выражение (4.3.9).

Обычно этот факт объясняется тем, что для достижения разрешения по частоте Δf при относительной точности $\Delta p/p$ необходима временная выборка продолжительности $\Delta t \geq 1/[\Delta f (\Delta p/p)^2]$. Хотя эта интерпретация является корректной, она непригодна для рядов с острым резонансом, записанных в течение короткого промежутка времени. В этом случае выборка данных оказывается предсказуемой на заметных расстояниях от ее концов, так что эффективный интервал Δt в некоторой степени (возможно, заметной) превосходит длину выборки.

Упражнения

1. Чтобы данные плавно сходились к нулю на концах выборки принято уменьшать амплитуды сигналов на ее концах. Выберите весовую функцию уменьшения амплитуд и приближенно определите ее влияние на величины Δt , Δf и $(\Delta p/p)^2$.
2. Ответьте на вопрос из первого упражнения при условии, что уменьшены амплитуды на концах автокорреляционной функции, а не у выборки данных.

4.5. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Центральная предельная теорема является, вероятно, самой важной теоремой в теории вероятностей и статистики. Вывод центральной предельной теоремы объясняет, почему гауссова функция вероятностей так часто встречается в практике; не только в физике, но также и в биологических, и в общественных науках. Не может быть ученого-экспериментатора, не знающего основных идей, вытекающих из этой теоремы. Хотя этот результат очень сложен и даже в настоящее время является предметом активных исследований, основную идею можно изложить достаточно просто.

Один из способов получения случайных целых чисел с помощью известной функции вероятности состоит в том, чтобы записать целые числа на длинных узких полосках бумаги и поместить их в шляпу. Затем нужно вытаскивать по одной бумажке за 1 раз. После этого каждый раз кладите бумажку обратно в шляпу. Вероятность вытащить целое число i задается отношением числа бумажек a_i , содержащих целое число i , к общему числу бумажек. Очевидно, что сумма по i чисел a_i должна быть единичной. Другой способ получения случайных чисел заключается в том, чтобы подбрасывать одну из пары игральных костей. Тогда все a_i равны нулю, кроме $a_1=a_2=a_3=a_4=a_5=a_6=\frac{1}{6}$. Вероятность выпадения после первого броска целого числа i , после второго броска — целого числа j равна $a_i a_j$. Если же вытаскивать по две бумажки или подбрасывать пару костей, то вероятность равенства суммы двух целых чисел i и j величине k определяется выражением

$$c_k = \sum_i a_i a_{k-i}. \quad (4.5.1)$$

Так как это уравнение является сверткой, то можно рассмотреть смысл следующего Z -преобразования:

$$A(Z) = \dots + a_{-1}Z^{-1} + a_0 + a_1Z + a_2Z^2 + \dots \quad (4.5.2)$$

В терминах Z -преобразования вероятность выполнения равенства $i+j=k$ служит коэффициентом при Z^k в выражении

$$C(Z) = A(Z)A(Z). \quad (4.5.3)$$

Очевидно, если сложить n случайных чисел, то вероятность того, что их сумма равняется k , задается коэффициентом при Z^k в следующем уравнении:

$$G(Z) = [A(Z)]^n. \quad (4.5.4)$$

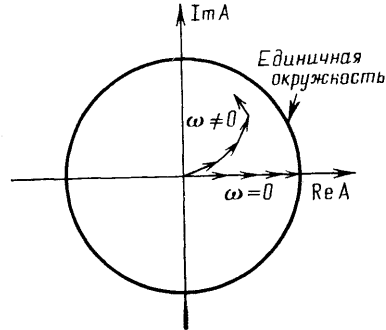
Центральная предельная теорема теории вероятностей гласит, что при $n \rightarrow \infty$ многочлен $G(Z)$ принимает особый вид. Практически независимо от характера многочлена $A(Z)$ график коэффициентов многочлена $G(Z)$ с ростом n все более приближается к огибающей гауссовой функции. Посмотрим, почему это происходит. Наше рассмотрение будет математически недостаточно строгим, так как теорема не всегда справедлива. Существуют аномальные функции A , которые не приводятся к функциям G , стремящимся к гауссовым функциям. Некоторые из аномальных функций иногда внезапно появляются в приложениях, однако не будем рассматривать такие примеры.

Рассмотрим функцию $A(Z)$ для действительных частот ω . Если $\omega=0$, то сумма членов функции $A(Z)$ может быть наглядно представлена на комплексной плоскости как сумма векторов $a_k \exp i\omega_k$, причем все они направлены вдоль оси действительных значений.

Если $\omega \neq 0$, то векторы указывают различные направления (рис. 4.8).

При возведении величины $A(e^{i\omega})$ в n -ю степень значения ω , представляющие наибольший интерес, располагаются около частоты $\omega=0$, в области которой A имеет наибольшее значение. Дело в том, что в любой области, где A мало, A^n будет крайне мало.

РИС. 4.8. Сложение комплексных чисел $a_k e^{i\omega k}$



Вблизи частоты $\omega=0$ или при $Z=1$ можно разложить $A(Z)$ в степенной ряд по степеням ω :

$$A(e^{i\omega}) = A|_0 + \frac{\partial A}{\partial \omega} \Big|_0 \omega + \frac{\partial^2 A}{\partial \omega^2} \Big|_0 \frac{\omega^2}{2!} + \dots \quad (4.5.5)$$

Заметим, что коэффициенты этого степенного ряда пропорциональны моментам m_i функции вероятностей, т. е.

$$A(e^{i\omega}) = \sum_k a_k e^{ik\omega}; \quad (4.5.6)$$

$$A(1) = \sum_k a_k = m_0 = 1; \quad (4.5.7)$$

$$\frac{\partial A}{\partial \omega} = \sum_k ika_k e^{i\omega k}; \quad (4.5.8)$$

$$\frac{\partial A}{\partial \omega} \Big|_0 = \sum_k ika_k = im_1; \quad (4.5.9)$$

$$\frac{\partial^2 A}{\partial \omega^2} \Big|_0 = - \sum_k k^2 a_k = -m_2. \quad (4.5.10)$$

При возведении многочлена $A(Z)$ в n -ю степень будем полагать, что только первые три члена заданного разложения в степенной ряд будут важными. (Это предположение становится несостоятельным, если любой из моментов функции вероятности равен бесконечности). Таким образом, к функции G применимы только наиболее важные свойства функции A , а именно: ее среднее значение $m=m_1$ и второй момент m_2 . Если это действительно так, то

можно вычислить функцию G , заменив функцию A некоторой функцией B , имеющей то же самое среднее значение и тот же второй момент, что и функция A . Можно использовать для этого простейшую найденную функцию. Примером такой функции может служить так называемая биномиальная функция вероятности, заданная в виде

$$B = \frac{Z^m (Z^\sigma + Z^{-\sigma})}{2}; \quad (4.5.11)$$

$$B = \frac{e^{i(m+\sigma)\omega} + e^{i(m-\sigma)\omega}}{2}. \quad (4.5.12)$$

Проверим ее первый и второй моменты:

$$\frac{\partial B}{\partial \omega} = \frac{i[(m+\sigma)e^{i(m+\sigma)\omega} + (m-\sigma)e^{i(m-\sigma)\omega}]}{2}; \quad (4.5.13)$$

$$\left. \frac{\partial B}{\partial \omega} \right|_0 = im; \quad (4.5.14)$$

$$\left. \frac{\partial^2 B}{\partial \omega^2} \right|_0 = -\frac{(m+\sigma)^2 + (m-\sigma)^2}{2}; \quad (4.5.15)$$

$$\left. \frac{\partial^2 B}{\partial \omega^2} \right|_0 = -(m^2 + \sigma^2). \quad (4.5.16)$$

Следовательно, σ следует выбрать так, чтобы удовлетворить условию

$$m_2 = m^2 + \sigma^2. \quad (4.5.17)$$

Нельзя ожидать, что m и σ обязательно окажутся целыми числами, поэтому выражение (4.5.11) не обязательно будет Z -преобразованием в обычном смысле. В действительности это не имеет значения, так как выражение (4.5.11) объясняется следующим образом:

1) вероятность получения числа $m+\sigma$ равна $1/2$;

2) вероятность получения числа $m-\sigma$ равна $1/2$;

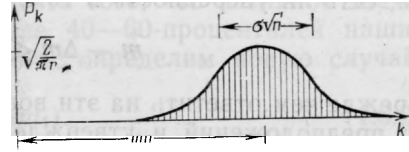
3) вероятность получения любого другого числа равна нулю.

Теперь, возводя выражение $(Z^\sigma + Z^{-\sigma})$ в n -ю степень, получим ряд степеней Z^σ , коэффициенты при которых распределены симметрично относительно Z в нулевой степени и имеют модули, заданные биномиальными коэффициентами. Коэффициенты многочлена $B^n(Z)$ представлены на рис. 4.9.

Покажем, что для больших n биномиальные коэффициенты асимптотически приближаются к гауссовым. Приближение к этому пределу является несколько сложным. Очевидно, сумма случайных целых чисел будет расходиться как \sqrt{n} . В то же время коэффициенты при степенях Z в выражении $\left(\frac{1}{2} + Z/2\right)^n$ уменьшаются по мере увеличения числа коэффициентов (с ростом n). Напом-

ним, что в анализе временных рядов использовалась подстановка $Z = \exp(i\omega\Delta t)$. Обычно мы выбираем $\Delta t = 1$, что имеет смысл квантования данных в целочисленные моменты времени. В случае применения Z -преобразования в теории вероятностей выбор $\Delta t = 1$ обусловлен тем, что выбранные случайно числа, записанные на полосках бумаги, были целыми. Теперь надо просуммировать n

РИС. 4.9. Коэффициенты многочлена $[Z^n (Z^z + Z^{-z})/2]^n$



этих случайных чисел. Целесообразно привести их к новому масштабу путем деления каждого из них на \sqrt{n} . Это можно сделать с помощью подстановки

$$Z = \exp(i\omega\Delta t) = \exp(i\omega/\sqrt{n}).$$

Коэффициент при Z^k дает вероятность получения числа k/\sqrt{n} . Возводя выражение $(Z^z + Z^{-z})/2$ в n -ю степень для определения функции распределения вероятностей суммы n независимо выбранных чисел, получим:

$$[B(Z)]^n = \left(\frac{Z^z + Z^{-z}}{2}\right)^n = \left(\cos \frac{\sigma\omega}{\sqrt{n}}\right)^n.$$

Ограничиваясь первым членом разложения косинуса в ряд, имеем:

$$[B(Z)]^n \approx \left(1 - \frac{\sigma^2\omega^2}{2n}\right)^n.$$

Используя хорошо известный факт, что $(1+x/n)^n \rightarrow e^x$, имеем для больших n :

$$[B(Z)]^n \approx e^{-\sigma^2\omega^2/2}. \quad (4.5.18)$$

Вероятность того, что число t будет результатом суммирования, теперь находится с помощью обратного преобразования Фурье функции (4.5.18). Результат преобразования Фурье гауссовой функции (4.5.18) можно найти по таблице интегралов. Он также является гауссовой функцией вида

$$P(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2\sigma^2}.$$

4.6. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

Всегда важно иметь некоторое представление о случайных погрешностях и об их влиянии на результаты исследований. При чем необходимо связать это представление с другими, например,

в форме следующего утверждения:

$$m = \hat{m} \pm \sigma/\sqrt{n}.$$

При спорном вопросе может возникнуть необходимость определения вероятности того, что правильное среднее значение случайной величины лежит в установленном интервале, или, другими словами, есть ли уверенность в том, что m лежит в интервале

$$\hat{m} - \Delta m < m < \hat{m} + \Delta m.$$

Прежде чем ответить на эти вопросы, необходимо сделать несколько предположений и утверждений относительно функции вероятности, которым подчиняются случайные погрешности. Утверждение, что погрешности независимы друг от друга, наиболее рискованное. Если они не являются независимыми, что бывает наиболее часто, то можно исправить числовую величину n , являющуюся оценкой числа независимых погрешностей. Мы уже делали нечто подобное при анализе временных рядов, когда выбирали в качестве величины n не число точек временного ряда, а число интервалов с длительностью Δt . Вторая большая опасность при определении доверительного интервала заключается в предположении, что вследствие центральной предельной теоремы теории вероятностей и недостатка лучшей информации погрешности распределены в соответствии с гауссовой функцией вероятности. Если в действительности погрешности данных включают грубые погрешности, обусловленные, например, просчетами человека или связанные с трудностями исключения влияния переходных процессов в электронном оборудовании, то предположение о распределении погрешностей в соответствии с гауссовой функцией вероятности является совершенно неправильным и может привести к серьезным ошибкам интерпретации геофизических данных. Некоторые полезные приемы устранения такого рода погрешностей найдены в области непараметрической статистики [18].

Полезно пересмотреть вопрос о среднем с целью замены понятия среднего более предпочтительным понятием о медиане. Медиана m_1 определяется как величина, которая предположительно будет меньше величин одной половины совокупности случайных чисел и больше величин другой половины этой совокупности. Во многих случаях медиана практически заменяет среднее арифметическое. Медиана нечувствительна к отсчетам данных, которые из-за грубой погрешности стремятся к бесконечности. Действительно, медиана и среднее случайной функции совпадают, когда функция вероятности симметрична. Для оценки медианы m_1 выборки из n чисел (x_i , $i=1, 2, \dots, n$) можно переписать эти числа, расставив их в порядке возрастания, а потом выбрать число из середины этой последовательности, считая его оценкой медианы \hat{m}_1 . В частности, пусть переупорядоченные случайные числа x_i обозначены через x'_i , где $x'_i \leq x'_{i+1}$. Тогда $\hat{m}_1 = x'_{n/2}$. Таким образом, без знания функции распределения плотности вероятности для

случайных переменных x_i , можно вычислить вероятность того, что истинное значение медианы m_1 находится в интервале

$$x'_{n/2-\alpha\sqrt{n}} < m_1 < x'_{n/2+\alpha\sqrt{n}}. \quad (4.6.1)$$

Например, пусть $\alpha=1$ и $n=100$. Можно утверждать, что теперь мы в состоянии рассчитать вероятность того, что истинное значение медианы m_1 лежит в интервале 40—60-перцентилей наших данных. Прием состоит в следующем: определим новую случайную переменную

$$y=f_{\text{ст}}(x-m_1). \quad (4.6.2)$$

Ступенчатая функция $f_{\text{ст}}=+1$, если $x>m_1$, и равна нулю, если $x<m_1$. Новая случайная переменная y с равной вероятностью принимает только два значения, равные нулю и единице. Таким образом, известна ее функция вероятности, хотя, возможно, неизвестна функция вероятности для случайной переменной x . Определим

$$s = \sum_{i=1}^n y_i.$$

Так как каждое y_i является нулем или единицей, то s должно быть целым числом, заключенным между нулем и n . Кроме того, вероятность того, что случайная переменная s принимает значение j , задается коэффициентом при Z^j полинома $\left(\frac{1}{2} + Z/2\right)^n$.

Теперь вероятность того, что s лежит в интервале $n/2-\alpha\sqrt{n} < s < n/2+\alpha\sqrt{n}$, легко определяется путем суммирования требуемых коэффициентов при степенях Z^j . Эта вероятность по определению равна вероятности того, что медиана m_1 лежит в интервале, заданном выражением (4.6.1). Для α , равных 1 и больших n , эта вероятность близка к 95%.

5. МАТРИЦЫ И МНОГОКАНАЛЬНЫЕ ВРЕМЕННЫЕ РЯДЫ

Знание матриц имеет большое значение для моделирования на ЭВМ как в физических, так и в общественных науках. Так как этому важному предмету посвящены прекрасные учебники¹, то наш обзор будет кратким. Обратим внимание лишь на те свойства матриц, которые понадобятся в последующих главах. Будем избегать доказательств, и, хотя приводимые данные должны быть полезны для наибольшего числа случаев, в них будут встречаться специальные матрицы (которые мы будем опускать как аномальные), для которых наши построения будут неприменимыми. На практике необходимо всегда проверять результаты вычислений. К сожалению, так называемые аномальные случаи возникают на практике довольно часто. При возникновении затруднений с матрицами первой реакцией исследователей является попытка увеличить точность вычислений. Однако опыт автора свидетельствует о том, что физически значимые расчеты редко требуют высокой точности. Иногда кажется, что необходима более высокая точность. Это часто является следствием того, что задача плохо поставлена физически. В корректных задачах при другой организации вычислений высокая точность может не требоваться. Так или иначе, в дальнейшем будем рассматривать нормальные случаи, но читатель должен помнить, что на практике могут встретиться аномальные случаи, которые станут серьезной проверкой его математических знаний и проницательности (интуиции).

5.1. ОБЗОР МАТРИЦ

Систему совместных уравнений можно записать в виде

$$\mathbf{Ax}=\mathbf{b}, \quad (5.1.1a)$$

где \mathbf{A} — квадратная матрица (прямоугольные матрицы рассматриваются в главе 6; \mathbf{x} и \mathbf{b} — векторы-столбцы. В случае матрицы \mathbf{A} размерностью 2×2 выражение (5.1.1a) имеет вид

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}. \quad (5.1.1b)$$

Уравнение (5.1.1b) имеет простую геометрическую интерпретацию. Оба столбца матрицы и столбец \mathbf{b} могут рассматриваться как векторы на плоскости. Уравнение (5.1.1b) показывает, что

¹ На русском языке по этому вопросу можно рекомендовать читателю книги: 1. Беллман Р. Введение в теорию матриц. М., Наука, 1976. 2. Гантмахер Ф. Р. Теория матриц. М., Наука, 1967. (Прим. перев.)

сумма произведений x_1 на первый вектор-столбец и x_2 на второй вектор-столбец равна вектор-столбцу \mathbf{b} . Затруднение возникает, когда оба вектор-столбца матрицы указывают одинаковое направление. Если вектор \mathbf{b} имеет другое направление, то решение для x_1 и x_2 найти невозможно. Это относится и к общему случаю. Вектор решения \mathbf{x} уравнения (5.1.1а) существует, если вектор \mathbf{b} лежит в промежутке, образуемом столбцами матрицы \mathbf{A} . В большинстве практических случаев матрица \mathbf{A} и вектор-столбец \mathbf{b} независимы, ввиду чего разумнее потребовать, чтобы столбцы матрицы \mathbf{A} образовывали пространство, которое включало бы произвольный вектор \mathbf{b} . Если матрица \mathbf{A} имеет размерность $n \times n$, то ее столбцы должны образовывать n -мерное пространство. В частности, n -размерный параллелепипед с ребрами, задаваемыми столбцами матрицы \mathbf{A} , не должен иметь объем, равный нулю. Этот объем равен определителю матрицы \mathbf{A} .

Другим примером системы совместных уравнений, которые часто возникают на практике, являются однородные уравнения:

$$\mathbf{Ax} = \mathbf{0}. \tag{5.1.2}$$

Эта система всегда имеет решение $\mathbf{x} = \mathbf{0}$, которое часто называют тривиальным. Чтобы уравнение (5.1.2) имело нетривиальное решение \mathbf{x} , определитель матрицы \mathbf{A} должен стремиться к нулю, показывая, что столбцы матрицы \mathbf{A} не образуют n -мерного пространства. Ниже мы вернемся к решению систем совместных уравнений.

Наиболее полезной особенностью матриц является то, что их элементы могут быть не только числами, но и матрицами. Поэтому большие матрицы можно расчленить на более мелкие подматрицы. Неожиданным является то, что произведение двух матриц не зависит от того, были ли матрицы расчленены или нет. Рассмотрим тождество

$$\left[\begin{array}{cc|c} a & b & c \\ d & e & f \end{array} \right] \left[\begin{array}{c} g & h \\ i & j \\ \hline k & l \end{array} \right] = \left[\begin{array}{cc} a & b \\ d & e \end{array} \right] \left[\begin{array}{cc} g & h \\ i & j \end{array} \right] + \left[\begin{array}{c} c \\ f \end{array} \right] [kl]. \tag{5.1.3}$$

В обозначениях суммирования левая часть выражения (5.1.3) имеет вид

$$C_{ik} = \sum_{j=1}^3 A_{ij} B_{jk}. \tag{5.1.4}$$

В то же время правая часть означает

$$C_{ik} = \sum_{j=1}^2 A_{ij} B_{jk} + \sum_{j=3}^3 A_{ij} B_{jk}. \tag{5.1.5}$$

Равенства (5.1.4) и (5.1.5) являются одинаковыми; это показывает, что расчленение матричного произведения просто изменяет порядок членов. Расчленение матриц реально ничего не дает с математической точки зрения, но оно крайне важно с точки зрения вычислений или обсуждения.

Теперь мы используем расчленение матриц для рассмотрения метода окаймления, применяемого для обращения матриц. Метод окаймления не самый быстрый и не самый точный, но он достаточно прост даже для несимметричных матриц с комплексными величинами. Кроме того, он позволяет найти определитель матрицы и применим к однородным уравнениям. Метод окаймления реализуется с помощью рекурсии. Пусть задана матрица, обратная матрице размерностью $k \times k$. Этот метод позволяет найти матрицу, обратную матрице размерностью $(k+1) \times (k+1)$, которая получена из предыдущей матрицы размерностью $k \times k$ путем присоединения дополнительных строки и столбца к ее границам. В частности, в приводимом ниже выражении (5.1.6) известными величинами являются \mathbf{A} , \mathbf{e} , \mathbf{f} , \mathbf{g} и \mathbf{A}^{-1} . Задача состоит в отыскании величин \mathbf{W} , \mathbf{x} , \mathbf{y} и \mathbf{z} :

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{f} \\ \hline \mathbf{e} & g \end{array} \right] \left[\begin{array}{c|c} \mathbf{W} & \mathbf{y} \\ \hline \mathbf{x} & z \end{array} \right] = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array} \right] \quad (5.1.6)$$

Прежде всего нужно перемножить выделенные в уравнении (5.1.6) части. Для произведения первого столбца имеем

$$\mathbf{AW} + \mathbf{fx} = \mathbf{I}; \quad (5.1.7)$$

$$\mathbf{eW} + \mathbf{gx} = 0. \quad (5.1.8)$$

При выборе в качестве \mathbf{W} функции

$$\mathbf{W} = \mathbf{A}^{-1}(\mathbf{I} - \mathbf{fx}) \quad (5.1.9)$$

выражение (5.1.7) становится тождеством. После этого \mathbf{x} все еще остается неизвестным, но можно найти его подстановкой формулы (5.1.9) в уравнение (5.1.8). Тогда

$$\mathbf{x} = \frac{\mathbf{eA}^{-1}}{\mathbf{eA}^{-1}\mathbf{f} - g}. \quad (5.1.10)$$

Для получения столбца с неизвестными \mathbf{y} и \mathbf{z} вычислим произведение для второго столбца выражения (5.1.6):

$$\mathbf{Ay} + \mathbf{fz} = 0; \quad (5.1.11)$$

$$\mathbf{ey} + \mathbf{gz} = 1. \quad (5.1.12)$$

Умножив выражение (5.1.11) на обратную матрицу \mathbf{A}^{-1} , получим:

$$\mathbf{y} = -\mathbf{A}^{-1}\mathbf{fz}. \quad (5.1.13)$$

Это дает вектор-столбец \mathbf{y} со скалярным множителем \mathbf{z} . Для определения этого множителя подставим выражение (5.1.13) в уравнение (5.1.12). Тогда

$$-\mathbf{eA}^{-1}\mathbf{fz} + \mathbf{gz} = 1.$$

Откуда

$$\mathbf{z} = \frac{1}{g - \mathbf{eA}^{-1}\mathbf{f}}. \quad (5.1.14)$$

Действительно, можно показать, что определитель обратной матрицы является произведением знаменателя выражения (5.1.14) на всех этапах метода окаймления. Следовательно, если в процессе выполнения рекурсии знаменатель выражения (5.1.14) в любой момент обращается в нуль, то матрица является сингулярной (вырожденной) и вычисления нельзя продолжать.

Рассмотрим кратко реализацию рекурсии: она начинается с члена, расположенного в верхнем левом углу матрицы. Это скалярная величина, и ее обращение простое. Затем рассматривается

```

SUBROUTINE CMAINE(N,B,A)
C   A=MATRIX INVERSE OF B
      COMPLEX B,A,C,R,DEL
      DIMENSION A(N,N),B(N,N),R(100),C(100)
      DO 10 I=1,N
      DO 10 J=1,N
10   A(I,J)=0.
      DO 40 L=1,N
      DEL=B(L,L)
      DO 30 I=1,L
      C(I)=0.
      R(I)=0.
      DO 20 J=1,L
      C(I)=C(I)+A(I,J)*B(J,L)
20   R(I)=R(I)+B(L,J)*A(J,I)
30   DEL=DEL-B(L,I)*C(I)
      C(L)=-1.
      R(L)=-1.
      DO 40 I=1,L
      C(I)=C(I)/DEL
      DO 40 J=1,L
40   A(I,J)=A(I,J)+C(I)*R(J)
      RETURN
      END

```

РИС. 5.1. Программа на Фортране для обращения матриц методом окаймления

часть, окаймленная строкой и столбцом, как это показано в выражении (5.1.6), и находится матрица, обратная к этой матрице размерностью 2×2 . Процесс продолжается, пока это необходимо. Типичным является вычисление величины z по формуле (5.1.14), а затем обратной матрицы A^{-1} с увеличенной на единицу размерностью:

$$A^{-1} \leftarrow \begin{bmatrix} A^{-1} \\ \text{нули} \end{bmatrix} + z \begin{bmatrix} A^{-1} f \\ -1 \end{bmatrix} [eA^{-1} - 1]. \quad (5.1.15)$$

Выражение (5.1.15) получено из формул (5.1.9), (5.1.10) и (5.1.13). Машинная программа на Фортране, выполняющая рассматриваемое преобразование, представлена на рис. 5.1.

Поучительно посмотреть, как ведет себя обратная матрица A^{-1} , если матрица A постепенно изменяется таким образом, что ее определитель вырождается. Если элемент g матрицы в формуле (5.1.6) приближается к $eA^{-1}f$, то из выражения (5.1.14) следует, что $z \rightarrow \infty$. Интересно отметить, что второй член выражения (5.1.15) при этом становится доминирующим над первым, а обратная матрица стремится к бесконечности, умноженной на произведение столбца c со строкой r .

Обычные формулы $\mathbf{A}\mathbf{A}^{-1}=\mathbf{I}$ или $\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$ для предельно малых z^{-1} можно представить в виде

$$\mathbf{A}c\mathbf{r}=z^{-1}\mathbf{I} \quad (5.1.16)$$

или

$$c\mathbf{r}\mathbf{A}=z^{-1}\mathbf{I}. \quad (5.1.17)$$

В обычном случае (если ранг матрицы $\mathbf{A}=n-1$, но не $\mathbf{A}<n-1$), когда ни c , ни r не стремятся тождественно к нулю, выражения (5.1.16) и (5.1.17) в пределе при $z^{-1}=0$ имеют вид

$$\mathbf{A}c=0; \quad (5.1.18)$$

$$c\mathbf{r}\mathbf{A}=0. \quad (5.1.19)$$

Для решения обычной системы совместных уравнений, подобных (5.1.1), можно сначала вычислить методом окаймления матрицу, обратную матрице \mathbf{A} , а затем умножить систему (5.1.1) на \mathbf{A}^{-1} . Тогда

$$\mathbf{x}=\mathbf{A}^{-1}\mathbf{b}. \quad (5.1.20)$$

Если $\mathbf{b}\rightarrow 0$, то найдем решение для однородных уравнений и будем ожидать, что z будет устремляться в бесконечность на последнем этапе процесса окаймления. (Если это случится раньше, то нужно уметь переупорядочить члены). Тогда решение дается столбцом c в выражении (5.1.18).

Ряд однородных уравнений (5.1.19) был введен потому, что такая система возникает естественным образом при отыскании собственных векторов-строк для несимметричных матриц. В следующем разделе рассмотрим свойства собственных векторов матриц подробнее. Собственный вектор-столбец c матрицы \mathbf{A} определяется решением уравнения

$$\mathbf{A}c=\lambda c, \quad (5.1.21)$$

где λ — собственное значение матрицы.

Одновременно рассмотрим также уравнение для собственных векторов-строк матрицы:

$$c\mathbf{A}=\lambda c. \quad (5.1.22)$$

Чтобы уравнение (5.1.21) или (5.1.22) имело решение, должно выполняться условие $\det(\mathbf{A}-\lambda\mathbf{I})=0$. После нахождения корней λ_j полинома $\det(\mathbf{A}-\lambda\mathbf{I})=0$ можно сформировать новую матрицу \mathbf{A}' для каждого λ_j :

$$\mathbf{A}'=\mathbf{A}-\lambda_j\mathbf{I}. \quad (5.1.23)$$

Тогда решение выражения

$$\mathbf{A}'\mathbf{x}=0 \quad (5.1.24)$$

получается из столбца c на последнем этапе метода окаймления. Это собственный вектор-столбец матрицы. Подобным же образом собственный вектор-строка матрицы является строкой последнего этапа метода окаймления.

Упражнения

1. Укажите размерности всех матриц в равенствах (5.1.7)—(5.1.14).

2. Покажите, как получить формулу (5.1.15) из выражений (5.1.9), (5.1.10), (5.1.13) и (5.1.14).

5.2. МАТРИЧНАЯ ТЕОРЕМА СИЛЬВЕСТРА

Теорема Сильвестра обеспечивает быстрый расчет матричных функций. Наиболее простыми, часто встречающимися матричными функциями являются \mathbf{A}^{-1} и \mathbf{A}^N (для больших N). Двумя другими матричными функциями, имеющими важное значение в теории распространения волн, являются $e^{\mathbf{A}}$ и $\mathbf{A}^{1/2}$. Прежде чем перейти к несколько абстрактному доказательству теоремы Сильвестра, приведем численный пример. Рассмотрим матрицу

$$\mathbf{A} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}. \quad (5.2.1)$$

Пусть требуется найти собственные векторы-столбцы и собственные значения этой матрицы; они определяются следующим образом:

$$\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad (5.2.2)$$

$$\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (5.2.3)$$

Поскольку матрица \mathbf{A} несимметрична, то ее собственные векторы-строки отличаются от векторов-столбцов и находятся следующим образом:

$$[-1 \ 2] \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} = 1 [-1 \ 2]; \quad (5.2.4)$$

$$[1 \ -1] \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} = 2 [1 \ -1]. \quad (5.2.5)$$

Запишем уравнения (5.2.2)—(5.2.5) в сокращенной форме:

$$\mathbf{A}\mathbf{c}_1 = \lambda_1 \mathbf{c}_1; \quad (5.2.6a)$$

$$\mathbf{A}\mathbf{c}_2 = \lambda_2 \mathbf{c}_2; \quad (5.2.6б)$$

$$\mathbf{r}_1 \mathbf{A} = \lambda_1 \mathbf{r}_1; \quad (5.2.6в)$$

$$\mathbf{r}_2 \mathbf{A} = \lambda_2 \mathbf{r}_2. \quad (5.2.6г)$$

Читатель видит, что при умножении \mathbf{r} или \mathbf{c} на произвольный масштабный коэффициент равенства (5.2.6) остаются справедливыми. Говорят, что собственные векторы матрицы нормализованы, если масштабные множители выбраны так, что $\mathbf{r}_1 \mathbf{c}_1 = 1$ и $\mathbf{r}_2 \mathbf{c}_2 = 1$. Впоследствии покажем, что $\mathbf{r}_1 \mathbf{c}_2 = 0$ и $\mathbf{r}_2 \mathbf{c}_1 = 0$. Этот общий результат будет дан в упражнениях.

Рассмотрим свойства матрицы

$$\mathbf{c}_1 \mathbf{r}_1 = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ & 2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}.$$

Эта матрица при возведении в любую степень равна сама себе. Например, ее квадрат

$$\begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}^2 = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}$$

Это свойство называется идемпотенцией. Это происходит потому, что $(\mathbf{c}_1 \mathbf{r}_1)(\mathbf{c}_1 \mathbf{r}_1) = \mathbf{c}_1(\mathbf{r}_1 \mathbf{c}_1) \mathbf{r}_1 = \mathbf{c}_1 \mathbf{r}_1$. То же самое, конечно, верно и для $\mathbf{c}_2 \mathbf{r}_2$. Отметим, что матрица $\mathbf{c}_1 \mathbf{r}_1$ «перпендикулярна» к матрице $\mathbf{c}_2 \mathbf{r}_2$:

$$\begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

так как векторы \mathbf{r}_2 и \mathbf{c}_2 перпендикулярны друг к другу.

Теорема Сильвестра гласит, что любую функцию f матрицы \mathbf{A} можно записать в виде

$$f(\mathbf{A}) = f(\lambda_1) \mathbf{c}_1 \mathbf{r}_1 + f(\lambda_2) \mathbf{c}_2 \mathbf{r}_2.$$

Рассмотрим простейший пример, когда $f(\mathbf{A}) = \mathbf{A}$:

$$\mathbf{A} = \lambda_1 \mathbf{c}_1 \mathbf{r}_1 + \lambda_2 \mathbf{c}_2 \mathbf{r}_2 = 1 \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} + 2 \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}. \quad (5.2.7)$$

Следующий пример

$$\mathbf{A}^2 = 1^2 \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} + 2^2 \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 7 & -6 \\ 3 & -2 \end{bmatrix}.$$

Обратная матрица

$$\mathbf{A}^{-1} = 1^{-1} \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} + 2^{-1} \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}.$$

Тождественность матрицы можно выразить через собственные векторы матрицы \mathbf{A} :

$$\mathbf{A}^0 = \mathbf{I} = 1^0 \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} + 2^0 \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Прежде чем рассматривать некоторые более сложные функции, посмотрим, что потребуется для доказательства теоремы Сильвестра. Нам понадобится один основной результат, который приводится во всех книгах по теории матриц и гласит, что большинство матриц можно привести к диагональному виду. В примере с матрицей размерностью 2×2 это утверждение принимает форму:

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \mathbf{A} [\mathbf{c}_1 | \mathbf{c}_2] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad (5.2.8)$$

где

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} [\mathbf{c}_1 | \mathbf{c}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (5.2.9)$$

Поскольку матрица перестановочна со своей обратной матрицей, то выражение (5.2.9) означает, что

$$[\mathbf{c}_1 | \mathbf{c}_2] \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (5.2.10)$$

Умножим выражение (5.2.8) справа на матрицу-строку, а слева на матрицу-столбец. Используя формулу (5.2.10), получим:

$$\mathbf{A} = [\mathbf{c}_1 | \mathbf{c}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix}. \quad (5.2.11)$$

Уравнение (5.2.11) представляет собой уравнение (5.2.7), но в скрытой форме, в чем легко убедиться, записав выражение (5.2.11) в следующем виде:

$$\begin{aligned} \mathbf{A} &= [\mathbf{c}_1 | \mathbf{c}_2] \left\{ \begin{bmatrix} \lambda_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \lambda_2 \end{bmatrix} \right\} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} = \\ &= [\mathbf{c}_1 | \mathbf{c}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} + [\mathbf{c}_1 | \mathbf{c}_2] \begin{bmatrix} 0 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} = \lambda_1 \mathbf{c}_1 \mathbf{r}_1 + \lambda_2 \mathbf{c}_2 \mathbf{r}_2. \end{aligned}$$

Теперь

$$\mathbf{A}^2 = (\lambda_1 \mathbf{c}_1 \mathbf{r}_1 + \lambda_2 \mathbf{c}_2 \mathbf{r}_2) (\lambda_1 \mathbf{c}_1 \mathbf{r}_1 + \lambda_2 \mathbf{c}_2 \mathbf{r}_2).$$

Учитывая ортонормированность величин $\mathbf{c}_1 \mathbf{r}_1$ и $\mathbf{c}_2 \mathbf{r}_2$, последнее выражение можно упростить:

$$\mathbf{A}^2 = \lambda_1^2 \mathbf{c}_1 \mathbf{r}_1 + \lambda_2^2 \mathbf{c}_2 \mathbf{r}_2.$$

Теперь ясно, как можно использовать выражение (5.2.11) для доказательства теоремы Сильвестра для любой полиномиальной функции матрицы \mathbf{A} . Понятно, что это справедливо для матриц не только размерностью 2×2 , но и размерностью $n \times n$. Аналогично можно рассмотреть функции матрицы \mathbf{A} в виде бесконечных рядов. Такое представление справедливо почти для любой функции. Рассмотрим трансцендентные функции, подобные синусоидальной, косинусоидальной, экспоненциальной функциям.

Экспоненциальные функции, естественно, возникают как решения дифференциальных уравнений. Рассмотрим дифференциальное уравнение в матричной форме

$$\frac{d}{dx} \mathbf{E} = \mathbf{A} \mathbf{E}. \quad (5.2.12)$$

Легко проверить, что решение этого уравнения может быть представлено в виде степенного ряда:

$$\mathbf{E} = \mathbf{I} + \mathbf{A}x + \frac{\mathbf{A}^2 x^2}{2!} + \dots$$

Это — степенной ряд, определяющий экспоненциальную функцию. Если матрица \mathbf{A} принадлежит к многочисленной группе матриц,

которые приводятся к диагональному виду, то экспоненциальную функцию значительно проще выразить с помощью теоремы Сильвестра. Для ранее рассмотренного числового примера имеем

$$\mathbf{E} = e^x \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} + e^{2x} \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}.$$

Эта экспоненциальная матрица является решением дифференциального уравнения (5.2.12) без учета граничных условий. В физике часто встречается дифференциальное уравнение

$$\frac{d}{dx} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{A} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad (5.2.13)$$

подчиненное двум граничным условиям, накладываемым на y_1 или на y_2 , либо на их сочетание. Можно проверить, что выражение

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = e^{\mathbf{A}x} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

является решением уравнения (5.2.13) при произвольных постоянных k_1 и k_2 , определяемых с помощью граничных условий. Отметим, что коэффициенты k_1 и k_2 являются величинами y_1 ($x=0$) и y_2 ($x=0$) соответственно.

Интересная ситуация возникает при определении квадратного корня матрицы. Матрица размерностью 2×2 , подобная матрице \mathbf{A} , будет иметь четыре квадратных корня, так как существует четыре возможных сочетания для выбора знаков плюс или минус для $\sqrt{\lambda_1}$ и $\sqrt{\lambda_2}$. В общем случае матрица размерностью $n \times n$ имеет 2^n квадратных корней. Важное применение этого факта будет описано в следующей главе, где рассматривается дифференциальный оператор $(k^2 + \partial^2/\partial x^2)^{1/2}$. Понятие квадратного корня из оператора редко объясняется в книгах, и немногие знают, что он означает. Лучший способ наглядно представить себе квадратный корень этого дифференциального оператора заключается в том, что нужно установить его связь с квадратным корнем матрицы \mathbf{M} , имеющей вид

$$\mathbf{M} = k^2 \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} + \frac{1}{(\Delta x)^2} \begin{bmatrix} ? & ? & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ & & & ? & ? \end{bmatrix}.$$

Здесь первая матрица является второй разностной аппроксимацией второй частной производной. Пусть

$$\mathbf{M} = k^2 \mathbf{I} + \mathbf{T}.$$

Желательно рассмотреть матрицу \mathbf{M} , обобщенную для случая очень большой размерности так, чтобы краевые эффекты можно было

считать минимальными. В общем, мы можем сделать матрицу \mathbf{M} сколь угодно большой размерности и для любой размерности получить 2^M квадратных корней. Практически нас будут интересовать только два основных квадратных корня: один — с положительными корнями всех собственных значений и второй — со всеми отрицательными корнями. Как найти эти «главные» квадратные корни? Значительный интерес представляет случай, когда можно использовать биномиальное разложение:

$$(k^2\mathbf{I} + \mathbf{T})^{1/2} = \pm k \left(\mathbf{I} + \frac{\mathbf{T}}{k^2} \right)^{1/2} = \pm k \left(\mathbf{I} + \frac{\mathbf{T}}{2k^2} - \frac{\mathbf{T}^2}{8k^4} + \dots \right).$$

Этот результат подтверждается простым возведением в квадрат предполагаемого квадратного корня. С другой стороны, он подтверждается теоремой Сильвестра. Следует отметить, что при возведении в квадрат предполагаемого квадратного корня используется тот факт, что матрицы \mathbf{I} и \mathbf{T} перестановочны. Таким образом, квадратный корень дифференциального оператора можно интерпретировать следующим образом:

$$\left(k^2 + \frac{\partial^2}{\partial x^2} \right)^{1/2} = k + \frac{1}{2k} \frac{\partial^2}{\partial x^2} + \dots,$$

если k не является функцией от x . Если k есть функция от x , то квадратный корень дифференциального оператора сохраняет тот же смысл, но не вычисляется так просто с помощью биномиального разложения.

Упражнения

1. Умножьте выражение (5.2.6в) справа на \mathbf{c}_2 и вычтите из результата уравнение (5.2.6б), слева умноженное на \mathbf{r}_1 . Является ли условие $\lambda_1 \neq \lambda_2$ необходимым для того, чтобы векторы \mathbf{r}_1 и \mathbf{c}_2 были перпендикулярными? Является ли это условие достаточным?

2. Докажите теорему Кейли — Гамильтона, состоящую в том, что если

$$0 = f(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = p_0 + p_1\lambda + p_2\lambda^2 + \dots + p_n\lambda^n,$$

то

$$f(\mathbf{A}) = p_0 + p_1\mathbf{A} + p_2\mathbf{A}^2 + \dots + p_n\mathbf{A}^n = 0.$$

3. Подтвердите, что в общем случае для матрицы \mathbf{A} размерностью 2×2 , для которой $\lambda_1 \neq \lambda_2$, выполняется условие $\mathbf{c}_1\mathbf{r}_1 = (\lambda_2\mathbf{I} - \mathbf{A})/(\lambda_2 - \lambda_1)$, где λ_1 и λ_2 — собственные значения матрицы \mathbf{A} . Какова общая формула для $\mathbf{c}_2\mathbf{r}_2$?

4. Можно показать, что у симметричных матриц всегда есть полный набор собственных векторов. Эта задача иногда возникает для несимметричных матриц. Исследуйте матрицу

$$\begin{bmatrix} 1 & 1 - \epsilon^2 \\ -1 & 3 \end{bmatrix}$$

при $\epsilon \rightarrow 0$, чтобы увидеть, почему у нее утерян один собственный вектор. Такая матрица называется неправильной. (Этот пример взят у Т. Р. Маддена.)

5. Многие задачи распространения волн в слоистых средах сводятся к следующему уравнению:

$$\frac{d}{dx} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & b \\ a & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Как зависит x от решения этого уравнения при положительном ab ? Когда ab отрицательно? Предположите, что a и b не зависят от x . Примените теорему Сильвестра. Что нужно сделать для получения вырожденной матрицы? Какое решение получится в случае вырожденной матрицы?

6. Рассмотрите матрицу следующей формы: $(\mathbf{I} + \mathbf{v}\mathbf{v}^T)^{-1}$, где \mathbf{v} — вектор-столбец, а \mathbf{v}^T — его транспонированный вектор. Найдите разложение в степенной ряд матрицы $(\mathbf{I} + \mathbf{v}\mathbf{v}^T)^{-1}$ по степеням $\mathbf{v}\mathbf{v}^T$. [Отметим, что $(\mathbf{v}\mathbf{v}^T)^N$ сжимается до $\mathbf{v}\mathbf{v}^T$, умноженного на масштабный множитель. Следовательно, этот степенной ряд значительно сокращается.]

7. Следующая матрица «векторного произведения» часто встречается в электродинамике. Пусть $\mathbf{B} = (B_x, B_y, B_z)$. Тогда

$$\mathbf{U} = \frac{1}{\sqrt{\mathbf{B} \cdot \mathbf{B}}} \begin{bmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{bmatrix}.$$

- Выпишите элементы $\mathbf{I} + \mathbf{U}^2$.
- Докажите, что $\mathbf{U}(\mathbf{I} + \mathbf{U}^2) = 0$ или $\mathbf{U}^3 = -\mathbf{U}$.
- Пусть \mathbf{v} — произвольный вектор. Куда направлены величины $\mathbf{U}\mathbf{v}$, $\mathbf{U}^2\mathbf{v}$ и $(\mathbf{I} + \mathbf{U}^2)\mathbf{v}$?
- Чему равны собственные значения матрицы \mathbf{U} ? [Указание. Используйте пункт б.]
- Почему соотношение $\mathbf{U}^3 = -\mathbf{U}$ нельзя сократить на \mathbf{U} ?
- Проверьте на идемпотентность следующие производные от матрицы \mathbf{U} :

$$\mathbf{c}_1\mathbf{r}_1 = [\mathbf{I} + \mathbf{U}^2]; \quad \mathbf{c}_2\mathbf{r}_2 = \frac{1}{2}(i\mathbf{U} - \mathbf{U}^2); \quad \mathbf{c}_3\mathbf{r}_3 = \frac{1}{2}(-i\mathbf{U} - \mathbf{U}^2).$$

5.3. МАТРИЧНЫЕ ФИЛЬТРЫ, СПЕКТРЫ И ФАКТОРИЗАЦИЯ

Два временных ряда значительно предпочтительнее одного вследствие возможности взаимодействия между ними. Общая линейная модель для двух рядов приведена на рис. 5.2.

Операцию фильтрации, представленную на этом рисунке, можно выразить как операцию перемножения матрицы на вектор, в которой элементы матрицы и вектора являются Z -преобразованиями полиномов. Тогда

$$\begin{bmatrix} Y_1(Z) \\ Y_2(Z) \end{bmatrix} = \begin{bmatrix} B_{11}(Z) & B_{12}(Z) \\ B_{21}(Z) & B_{22}(Z) \end{bmatrix} \begin{bmatrix} X_1(Z) \\ X_2(Z) \end{bmatrix}. \quad (5.3.1)$$

Один факт, который очевиден, но мало известен, заключается в том, что матрица, элементами которой являются полиномы, есть не что иное, как полином с матричными коэффициентами. Это иллюстрируется следующим примером:

$$\begin{bmatrix} 1 + Z + 2Z^2 & Z \\ 1 & 1 + Z^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} Z + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} Z^2.$$

Теперь мы можем обратиться к обратной задаче: если заданы фильтр \mathbf{B} и выходной сигнал \mathbf{Y} , то как найти входной сигнал \mathbf{X} ? Решение этой задачи аналогично решению задачи для отдельно

взятого временного ряда. Будем считать фильтр $\mathbf{V}(Z)$ полиномиальной матрицей. Известно, например, что обратной для любой матрицы размерностью 2×2 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ является матрица

$$\frac{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}{ad-bc}.$$

Следовательно, уравнение $\mathbf{Y}=\mathbf{VX}$ можно решить относительно \mathbf{X} следующим образом:

$$\mathbf{X}=\mathbf{V}^{-1}\mathbf{Y},$$

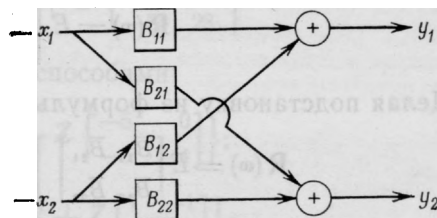
где

$$\mathbf{V}^{-1}=\begin{bmatrix} B_{22}(Z) & -B_{12}(Z) \\ -B_{21}(Z) & B_{11}(Z) \end{bmatrix} \frac{1}{B_{11}(Z)B_{22}(Z)-B_{12}(Z)B_{21}(Z)}.$$

Знаменатель последнего выражения является скалярной величиной. Скалярные знаменатели мы уже рассматривали раньше. Если все нули знаменателя лежат вне единичного круга, то для получения обратной матрицы можно использовать обычный степенной ряд. В противном случае решение не будет минимально фазовой функцией. Тогда необходимо использовать ряд Лорана.

При обобщении на многоканальные временные ряды матрица в числителе является сопряженной матрицей, а знаменатель — определителем. Сопряженная матрица может быть сформирована

РИС. 5.2. Общая модель линейной многоканальной фильтрации (на примере двух временных рядов x_1 и x_2 , вводимых в матрицу из четырех фильтров)



без использования каких-либо операций деления. Другими словами, элементы сопряженной матрицы представляются в форме сумм произведений. Поэтому можно утверждать, что критерием существования минимально фазового матричного сейсмического сигнала является то, что детерминант его Z -преобразования не имеет нулей внутри единичного круга.

Уравнение (5.3.1) описывает модель, представленную на рис. 5.2. Однако в некоторых случаях (когда фильтр неизвестен и его необходимо определить) более подходящей является транспонированная форма уравнения (5.3.1). Если b_{12} и b_{21} на рис. 5.2 поменять местами, то можно использовать выражение

$$[Y_1(Z) \ Y_2(Z)] = [X_1(Z) \ X_2(Z)] \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \quad (5.3.2)$$

Чтобы обобщить понятие фильтрации, установленное для скалярных временных рядов, на векторные временные ряды, нужно сначала обобщить понятие спектра. Для функции времени с векторными значениями спектр определяется так называемой *спектральной матрицей*, задаваемой следующим образом:

$$\mathbf{R}(\omega) = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \left(\frac{1}{Z} \right) \\ \bar{Y}_2 \left(\frac{1}{Z} \right) \end{bmatrix} [Y_1(Z) \ Y_2(Z)] = \begin{bmatrix} \bar{Y}_1 Y_1 & \bar{Y}_2 Y_2 \\ \bar{Y}_2 Y_1 & \bar{Y}_1 Y_2 \end{bmatrix}. \quad (5.3.3)$$

Следует отметить, что операция умножения вектора на вектор, определяемая выражением (5.3.3), больше похожа на «внешнее» (векторное) произведение, чем на «внутреннее» (скалярное). Диагонали спектральной матрицы \mathbf{R} содержат автоспектр каждого канала. Внедиагональные элементы матрицы содержат взаимные спектры двух каналов. Так как выражение (5.3.3) является внешним произведением, то полученная матрица является вырожденной. Теперь вместо величины $[Y_1(Z)Y_2(Z)]$ для получения функции времени с конечной энергией предположим, что входные сигналы фильтра $[x_1(t), x_2(t)]$ в соотношении (5.3.2) представлены случайными числами, полученными независимо для каждой точки во времени из некоторой функции вероятности. В этом случае $y_1(t)$ и $y_2(t)$ являются случайными временными рядами, а их спектральная матрица определяется математическим ожиданием (средним по ансамблю) выражения, подобного формуле (5.3.3). Тогда

$$\mathbf{R}(\omega) = E \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix} [Y_1 \ Y_2].$$

Делая подстановку из формулы (5.3.2), получим:

$$\mathbf{R}(\omega) = E \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} [X_1 X_2] \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Группируя суммирование по ансамблю со случайными переменными, найдем:

$$\mathbf{R} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \left\{ E \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} [X_1 X_2] \right\} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \quad (5.3.4)$$

Введем в явном виде предположение, что случайные числа $x_1(t)$ выбраны независимо от чисел $x_2(t)$. Следовательно, $E[X_2(1/Z) \times X_1(Z)] = 0$. Предположим также, что случайные числа $x_i(t)$ являются «белым светом», т. е. $E[x_i(t)x_i(t+s)] = 0$, если $s \neq 0$, и имеют единичную дисперсию $E[x_i^2(t)] = 1$. Поэтому выражение (5.3.4) приобретает вид

$$\mathbf{R} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{21} \\ \bar{B}_{12} & \bar{B}_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \quad (5.3.5)$$

На практике спектральные матрицы вычисляются с помощью величины \hat{R} на основании конечных выборок данных. Это значит, что суммирование по ансамблю должно моделироваться. Если сумма по ансамблю в выражении (5.3.4) подобна сумме по одной точке (суммирование отсутствует), то матрица (5.3.4) оказывается вырожденной, подобно матрице (5.3.3). Как указывалось ранее, точность элементов спектральной матрицы растет пропорционально корню квадратному из числа элементов ансамбля, входящих в сумму.

Одноканальная факторизация спектра дала возможность понять многие важные задачи математической физики. Мы видели также, что достаточно расширенные понятия фильтра и спектра применимы в некоторой степени к многоканальным записям. Совершенно естественно, что много усилий понадобилось для развития методов факторизации спектров для многоканальных записей. Эти усилия оказались успешными. Однако следует признать, что с точки зрения машинного моделирования и интерпретации наблюдаемых волн спектральное разложение многоканальных записей оказалось не особенно полезным. Тем не менее кратко изложим полученные результаты.

Метод корня. Автором метод корня для одиночных записей был обобщен на многоканальные записи [19]. Этот метод более громоздок, чем для одиночных записей. Самым поразительным оказалось то, что полиномы с матричными коэффициентами можно разложить на множители. Например, выражение

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + Z \begin{bmatrix} -3 & -1 \\ 14 & -11 \end{bmatrix} + Z^2 \begin{bmatrix} -4 & 4 \\ -58 & 28 \end{bmatrix}$$

разлагается на множители шестью способами:

$$\begin{aligned} & \left\{ I + Z \begin{bmatrix} 20 & -1 \\ 20 & -7 \end{bmatrix} \right\} \left\{ I + Z \begin{bmatrix} -5 & 0 \\ -6 & -4 \end{bmatrix} \right\}; \\ & \left\{ I + Z \begin{bmatrix} -4 & 0 \\ -10 & -2 \end{bmatrix} \right\} \left\{ I + Z \begin{bmatrix} 1 & -1 \\ 24 & -9 \end{bmatrix} \right\}; \\ & \left\{ I + Z \begin{bmatrix} 0 & -1 \\ 10 & -7 \end{bmatrix} \right\} \left\{ I + Z \begin{bmatrix} -3 & 0 \\ 4 & -4 \end{bmatrix} \right\}; \\ & \left\{ I + Z \begin{bmatrix} -4 & 9 \\ -4 & -3 \end{bmatrix} \right\} \left\{ I + Z \begin{bmatrix} 1 & -1 \\ 18 & -8 \end{bmatrix} \right\}; \\ & \left\{ I + Z \begin{bmatrix} -1 & -1 \\ 8 & -7 \end{bmatrix} \right\} \left\{ I + Z \begin{bmatrix} -2 & 0 \\ 6 & -4 \end{bmatrix} \right\}; \\ & \left\{ I + Z \begin{bmatrix} -4 & 0 \\ 2 & -5 \end{bmatrix} \right\} \left\{ I + Z \begin{bmatrix} 1 & -1 \\ 12 & -6 \end{bmatrix} \right\}. \end{aligned}$$

Метод Теплица. Этот метод является единственным практическим методом для нахождения обращаемых матриц многока-

нальных сейсмических сигналов с заданным спектром. Необходимые алгебраические выкладки будут приведены позже в разделе о предсказании многоканальных временных рядов.

Экспоненциально-логарифмический метод и метод преобразования Гильберта. Многие известные математики, в том числе Норберт Винер, работали над проблемой обобщения экспоненциально-логарифмического метода или метода преобразования Гильберта. Главным затруднением здесь является то, что $\exp(A+B) \neq \exp(A)\exp(B)$, если только величины A и B не оказываются коммутативными, т. е. $AB=BA$. Обычно это условие не выполняется. Хотя этому предмету посвящено много статей (в некоторых из которых утверждалось, что их авторами задача решена), автор этой книги не знает никого, кто когда-нибудь написал бы программу для ЭВМ, работающую со скоростью быстрого преобразования Фурье, как это осуществляется по методу преобразования Гильберта для отдельного канала.

Упражнения

1. Придумайте матричный фильтр, у которого два выходных сигнала $y_1(t)$ и $y_2(t)$ одинаковы и отличаются только масштабным множителем. Очевидно, вход X не может быть восстановлен по выходу Y . Покажите, что определитель этого фильтра стремится к нулю. Подберите другой пример, в котором определитель равен нулю на одной частоте и не равен нулю на какой-нибудь другой. Объясните, в каком смысле входной сигнал нельзя получить из выходного (во временной области).

2. Даны термометр, измеряющий температуру плюс коэффициент α , умноженный на давление, и манометр, измеряющий давление плюс коэффициент β , умноженный на скорость изменения температуры во времени; найдите матричный фильтр, который преобразует наблюдаемые ряды в значения температуры и давления. [Указание. Используйте одну из следующих двух аппроксимаций производной по времени: либо $1-Z$ либо $2(1-Z)/(1+Z)$.]

3. Пусть

$$\mathbf{B}(Z) = \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} + Z \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}.$$

Приравняйте коэффициенты при степенях Z в выражении $\mathbf{B}(Z)\mathbf{A}(Z) = \mathbf{I}$ с целью рекурсивного определения коэффициентов полинома $\mathbf{A}(Z) = [\mathbf{B}(Z)]^{-1}$.

4. Найдите матрицу, обратную матрице $\begin{bmatrix} 1+2Z & Z \\ 3 & 2 \end{bmatrix}$, в виде ряда Тейлора или

Лорана в зависимости от того, какое представление необходимо.

5. Определитель полинома с матричными коэффициентами может не зависеть от Z . Применительно к матричным фильтрам это может означать, что обратный (инверсный) фильтр может иметь только конечное число степеней Z вместо бесконечного ряда степеней, которым всегда обладают скалярные фильтры. Попытайтесь найти наиболее сложный пример.

5.4. МАРКОВСКИЕ ПРОЦЕССЫ

Марковский процесс является еще одной математической моделью временного ряда. До настоящего времени он ограниченно применялся в геофизике, но мы все же его рассмотрим, так как он может быть полезен, и в то же время объясним с помощью ранее рассмотренных методов.

Предположим, что x_t может принимать только целые значения, которые будем называть состояниями. Переходы во времени от j -го состояния к i -му осуществляются в соответствии с матрицей вероятностей P_{ij} . Система не имеет памяти. Состояние в следующий момент времени зависит от текущего состояния по вероятностному закону, но не зависит от предыдущих состояний. Классическим является пример с лягушкой в водоеме с лилиями. С течением

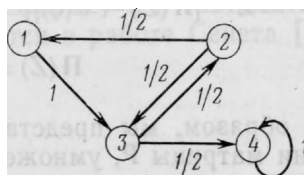


РИС. 5.3. Диаграмма состояний-переходов

времени лягушка прыгает с одного плавающего листа лилии на другой. Она может прыгать на ближний лист с большей вероятностью, чем на дальний. Она может предпочитать большие листы маленьким, и она не помнит тот лист, на котором была последний раз. Состояниями системы являются номера листов, которые лягушка постепенно занимает, а переходами — ее прыжки.

Для начала определим вероятность состояния $\pi_i(k)$, означающую вероятность того, что система придет в состояние i после k переходов, если ее состояние известно при $k=0$. Определим также матрицу перехода P_{ij} . Тогда

$$\pi(k+1) = \mathbf{P}\pi(k). \quad (5.4.1)$$

Вектор вероятностей начального состояния имеет вид $\pi(0)$. Так как начальное состояние известно, то все компоненты $\pi(0)$ равны нулю, за исключением одной (равной 1) в положении, соответствующем начальному состоянию. Диаграмма состояний-переходов приведена на рис. 5.3.

Эта диаграмма соответствует следующей матрице вероятностей:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{bmatrix}_{k+1} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{bmatrix}_k$$

Так как в каждый момент времени должен быть переход, то считаем, что сумма элементов каждого столбца должна быть единичной. Другими словами, вектор-строка $[1111]$ является собственным вектором матрицы \mathbf{P} с единичным собственным значением. Определим Z -преобразование вектора вероятностей следующим

образом:

$$\Pi(Z) = \pi(0) + Z\pi(1) + Z^2\pi(2) + \dots \quad (5.4.2)$$

В обозначениях Z -преобразования выражение (5.4.1) принимает следующий вид:

$$\begin{aligned} [\pi(1) + Z\pi(2) + \dots] &= P[\pi(0) + Z\pi(1) + \dots] = \\ &= Z^{-1}[\Pi(Z) - \pi(0)] = P\Pi(Z); \quad (I - ZP)\Pi(Z) = \pi(0); \quad (5.4.3) \\ \Pi(Z) &= (I - ZP)^{-1}\pi(0). \end{aligned}$$

Таким образом, мы представили общее решение задачи в виде функции матрицы P , умноженной на вектор начального состояния. Очевидно, могут встретиться значения Z , для которых матрица, обратная матрице $(I - ZP)$, не существует. Эти значения Z_j задаются одним из двух уравнений $\det(I - Z_j P) = 0$ или $\det(P - Z^{-1}_j I) = 0$. Очевидно, что значения Z^{-1}_j являются собственными значениями матрицы P . Тогда, применяя теорему Сильвестра, имеем:

$$(I - ZP)^{-1} = \sum_j \frac{c_j r_j}{1 - Z/Z_j}. \quad (5.4.4)$$

В случае повторяющихся собственных значений требуется некоторая модификация выражения (5.4.4). Уравнение (5.4.4) является, по существу, частным разложением на дроби. Типичный член этого разложения имеет вид

$$\frac{1}{1 - Z/Z_j} = 1 + \frac{Z}{Z_j} + \left(\frac{Z}{Z_j}\right)^2 + \dots$$

Таким образом, коэффициенты при последовательных степенях Z убывают со временем по закону $(Z^{-1}_j)^t$. Очевидно, что если вероятности должны быть ограниченными, то корни $1/Z_j$ должны лежать внутри единичного круга (см. минимально фазовые функции). Мы уже показали, что один из корней Z_1 всегда единичный. Это приводит к установившемуся (стационарному) решению $1^t = 1$. В нашем частном примере вектор вероятности «установившегося состояния» имеет вид $[0 \ 0 \ 0 \ 1]^T$. Следовательно, общее решение имеет форму

$$\pi(t) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} [1 \ 1 \ 1 \ 1] + \sum_{j=2}^4 Z_j^{-1} c_j r_j \right\} \pi(0).$$

В заключение отметим, что на практике изредка встречаются неправильные матрицы (с неполным набором собственных векторов),

к которым не применима теорема Сильвестра. В таких случаях решения содержат не только члены Z_j^{-t} , но также и члены tZ^{-t} и i^2Z^{-t} . Такая же ситуация наблюдается и для обыкновенных дифференциальных уравнений с постоянными коэффициентами. Обычно решения получаются в форме $(r_i)^t$, где r_i — i -й корень *показательного* уравнения. Присутствие повторяющихся корней приводит к решениям, подобным (по форме) tr_i^t . Математическое исследование этой проблемы приводится в работе Сенета [20].

6. МОДЕЛИРОВАНИЕ ДАННЫХ НАБЛЮДЕНИЙ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

Согласование теории и данных наблюдений составляет сущность науки. Наиболее часто для решения этой задачи применяется сглаживание с помощью метода наименьших квадратов. В книгах по элементарным вычислениям обычно рассматривается сглаживание разбросанных точек, характеризующих данные наблюдения, с помощью прямой. Это не дает достаточного представления о множестве проблем, которые могут быть решены с помощью метода наименьших квадратов. Некоторые геофизические примеры, которые мы рассмотрим, включают определение местоположения очага землетрясений, анализ морских приливов и отливов, разложение силы земного притяжения и магнитных полей на сферические гармоники и интересные преобразования временных рядов. Если предыдущие значения временного ряда известны, то метод наименьших квадратов позволяет определить фильтр, который предсказывает некоторые будущие значения этого временного ряда. Когда временной ряд, который был очень хорошо предсказуем для большого промежутка времени, вдруг становится намного хуже предсказуемым, то говорят, что произошло «событие». Фильтр, который подчеркивает такие события, называется *фильтром, предсказывающим ошибку*. При поиске специфически распределенной во временном ряду волны он может помочь в конструировании фильтра, который сжимает эту волну в некоторую легко распознаваемую форму, например в импульс определенного вида. Такой формирующий волны фильтр можно сконструировать с помощью метода наименьших квадратов. В случае многоканальных временных рядов, получаемых несколькими датчиками, регистрирующими волны в пространстве, метод наименьших квадратов можно использовать для построения фильтров, которые реагируют только на определенные направления и скорости распространения волн.

Прежде чем излагать общую теорию, рассмотрим простой пример из области анализа временных рядов. Например, пусть задан входной сигнал $\mathbf{x}=(2, 1)$ для фильтра $\mathbf{f}=(f_0, f_1)$. Тогда выходной сигнал должен непременно иметь вид $\mathbf{c}=(2f_0, f_0+2f_1, f_1)$. Для конструирования обратного фильтра желательно, чтобы выходной сигнал \mathbf{c} по возможности приближался к величине $(1, 0, 0)$. Чтобы уменьшить разность между фактическим и требуемым выходными сигналами, минимизируем выражение:

$$E(f_0, f_1) = (2f_0 - 1)^2 + (f_0 + 2f_1)^2 + (f_1)^2.$$

Сумма E квадратов погрешностей будет минимальной, если величины f_0 и f_1 выбраны так, чтобы выполнялись условия

$$0 = \frac{\partial E}{\partial f_0} = 2(2f_0 - 1)2 + 2(f_0 + 2f_1);$$

$$0 = \frac{\partial E}{\partial f_1} = 2(f_0 + 2f_1)2 + 2f_1.$$

Сокращая предыдущие выражения на 2 и приводя их к стандартной для совместных уравнений форме, получим

$$\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Решение этой системы имеет вид

$$\begin{bmatrix} f_0 \\ f_1 \end{bmatrix} = \frac{1}{21} \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{10}{21} \\ -\frac{4}{21} \end{bmatrix}.$$

Фактический выходной сигнал c фильтра имеет вид $\left(\frac{20}{21}, +\frac{2}{21}, -\frac{4}{21}\right)$,

что является неплохим приближением для ожидаемого сигнала $(1, 0, 0)$.

6.1. УРАВНЕНИЙ БОЛЬШЕ, ЧЕМ НЕИЗВЕСТНЫХ

Если число линейных уравнений превышает число неизвестных, то невозможно найти решение, которое точно удовлетворяло бы всем этим уравнениям¹. Тогда ищут решение, которое приближенно удовлетворяет всем уравнениям. Рассмотрим следующую систему уравнений, в которой уравнений больше, чем неизвестных; величины a и c — известны, а x — неизвестна:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & & \cdot \\ a_{31} & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \approx \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_n \end{bmatrix}. \quad (6.1.1)$$

Обычно не существует набора величин x_i , точно удовлетворяющих уравнениям (6.1.1). Определим вектор погрешностей e_i сле-

¹ Такая система уравнений называется переопределенной. (Прим. перев.)

дующим образом:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & & \cdot \\ a_{31} & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_m \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_n \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_n \end{bmatrix}. \quad (6.1.2)$$

Выкладки упростятся, если переписать это уравнение, как показано ниже (этот прием взят у Джона Бурга):

$$\begin{bmatrix} -c_1 & a_{11} & \dots & a_{1m} \\ -c_2 & a_{21} & & \cdot \\ -c_3 & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ -c_n & a_{n1} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ \cdot \\ \cdot \\ x_m \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_n \end{bmatrix}. \quad (6.1.3)$$

Эти уравнения можно записать сокращенно в виде

$$\mathbf{Bx} = \mathbf{e}, \quad (6.1.4)$$

где \mathbf{B} — матрица, содержащая \mathbf{c} и \mathbf{a} . Для i -й погрешности можно записать скалярное произведение, а каждый из векторов представить как столбец. Тогда

$$e_i = [b_{i1} b_{i2} \dots] \begin{bmatrix} 1 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = [1 \ x_1 \ \dots] \begin{bmatrix} b_{i1} \\ b_{i2} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}.$$

Теперь будем минимизировать сумму квадратов погрешностей E , равную величине

$$E = \sum_{i=1}^n [1 \ x_1 \ \dots] \begin{bmatrix} b_{i1} \\ b_{i2} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} [b_{i1} \ b_{i2} \ \dots] \begin{bmatrix} 1 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}. \quad (6.1.5)$$

Вынося постоянные из-под знака суммирования, получим:

$$E = [1 \ x_1 \ x_2 \ \dots] \left\{ \sum_{i=1}^n \begin{bmatrix} -c_i \\ a_{i1} \\ a_{i2} \\ \cdot \\ \cdot \end{bmatrix} [-c_i a_{i1} a_{i2} \ \dots] \right\} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \cdot \\ \cdot \end{bmatrix}. \quad (6.1.6)$$

Оставшаяся в середине выражения (6.1.6) матрица, которую обозначим через r_{ij} , является симметричной и положительно (точнее,

неотрицательно) определенной матрицей, так как не существует таких x , для которых величина E отрицательна, поскольку E представляет собой сумму квадратов погрешностей e_i . Величина x находится из условия минимума функции E , что достигается при $\partial E/\partial x_1=0$, $\partial E/\partial x_2=0$, ..., $\partial E/\partial x_m=0$. Отметим, что при этом получается точно по одному уравнению для каждой неизвестной величины. Чтобы внести ясность в эти представления, рассмотрим выражение (6.1.6) для случая двух неизвестных:

$$E = [1 \ x_1 \ x_2] \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}. \quad (6.1.7)$$

Приравнивая к нулю производную по x_1 , получим:

$$0 = \frac{\partial E}{\partial x_1} = [0 \ 1 \ 0] R \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} + [1 \ x_1 \ x_2] R \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \quad (6.1.8)$$

Так как $r_{ij}=r_{ji}$, то оба члена суммы в правой части равны и, следовательно, выражение (6.1.8) можно переписать в виде

$$0 = \frac{\partial E}{\partial x_1} = 2 [r_{10} \ r_{11} \ r_{12}] \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}. \quad (6.1.9)$$

Аналогично, дифференцирование по x_2 дает

$$0 = \frac{\partial E}{\partial x_2} = 2 [r_{20} \ r_{21} \ r_{22}] \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}. \quad (6.1.10)$$

Уравнения (6.1.9) и (6.1.10) можно объединить

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}. \quad (6.1.11)$$

Последнее выражение представляет собой два уравнения с двумя неизвестными. Его можно записать в более традиционной форме

$$\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = - \begin{bmatrix} r_{10} \\ r_{20} \end{bmatrix}. \quad (6.1.12)$$

В матрице уравнения (6.1.11) не хватает только верхней строки, чтобы она стала равной матрице из уравнения (6.1.7). Чтобы ввести в нее эту строку, можно дополнить выражение (6.1.11) следующим равенством:

$$v = r_{00} + r_{01}x_1 + r_{02}x_2, \quad (6.1.13)$$

которое можно рассматривать как определение новой переменной v . Прибавляя равенство (6.1.13) сверху к выражению (6.1.11), получим:

$$\begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}. \quad (6.1.14)$$

Решением x уравнения (6.1.12) или (6.1.14) является такая совокупность x_k , для которой величина E будет минимальной. Для понимания интерпретации величины v можно умножить обе части выражения (6.1.14) на $[1 \ x_1 \ x_2]$. Тогда получим

$$v = [1 \ x_1 \ x_2] \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix} = [1 \ x_1 \ x_2] R \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}. \quad (6.1.15)$$

Сравнивая выражения (6.1.15) и (6.1.7), видим, что v представляет собой минимальное значение E .

Иногда удобнее представлять наиболее важные уравнения с помощью расчлененных матриц. В этом случае формула погрешности (6.1.6) записывается в виде

$$E = [1 \ | \ x]^T \begin{bmatrix} -c^T \\ A^T \end{bmatrix} [-c \ | \ A] \begin{bmatrix} 1 \\ x \end{bmatrix}. \quad (6.1.16)$$

Последнее уравнение расщепляется на два:

$$V = c^T c - c^T A x; \quad (6.1.17)$$

$$0 = -A^T c + A^T A x, \quad (6.1.18)$$

где формула (6.1.18) представляет собой систему уравнений, решаемую относительно x . Уравнение (6.1.18) вводится в ЭВМ. Оно легко запоминается благодаря быстрому и нестрогому (очень нестрогому) выводу. Поэтому начнем с переопределенного уравнения $Ax \approx c$; умножение его на A^T дает выражение $(A^T A)x = A^T c$, которое совпадает с формулой (6.1.18).

В физических приложениях переменная z_j часто является комплексной величиной, например $z_j = x_j + iy_j$. К переменной z_j всегда можно применять предшествующий анализ, трактуя задачу таким образом, как если бы величины x_j и y_j были действительными независимыми переменными. Рассмотрение величин z_j в качестве единой комплексной переменной дает значительный выигрыш за счет простоты и экономии вычислений. Погрешность E может рассматриваться либо как функция x_j и y_j , либо как функция z_j и \bar{z}_j . Вообще $j=1, 2, \dots, N$, но мы ограничимся обсуждением случая $N=1$, оставив общий случай для упражнений. Минимум погреш-

ности находится из следующих соотношений:

$$0 = \frac{dE}{dx} = \frac{\partial E}{\partial z} \frac{dz}{dx} + \frac{\partial E}{\partial \bar{z}} \frac{d\bar{z}}{dx} = \frac{\partial E}{\partial z} + \frac{\partial E}{\partial \bar{z}}; \quad (6.1.19)$$

$$0 = \frac{dE}{dy} = \frac{\partial E}{\partial z} \frac{dz}{dy} + \frac{\partial E}{\partial \bar{z}} \frac{d\bar{z}}{dy} = i \left(\frac{\partial E}{\partial z} - \frac{\partial E}{\partial \bar{z}} \right). \quad (6.1.20)$$

Умножив формулу (6.1.20) на i и произведя сложение и вычитание этих уравнений, можно получить более простые условия для минимума:

$$0 = \frac{\partial E}{\partial z}; \quad (6.1.21)$$

$$0 = \frac{\partial E}{\partial \bar{z}}. \quad (6.1.22)$$

Однако погрешность E обычно является положительной действительной квадратичной функцией z и \bar{z} , а $\partial E/\partial z$ — просто комплексно сопряженным с $\partial E/\partial \bar{z}$. Следовательно, два условия (6.1.21) и (6.1.22) можно заменить любым из них. При работе с комплексными переменными обычно минимизируют положительную квадратичную форму следующего вида:

$$E(z^*, z) = |\mathbf{A}z - \mathbf{c}|^2 = (z^* \mathbf{A}^* - \mathbf{c}^*) (\mathbf{A}z - \mathbf{c}), \quad (6.1.23)$$

где индекс (*) обозначает комплексно-сопряженные транспонированные векторы и матрицы. Тогда формулу (6.1.22) приводим к выражению

$$0 = \frac{\partial E}{\partial z^*} = \mathbf{A}^* (\mathbf{A}z - \mathbf{c}), \quad (6.1.24)$$

которое и является комплексной формой уравнения (6.1.18).

Рассмотрим пример. Предположим, что последовательность моментов времени t_i прихода фронта волны измеряется датчиками, расположенными на оси x в точках x_i . Пусть фронт волны приближенно имеет форму параболы $t_i \approx a + bx_i + cx_i^2$, где x_i — известные величины, а a, b, c — неизвестные коэффициенты. Для каждого датчика i имеем уравнение

$$[-t_i \ 1 \ x_i \ x_i^2] \begin{bmatrix} 1 \\ a \\ b \\ c \end{bmatrix} \approx 0. \quad (6.1.25)$$

Когда $i > 3$, то уравнений будет больше, чем неизвестных. Для этого примера выражение (6.1.14) принимает следующий вид:

$$\left\{ \sum_{i=1}^n \begin{bmatrix} -t_i \\ 1 \\ x_i \\ x_i^2 \end{bmatrix} \right\} [-t_i \ 1 \ x_i \ x_i^2] \left\{ \begin{bmatrix} 1 \\ a \\ b \\ c \end{bmatrix} \right\} = \begin{bmatrix} v \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (6.1.26)$$

Эту систему уравнений можно решить стандартными методами относительно величин a , b и c .

Последние три строки в уравнении (6.1.26) можно записать в виде

$$\sum_{i=1}^n \begin{bmatrix} 1 \\ x_i \\ x_i^2 \end{bmatrix} e_i = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (6.1.27)$$

Это говорит о том, что вектор погрешностей e_i является перпендикулярным (или нормальным) к функциям $1, x$ и x^2 , которыми приближаются наши данные. По этой причине такие уравнения часто называют нормальными уравнениями.

Упражнения

1. Обобщите выражение (6.1.24) для волн, наблюдаемых в плоскости (x, y) , т. е. для двухмерной квадратичной формы.

2. Пусть $y(t)$ представляет собой функцию комплексных значений для последовательных целых значений t . Постройте для значений $y(t)$ по методу наименьших квадратов аппроксимирующую прямую $y(t) \approx a + \beta t$, где $a = a_r + i a_i$, $\alpha = \beta_r + i \beta_i$. Сделайте это двумя способами: а) предположив, что a_r , a_i , β_i и β_r четыре независимые переменные; б) предположив, что a , α и β — независимые переменные.

Представьте ответ в виде ряда $\left(s_n = \sum_t t^n \right)$

3. Уравнение (6.1.14) предполагает все величины действительными. Обобщите это уравнение для случая, когда все величины комплексные. Проверьте, является ли матрица эрмитовой.

4. Волны наблюдаемого землетрясения достигают j -й сейсмической обсерватории (широта x_j , долгота y_j) в момент времени t_j . Было высказано предположение о том, что землетрясение началось в момент времени t в очаге, расположенном на широте x и долготе y . Теоретическое время распространения волны можно найти по таблице времен распространения волн $T(\Delta)$, где T — время распространения волны, а Δ — наибольший сферический угол, определяемый выражением

$$\cos \Delta = \sin y \sin y_j + \cos y \cos y_j \cos (x - x_j).$$

Разность наблюдаемого и теоретического времен распространения волны для j -й станции в предположении, что землетрясение зарегистрировано в точке с координатами (x, y, t) ,

$$e_j = t + T(\Delta_j) - t_j.$$

Разность времен в предположении, что землетрясение произошло в точке $(x + dx, y + dy, t + dt)$,

$$e_j = t + dt + T(\Delta_j) + \left(\frac{\partial T}{\partial \Delta} \frac{\partial \Delta}{\partial x} \right)_j dy - t_j.$$

Получите уравнения, разрешенные относительно dx , dy и dt , которые минимизировали бы сумму квадратов разностей времен.

5. Сила тяжести g_j измерена в N беспорядочно расположенных точках на поверхности Земли (широта x_j , долгота y_j , $j = 1, N$). Покажите, что матрица нормального уравнения, которое аппроксимирует эти данные сферическими гармониками, может быть записана как сумма столбцов, умноженных на их транспонированные матрицы, как в предыдущем упражнении. Как можно было бы упростить эту матрицу, если бы она имела бесконечно много одинаково расположенных точек данных? [Замечание. Сферические гармоники S составляют класс функций вида

$$S^m_n(x, y) = P^m_n(\cos x) \exp(imy)$$

для $(m = -n, \dots, -1, 0, 1, \dots, n)$ и $(n = 0, 1, \dots, \infty)$, где P_n^m — присоединенный полином Лежандра степени n и порядка m].

6. Морские приливы и отливы достаточно точно аппроксимируются синусоидальными функциями известных частот. С приливом и отливом связаны наклоны поверхности Земли. Комплексный временной ряд можно составить из наклона поверхности в направлении север—юг плюс $\sqrt{-1}$, умноженный на наклон в направлении восток—запад. Рассматриваемый комплексный временной ряд можно представить в аналитической форме $\sum_{j=1}^N A_j e^{i\omega_j t}$. Найдите систему урав-

нений, которую можно решить относительно величины A_j , дающей наилучшее приближение этой формулы к наблюдаемым данным. Покажите, что некоторые элементы матрицы нормального уравнения являются суммами, которые можно вычислить аналитически.

7. Общее решение уравнения Лапласа в цилиндрических координатах (r, θ) для потенциального поля P , которое стремится к нулю при $r \rightarrow \infty$, дается выражением

$$P(r, \theta) = \operatorname{Re} \sum_{m=0}^{\infty} A_m \frac{e^{im\theta}}{r^{m+1}}.$$

Найдите потенциальное поле, окружающее квадратный объект, для источника, имеющего единичный потенциал. Прodelайте это путем нахождения N коэффициентов A_m с помощью минимизации квадрата разности между $P(r, \theta)$ и единицей, проинтегрированной вокруг квадрата. Дайте ответ в обозначениях обратной матрицы интегралов. Какие коэффициенты A_m исчезают именно вследствие симметрии?

6.2. ВЕСОВЫЕ ФУНКЦИИ И ОГРАНИЧЕНИЯ

Часто данные некоторых наблюдений считаются надежнее других. При этом может возникнуть необходимость придать более надежным данным большие веса. Другими словами, можно умножить i -е уравнение на весовую функцию $\sqrt{\omega_i}$:

$$\sqrt{\omega_i} [-c_i \ a_{i1} \ a_{i2} \ \dots] \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ \vdots \end{bmatrix} = \sqrt{\omega_i} e_i. \quad (6.2.1)$$

Тогда взвешенная суммарная квадратическая погрешность

$$E = \sum_{i=1}^n \omega_i e_i^2. \quad (6.2.2)$$

Следуя методу, изложенному в предыдущем разделе, легко показать, что величина \mathbf{x} , которая минимизирует взвешенную погрешность E в соотношении (6.2.2), одновременно удовлетворяет си-

стеме уравнений

$$\left\{ \sum_{i=1}^n \omega_i \begin{bmatrix} -c_i \\ a_{i1} \\ a_{i2} \\ \vdots \end{bmatrix} \begin{bmatrix} -c_i & a_{i1} & a_{i2} & \dots \end{bmatrix} \right\} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \\ \vdots \end{bmatrix}. \quad (6.2.3)$$

Выбор весовых коэффициентов часто является субъективным, чем объективным, однако если качество данных неодинаково, то этого невозможно избежать. Отсутствие весовых коэффициентов ω означает, что они равны единице.

Значительный интерес представляет случай, когда некоторые уравнения должны быть решены точно. Такие уравнения называются уравнениями ограничений. Они часто возникают из теоретических представлений и поэтому могут в принципе не допускать какой-либо погрешности. Остальные уравнения часто включают результаты измерений. Поскольку измерение можно производить много раз, то легко получить намного больше уравнений, чем неизвестных. Так как измерение всегда включает погрешность, то метод наименьших квадратов используется для того, чтобы минимизировать среднюю погрешность. Чтобы удостовериться в том, что уравнения ограничений решены точно, можно использовать очень большие весовые коэффициенты для этих уравнений. Проблема состоит в том, что понятие «очень большие» недостаточно хорошо определено. Весовой коэффициент, равный 10^{10} , может оказаться недостаточным для гарантии, что уравнения ограничений удовлетворяются с достаточной точностью. С другой стороны, $\omega = 10^{10}$ может привести к опасному округлению при решении системы уравнений на восьмиразрядной ЭВМ. Наилучший подход заключается в том, чтобы проанализировать ситуацию теоретически при $\omega \rightarrow \infty$.

Примером уравнения ограничения является равенство суммы x_i величине M . Другим ограничением может быть условие $x_1 = x_2$. Рассматриваемые два уравнения ограничения, приведенные к матричной форме, имеют вид

$$\begin{bmatrix} -M & 1 & 1 & 1 & \dots \\ 0 & 1 & -1 & 0 & \dots \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (6.2.4)$$

Запишем общую систему k уравнений ограничений в виде

$$\mathbf{G} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \end{bmatrix}. \quad (6.2.5)$$

Минимизация погрешности при $\omega \rightarrow \infty$ в уравнениях

$$\sqrt{\omega} \mathbf{G} \mathbf{x} \approx \mathbf{0};$$

$$\mathbf{B} \mathbf{x} \approx \mathbf{0}$$

алгебраически подобна минимизации погрешности в уравнении $\mathbf{B} \mathbf{x} \approx \mathbf{0}$. Строки в $\sqrt{\omega} \mathbf{G}$ похожи на некоторые дополнительные строки для \mathbf{B} . Результирующее уравнение для \mathbf{x} имеет вид

$$\left\{ \sum_{i=1}^n \begin{bmatrix} -c_i \\ a_{i1} \\ \vdots \\ \vdots \end{bmatrix} [-c_i \ a_{i1} \ \dots] + \sum_{i=1}^k \omega_i \begin{bmatrix} g_{i0} \\ g_{i1} \\ \vdots \\ \vdots \end{bmatrix} [g_{i0} \ g_{i1} \ \dots] \right\} \times \\ \times \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \\ \vdots \\ \vdots \end{bmatrix}. \quad (6.2.6)$$

Теперь возьмем все ω_i , равные $1/\varepsilon$, и устремим ε к нулю. Пусть также

$$\mathbf{x} = \mathbf{x}^{(0)} + \varepsilon \mathbf{x}^{(1)} + \varepsilon^2 \mathbf{x}^{(2)} + \dots; \quad (6.2.7a)$$

$$\mathbf{v} = \mathbf{v}^{(0)} + \varepsilon \mathbf{v}^{(1)} + \varepsilon^2 \mathbf{v}^{(2)} + \dots \quad (6.2.7b)$$

Учитывая это, запишем формулу (6.2.6) в виде

$$\left(\mathbf{B}^T \mathbf{B} + \frac{1}{\varepsilon} \mathbf{G}^T \mathbf{G} \right) (\mathbf{x}^{(0)} + \varepsilon \mathbf{x}^{(1)} + \dots) = \mathbf{v}^{(0)} + \mathbf{v}^{(1)} \varepsilon + \dots \quad (6.2.8)$$

Приравняем коэффициенты при степенях ε :

$$\varepsilon^{-1} : \mathbf{G}^T \mathbf{G} \mathbf{x}^{(0)} = \mathbf{0}; \quad (6.2.9a)$$

$$\varepsilon^0 : \mathbf{B}^T \mathbf{B} \mathbf{x}^{(0)} + \mathbf{G}^T \mathbf{G} \mathbf{x}^{(1)} = \mathbf{v}^{(0)}; \quad (6.2.9b)$$

$\varepsilon^1, \varepsilon^2$: не требуются.

Равенство (6.2.9a) является системой из m уравнений с m неизвестными. Оно будет автоматически удовлетворяться, если удовлетворяются k уравнений в формуле (6.2.5). Уравнение (6.2.9b) должно включать m неизвестных в $\mathbf{x}^{(0)}$ плюс еще m неизвестных в $\mathbf{x}^{(1)}$. На самом деле, нет необходимости в m величинах из $\mathbf{x}^{(1)}$; поскольку k неизвестных из

$$\boldsymbol{\lambda} = \mathbf{G} \mathbf{x}^{(1)} \quad (6.2.10)$$

будет достаточно.

Объединяя уравнения (6.2.9b) и (6.2.5) и опуская верхние индексы, получим квадратную матрицу с $m+k$ неизвестными:

$$\left[\begin{array}{c|c} \mathbf{B}^T \mathbf{B} & \mathbf{G}^T \\ \hline \mathbf{G} & 0 \end{array} \right] \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ \vdots \\ \vdots \end{bmatrix}. \quad (6.2.11)$$

Система уравнений (6.2.11) служит одновременно для определения неизвестных \mathbf{x} и λ . Она может рассматриваться и как решение задачи минимизации квадратичной формы:

$$E = [\mathbf{x}^T \lambda^T] \begin{bmatrix} \mathbf{B}^T \mathbf{B} & \mathbf{G}^T \\ \mathbf{G} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \mathbf{x}^T \mathbf{B}^T \mathbf{B} \mathbf{x} + \lambda^T \mathbf{G} \mathbf{x} + \mathbf{x}^T \mathbf{G}^T \lambda,$$

а так как скаляр всегда можно переместить, то

$$E = \mathbf{x}^T \mathbf{B}^T \mathbf{B} \mathbf{x} + 2\lambda^T \mathbf{G} \mathbf{x}. \quad (6.2.12)$$

В соответствии с методом множителей Лагранжа можно минимизировать квадратичную форму, подлежащую ограничению, путем минимизации взамен этого суммы квадратичной формы и ограничивающих членов, где каждый ограничивающий член является произведением уравнения ограничения на множитель Лагранжа λ_i . Это точно совпадает с выражением (6.2.12). Решение при этом описывается уравнением (6.2.11). Множители Лагранжа часто возникают в связи с интегральными уравнениями. Это понятие легко трансформируется в матричную форму путем приближения интегрирования с помощью суммирования.

Упражнения

При определении распределения плотности для глубинного разреза земной коры можно было бы минимизировать квадрат разности между некоторыми теоретическими величинами (например, частотами свободного колебания Земли) и наблюдаемыми данными. Общая масса и момент инерции Земли хорошо известны по астрономическим данным. Если разделить Землю на произвольные тонкие слои равной толщины, то каковы будут два астрономических уравнения ограничения плотности этих слоев ρ_i ? Если задача метода наименьших квадратов является нелинейной (как это часто бывает), то она может быть линеаризована с помощью предположения, что заданный набор плотностей ρ_i является хорошей гипотезой, удовлетворяющей ограничениям и дающей возможность применить метод наименьших квадратов для возмущения $d\rho_i$. Каковы уравнения ограничения для $d\rho_i$?

6.3. УРАВНЕНИЙ МЕНЬШЕ, ЧЕМ НЕИЗВЕСТНЫХ

Что следует делать, когда уравнений¹ меньше, чем неизвестных, отказаться от решения? Конечно же нет, следует использовать принцип упрощений. Найдем простейшее решение, удовлетворяющее всем уравнениям. Такая ситуация возникает часто. Предположим, что после проведения конечного числа измерений делается попытка определить непрерывную функцию, например плотность $\rho(r)$ как функцию глубины в Земле. Тогда в ЭВМ $\rho(r)$ можно было бы представить дискретно в N точках глубины r_i , $i=1, 2, \dots, N$. Но только из-за того, что величину N принимают большой, получают больше неизвестных, чем уравнений.

Мерой упрощения неизвестной функции x_i можно считать минимальное отклонение ее от некоторой известной функции. Другими

¹ Такая система уравнений называется недоопределенной. (Прим. перев.)

словами, минимизируем величину

$$E = \sum (x_i - x_{i-1})^2 \quad (6.3.1)$$

при точном удовлетворении наблюдаемых данных или уравнений ограничения:

$$\mathbf{Gx} = \mathbf{0}. \quad (6.3.2)$$

Другой, более популярной мерой упрощения (которая не учитывает порядок переменных x_i) является минимизация величины

$$E = \sum_{i=1}^n x_i^2. \quad (6.3.3)$$

Если минимизировать выражение (6.3.3) без каких-либо ограничений, то x будет удовлетворять следующей системе уравнений:

$$\begin{array}{c} \left[\begin{array}{cccc} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \end{array} \right] \begin{array}{c} \text{Нули} \\ \\ \\ \\ \\ \text{Нули} \end{array} \left[\begin{array}{c} 1 \\ x_1 \\ x_2 \\ \vdots \\ \vdots \end{array} \right] = \left[\begin{array}{c} v \\ 0 \\ 0 \\ 0 \end{array} \right]. \end{array}$$

Из последнего выражения следует, что $x_i = 0$. Теперь введем два уравнения ограничений и для определенности возьмем три неизвестных. Тогда метод, описанный в предыдущем разделе, дает

$$\begin{array}{c} \left[\begin{array}{cccc|cc|c|c} 1 & & & & -d_1 & -d_2 & 1 & v \\ & 1 & & & g_{11} & g_{21} & x_1 & 0 \\ & & 1 & & g_{12} & g_{22} & x_2 & 0 \\ & & & 1 & g_{13} & g_{23} & x_3 & 0 \\ \hline -d_1 & g_{11} & g_{12} & g_{13} & 0 & 0 & \lambda_1 & 0 \\ -d_2 & g_{21} & g_{22} & g_{23} & 0 & 0 & \lambda_2 & 0 \end{array} \right] \end{array} \quad (6.3.4)$$

Уравнение (6.3.4) имеет размерность, равную сумме числа переменных и числа ограничений. Его можно решить численно или вначале свести к матрице, размерность которой определяется числом ограничений. Разобьем выражение (6.3.4) на два уравнения:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \\ g_{13} & g_{23} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (6.3.5)$$

и

$$\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}. \quad (6.3.6)$$

Введем сокращенную запись этих уравнений в виде

$$\mathbf{x} + \mathbf{G}^T \boldsymbol{\lambda} = \mathbf{0} \text{ и } \mathbf{Gx} = \mathbf{d}.$$

Умножив выражение (6.3.5) слева на \mathbf{G} , получим

$$\mathbf{G}\mathbf{x} + \mathbf{G}\mathbf{G}^T\boldsymbol{\lambda} = \mathbf{0}.$$

Подставим в него уравнение (6.3.6)

$$\mathbf{d} + \mathbf{G}\mathbf{G}^T\boldsymbol{\lambda} = \mathbf{0}.$$

Решим это выражение относительно $\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda} = -(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{d}$$

и подставим полученный результат снова в уравнение (6.3.5). Тогда

$$\mathbf{x} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{d}.$$

Переписав в развернутом виде последнее выражение, имеем:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [\mathbf{G}^T] [\mathbf{G}\mathbf{G}^T]^{-1} [\mathbf{d}]. \quad (6.3.7)$$

Это и есть конечный результат, т. е. минимально отклоняющееся решение \mathbf{x} , удовлетворяющее недоопределенной системе, образуемой уравнениями ограничения.

Упражнения

1. Если определить отклонение с помощью выражения (6.3.1) вместо формулы (6.3.3), то какой вид примет формула (6.3.6)?

2. Заданы масса и момент инерции Земли. Вычислите плотность массы как функцию глубины, используя принцип минимального отклонения, выражаемой формулой (6.3.7). Какие критические замечания есть у вас относительно этой процедуры? (Указание: наилучшее решение использует интегралы вместо бесконечных сумм.)

3. Используйте методы этого раздела с целью уменьшения в формуле (6.2.11) размерности матриц, которые должны обращаться.

6.4. ПРЕОБРАЗОВАНИЯ ХАУСХОЛДЕРА И МЕТОД ГОЛУБА [21]

При обсуждении метода наименьших квадратов мы всегда приходили к матрицам в форме $\mathbf{A}^T\mathbf{A}$, которые затем необходимо было обращать. Метод Голуба, использующий преобразования Хаусхолдера, касается непосредственно матрицы \mathbf{A} и имеет то преимущество, что он значительно точнее методов, обращающих выражение $\mathbf{A}^T\mathbf{A}$. По-видимому, требуется почти вдвое большая точность для обращения выражения $\mathbf{A}^T\mathbf{A}$, чем та, которая необходима при работе непосредственно с матрицей \mathbf{A} . Другой причиной для изучения метода Голуба является то, что вычисление организуется в нем совершенно другим способом. Поэтому он часто будет иметь те или иные преимущества или недостатки, которые отличаются для различных случаев применения.

Отражающее преобразование, включающее преобразование Хаусхолдера, описывается матрицей $\mathbf{R} = (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T / \mathbf{v}^T\mathbf{v})$, где \mathbf{v} — произвольный вектор. Очевидно, что матрица \mathbf{R} симметрична, т. е.

$\mathbf{R}=\mathbf{R}^T$. Оказывается, что отражающее преобразование является своим собственным обратным преобразованием, т. е. $\mathbf{R}=\mathbf{R}^{-1}$. Подтвердим это на примере $\mathbf{R}^2=\mathbf{I}$:

$$\left(\mathbf{I}-\frac{2\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right)^2=\mathbf{I}-\frac{4\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}+\frac{4\mathbf{v}(\mathbf{v}^T\mathbf{v})\mathbf{v}^T}{(\mathbf{v}^T\mathbf{v})^2}=\mathbf{I}. \quad (6.4.1)$$

Говорят, что матрица преобразования \mathbf{M} унитарна, если $\mathbf{M}^T\mathbf{M}=\mathbf{I}$. Если матрица \mathbf{M} является унитарной, то это означает, что вектор \mathbf{x} имеет ту же длину, что и вектор $\mathbf{M}\mathbf{x}$. Эти длины равны $\mathbf{x}^T\mathbf{x}$ и $(\mathbf{M}\mathbf{x})^T(\mathbf{M}\mathbf{x})=\mathbf{x}^T\mathbf{M}^T\mathbf{M}\mathbf{x}=\mathbf{x}^T\mathbf{I}\mathbf{x}=\mathbf{x}^T\mathbf{x}$, т. е. выражаются одной и той же величиной. Отражающие преобразования являются унитарными, так как $\mathbf{R}^{-1}=\mathbf{R}^T$. Они имеют простую физическую интерпретацию. Рассмотрим ортогональную систему координат, в которой одна из координатных осей направлена вдоль вектора \mathbf{v} . Отражающее преобразование изменяет знак этой координатной оси вектора на противоположный (поскольку $\mathbf{R}\mathbf{v}=-\mathbf{v}$), но оно оставляет без изменений все другие координатные оси вектора. Таким образом, из геометрических соображений очевидно, что отражающее преобразование сохраняет длину оси, а применение его дважды преобразует первоначальный вектор в самого себя. Найдем теперь специальное отражающее преобразование — преобразование Хаусхолдера:

$$\mathbf{H} \begin{bmatrix} a & a & a & a \\ 0 & a & a & a \\ 0 & 0 & a & a \\ 0 & 0 & a & a \\ 0 & 0 & a & a \end{bmatrix} = \begin{bmatrix} a & a & a & a \\ 0 & a & a & a \\ 0 & 0 & a & a \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & a \end{bmatrix}, \quad (6.4.2)$$

где a — произвольный элемент.

Определив требуемое преобразование, узнаем, как трансформировать любую матрицу в матрицу в форме верхнего треугольника

$$\begin{bmatrix} a & a & a & a \\ 0 & a & a & a \\ 0 & 0 & a & a \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (6.4.3)$$

с помощью последовательности преобразований Хаусхолдера. Голуб оценил значение этого метода для решения переопределенных систем совместных уравнений. Он отметил, что если вектор погрешности $\mathbf{e}=\mathbf{A}\mathbf{x}-\mathbf{b}$ преобразуется унитарной матрицей $\mathbf{U}\mathbf{e}$, то задача минимизации длины $(\mathbf{e}^T\mathbf{U}^T\mathbf{U}\mathbf{e})^{1/2}$ вектора $\mathbf{U}\mathbf{e}$ путем изменения величины \mathbf{x} сводится точно к такой же задаче, как и минимизация длины $(\mathbf{e}^T\mathbf{e})^{1/2}$ вектора ошибок \mathbf{e} по отношению к измене-

нию вектора x . Таким образом, можно найти последовательность преобразований Хаусхолдера, приводящих вектор $e = Ax - b$ к виду

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} a & a & a \\ 0 & a & a \\ 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} a \\ a \\ a \\ a \\ a \\ a \end{bmatrix}. \quad (6.4.4)$$

Отметим, что из-за нулей в нижней части преобразованной матрицы A нет возможности выбрать какие-либо значения x_i , которые изменяли бы величину e_2 любым из способов. Верхняя часть преобразованной матрицы A является верхней треугольной матрицей, которая для любого значения e_1 может быть точно решена относительно x_i . Решение x_i , полученное методом наименьших квадратов, является единственным, для которого e_1 будет равно нулю.

Вернемся теперь к задаче отыскания специального отражающего преобразования, называемого преобразованием Хаусхолдера, которое выполняет операцию, описываемую с помощью выражения (6.4.2). Заметим, что оператор в левой части соотношения, приводимого ниже, является отражающим преобразованием для любого выбранного числового значения s :

$$\left\{ \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} - \frac{2}{(a_3 - s)^2 + a_4^2 + a_5^2} \begin{bmatrix} 0 \\ 0 \\ a_3 - s \\ a_4 \\ a_5 \end{bmatrix} \begin{bmatrix} 0 & 0 & (a_3 - s) & a_4 & a_5 \end{bmatrix} \right\} \times$$

$$\times \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ a_3 - s \\ a_4 \\ a_5 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ s \\ 0 \\ 0 \end{bmatrix} \quad (6.4.5)$$

Соответственно, если условие (6.4.5) выполняется, величина s должна принять определенное значение:

$$1 = \frac{2}{(a_3 - s)^2 + a_4^2 + a_5^2} \begin{bmatrix} 0 & 0 & (a_3 - s) & a_4 & a_5 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}. \quad (6.4.6)$$

или

$$1 = \frac{-2sa_3 + 2(a_3^2 + a_4^2 + a_5^2)}{s^2 - 2sa_3 + (a_3^2 + a_4^2 + a_5^2)}. \quad (6.4.7)$$

Это будет справедливо только для s , заданного выражением

$$s = \pm (a_3^2 + a_4^2 + a_5^2)^{1/2}. \quad (6.4.8)$$

```

SUBROUTINE GOLUB (A,X,B,M,N)
C
C A(M,N) ; B(M) GIVEN WITH M>N SOLVES FOR X(N) SUCH THAT
C || B - AX || = MINIMUM
C METHOD OF G.GOLUB, NUMERISCHE MATHEMATIK 7,206-216 (1965)
C
IMPLICIT DOUBLE PRECISION (D)
REAL A(M,N),X(N),B(M),U(50)
C.....DIMENSION U(M)
C.....PERFORM N ORTHOGONAL TRANSFORMATIONS TO A(.,.) TO
C.....UPPER TRIANGULARIZE THE MATRIX
DO 3010 K=1,N
DSUM=0.0D0
DO 1010 I=K,M
DAJ=A(I,K)
1010 DSUM=DSUM+DAJ**2
DAI=A(K,K)
DSIGMA=DSIGN(DSQRT(DSUM),DAI)
DBI=DSQRT(1.0D0+DAI/DSIGMA)
DFACT=1.0D0/(DSIGMA*DBI)
U(K)=DBI
FACT=DFACT
KPLUS=K+1
DO 1020 I=KPLUS,M
1020 U(I)=FACT*A(I,K)
C.....I - UU' IS A SYMMETRIC, ORTHOGONAL MATRIX WHICH WHEN APPLIED
C..... TO A(.,.) WILL ANNIHILATE THE ELEMENTS BELOW THE DIAGONAL K
DO 2030 J=K,N
C.....APPLY THE ORTHOGONAL TRANSFORMATION
FACT=0.0
DO 2010 I=K,M
2010 FACT=FACT+U(I)*A(I,J)
DO 2020 I=K,M
2020 A(I,J)=A(I,J)-FACT*U(I)
2030 CONTINUE
FACT=0.0
DO 2040 I=K,M
2040 FACT=FACT+U(I)*B(I)
DO 2050 I=K,M
2050 B(I)=B(I)-FACT*U(I)
3010 CONTINUE
C.....BACK SUBSTITUTE TO RECURSIVELY YIELD X(.)
X(N)=B(N)/A(N,N)
LIM=N-1
DO 4020 I=1,LIM
IROW=N-I
SUM=0.0
DO 4010 J=1,I
4010 SUM=SUM+X(N-J+1)*A(IROW,N-J+1)
4020 X(IROW)=(B(IROW)-SUM)/A(IROW,IROW)
RETURN
END

```

РИС. 6.1. Подпрограмма сглаживания данных методом наименьших квадратов (составлена Доном С. Райли)

Теперь посмотрим, почему оператор в левой части выражения (6.4.5) может реализовать преобразование, выражаемое формулой (6.4.2). Выбор в качестве вектора \mathbf{a} третьего столбца матрицы в формуле (6.4.2) вводит требуемые нули в правую часть. В заключение необходимо отметить, что такой выбор оператора \mathbf{H} не уничтожает каких-либо нулей, которые уже есть в левой части выражения (6.4.2). Подпрограмма для решения этой задачи на ЭВМ представлена на рис. 6.1. Эта программа не рассчитана на

квадратные матрицы. Необходимо, чтобы выполнялось условие $M > N$. Преобразования Хаусхолдера могут применяться также в задачах с ограничениями. В системе

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{A} \end{bmatrix} [\mathbf{x}] = \begin{bmatrix} \mathbf{d} \\ \mathbf{b} \end{bmatrix} \quad (6.4.9)$$

можно потребовать точного удовлетворения для верхнего блока, а для нижнего блока — только в смысле наименьших квадратов. Определим \mathbf{y} как результат последовательности преобразований Хаусхолдера по \mathbf{x} . Например, $\mathbf{y} = \mathbf{H}_2 \mathbf{H}_1 \mathbf{x}$. Затем подставим $\mathbf{x} = \mathbf{H}_1 \mathbf{H}_2 \mathbf{H}_2 \mathbf{H}_1 \mathbf{x} = \mathbf{H}_1 \mathbf{H}_2 \mathbf{y}$ в выражение (6.4.9). Преобразования Хаусхолдера, применяемые как множители справа матрицы в формуле (6.4.9), могут быть выбраны таким образом, чтобы ввести нули в верхние две строки в выражении (6.4.9), например:

$$\begin{bmatrix} a & 0 & 0 & 0 \\ a & a & 0 & 0 \\ a & a & a & a \\ a & a & a & a \\ a & a & a & a \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \approx \begin{bmatrix} d_1 \\ d_2 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad (6.4.10)$$

Теперь можно использовать преобразования Хаусхолдера, умножив выражение (6.4.10) на оператор \mathbf{H} слева, чтобы привести его к виду

$$\begin{bmatrix} a & 0 & 0 & 0 \\ a & a & 0 & 0 \\ a & a & a & a \\ a & a & 0 & a \\ a & a & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \approx \begin{bmatrix} d_1 \\ d_2 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad (6.4.11)$$

Так как верхние два уравнения в формуле (6.4.10) или (6.4.11) должны удовлетворяться точно, то величины y_1 и y_2 определены однозначно. Поэтому их нельзя использовать для достижения минимальной погрешности в нижних трех уравнениях. В то же время верхние два уравнения не накладывают ограничений на величины y_3 и y_4 . Следовательно, их можно использовать для получения минимальной погрешности в трех нижних уравнениях. Никакое регулирование величин y_3 и y_4 не может изменить погрешности в последнем уравнении. Поэтому это уравнение можно не учитывать при определении величин y_3 и y_4 . Третье и четвертое уравнения можно удовлетворить с нулевой погрешностью, выбрав подходящим образом величины y_3 и y_4 . Это должно быть решение с минимальным квадратом погрешности. Имея \mathbf{y} , можно вернуться назад и найти \mathbf{x} с помощью выражения $\mathbf{x} = \mathbf{H}_1 \mathbf{H}_2 \mathbf{y}$.

6.5. ВЫБОР НОРМЫ МОДЕЛИ

Еще недавно точка зрения на моделирование геофизических данных состояла в том, что Земля представлялась непрерывной и поэтому нужно было считать число неизвестных бесконечным, а число наблюдений ограниченным. Поэтому в ЭВМ такая ситуация аппроксимируется сильно недоопределенной системой совместных

уравнений. Для получения единственного ответа решение должно экстремализовать значение некоторого интеграла. На практике сумма квадратов часто минимизируется таким способом, что получают гладкое решение. Типичная математическая формулировка рассматриваемой задачи заключается в сглаживании методом наименьших квадратов системы уравнений следующего вида:

$$\begin{bmatrix} \mathbf{A} \\ \dots \\ \dots \\ \dots \end{bmatrix} \mathbf{x} \approx \begin{bmatrix} \mathbf{d} \\ 0 \end{bmatrix} \quad (6.5.1)$$

где верхний блок \mathbf{A} обозначает недоопределенные уравнения ограничения с вектором данных \mathbf{d} ; неизвестные представляет вектор \mathbf{x} ; а нижний блок является полосовой матрицей (все нулевые элементы которой группируются около диагонали). Это говорит о том, что некоторые значения \mathbf{x} при фильтрации должны стремиться к нулю. Фильтром часто является грубый фильтр, подобный, например, оператору первой разности. При отсутствии данных оператор первой разности приводит к решению в виде постоянной, которая имеет лишь математический смысл. В то же время этот вынужденный результат не имеет геофизического смысла, так как решение оказывается сильно сглаженным даже в условиях, когда реальные модели Земли содержат ступенчатые разрывы (там, где контактируют две однородные среды с различными свойствами).

Фильтр часто выбирается субъективно на основании результата, к которому приводит его применение. К сожалению, решение часто оказывается функцией, чувствительной к субъективно взятым весам и фильтрам. Существуют общие теории решения обратных задач, но они не подготавливают геофизика к использованию особенностей конкретной ситуации или набора данных.

Одной из полезных концепций в теории обратных задач является понятие об инвариантности системы координат. Суть его заключается в том, что должно быть получено одинаковое решение задачи, например, об электрической проводимости независимо от того, введены ли параметры Земли с помощью задания неизвестной электрической проводимости в каждой точке достаточно густой сетки, сопротивлений на сетке либо коэффициентов некоторого разложения при полном наборе базисных функций. Очевидно, что сглаживание коэффициентов низкого порядка в некотором разложении путем приравнивания к нулю коэффициентов высокого порядка не является координатно-инвариантным подходом. Смещение начала полиномиального разложения или применение другого набора базисных функций может все изменить. Конечно, при решении обратных задач геофизики необязательно использовать инвариантный к системе координат метод. Но если он не применяется, то следует остерегаться влияния на наше решение изменений системы координат.

Рассмотрим некоторые процедуры обращения данных (решения обратных задач), инвариантные по отношению к системе координат. Ограничимся физическими задачами, в которых можем определить положительную функцию плотности p как энергию рассеяния на единицу объема. Обозначим через \bar{p} величину этой энергии как функцию расстояния в частной модели Земли. Частной моделью является та, которую мы хотим найти, не имея данных

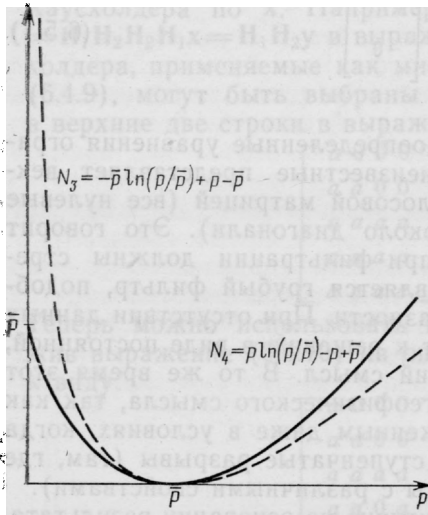


РИС. 6.2. Минимизация любой из двух функций при движении величины p в направлении \bar{p}

измерений. В такой модели свойства вещества не зависят от расстояния. Теперь нам понадобятся некоторые функции, которые мы назовем *нормами модели*. Они обладают свойствами положительности для всех (положительных) p и \bar{p} и минимальности при $p = \bar{p}$. Некоторые примеры таких функций приведены на рис. 6.2, где

$$N_1(p, \bar{p}) = |p - \bar{p}|; \quad N_2(p, \bar{p}) = \frac{(p - \bar{p})^2}{\bar{p}};$$

$$N_3(p, \bar{p}) = -\bar{p} \ln\left(\frac{p}{\bar{p}}\right) + p - \bar{p}; \quad N_4(p, \bar{p}) = p \ln\left(\frac{p}{\bar{p}}\right) - p + \bar{p}.$$

Обозначим теперь подбираемые при минимизации свойства Земли через x , т. е. функцию расстояния. Можно выбрать x так, чтобы минимизировать некоторый объемный интеграл для одной из норм модели при условии (ограничении), что эта модель дает результаты, совпадающие с данными наблюдений. Иногда мы предполагаем результатами наблюдений от $j=1, 2, 3, \dots, n_s$ локализованных источников. Тогда нужно вычислить частное распределение энергии \bar{p}_j для каждого источника. После этого можно минимизировать сумму объемных интегралов

$$\min_x \sum_{j=1}^{n_g} \int N(p_j, \bar{p}_j) dV$$

при условии сглаживания всех значений данных.

Следует отметить, что все функции нормы модели являются однородными функциями первого порядка. При этом $N(ap, a\bar{p}) = aN(p, \bar{p})$ для $a > 0$. Это означает, что N есть объемная плотность. Не будь этого свойства, у нас возникли бы трудности, связанные с тем, что сумма $N_k(p, \bar{p})$ по набору частей объема ΔV_k изменялась бы по мере уменьшения шага сетки. Инвариантность по отношению к выбору системы координат обеспечивается обычными правилами преобразования объемных интегралов из одной системы координат в другую.

Теперь рассмотрим пример из теории фильтров, которая оказывается связанной с вычислением спектра методом максимально-энтропийного спектрального анализа. Возьмем известный входной спектр $R(Z)$ и найдем фильтр конечной длины $X(Z) = x_0 + x_1 Z + x_2 Z^2$, выходной сигнал которого является, насколько это только возможно, «белым» в смысле минимизации интеграла N_3 по спектру. Пусть спектр фильтра $S(Z) = \bar{X}(1/Z)X(Z)$. Мы имеем $\bar{p} = 1$, $p = R(Z)S(Z)$. Таким образом, минимизация заключается в том, чтобы получить

$$\min \int (-\ln RS + RS) d\omega.$$

Приравнивая производную этого выражения по \bar{x}_k нулю, имеем:

$$\begin{aligned} 0 &= \int \left(-\frac{1}{S} + R \right) \frac{\partial S}{\partial \bar{x}_k} d\omega = \int \left(-\frac{1}{S} + R \right) Z^{-k} X(Z) d\omega = \\ &= \int Z^{-k} \left[-\frac{1}{\bar{X}(1/Z)} + R(Z) X(Z) \right] d\omega. \end{aligned}$$

Поскольку минимально фазовые функции могут представлять любой спектр, то выберем функцию $\bar{X}(1/Z)^{-1}$ такой, чтобы она разлагалась в ряд вида $(\bar{b}_0 + \bar{b}_1/Z + \bar{b}_2/Z^2 + \dots)$. Тогда имеем

$$0 = \int Z^{-k} \left[-\left(\bar{b}_0 + \frac{\bar{b}_1}{Z} + \frac{\bar{b}_2}{Z^2} + \dots \right) + RX \right] d\omega.$$

Вспомним, что этот интеграл использует коэффициент при Z^0 аргумента. Если предположить, что этот фильтр имеет ограничение $x_k = 0$ при $k \geq 3$, то получим хорошо знакомую систему Теплица:

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_0 \\ 0 \\ 0 \end{bmatrix}.$$

6.6. НАДЕЖНОЕ МОДЕЛИРОВАНИЕ

Медиана и среднее значение — это два вида статистического среднего. В нормальной ситуации они ведут себя почти одинаково. В настоящее время ученые-физики используют среднее значение и, следовательно, не знают о «драматичной» возможности медианы сглаживать действие грубых погрешностей наблюдаемых данных. В качестве примера рассмотрим дорогостоящий многодневный эксперимент, производительность которого — одно значение на результат. В первый день результат был 2,17, на второй день — 2,14, а на третий, последний день — 1638,03. Среднее этих результатов 547,78, а медиана — 2,17. Если «сомнительный» результат, полученный на третий день, является грубой ошибкой, то определение медианы может предохранить от нее. Статистики называют это свойство медианы «надежностью». В этом разделе покажем, как многие виды сглаживания геофизических данных можно сделать надежными. В частности, все выполняемые нами расчеты, которые равнозначны решению переопределенной системы линейных уравнений с помощью минимизации суммы квадратов погрешностей, могут быть сделаны надежными путем минимизации суммы абсолютных погрешностей. Затраты на вычисления этим методом часто сравнимы с затратами на вычисления по методу наименьших квадратов. Положительные погрешности могут быть уменьшены с помощью различных весовых множителей в большей степени, чем отрицательные. Этот вариант называется несимметричной линейной нормой. Особым случаем несимметричной линейной нормы является неравенство. Неудивительно, если окажется, что все задачи линейного программирования являются частными случаями задач несимметричной линейной нормы, а методика решения задач методом несимметричной линейной нормы проще, чем методом линейного программирования.

Во-первых, покажем, почему средние значения и медианы случайных величин соответствуют квадратам и абсолютным значениям. Пусть x_i любое число. Определим m_2 путем минимизации суммы квадратов разностей (называемой L_2 нормой) между m_2 и x_i :

$$m_2: \min \sum_{i=1}^N (m_2 - x_i)^2. \quad (6.6.1)$$

Прямой путь нахождения минимума состоит в приравнивании к нулю частной производной суммы по отношению к m_2 . Тогда

$$0 = \sum_{i=1}^N 2(m_2 - x_i)$$

или

$$m_2 = \frac{1}{N} \sum_{i=1}^N x_i. \quad (6.6.2)$$

Очевидно, что m_2 получено в таком же виде, как обычно определяется *среднее* значение случайной величины. Далее, определим m_1 путем минимизации суммы абсолютных погрешностей (называемой L_1 нормой). Тогда

$$m_1 : \min \sum_{i=1}^N |m_1 - x_i|. \quad (6.6.3)$$

Для определения минимума этого выражения необходимо приравнять к нулю частную производную по m_1 :

$$0 = \sum_{i=1}^N \operatorname{sgn}(m_1 - x_i). \quad (6.6.4)$$

Здесь sgn — сигнум-функция, равная $+1$, когда аргумент положителен, равная -1 , когда аргумент отрицателен, и лежащая между ними, если аргумент равен нулю. Уравнение (6.6.4) говорит о том, что величина m_1 должна выбираться так, чтобы она превышала x_i для $N/2$ членов и была меньше x_i для $N/2$ членов, и что если x_i лежит посередине, то $m_1 = x_i$. Это определяет величину m_1 как *медиану*. [Для четного числа N определение (6.6.3) требует только, чтобы значение m_1 лежало где-нибудь посередине между двумя значениями x_i].

Затраты на вычисление среднего пропорциональны числу точек N . Затраты на полное определение перечисленных чисел равны $N \ln N$ [22], но при нахождении медианы полное определение не требуется. В работе Хора [23] описан алгоритм для нахождения медианы, который требует $3N$ операций. Вычислительный алгоритм, основанный на алгоритме Хора, можно применять для определения взвешенных медиан, которые являются аналогами взвешенных сумм. Обычно величина 2,17 рассматривается как медиана чисел (2,14; 2,17; 1638,03), так как подразумевается применение одинаковых весов (1; 1; 1). Если применить веса (3; 1; 1), то это будет эквивалентно заданию чисел 2,14; 2,14; 2,14; 2,17; 1638,03. Тогда медиана будет равна 2,14. Формально взвешенную медиану можно определить с помощью минимизации выражения

$$m_1 : \min \sum |w_i| |m_1 - x_i|. \quad (6.6.5)$$

Очевидно, что если все весовые множители равны единице, то последнее выражение сводится к ранее приведенному. Используя, например, весовой множитель, равный трем, получим такой же результат, как при включении одного и того же члена трижды с весом, равным единице. Рис. 6.3 иллюстрирует определение выражения (6.6.5) для простого случая. Из него следует, что медиана всегда равна одному из значений x_i , даже если весовые множители не являются целыми числами. Если все весовые множители единичны, а числа последовательности выравнены (расположены в порядке возрастания чисел), то погрешность нормы будет лежать на прямой между двумя средними числами. Тогда любая величина,

лежащая в этом промежутке, удовлетворяет нашему определению медианы, основанному на минимизации суммы.

Запишем снова выражение (6.6.5), внося величину $|\omega_i|$ внутрь другой функции абсолютных величин. Тогда

$$m_1: \min \sum_i |\omega_i| m - |\omega_i| x_i = \min \sum_i |\omega_i m - \omega_i x_i|. \quad (6.6.6)$$

Заменим теперь обозначения, принятые в статистике, обозначениями, обычно употребляемыми в совместных уравнениях и линейном программировании. Пусть

$$a_i := \omega_i; \quad b_i := \omega_i x_i; \quad x := m. \quad (6.6.7a, б, в)$$

С этими новыми обозначениями выражение (6.6.6) приобретает вид

$$x: \min \sum_i |a_i x - b_i|. \quad (6.6.7)$$

Выражение (6.6.7) позволяет решить систему переопределенных уравнений вида

$$ax \approx b \quad (6.6.8)$$

для x путем минимизации нормы L_1 . В сущности, это задача взвешенной медианы. Если уравнение (6.6.8) решается путем мини-

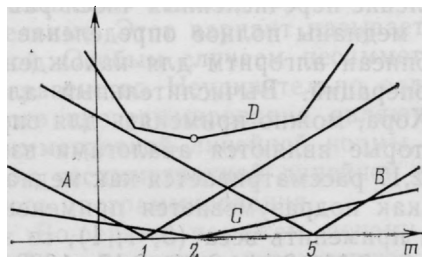


РИС. 6.3. Сумма взвешенных абсолютных значений норм.

Функции, обозначенной буквой А, соответствует $0,5|m-1|$; В — $0,5|m-5|$; С — $0,1|m-2|$, а D является суммой А, В и С. Сумма D минимизируется при $m=2$, т. е. в точке, которой отвечает решение $C=0=-0,1|m-2|$

мизации нормы L_2 (методом наименьших квадратов), то величина x должна быть средневзвешенной $x = (a \cdot b) / (a \cdot a)$.

Рассмотрим теперь технику решения при минимизации выражения (6.6.5). Здесь необходим алгоритм Хора. На основе опыта выделим случайные уравнения из системы (6.6.8), которые точно удовлетворяются. Эти уравнения, называемые базисными (основными) уравнениями, можно обозначить $a_k x_{\text{опытн}} = b_k$. Подставляя $x_{\text{опытн}}$ в уравнение (6.6.8), получим уравнения с положительными, отрицательными и нулевыми погрешностями. Если выбрать случайные уравнения с $x_{\text{опытн}}$, то может оказаться, что группа уравнений с нулевой погрешностью имеет достаточно большую вес, чтобы уравновесить веса положительных и отрицательных отклонений во всех направлениях. В противном случае мы должны выбрать новое базисное уравнение из наиболее сильной представительной положительной или отрицательной группы. К счастью, нет необ-

ходимости далее рассматривать слабые группы, так как эти остатки не могут изменить знаки при минимизации. В этом можно убедиться на основе геометрических представлений (см. рис. 6.3). Мы всегда хотим двигаться вниз по склону, так что следует еще раз убедиться в том, что если следующая точка данных находится выше настоящей точки, то никогда не возникает необходимость проверить верхнюю точку. Таким образом, размер группы уравнений, подвергаемых исследованию, быстро уменьшается. На рис. 6.4 приведена программа для ЭВМ, осуществляющая эту операцию. Эта подпрограмма является некоторой частью более сложной программы, которая специально предназначена для корректирования данных, когда весовые множители равны нулю, и предусматривает указание всех уравнений (иногда их больше одного), которые удовлетворяются в конце процесса минимизации.

Следующим шагом по сложности является рассмотрение этой задачи для двух неизвестных. Очевидное обобщение выражения (6.6.8) дает

$$\begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ a_k & c_k \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \approx \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_k \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}. \quad (6.6.9)$$

Допустим, что читатель легко решит уравнение (6.6.9) методом наименьших квадратов. Решение путем минимизации суммы абсолютных погрешностей начинается также. Начнем с определения погрешностей

$$E = \sum_{k=1}^N |b_k - a_k x - c_k y|. \quad (6.6.10)$$

Затем приравняем к нулю частные производные выражения (6.6.10) по x и y :

$$0 = \frac{\partial E}{\partial x} = \sum_{k=1}^N -a_k \operatorname{sgn}(b_k - a_k x - c_k y); \quad (6.6.11a)$$

$$0 = \frac{\partial E}{\partial y} = \sum_{k=1}^N -c_k \operatorname{sgn}(b_k - a_k x - c_k y). \quad (6.6.11b)$$

Теперь перейдем к трудностям. Если сигнум-функция всегда принимает значения $+1$ или -1 , то выражение (6.6.11a) подразумевает, что множители a_k можно разделить на две группы равного веса. Ясно, что в действительности большая часть массивов чисел

не может быть так сбалансирована (в частности, если все множители a_i , за исключением одного, являются целыми числами). Трудности могут быть исключены, если хотя бы одно из уравнений (6.6.9) решается точно также, как сигнум-функция становится для этого члена неопределенной величиной. Любое алгебраическое

```

SUBROUTINE SKEWER(ND,N,W,F,GU,GD,SMALL,K,T,ML,MH)
C SOLVE RANK 1 OVERDETERMINED EQUATIONS WITH SKEW NORM
C INPUTS- N,W,F,GU,GD,SMALL,K. OUTPUTS- K,T,ML,MH.
C FIND T TO MINIMIZE
C N
C LS = SUM SKEWNORM(K,F(K)-W(K)*T)
C K=1
C WHERE (GU(K)*(ER-SMALL) IF ER.GT.+SMALL GU.GT.0
C SKEWNORM(K,ER) = (GD(K)*(ER+SMALL) IF ER.LT.-SMALL GD.LT.0
C (0. IF ABS(ER).LE.SMALL.GE.0.
C GU,GD,W,AND F ARE REFERENCED INDIRECTLY AS W(K(I)),I=1,N ETC
C MINIMA WILL BE AT EQUATIONS K(ML),K(ML+1),...K(MH).
DIMENSION W(ND),F(ND),K(ND),GU(ND),GD(ND)
DIMENSION G(1000)
LOW=1
LARGE=N
ML=N
MH=1
GN=0.
GP=0.
DO 50 ITRY=1,N
L=K(LOW+MOD((LARGE-LOW)/3+ITRY,LARGE-LOW+1))
IF(ABS(W(L)).EQ.0.) GO TO 50
T=F(L)/(W(L))
F(L)=W(L)*T
DO 10 I=LOW,LARGE
L=K(I)
ER=F(L)-W(L)*T
G(L)=0.
IF(ER.GT.SMALL) G(L)=-W(L)*GU(L)
10 IF(ER.LT.-SMALL)G(L)=-W(L)*GD(L)
CALL SPLIT(LOW,LARGE,K,G,MLT,MHT)
GNT=GN
DO 20 I=LOW,MLT
20 GNT=GNT+G(K(I))
GPT=GP
DO 30 I=MHT,LARGE
30 GPT=GPT+G(K(I))
GPLX=0.
GMIX=0.
DO 40 I=MLT,MHT
L=K(I)
IF(W(L).LT.0.)GPLX=GPLX-W(L)*GU(L)
IF(W(L).GT.0.)GPLX=GPLX-W(L)*GD(L)
.F(W(L).GT.0.)GMIX=GMIX-W(L)*GU(L)
40 IF(W(L).LT.0.)GMIX=GMIX-W(L)*GD(L)
GRAD=GNT+GPT
IF((GRAD+GPLX)*(GRAD+GMIX).LT.0.) GO TO 60
IF(GRAD.GE.0.)LOW=MHT+1
IF(GRAD.LE.0.)LARGE=MLT-1
IF(LOW.GT.LARGE) GO TO 60
IF(GRAD.GE.0.)GN=GNT+GMIX
IF(GRAD.LE.0.)GP=GPT+GPLX
IF((GRAD+GPLX).EQ.0.)ML=MLT
IF((GRAD+GMIX).EQ.0.)MH=MHT
50 CONTINUE

```

РИС. 6.4. Подпрограмма вычисления взвешенных и несимметричных медиан (не-

несоответствие можно быстро устранить с помощью рис. 6.3. Этот результат, сходный с определением минимума для одного неизвестного, обычно наблюдается на краях, где первая производная разрывна. Аналогичная ситуация возникает и для уравнения (6.6.116). Обычно для N уравнений и M неизвестных именно M из N уравнений точно удовлетворяются, чтобы дать возможность уменьшить градиент погрешности до минимума. В области линейного программирования обычно принято рассматривать любое невырожденное подмножество M из N уравнений как систему *базисных уравнений*. Частная система из M уравнений, которая решается, когда погрешности минимизированы, называется *оптимальным базисом*.

Хотя линейное программирование развилось в двадцатом веке, основные его идеи были хорошо известны до Лапласа в восемнадцатом веке. Например, в статье Гаусса «Theoria Motus Corporum Coelestium», которая была опубликована в 1809 г. [24], указывается, что Лаплас использовал другой способ решения линейных уравнений, число которых больше числа неизвестных величин, чем ранее предложенный Босковичем, а именно, способ с вычислением разностей, которые считались положительными и должны были быть столь малыми, насколько это возможно. Можно легко пока-

```

60  ML=MINO(ML,MLT)
    MH=MAXO(MH,MHT)
    RETURN
    END

SUBROUTINE SPLIT(LOW,LARGE,K,G,ML,MH)
C   GIVEN G(K(I)),I=LOW,LARGE
C   THEN REARRANGE K(I),I=LOW,LARGE AND FIND ML,MH SO THAT
C   (G(K(I)),I=LOW,(ML-1)) .LT. 0 AND
C   (G(K(I)),I=ML,MH)=0. AND
C   (G(K(I)),I=(MH+1),LARGE) .GT. 0.
    DIMENSION K(LARGE),G(41)
    ML=LOW
    MH=LARGE
10  ML=ML-1
20  ML=ML+1
    IF(G(K(ML)))20,30,30
30  MH=MH+1
40  MH=MH-1
    IF(G(K(MH)))50,50,40
50  KEEP=K(MH)
    K(MH)=K(ML)
    K(ML)=KEEP
    IF(G(K(ML)).NE.G(K(MH)))GO TO 10
    DO 60 I=ML,MH
    I7=I
    IF(G(K(I)).NE.0.0) GO TO 70
60  CONTINUE
    RETURN
70  KEEP=K(MH)
    K(MH)=K(I7)
    K(I7)=KEEP
    GO TO 30
    END

```

симметричная медиана часто называется квантилем)

зять, что система определения неизвестных величин, основанная только на этом принципе, неизбежно должна (за исключением специальных случаев, в которых этот способ остается в некоторой степени неопределенным) точно удовлетворяться для многих уравнений из числа предложенных, имеющих неизвестные величины. Остальные уравнения вводятся в рассмотрение только до тех пор, пока они помогают *определить решение*.

Дальнейшее исследование этой проблемы и многочисленные геофизические приложения можно найти в работе [25].

Теперь опишем простую, но эффективную методику уменьшения многомерной поверхности погрешностей. Положение точки \mathbf{x} на линии, проходящей через точку \mathbf{x}_0 , можно определить с помощью скалярного параметра t . Направление линии определено M компонентами вектора \mathbf{g} . Тогда любую точку \mathbf{x} на линии можно представить в виде

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{g}t. \quad (6.6.12)$$

Подставляя выражение (6.6.12) в переопределенную систему

$$\mathbf{A}\mathbf{x} \approx \mathbf{b}, \quad (6.6.13)$$

получим

$$\mathbf{A}(\mathbf{x}_0 + \mathbf{g}t) \approx \mathbf{b}; \quad (6.6.14a)$$

$$(\mathbf{A}\mathbf{g})t \approx \mathbf{b} - \mathbf{A}\mathbf{x}_0. \quad (6.6.14b)$$

Определим величины \mathbf{w} и \mathbf{e} с помощью выражений

$$\mathbf{w} = \mathbf{A}\mathbf{g}; \quad (6.6.15a)$$

$$\mathbf{e} = \mathbf{b} - \mathbf{A}\mathbf{x}_0. \quad (6.6.15b)$$

Тогда формула (6.6.14b) дает

$$\mathbf{w}t \approx \mathbf{e}. \quad (6.6.16)$$

Решение уравнения (6.6.16) путем минимизации суммы абсолютных погрешностей дает также минимум погрешности вдоль линии в выражении (6.6.14a). Но выражение (6.6.16) представляет собой ранее обсуждавшуюся задачу для взвешенной медианы. Напомним, что решение уравнения (6.6.16) для t , дающее минимум абсолютной погрешности, будет точно удовлетворять одному из уравнений в системе (6.6.16). Пусть $t = e_k / w_k$. Для этой величины t k -е уравнение из системы (6.6.16) также будет точно удовлетворяться. Теперь оно может быть введено в число базисных уравнений. Покажем далее, как выбрать вектор \mathbf{g} так, чтобы он постоянно удовлетворял k -му уравнению (оставался в k -й гиперплоскости), когда мы подбираем t в следующей итерации.

Теперь необходима система базисных уравнений. Это система из M уравнений, которые временно удовлетворяются. Тогда, поскольку в базис вводятся новые уравнения путем определения взвешенной медианы, прежние уравнения исключаются. Стратегия настоящего алгоритма проста: исключить то из уравнений, которое имеет большую длину. Обозначим наши базисные уравнения с по-

мощью выражения

$$\mathbf{A}'\mathbf{x}=\mathbf{d}', \quad (6.6.17)$$

где \mathbf{A}' — квадратная матрица.

Обратную матрицу по отношению к матрице \mathbf{A}' , которая нам понадобится, обозначим через \mathbf{B} . Теперь предположим, что мы решили отбросить p -е уравнение из базиса матрицы \mathbf{A}' . Тогда для вектора \mathbf{g} выделим p -й столбец матрицы \mathbf{B} . Для наглядности этих вычислений отметим, что так как $\mathbf{A}'\mathbf{B}=\mathbf{I}$, то M — вектор $\mathbf{A}'\mathbf{g}$ — является p -м столбцом тождественной матрицы. Поэтому у N вектора $\mathbf{w}=\mathbf{A}\mathbf{g}$ имеются компоненты, равные $+1$, $M-1$ компонент, равных нулю, и $N-M$ других, неопределенных элементов. Если k -е уравнение в системе (6.6.13) или (6.6.16) сохраняется в базисе системы (6.6.17), то k -е уравнение в системе $\mathbf{A}\mathbf{g}t=\mathbf{d}-\mathbf{A}\mathbf{x}$ теперь имеет вид

$$0 \cdot t=0. \quad (6.6.18)$$

В последнем выражении нуль слева является элементом из тождественной матрицы, а нуль справа следует из утверждения, что k -е уравнение точно удовлетворяется. Очевидно, теперь можно по желанию выбрать t таким большим, чтобы получить новый локальный минимум и k -е уравнение всегда точно удовлетворялось. Имеется также уравнение вида

$$1 \cdot t=0. \quad (6.6.19)$$

Последнее равенство удовлетворяется только при условии, что t равно нулю. Геометрически это означает, что если мы должны двигаться в направлении минимума, то это уравнение не удовлетворяется и, следовательно, мы выходим из соответствующей гиперплоскости. Это уравнение одно из тех, которые не входят в базис. Конечно, если $t \rightarrow 0$, то это уравнение включается в базис. Предыдущий шаг повторяется до тех пор, пока после M последовательных итераций уравнение исключается из базиса благодаря его старению и немедленно восстанавливается при $t=0$. Это означает, что базис не может быть более улучшен и что мы нашли оптимальный базис и окончательное решение.

7. ПРИМЕНЕНИЕ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ ДЛЯ ОПРЕДЕЛЕНИЯ ФОРМЫ ВОЛНОВОГО ИМПУЛЬСА

Как известно, методы вычислительной математики позволяют находить координаты экстремальных точек на кривой. Вариационное исчисление дает возможность находить экстремальные функции. При этом непрерывное пространство может быть аппроксимировано сеткой с нерезкими границами. Пространственная задача описывается неограниченными матрицами, являющимися проблемой в вариационном исчислении. Однако этот недостаток частично компенсируется регулярной формой матриц. В этой главе будут рассмотрены примеры использования метода наименьших квадратов при определении формы волнового импульса и установлении соотношений между группами сигналов различной формы. Эти задачи приводят к массивным полным матрицам, называемым блоком Теплица, для решения которых разработан особый метод.

7.1. ПРОГНОЗИРУЮЩИЕ И ФОРМИРУЮЩИЕ ФИЛЬТРЫ

Предположим, что сейсмический сигнал описывается данными $\mathbf{b}=(b_0, b_1, \dots, b_n)$. Необходимо построить фильтр вида $\mathbf{f}=(f_0, f_1, \dots, \dots, f_n)$. Определим фильтрацию таким образом, что, когда данные \mathbf{b} поступают на вход фильтра \mathbf{f} , выходной сигнал \mathbf{c} описывается произведением матриц

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n+m} \end{bmatrix} = \begin{bmatrix} b_0 & 0 & \dots & 0 \\ b_1 & b_0 & & 0 \\ b_2 & b_1 & b_0 & \\ \vdots & \vdots & \vdots & \vdots \\ b_n & \vdots & b_2 & \\ \vdots & \vdots & \vdots & \\ 0 & b_n & \vdots & \\ \vdots & \vdots & \vdots & b_{n-1} \\ \vdots & \vdots & \vdots & \\ 0 & 0 & \dots & b_n \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_m \end{bmatrix} \quad (7.1.1)$$

Эту операцию часто называют *полной сверткой импульсной характеристики фильтра и входной последовательности*. Это то же самое, что и идентификация коэффициентов при перемножении многочленов.

Рассмотрим теперь еще один сейсмический сигнал \mathbf{d} , имеющий несколько компонент вида \mathbf{c} . Будем считать величину \mathbf{d} требуемым выходом фильтра, а \mathbf{c} действительным выходом, являющимся функцией входных данных \mathbf{b} и реакции фильтра \mathbf{f} . Необходимо определить характеристику фильтра \mathbf{f} таким образом, чтобы величины \mathbf{c} и \mathbf{d} были практически неотличимы друг от друга. Для этого выберем \mathbf{f} так, чтобы квадрат длины разностного вектора $\mathbf{c} - \mathbf{d}$ имел минимум (в $n+m+1$ -мерном пространстве). Воспользуемся методом наименьших квадратов для решения системы переопределенных уравнений:

$$\begin{bmatrix} b_0 & 0 & \dots & 0 \\ b_1 & b_0 & 0 & \\ b_2 & b_1 & b_0 & \\ \vdots & \vdots & \vdots & \\ b_n & \vdots & b_1 & \\ 0 & b_n & \vdots & \\ \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \\ 0 & 0 & b_{i-1} & \\ & & b_n & \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_m \end{bmatrix} \approx \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{n+m} \end{bmatrix} \quad (7.1.2)$$

Используя «быстрый приближенный метод», рассмотренный в предыдущей главе, можно умножить выражение (7.1.2) слева на транспонированную матрицу. В результате получим матрицу Теллица:

$$\begin{bmatrix} r_0 & r_1 & r_2 & \dots & r_m \\ r_1 & r_0 & r_1 & & \\ r_2 & r_1 & r_0 & & \\ \vdots & & \vdots & \ddots & \\ \vdots & & \vdots & & \\ r_m & & r_0 & & \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_m \end{bmatrix}, \quad (7.1.3)$$

где r_k — автокорреляционная функция входных данных x_k ; g_k — взаимокорреляционная функция входных данных x_k с требуемым выходным сигналом d_k . О технике вычислений см. раздел 7.5.

Формулы этого раздела можно попытаться использовать также для предсказания будущих значений временного ряда по его про-

шлым значениям. Например, последовательность f_1, f_2, \dots, f_m можно рассматривать как фильтр, прогнозирующий x_{t+10} значение временного ряда по его прошлым значениям $x_t, x_{t-1}, \dots, x_{t-m+1}$, если с помощью метода наименьших квадратов решить уравнение

$$\begin{bmatrix} x_{10} \\ x_{11} \\ x_{12} \\ \vdots \\ \vdots \end{bmatrix} \approx \begin{bmatrix} x_0 & x_{-1} & \dots & x_{-m+1} \\ x_1 & x_0 & & x_{-m+2} \\ x_2 & x_1 & & \\ x_3 & x_2 & & \\ \vdots & \vdots & & \\ \vdots & \vdots & & \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \end{bmatrix}. \quad (7.1.4)$$

Матрицу (7.1.4) можно продолжить вниз до тех пор, пока имеются данные. В приложениях величина t в формуле (7.1.4) должна прогнать все предыдущие значения t . Решение уравнений для фильтра \mathbf{f} характеризует такой временной ряд, для которого фильтр \mathbf{f} можно использовать с целью предсказания будущих значений того же ряда, не входивших в уравнение, определяющее фильтр \mathbf{f} .

Если высота матрицы (7.1.4) намного больше ее ширины, то можно попытаться получить требуемый конечный результат другим путем. Используя замену

$$\begin{bmatrix} x_{10} \\ x_{11} \\ \vdots \\ \vdots \end{bmatrix} \approx \begin{bmatrix} x_0 & 0 & 0 \\ x_1 & x_0 & 0 \\ x_2 & x_1 & \\ x_3 & x_2 & \\ \vdots & \vdots & \\ \vdots & \vdots & \\ \text{Нули} & \text{Нули} & \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ \vdots \\ f_m \end{bmatrix}, \quad (7.1.5)$$

найдем, что наименьшие квадраты нормального уравнения образуют матрицу Тейлица, тогда как для выражения (7.1.4) матрицы Тейлица, обладающей многими преимуществами как в теоретическом, так и в вычислительном плане, не существует.

Особый интерес представляет фильтр, описываемый уравнениями вида

$$\begin{bmatrix} x_0 & \text{Нули} \\ x_1 & x_0 \\ x_2 & x_1 & x_0 & \cdot \\ x_3 & x_2 & x_1 & \cdot \\ \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \\ x_n & x_n \\ \text{Нули} & x_n \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ \vdots \\ a_m \end{bmatrix} \approx \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}. \quad (7.1.6)$$

Такие фильтры называются *фильтрами ошибки предсказания в единичном интервале*, поскольку величины a_k , действуя на последовательность $(x_{t-1}, x_{t-2}, \dots)$, стремятся ослабить влияние данных x_t . Таким образом, действие a_k на последовательность $(x_{t-1}, x_{t-2}, \dots)$ даст взятое с обратным знаком наилучшее предсказанное значение временного ряда x_t на основе последовательности $(x_{t-1}, x_{t-2}, \dots)$. Как следствие выражения (7.1.6), нормальные уравнения образуют квадратную систему

$$\begin{bmatrix} r_0 & r_1 & r_2 & \dots \\ r_1 & r_0 & r_1 & \\ r_2 & r_1 & r_0 & \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (7.1.7)$$

Следует отметить, что для расчета фильтра ошибки предсказания используется только функция автокорреляции временного ряда, а не сам ряд. Как видно из формулы (3.3.6), решениями этих уравнений являются коэффициенты минимально фазового многочлена.

Решения уравнений Теплица (7.1.3) при достаточно произвольной правой части необязательно минимально фазовые. Однако для ускорения вычислений в рассматриваемых условиях можно соответствующим образом обобщить рекурсивный фильтр Левинсона. Это показано в разд. 7.5 для многоканального рекурсивного фильтра Левинсона.

Упражнения

1. Найдите трехчленный фильтр с нулевой задержкой, обращающий сейсмический импульс вида (1,2). Сопоставьте полученную погрешность с погрешностью того же выражения для сигнала вида (2,1). Сравните форму сигналов. Достаточно полное обсуждение погрешностей минимально квадратичных инверсных фильтров приведено в работе [26]. В процессе расчета может оказаться, что длина фильтра становится бесконечной. Это возможно, хотя сумма квадратов погрешностей стремится к нулю в двух случаях.

а. При обращении с нулевой задержкой, но только в том случае, когда сейсмический импульс, который должен быть обращен, является минимально фазовым.

б. Если обращаемый сейсмический импульс не является минимально фазовым. Тогда погрешность стремится к нулю, если выходной сигнал запаздывает так, что $d = (\dots, 0, 0, 1, 0, 0, \dots)$. Рассчитайте трехчленный фильтр с запаздыванием для обращения сигнала вида (1,2), когда $d = (0, 1, 0, 0)$ или $d = (0, 0, 1, 0)$.

2. Датчик давления, помещенный в глубокую скважину, записывает выходящие сейсмические волны, а через некоторый интервал времени t_0 такие же волны, идущие вниз, с противоположным знаком. Определите запаздывание и найдите минимально квадратичный фильтр без задержки длиной в m элементов, подавляющий повторный импульс. (Вы можете предполагать существование решения больших матриц рассматриваемого типа. Попытайтесь использовать фильтры вида $f_k = \alpha + \beta k$, где α и β — скаляры.) Какова погрешность фильтра в зависимости от его длины?

3. Пусть имеются данные $b_t = (\dots, 1, 1, -2, 1, 1, -2, \dots)$. Найдите с помощью метода наименьших квадратов наилучший одночленный фильтр, предсказывающий значения b_t с использованием только величины b_{t-1} . Найдите также наилучший двучленный фильтр, использующий величины b_{t-1} и b_{t-2} . Аналогичным образом найдите наилучший трехчленный фильтр. Какова погрешность фильтра в зависимости от времени в каждом случае?

7.2. ВЫЧИСЛЕНИЕ СПЕКТРА МЕТОДОМ БУРГА [27]

Принцип неопределенности гласит, что если функция времени несет большую часть своей энергии во временном интервале Δt , то ее Фурье-преобразование (спектр) содержит большую часть своей энергии в полосе частот $\Delta f \geq 1/\Delta t$. С другой стороны, часто считают, что, имея простой стационарный временной ряд длительностью Δt , можно получить наилучшее разрешение по частоте при выполнении условия $\Delta t = 1/\Delta f$. Различие между этими формулировками заключается в том, что исходная функция в первом случае равна нулю вне интервала Δt , а во втором она продолжается вне данного интервала. Если выборка данных может быть продолжена указанным образом за пределы заданного интервала, то разрешение Δf по частоте значительно уменьшится по сравнению с $1/\Delta t$. Точнее, разрешающая способность зависит от предсказуемости данных на концах выборки. Если имеющийся сегмент стационарного процесса короче автокорреляционной функции стационарного процесса, то процедура определения спектра, предложенная Бургом, значительно лучше метода усеченного Фурье-преобразования (с конечными пределами). Этот случай часто встречается в физических задачах, когда имеют дело с резонансными явлениями, время затухания которых значительно больше времени наблюдения, или при изучении пространственных функций, когда каждая точка в пространстве отображается (представляется) другим прибором.

Если спектр $R(Z)$ вычисляется с помощью выражения $\bar{X}(1/Z)X(Z)$, где $X(Z)$ — многочлен, составленный для $N+1$ точек с известными данными, то коэффициенты функции $R(Z)$ рассчитываются по формуле

$$r_k = \sum_{j=0}^{N-k} \bar{x}_{j+k} x_j. \quad (7.2.1)$$

Отметим, что r_0 рассчитывается из $N+1$ членов, а r_1 из N членов и т. д. Если N недостаточно большое число, то образуется нежелательный эффект смещения. Для устранения этого эффекта r_k следует вычислять по измененной формуле

$$r_k = \frac{1}{N-k+1} \sum_{j=0}^{N-k} \bar{x}_{j+k} x_j. \quad (7.2.2)$$

Трудности применения формулы (7.2.2), заключаются в том, что используемые для определения значений r_k выборки данных не являются точными автокорреляционными функциями. Например,

спектр сигнала не является положительным для всех частот, решение уравнения Теплица имеет разрыв и т. п.

Приближение Бурга устраняет краевые эффекты при вычислениях по формуле (7.2.1), а также снижает вероятность получения неправдоподобных результатов по формуле (7.2.2). Вместо вычисления функции автокорреляции r_k данных он вычисляет минимально фазовый фильтр ошибки предсказания непосредственно по исходным данным. Выход этого фильтра имеет «белый» спектр. (Если это не так, то цвет спектра¹ можно использовать для уточнения предсказания). Так как спектр выходного сигнала равен произведению спектра входного сигнала на спектр фильтра, то спектр сигнала на входе может быть вычислен как обращенный спектр фильтра ошибки предсказания. Узкие спектральные максимумы значительно легче представляются с помощью знаменателя, чем с помощью числителя.

Предположим, что задан сегмент данных в виде последовательности x_0, x_1, \dots, x_n . Тогда двучлен, описывающий фильтр ошибки предсказания вида (1, a) для временного ряда x_t , задается путем выбора величины a , минимизирующей выражение

$$E(a) = \sum_{t=1}^N |x_t + ax_{t-1}|^2. \quad (7.2.3)$$

К сожалению, рассмотрение различных примеров показало, что существуют временные ряды [подобные (1,2)], для которых величина $|a|$ может превышать единицу. Это неприемлемо, так как фильтры ошибки предсказания не являются минимально фазовыми, спектры неположительны и т. д. Напомним, что рассматриваемые фильтры определены в предыдущем разделе как зависящие только от корреляционной функции данных и не зависящие от самих данных. Это означает, что один и тот же фильтр получается как для временного ряда, так и для (комплексно-сопряженного) обращенного во времени временного ряда. Следовательно, ошибка предсказания будущих значений данных по формуле (7.2.3) будет увеличена за счет ошибки предсказания прошлых значений. Тогда

$$E(a) = \sum_{t=1}^N \{|x_t + ax_{t-1}|^2 + |\bar{x}_{t-1} + \bar{a}\bar{x}_t|^2\}. \quad (7.2.4)$$

Позже мы покажем, что минимизация выражения (7.2.4) всегда приводит к значениям величины $|a|$, меньшим единицы. Спектр мощности, вычисленный с учетом величины a , имеет вид $R=1/[1+(a/Z)(1+\bar{a}Z)]$. Полоса частот Δf может быть очень узкой, если величина a очень близка к единичной окружности.

Естественное обобщение формулы (7.2.4) для фильтров с боль-

¹ Точнее, отклонение от «белого» спектра. (Прим. перев.)

шим числом членов (более двух) приводит к минимизации выражения

$$E(a_1, a_2) = \sum_{t=2}^N \{|x_t + a_1 x_{t-1} + a_2 x_{t-2}|^2 + |\bar{x}_{t-2} + a_1 \bar{x}_{t-1} + a_2 \bar{x}_t|^2\}. \quad (7.2.5)$$

К сожалению, Бургу удалось найти временные ряды, для которых рассчитанный фильтр вида $A(Z) = 1 + a_1 Z + a_2 Z^2$ не является минимально фазовым. При этом величина $R = 1/[A(1/Z)A(Z)]$ не может применяться для вычисления спектра, так как значение $R(Z)$ должно определяться на единичной окружности, в то время как величина $1/A(Z)$ в пределе к ней не стремится.

Бург отметил, что рекурсия Левинсона всегда дает минимально фазовый фильтр. С помощью рекурсии Левинсона фильтр третьего порядка может быть построен из фильтра 2-го порядка с помощью выражения

$$\begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ a \\ 0 \end{bmatrix} - c \begin{bmatrix} 0 \\ a \\ 1 \end{bmatrix}.$$

Таким образом, Бург показал, что вместо применения метода наименьших квадратов для определения величин a_1 и a_2 , входящих в выражение (7.2.5), величину a можно найти из формулы (7.2.4), а затем методом наименьших квадратов определить величину c . Это должно быть сделано таким образом, чтобы обеспечить выполнение условия, при котором величина $|c|$ меньше единицы. Это гарантирует, что фильтр $A(Z) = 1 + a_1 Z + a_2 Z^2$ будет минимально фазовым. Таким образом, Бург предлагает преобразовать выражение (7.2.5) к виду

$$E(c) = \sum_{t=2}^N \{|x_t + a x_{t-1} - c(\bar{a} x_{t-1} + x_{t-2})|^2 + |\bar{x}_{t-2} + a \bar{x}_{t-1} - c(\bar{a} \bar{x}_{t-1} + \bar{x}_t)|^2\}. \quad (7.2.6)$$

Теперь выражение (7.2.6), являющееся суммой ошибок предсказания будущих и прошлых данных, минимизируется с учетом вариации величины c . (В следующей главе будет показано, что величина c является коэффициентом отражения.) Неизвестная величина a находится при минимизации выражения (7.2.4). Предположим, что величина $|c| < 1$. Обозначим через e_+ временной ряд $x_t + a x_{t-1}$, который является ошибкой предсказания будущих значений x_t , а через e_- — временной ряд $x_{t-2} + \bar{a} x_{t-1}$ ошибки предсказания прошлых значений. С учетом этого формула (7.2.6) принимает вид

$$E = \sum_t \{|e_+ - c e_-|^2 + |\bar{e}_- - c \bar{e}_+|^2\} = \sum_t \{(\overline{e_+ - c e_-})(e_+ - c e_-) + (\overline{\bar{e}_- - c \bar{e}_+})(\bar{e}_- - c \bar{e}_+)\}. \quad (7.2.7)$$

Находя производную последнего выражения по c и приравнявая ее к нулю, имеем:

$$0 = \sum_t [\bar{e}_- (e_+ - ce_-) + e_+ (\bar{e}_- - c\bar{e}_+)].$$

Откуда

$$c = \frac{+\sum_t \bar{e}_- e_+}{\sum_t (\bar{e}_+ e_+ - \bar{e}_- e_-)}. \quad (7.2.8)$$

(Отметим, что вычисление по формуле $\partial E/\partial c=0$ приводит к тому же результату.) То, что величина $|c|$ всегда меньше единицы, можно видеть из того, что длина вектора $e_+ \pm e_-$ всегда положительна. В частности,

$$\begin{aligned} \sum_t |e_+ \pm e_-|^2 &\geq 0; \\ \sum_t (\bar{e}_+ e_+ \pm \bar{e}_+ e_- \pm \bar{e}_- e_+ + \bar{e}_- e_-) &\geq 0; \\ \sum_t (\bar{e}_+ e_+ + \bar{e}_- e_-) &> 2|\bar{e}_- e_+|; \\ |c| &\leq 1. \end{aligned} \quad (7.2.9)$$

Если мы переопределим величины e_+ и e_- следующим образом:

$$e_+ \leftarrow e_+ - ce_-; \quad (7.2.10a)$$

$$e_- \leftarrow e_- - c\bar{e}_+; \quad (7.2.10б)$$

то получим трехчленный фильтр ошибки предсказания будущих и прошлых данных вида $(1, a'_1, a'_2) = (1, a - c\bar{a}_1, -c)$. Теперь можно вернуться к формуле (7.2.7) и продолжить ее рекурсивно. Величины e_- и e_+ как рекурсивные продолжения постепенно становятся непредсказуемыми случайными числами. Поэтому необходим фильтр $A(Z)$, который может отфильтровать в последовательности $X(Z)$ как будущие, так и прошлые значения и сделает спектр выхода «белым». Поскольку выход имеет постоянный спектр, то спектр входной последовательности должен соответствовать обращенному спектру фильтра.

На рис. 7.1 приведена машинная программа, реализующая алгоритм Бурга. Данные X являются вектором размерности LX . Выбор условия $LC \leq LX$ является компромиссом между высоким разрешением и большими затратами машинного времени. Плотность точек на оси частот, которая контролируется условием $N2048 \gg LX$, выбирается исходя из удобства их нанесения и должна быть достаточно большой для выделения узких спектральных линий.

В последующих главах мы рассмотрим интерпретацию процесса распространения волны на основе алгоритма Бурга. В слоистых средах параметры c_k интерпретируются как коэффициенты отраже-

ния; векторы e_+ и e_- рассматриваются соответственно как восходящие и нисходящие волны. Будет рассмотрен также весь процесс расчета последовательности коэффициентов c_k по сейсмограммам для все более глубоко залегающих внутри Земли непрерывных поверхностей и определения коэффициентов отражения c_k для модели Земли.

Упражнения

1. Рассмотрите временной ряд из десяти точек $(1, 1, 1, -1, -1, -1, 1, 1, 1, -1)$. Рассчитайте полиномы третьей степени по Z для величин C и A . Сравните автокорреляционную функцию l_t , рассчитанную по методу Бурга, с функцией $R(Z)$,

```

      SUBROUTINE BURGC(LX,X,EP,EM,LC,C,A,N2048,S)
C GIVEN A TIME SERIES X(1...LX) GET ITS LOG SPECTRUM S(1...N2048)
      DIMENSION X(LX),EP(LX),EM(LX),C(LC),A(LC),S(N2048)
      COMPLEX X,EP,EM,C,A,S,TOP,BOT,EPI,CONJG,CLOG
      DO 10 I=1,N2048
10    S(I)=0.
      A(1)=1.
      DO 20 I=1,LX
      EM(I)=X(I)
20    EP(I)=X(I)
      DO 60 J=2,LC
      TOP=0.
      BOT=0.
      DO 30 I=J,LX
      BOT=BOT+EP(I)*CONJG(EP(I))+EM(I-J+1)*CONJG(EM(I-J+1))
30    TOP=TOP+EP(I)*CONJG(EM(I-J+1))
      C(J)=2*TOP/BOT
      DO 40 I=J,LX
      EPI=EP(I)
      EP(I)=EP(I)-C(J)*EM(I-J+1)
40    EM(I-J+1)=EM(I-J+1)-CONJG(C(J))*EPI
      A(J)=0.
      DO 50 I=1,J
50    S(I)=A(I)-C(J)*CONJG(A(J-I+1))
      DO 60 I=1,J
60    A(I)=S(I)
      CALL FORK(N2048,S,+1.)
      DO 70 I=1,N2048
70    S(I)=-CLOG(S(I))*2.
      RETURN
      END

```

РИС. 7.1. Машинная программа, реализующая алгоритм Бурга

вычисленной по ограниченной выборке данных, и с функцией $R(Z)$, вычисленной с помощью интуитивно заданной выборки данных, растянутой во времени от минус до плюс бесконечности.

2. Модифицируйте программу, приведенную на рис. 7.1, таким образом, чтобы рассчитать и включить в нее скалярный множитель V , относящийся к спектру.

7.3. АДАПТИВНЫЕ ФИЛЬТРЫ

Адаптивным фильтром называется фильтр, параметры которого изменяются во времени в соответствии с изменением во времени фильтруемого временного ряда. Предположим, что предсказывается фильтр по одной точке впереди временного ряда. Тогда, исполь-

зую набор прошлых значений временного ряда, можно сконструировать фильтр, применение которого для вводимых в данный момент данных позволяет предсказать будущие значения входных данных. По прошествии некоторого времени, естественно, возникает необходимость пересчитать фильтр на основе новых входных данных. Как часто должен пересчитываться фильтр? В принципе нет причин, мешающих пересчитывать фильтр достаточно часто, например после поступления каждой новой точки входных данных. На практике частота пересчета фильтра обычно ограничивается затратами машинного времени. Чтобы рассчитать фильтр длиной n операторов требуется затратить n умножений и сложений на каждую новую точку входной последовательности. Для пересчета фильтра с помощью рекурсии Левинсона требуется свыше n^2 умножений и сложений. Однако часто полагают, что фильтр необходимо изменять лишь после поступления некоторого небольшого числа входных данных. По этой причине воспользуемся алгоритмом адаптивного фильтра Видрова [28], который позволяет изменить фильтр с помощью лишь n арифметических операций. Тогда новый фильтр может рассчитываться для каждой точки входной последовательности данных.

Рассмотрим для определенности двучленное предсказание ситуации, когда ошибка предсказания e_t временного ряда x_t , состоящего из двух его прошлых значений,

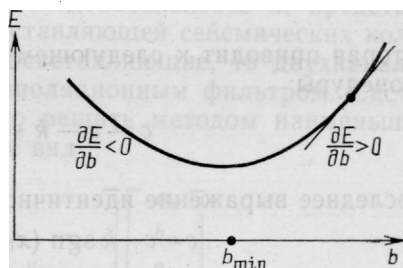
$$e_t = x_t - bx_{t-1} - cx_{t-2}. \quad (7.3.1)$$

Тогда сумма квадратов ошибок предсказания

$$E = \sum_t e_t^2 = \sum_t (x_t - bx_{t-1} - cx_{t-2})^2. \quad (7.3.2)$$

Если параметр b выбран правильно, то должно выполняться условие $\partial E / \partial b = 0$. Однако если свойства временного ряда x_t изменяют-

РИС. 7.2. Поведение знака частной производной, когда $b > b_{\min}$ или $b < b_{\min}$



ся во времени, то производная $\partial E / \partial b$ может отличаться от нуля, например, когда в сумму (7.3.2) включаются новые данные. Величина E является положительной квадратичной функцией параметра b . Если величина $\partial E / \partial b$ становится положительной, то параметр b должен быть уменьшен, и, наоборот, если производная становится отрицательной, то величина b должна быть увеличена (рис. 7.2).

Из формулы (7.3.2) после дифференцирования по b имеем

$$\frac{\partial E}{\partial b} = - \sum_{i=0}^t 2e_i x_{i-1}. \quad (7.3.3)$$

Как видно, приращение производной за счет добавления отсчета в точке x_t равно $-2e_t x_{t-1}$. Таким образом, появляется необходимость изменить величины b и c следующим образом:

$$\begin{bmatrix} b \\ c \end{bmatrix} \leftarrow \begin{bmatrix} b \\ c \end{bmatrix} + k e_t \begin{bmatrix} x_{t-1} \\ x_{t-2} \end{bmatrix}. \quad (7.3.4)$$

Здесь k — коэффициент, масштабирующий результат изменений, вносимых в параметры b и c за один такт (дискрет). Если коэффициент k выбран очень малым, то настройка осуществляется очень медленно. При слишком больших значениях k минимум легко перейти; однако возможно, что при следующем шаге система перебросятся обратно. Изменение коэффициента k определяется, в частности, на основе изучения свойств временного ряда x_t .

Известны многочисленные вариации рассматриваемого подхода. Можно использовать норму L_1 и минимизировать, например,

$$E(c) = \sum_t |c x_t - y_t|. \quad (7.3.5)$$

Окончательное выражение для адаптирующейся процедуры имеет вид

$$c \leftarrow c - k x_t \operatorname{sgn}(c x_t - y_t). \quad (7.3.6)$$

Уравнение (7.3.5), безусловно, является взвешенной медианой. Такой же, но более надежной процедурой является определение однородной взвешенной медианы

$$E(c) = \sum_t \left| c - \frac{y_t}{x_t} \right|, \quad (7.3.7)$$

которая приводит к следующему выражению для адаптирующейся процедуры:

$$c \leftarrow c - k \operatorname{sgn}\left(c - \frac{y_t}{x_t}\right). \quad (7.3.8a)$$

Последнее выражение идентично следующему:

$$c \leftarrow c - k \operatorname{sgn}(x_t) \operatorname{sgn}(c x_t - y_t). \quad (7.3.8b)$$

Примеры, описываемые выражениями (7.3.5) и (7.3.7), можно применить для стационарных процессов способом, подобным использованным в алгоритме Бурга. Наряду с выражением (7.3.7) можно минимизировать функцию

$$E = \sum_t \left\{ \left| c - \frac{y_t}{x_t} \right| + \left| c - \frac{x_t}{y_t} \right| \right\}. \quad (7.3.9)$$

Это приводит к подбору величины c внутри соответствующей области, потому что

$$-1 \leq \text{медиана} \left(0, \frac{y_t}{x_t}, \frac{x_t}{y_t} \right) \leq +1$$

для всех t .

Упражнения

1. Если величина x_t выражена в вольтах, то какова должна быть размерность коэффициента k ? Если величина x_t имеет среднюю квадратичную величину 100 В и интервал дискретизации во времени $\Delta t = 1$ мс, то каково должно быть значение коэффициента k , позволяющее адаптировать фильтр Видрова к новым условиям в течение примерно 1 с?

2. Задан временной ряд $x_t = (\dots, 1, 1, 1, 1, -4, 1, 1, 1, 1, -4, 1, 1, 1, 1, -4, 1, 1, \dots)$. Рассмотрите предсказание последовательности вида $x_{t+1} = cx_t$. Каков результат предсказания методом наименьших квадратов? Каков результат предсказания данных для нормы L_1 с помощью взвешенной и однородной взвешенной медианы?

7.4. РАСЧЕТ МНОГОКАНАЛЬНЫХ ФИЛЬТРОВ

Многоканальные фильтры применяются довольно часто. Например, с помощью вектор-предсказующего фильтра можно предсказать значения временного ряда, используя его прошлые значения и прошлые значения ряда других последовательностей. С помощью матричного предсказующего фильтра можно предсказать ряд последовательностей, используя их прошлые значения. Если последовательности в группе связаны между собой (взаимозависимы), то групповое предсказание лучше независимых предсказаний каждой из последовательностей в отдельных каналах. Возьмем для определенности два временных ряда x_t и y_t и попытаемся найти векторный фильтр, который преобразовал бы их в третий временной ряд d_t . Если последовательности d_t соответствует последовательности x_{t+1} , то для x_t необходим фильтр с предсказанием значений через один временной интервал. Если последовательность d_t представляет собой запись вертикальной составляющей сейсмических колебаний, а x_t и y_t — горизонтальные составляющие, то двухканальный фильтр можно назвать экстраполяционным фильтром. Система уравнений, которую необходимо решить методом наименьших квадратов, принимает в этом случае вид

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 & x_1 & y_1 \\ x_3 & y_3 & x_2 & y_2 & x_1 & y_1 \\ x_4 & y_4 & x_3 & y_3 & x_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \vdots \\ \vdots \\ a_m \\ b_m \end{bmatrix} \quad (7.4.1)$$

Запишем систему (7.4.1) в сокращенном виде

$$\mathbf{d} \approx \mathbf{Bf}. \quad (7.4.2)$$

Тогда, как следует из предыдущей главы, решением системы (7.4.1) является

$$\mathbf{f} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{d}. \quad (7.4.3)$$

Рассмотрим обращаемую матрицу \mathbf{R} . Для фильтра с тремя временными задержками

$$\mathbf{R} = \sum_t \begin{bmatrix} x_t \\ y_t \\ x_{t-1} \\ y_{t-1} \\ x_{t-2} \\ y_{t-2} \end{bmatrix} [x_t \ y_t \ x_{t-1} \ y_{t-1} \ x_{t-2} \ y_{t-2}]. \quad (7.4.4)$$

Вводя обозначения

$$r_{xx}(i) = \sum_t x_t x_{t+i};$$

$$r_{xy}(i) = \sum_t x_t y_{t+i}$$

и подобные им $r_{yx}(i)$ и $r_{yy}(i)$, приведем матрицу (7.4.4) к виду

$$\mathbf{R} = \begin{bmatrix} \begin{array}{cc|cc} r_{xx}(0) & r_{xy}(0) & r_{xx}(-1) & r_{xy}(-1) \\ r_{yx}(0) & r_{yy}(0) & r_{yx}(-1) & r_{yy}(-1) \end{array} & \begin{array}{cc} r_{xx}(-2) & r_{xy}(-2) \\ r_{yx}(-2) & r_{yy}(-2) \end{array} \\ \hline \begin{array}{cc|cc} r_{xx}(1) & r_{xy}(1) & r_{xx}(0) & r_{xy}(0) \\ r_{yx}(1) & r_{yy}(1) & r_{yx}(0) & r_{yy}(0) \end{array} & \begin{array}{cc} r_{xx}(-1) & r_{xy}(-1) \\ r_{yx}(-1) & r_{yy}(-1) \end{array} \\ \hline \begin{array}{cc|cc} r_{xx}(2) & r_{xy}(2) & r_{xx}(1) & r_{xy}(1) \\ r_{yx}(2) & r_{yy}(2) & r_{yx}(1) & r_{yy}(1) \end{array} & \begin{array}{cc} r_{xx}(0) & r_{xy}(0) \\ r_{yx}(0) & r_{yy}(0) \end{array} \end{bmatrix}. \quad (7.4.5)$$

Теперь можем расчлнить матрицу 6×6 (7.4.5) на блоки матриц 3×3 или на субматрицы 2×2 . Если задать блоки субматриц в виде

$$\mathbf{R}(\tau) = \begin{bmatrix} r_{xx}(\tau) & r_{xy}(\tau) \\ r_{yx}(\tau) & r_{yy}(\tau) \end{bmatrix} = \mathbf{R}^T(-\tau), \quad (7.4.6)$$

то матрица (7.4.5), определенная через блоки вида (7.4.6), примет форму

$$\mathbf{R} = \begin{bmatrix} R(0) & R(-1) & R(-2) \\ R(1) & R(0) & R(-1) \\ R(2) & R(1) & R(0) \end{bmatrix}. \quad (7.4.7)$$

Матрица (7.4.7) называется *блоком Тейлица (или мультиканалом Тейлица)*. Для матрицы Тейлица разработан особый метод решения, который рассматривается в следующем разделе.

Читатель, возможно, заметил, что матрица \mathbf{R} не зависит от ожидаемой последовательности \mathbf{d} на выходе. Это экономит вычисления в случаях, когда требуется найти более одного возможного выхода, что наблюдается, например, при необходимости предсказания нескольких различных последовательностей или интервалов будущих значений некоторой последовательности.

Упражнение

1. В упражнениях к главе 2 функции $B(Z)$ и $A(Z)$ были определены таким образом, что некоторый заданный степенной ряд $C(Z)$ выражался зависимостью $C(Z) = B(Z)/A(Z)$. Напишите нормальные уравнения (не решая их), чтобы сделать это приближенно с помощью минимизации выражения

$$\min(A, B) = \sum_t \left(B_t - \sum_{\tau} C_{t-\tau} A_{\tau} \right)^2,$$

где $A = (A_0, A_1, A_2)$, $B = (B_0, B_1, B_2)$, при условии $A_0 = 1$. (Можно показать с помощью рекурсии Левинсона, что функция $A(Z)$ не является минимально фазовой.)

7.5. РЕКУРСИЯ ЛЕВИНСОНА

Рекурсия Левинсона является простейшим методом решения нормальных уравнений. Можно показать, что она эквивалентна рекуррентному соотношению в теории ортогональных полиномов. Простота метода Левинсона объясняется тем, что матрица \mathcal{R} имеет в действительности только N различных элементов, тогда как общая матрица может иметь N^2 различных элементов.

Левинсон развил свой метод для отдельных временных рядов (основная идея описана в разделе 3.3). Очень небольшие дополнительные изменения исходных допущений позволяют осуществить рекурсию для множества временных рядов. Начнем с нормального уравнения ошибки предсказания. Для множества временных рядов так же, как и для одного временного ряда, проблема предсказания изменяется при обращении времени. Можно записать оба нормальных уравнения ошибки предсказания как для будущих, так и для прошлых значений в виде одного уравнения (7.5.1).

Поскольку краевые эффекты играют важную роль, покажем, что зная \mathcal{A} и \mathcal{B} и решение для трехчленного фильтра

$$\begin{bmatrix} R_0 & R_{-1} & R_{-2} \\ R_1 & R_0 & R_{-1} \\ R_2 & R_1 & R_0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & B_2 \\ A_1 & B_1 \\ A_2 & \mathbf{I} \end{bmatrix} = \begin{bmatrix} V_A & 0 \\ 0 & 0 \\ 0 & V_b \end{bmatrix}. \quad (7.5.1)$$

Можно найти решения \mathcal{A}' и \mathcal{B}' для четырехчленного фильтра:

$$\begin{bmatrix} R_0 & R_{-1} & R_{-2} & R_{-3} \\ R_1 & R_0 & R_{-1} & R_{-2} \\ R_2 & R_1 & R_0 & R_{-1} \\ R_3 & R_2 & R_1 & R_0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & B_3 \\ A_1 & B_2 \\ A_2 & B_1 \\ A_3 & \mathbf{I} \end{bmatrix}' = \begin{bmatrix} V_A & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & V_b \end{bmatrix}' \quad (7.5.2)$$

путем формирования линейных комбинаций величин \mathcal{A} и \mathcal{B} . Это можно сделать путем выбора постоянных матриц α и β в выражении

$$\begin{bmatrix} R_0 & R_{-1} & R_{-2} & R_{-3} \\ R_1 & R_0 & R_{-1} & R_{-2} \\ R_2 & R_1 & R_0 & R_{-1} \\ R_3 & R_2 & R_1 & R_0 \end{bmatrix} \left\{ \begin{bmatrix} I \\ A_1 \\ A_2 \\ 0 \end{bmatrix} \alpha + \begin{bmatrix} 0 \\ B_2 \\ B_1 \\ I \end{bmatrix} \beta \right\} = \begin{bmatrix} V_A \\ 0 \\ 0 \\ E_A \end{bmatrix} \alpha + \begin{bmatrix} E_B \\ 0 \\ 0 \\ V_B \end{bmatrix} \beta. \quad (7.5.3)$$

Изменим матрицы α и β таким образом, чтобы нижний элемент \mathcal{A} в правой части выражения (7.5.3) стремился к нулю. Тогда $\alpha = I$; $\beta = -V_B^{-1}E_A$. Аналогично устремим к нулю верхний элемент в \mathcal{B} правой части. Тогда имеем $\beta = I$; $\alpha = -V_A^{-1}E_B$.

Конечно, желательно решить не только задачу предсказания ошибки. Например, нужно перейти от матриц 3×3 к матрицам 4×4 в решении уравнений фильтров с произвольной правой частью \mathcal{G} . Это достигается с помощью подбора коэффициента γ в выражении (7.5.4) таким образом, чтобы выполнялось условие $E_f + V_B \gamma = G_3$:

$$\begin{bmatrix} \mathcal{R} \end{bmatrix} \left\{ \begin{bmatrix} F_2 \\ F_1 \\ F_2 \\ 0 \end{bmatrix} + \begin{bmatrix} B_3 \\ B_2 \\ B_1 \\ I \end{bmatrix} \gamma \right\} = \begin{bmatrix} G_0 \\ G_1 \\ G_2 \\ E_f \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ V \end{bmatrix} \gamma. \quad (7.5.4)$$

7.6. ФИЛЬТРЫ С ОГРАНИЧЕНИЯМИ

В геофизической теории часто рассматривают плоскую волну, регистрируемую группой приемников. При этом хотят увидеть форму сигнала на выходе каждого приемника. Однако в каждом приемнике присутствуют шумы, которые искажают сигнал и могут коррелироваться или не коррелироваться от одного приемника к другому. Очевидно, можно допустить наличие в каждом приемнике такого сильного шума, при котором сигнал одиночным приемником не обнаруживается вообще. Такая ситуация возникла у М. Левина [29], когда он пытался обнаружить слабый подземный ядерный взрыв группой сейсмоприемников. Для получения сигнала он предложил фильтр с подавлением шумов. Для начала предположим, что все сигналы приходят одновременно или, если времена их прихода различны, они по крайней мере известны. Так что данные разных каналов могут быть выравнены с помощью сдвига во времени. Теперь задача состоит в том, чтобы профильтровать каждый канал и затем сложить результаты всех каналов; при этом шумы должны ослабиться, а форма сигнала сохраниться. Пусть $\hat{i}_i(j)$

представляет собой весовую функцию фильтра для i -го канала и j -й задержки. Для примера рассмотрим эту функцию для случая двух каналов и трех временных задержек. Тогда фильтр Левина с ограничением, который предохраняет сигнал от искажений, имеет вид

$$\begin{aligned} 1 &= f_1(0) + f_2(0); \\ 0 &= f_1(1) + f_2(1); \\ 0 &= f_1(2) + f_2(2). \end{aligned} \quad (7.6.1)$$

В рассматриваемом случае сигнал не искажается, так как если один и тот же сигнал $s(Z)$ поступает в каждый канал, то выходом будет $s(Z) [f_1(Z) + f_2(Z)]$. Но $f_1 + f_2$ в этом случае равна величине $(1, 0, 0)$, а в общем случае — дельта-функции. Назовем систему уравнений (7.6.1) уравнениями с ограничениями, так как здесь меньше уравнений, чем неизвестных. Рассматриваемые уравнения можно записать в обычной форме:

$$\begin{bmatrix} -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ f_1(0) \\ f_2(0) \\ f_1(1) \\ f_2(1) \\ f_1(2) \\ f_2(2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (7.6.2)$$

Последнее выражение можно сокращенно представить в виде $\mathbf{Gf} = \mathbf{0}$. Применяв метод наименьших квадратов для минимизации общей энергии в фильтре, можно попытаться подавить как сигнал, так и помеху. Но, поскольку уравнения с ограничениями предотвращают подавление сигнала, следовательно, ослаблены будут лишь шумы. Если обозначить через \mathbf{R} спектральную матрицу входных данных, то фильтр \mathbf{f} определится решением уравнения, подобного следующему:

$$\begin{bmatrix} \mathbf{R} & \mathbf{G}^T \\ \mathbf{G} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \lambda \end{bmatrix} = \begin{bmatrix} V \\ 0 \\ \vdots \\ \vdots \end{bmatrix}. \quad (7.6.3)$$

Мы уже решали уравнения этого типа в предыдущих разделах.

Упражнения

1. В некоторых случаях, когда усиление каналов точно не контролируется, главные члены фильтра равны $f_1(0) = 100$ и $f_2(0) = -99$. Несмотря на то что этот фильтр удовлетворяет всем предъявленным к нему при проектировании требо-

ваниям, его нельзя считать удовлетворительным, так как положенное в его основу предположение об идентичности сигналов во всех каналах неточно. Это лишь допустимое приближение. Можете ли Вы предложить более подходящую матрицу для подавления шума?

2. Рассмотрите три сейсмоприемника, расставленных в ряд на поверхности Земли. Рассматривая фильтр с ограничениями, изучите какое значение имеет тот факт, что все сигналы приходят одновременно, т. е. волна падает (выходит) на поверхность Земли вертикально. Определите матрицу фильтра с ограничениями, пропускающего две волны: вертикально падающую и волну, которая образует временной ряд $x_1(t) = x_2(t+1) = x_3(t+2)$. Какой самый короткий фильтр может удовлетворительно передать обе эти волны и одновременно, насколько это возможно, подавить шумы?

3. Рассмотрите фильтр Левина для m каналов с фильтрами, содержащими k задержек. Какой размер будет иметь при этом матрица в выражении (7.6.3)?

8. ОБНАРУЖЕНИЕ СЛОЕВ С ПОМОЩЬЮ ФИЛЬТРАЦИИ РАССЕЯННЫХ ВОЛН

Волны встречаются почти во всех областях физики. Мы будем изучать здесь волны, не предполагая у читателя знания физики и дифференциальных уравнений. Для решения наших задач достаточно использовать только общие принципы запаздывания волн, их непрерывности и сохранения энергии. Полученные при этом результаты будут непосредственно применимы к звуковым волнам, волнам в воде, световым волнам в тонких пленках, нормально падающим на поверхность продольным и поперечным упругим волнам, электромагнитным волнам, линиям передач, электрическим многозвенным сетям и другим, подобным объектам. Рассматриваемые методы могут быть применены также в задачах диффузии. Наша главная цель — научить решать задачи расчета волновых полей по заданным коэффициентам отражения на границах и вычислять коэффициенты отражения границ по наблюдаемым волновым полям.

8.1. КОЭФФИЦИЕНТЫ ОТРАЖЕНИЯ И ПРОХОЖДЕНИЯ ВОЛН НА ГРАНИЦАХ РАЗДЕЛА

Рассмотрим два полупространства (воздух — земля). Если некоторая волна единичной амплитуды падает на эту границу раздела, то в соответствии с рис. 8.1 будут наблюдаться проходящая волна с амплитудой t и отраженная волна с амплитудой c .

Существует очень простая связь между t и c . Амплитуды волн имеют физический смысл, схожий с физическим смыслом давления, перемещения вещества или тангенциальных электрических или магнитных полей, причем эти физические переменные должны быть одинаковыми по обе стороны границы. Таким образом, на границе должно выполняться условие

$$t = 1 + c; \quad (8.1.1)$$

$$t' = 1 + c'. \quad (8.1.2)$$

Вероятно, удивительно, что t может быть больше единицы. Однако это явление легко наблюдается в океане, где волны становятся больше по мере приближения к берегу (пока не разбиваются). Энергия волны не определяется только ее высотой. Энергия равна квадрату амплитуды волны, умноженному на коэффициент пропорциональности Y , зависящий от свойств среды, в которой распространяется и измеряется волна. Если обозначить этот коэффициент для верхней среды через Y_1 , а для нижней среды через Y_2 , то

утверждение о равенстве энергии до падения волны на границу и после него записывается в виде

$$Y_2 l^2 = Y_2 c^2 + Y_1 t^2. \quad (8.1.3)$$

Решая последнее выражение относительно c , с учетом равенства (8.1.1) получим:

$$\begin{aligned} 0 &= -Y_2 + Y_2 c^2 + Y_1 (1 + c)^2; \\ 0 &= (c - 1)Y_2 + (c + 1)Y_1; \\ 0 &= c(Y_2 + Y_1) + (Y_1 - Y_2); \\ c &= \frac{Y_2 - Y_1}{Y_2 + Y_1}. \end{aligned} \quad (8.1.4)$$

В акустике переменные восходящие и нисходящие волны могут быть нормированы по давлению или скорости. При измерении ско-

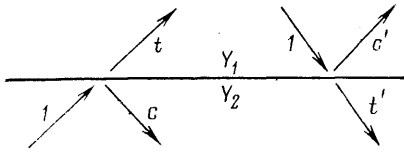


РИС. 8.1. Прохождение волн через границу раздела.

l — падающая волна; c — отраженная волна; t — проходящая волна; Y — проводимость среды

рости используют масштабный коэффициент, умноженный на квадрат скорости, который называют импедансом l . При измерении давления используют масштабный множитель, называемый полной проводимостью Y .

Когда энергия падает на границу с другой стороны, отраженная волна c' получается из формулы (8.1.4), если Y_1 и Y_2 поменять в ней местами, т. е.

$$c' = -c. \quad (8.1.5)$$

Полностью отражающей границей раздела является такая, которая не пропускает через себя энергию. Это наблюдается не только при $t=0$ или $c=-1$, но и при $t=2$ или $c=+1$. Чтобы уяснить это, обратим внимание на левую часть рис. 8.1, для которой

$$\frac{E_{\text{пр}}}{E_{\text{пад}}} = \frac{Y_1 t^2}{Y_2 l^2} = \frac{Y_1}{Y_2} \left(1 + \frac{Y_2 - Y_1}{Y_1 + Y_2} \right)^2 = \frac{Y_1}{Y_2} \left(\frac{2Y_2}{Y_1 + Y_2} \right)^2 = \frac{4Y_1 Y_2}{(Y_1 + Y_2)^2}, \quad (8.1.6)$$

где $E_{\text{пр}}$ — энергия проходящей волны; $E_{\text{пад}}$ — энергия падающей волны.

Уравнение (8.1.6) показывает, что энергия полностью проходит через границу, когда $Y_1 = Y_2$, но процент пропускаемой через границу энергии очень мал, если Y_1 сильно отличается от Y_2 .

Следует подчеркнуть значение правильного выбора переменных при рассмотрении отражения и пропускания волн на границах. Необходимо применять переменные, непрерывные на границах. Например, плотность среды является неподходящей переменной в аку-

стике, так как она претерпевает скачки на границе раздела сред. Давление или нормальная составляющая скорости больше подходят для описания интенсивности волны.

Для описания волн обычно используются два вида переменных, являющихся непрерывными функциями в области разрыва среды. Одна из этих переменных является скалярной величиной, подобно давлению, натяжению, электрическому напряжению, потенциалу, напряжению или температуре. Другая является вектором, вертикальной составляющей которого мы пользуемся. Примерами последней величины могут служить скорость, усилие, электрический ток, смещение и тепловой поток. Иногда волновая переменная является тензором. Если граничное условие заключается в обращении в нуль одной из составляющих смещения, то обычно говорят, что граница жесткая (недеформируемая). При обращении на границе в нуль давления или потенциала границу называют свободной (мягкой). Жесткая и свободная границы отражают волны с коэффициентами отражения, модули которых равны единице.

В главе рассматриваются основные математические свойства волн в слоях без каких-либо ограничений их частными физическими типами волн. Однако, чтобы не маскировать физический аспект математикой, дадим точное определение восходящих и нисходящих волн U и D в рамках классической акустики. В акустике рассматривают давление P и вертикальную составляющую скорости смещения частиц среды W (не путать со скоростью распространения волны v). Величины U и D можно определить из выражений

$$D = \frac{P + W/Y}{2}; \quad (8.1.7a)$$

$$U = \frac{P - W/Y}{2}; \quad (8.1.7b)$$

с обратными соотношениями

$$P = D + U; \quad (8.1.8a)$$

$$W = (D - U)Y. \quad (8.1.8b)$$

Возможны и другие способы определения U и D с различными масштабными множителями и знаками. В рассматриваемом случае справедливо соотношение $t=1+c$, которое ассоциируется с выражением (8.1.8a) и непрерывностью давления на границе раздела. Знаки минус в формулах (8.1.7) и (8.1.8) связаны с направлением оси z . Изменение направления оси z на противоположное изменяет W на $-W$. При этом меняются ролями величины U и D .

Отметим, что нисходящая волна D при $U \rightarrow 0$ сама полностью описывает процесс распространения как давления P , так и скорости W . Обращение в нуль функции U убеждает нас в том, что соотношение между обеими рассматриваемыми величинами $W/P=Y$ характеризует проводимость (волновую) материала среды. Энергия, как уже отмечалось, пропорциональна величине YP^2 при $1/W^2$, ввиду чего соотношение $W/P=Y$ позволяет заключить, что импе-

данс среды является обратной величиной ее полной проводимости $I=1/Y$.

Для звуковых волн в океане поверхность моря является почти идеальным отражателем вследствие сильного контраста между упругими свойствами воды и воздуха. Если эту границу раздела считать идеальным отражателем, то она является свободной поверхностью (рис. 8.2). Давление на свободной поверхности обра-

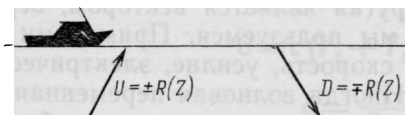


РИС. 8.2. Отражение волнового импульса $R(Z)$ от поверхности моря. Волны: U — восходящая, D — нисходящая

щается в нуль ($U+D=0$). Следовательно, коэффициент отражения на этой поверхности равен -1 . Вертикальная скорость на поверхности пропорциональна величине $D-U$. Теоретически для наблюдения волн на поверхности следовало бы измерять величину W . Однако исходя из практических соображений обычно наблюдают давление P в воде на глубине нескольких долей длины волны от

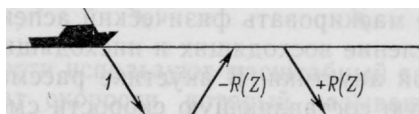


РИС. 8.3. Возбуждение импульса на поверхности моря.

1 — начальное (первичное) возмущение, распространяющееся вниз; $R(Z)$ — отраженные волны

поверхности моря. Наблюдения проводят с помощью чувствительных к давлению приемников (гидрофонов), удерживаемых на заданной глубине ниже поверхности моря. Давление на поверхности моря обычно стремится к нулю, но при возбуждении импульсного возмущения давление может на мгновение принять некоторое другое значение, например 1 . Это показано на рис. 8.3. Начальное возмущение 1 , распространяющееся вниз, порождает отраженную в среде волну $-R(Z)$, которая снова отражается от поверхности моря, образуя волну $R(Z)$. Давление на поверхности равно $D+U=1+R-R=1$. Полная вертикальная составляющая скорости смещения частиц воды на поверхности моря, обусловленная источником и результирующей акустической волной, равна $D-U=1++2R(Z)$.

Упражнения

1. Выразите амплитуду проходящей через границу волны с помощью величин Y_1 и Y_2 .
2. В некоторых приложениях непрерывность волнового процесса на границе выражается в том, что разность $D-U$ считается одинаковой с обеих сторон границы раздела. Это предполагает, что $t=1-c$. Получите для этого случая уравнение для коэффициента отражения, подобное уравнению (8.2.4) и выраженное через полную проводимость Y .
3. Как выражаются коэффициенты отражения и пропускания через импеданс I ? (исключите дроби в полученных конечных формулах).
4. На основании принципа сохранения энергии ранее мы показали, что при прохождении волны через границу раздела двух сред в противоположных направлениях должно выполняться условие $c'=-c$. Это условие можно получить

также путем обращения времени. Чтобы сделать это, начертите рис. 8.1 с противоположными направлениями стрелок. Выбрав масштаб и линейно накладывая различные рисунки, попытайтесь найти положение, при котором изображение, подобное представленному в правой части рис. 8.1, характеризуется величиной $-c'$ для отраженной волны. [Указание. Нарисуйте стрелки для случая нормального падения волны на границу.]

8.2. ПОТОК ЭНЕРГИИ В СЛОИСТЫХ СРЕДАХ

Рассмотрим сначала явление резонанса в некотором слое при прохождении через него волны. Зададим время прохождения волны через этот слой в прямом и обратном направлениях с помощью оператора задержки Z . Рассматриваемая ситуация показана на

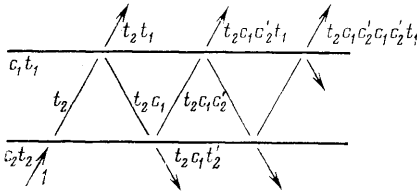


РИС. 8.4. Распространение лучей волн в слое при резонансе

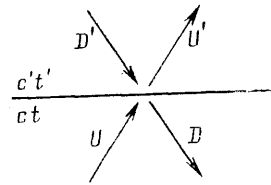


РИС. 8.5. Падающие и отраженные волны на границе раздела

рис. 8.4. Волну, наблюдаемую над этим слоем, можно представить в виде

$$t_2 t_1 [1 + c_1 c_2' Z + (c_1 c_2')^2 Z^2 + \dots] = \frac{t_2 t_1}{1 - c_1 c_2' Z}.$$

Как видно, бесконечный ряд в этой формуле имеет конечную сумму, что не случайно. Ниже будет показано, что в случае n слоев волны бесконечной длительности могут быть выражены простыми полиномами степени n . Рассмотрим общую задачу определения волнового поля по заданным коэффициентам отражения границ в среде и задачу определения коэффициентов отражения границ по заданному волновому полю в случае многослойной среды.

Коэффициенты отражения и пропускания показывают как вычислять результирующие волны, возникающие при падении волны на слой. Ниже приводятся уравнения (8.2.1), соответствующие рис. 8.5 и показывающие как экстраполировать волны U и D' в будущее время, чтобы получить волны U' и D :

$$\begin{aligned} U' &= tU + c'D'; \\ D &= cU + t'D'. \end{aligned} \quad (8.2.1)$$

Преобразуем уравнения (8.2.1) таким образом, чтобы величины U' и D' оказались в правой части, а величины U и D — в левой. Тогда получим уравнения, которые трансформируют среду, обозна-

ченную штрихом, в среду без штриха. Мы имеем

$$\begin{aligned} -tU &= -U' + c'D'; \\ -cU + D &= t'D', \end{aligned}$$

которые можно записать в матричной форме

$$\begin{bmatrix} -t & 0 \\ -c & 1 \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix} = \begin{bmatrix} -1 & c' \\ 0 & t' \end{bmatrix} \begin{bmatrix} U' \\ D' \end{bmatrix}.$$

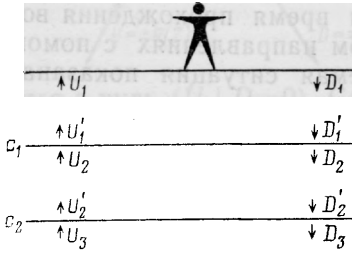


РИС. 8.6. Слоистая среда типа Гоупиллауда (слои характеризуются одинаковым временем пробега волны)

Теперь, умножив последнее выражение слева на матрицу, обратную матрице в левой части

$$\begin{bmatrix} U \\ D \end{bmatrix} = \frac{1}{-t} \begin{bmatrix} 1 & 0 \\ c & -t \end{bmatrix} \begin{bmatrix} -1 & c' \\ 0 & t' \end{bmatrix} \begin{bmatrix} U' \\ D' \end{bmatrix} = \frac{1}{-t} \begin{bmatrix} -1 & c' \\ -c & cc' - tt' \end{bmatrix}'$$

получим окончательный результат в виде общего уравнения для экстраполяции среды со штрихом в среду без штриха:

$$\begin{bmatrix} U \\ D \end{bmatrix} = \frac{1}{t} \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \begin{bmatrix} U' \\ D' \end{bmatrix}, \quad (8.2.2)$$

Рассмотрим теперь показанную на рис. 8.6 слоистую среду типа Гоупиллауда [30], которая характеризуется равенством времен пробега во всех слоях. Для рассматриваемого случая уравнение (8.2.2) может быть записано в виде

$$\begin{bmatrix} U \\ D \end{bmatrix}_{k+1} = \frac{1}{t_k} \begin{bmatrix} 1 & c_k \\ c_k & 1 \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}'_k.$$

Пусть $Z = e^{+i\omega T}$ (T — двойное время распространения волны в слое) равно интервалу выборки данных. Очевидно, что умножение на \sqrt{Z} эквивалентно запаздыванию функции на $T/2$, т. е. времени прохождения волны через слой. Это дает соотношение между волнами со штрихами и без штриха в k -м слое:

$$\begin{bmatrix} U \\ D \end{bmatrix}'_k = \begin{bmatrix} 1/\sqrt{Z} & 0 \\ 0 & \sqrt{Z} \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k. \quad (8.2.3)$$

Подставив формулу (8.2.3) в выражение (8.2.2), получим выражение для матрицы слоя:

$$\begin{aligned} \begin{bmatrix} U \\ D \end{bmatrix}_{k+1} &= \frac{1}{t_k} \begin{bmatrix} 1 & c_k \\ c_k & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{Z} & 0 \\ 0 & \sqrt{Z} \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k = \\ &= \frac{1}{t_k} \begin{bmatrix} 1/\sqrt{Z} & c_k \sqrt{Z} \\ c_k \sqrt{Z} & \sqrt{Z} \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k = \frac{1}{\sqrt{Z} t_k} \begin{bmatrix} 1 & c_k Z \\ c_k & Z \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k. \end{aligned} \quad (8.2.4)$$

Если энергия волнового процесса распространяется через пачку слоев, то общий поток энергии, проходящей через k -й слой, должен быть таким же, как и поток энергии, проходящей через

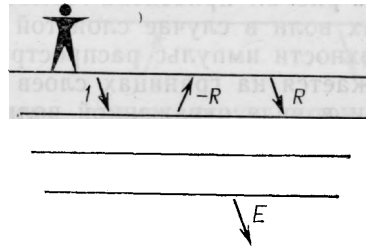


РИС. 8.7. Геометрия отраженных сейсмических волн.

Волны: I — падающая, R — отраженная, E — уходящая в глубь Земли

$(k+1)$ -й слой. В противном случае должен существовать поглотитель или источник энергии на границе слоя. Результирующий восходящий поток энергии $J(\omega)$ на любой частоте ω в k -м слое определяется выражением

$$J(\omega) = Y_k \left[U(Z) \bar{U}\left(\frac{1}{Z}\right) - D(Z) \bar{D}\left(\frac{1}{Z}\right) \right]_k. \quad (8.2.5)$$

Чтобы установить, что поток энергии действительно не зависит от k , рассмотрим уравнение Эрмита, сопряженное с уравнением (8.2.4) (транспонированное и сопряженное по отношению к действительному ω):

$$\left[U\left(\frac{1}{Z}\right) D\left(\frac{1}{Z}\right) \right]_{k+1} = \frac{1}{t_k} \left[U\left(\frac{1}{Z}\right) D\left(\frac{1}{Z}\right) \right]_k \begin{bmatrix} \sqrt{Z} & c_k \sqrt{Z} \\ c_k/\sqrt{Z} & 1/\sqrt{Z} \end{bmatrix}. \quad (8.2.6)$$

Объединив уравнения (8.2.4) и (8.2.6), получим

$$\begin{aligned} \left[U\left(\frac{1}{Z}\right) D\left(\frac{1}{Z}\right) \right]_{k+1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} U(Z) \\ D(Z) \end{bmatrix}_{k+1} &= \\ = \frac{1}{t_k^2} [\bar{U} \bar{D}]_k \begin{bmatrix} \sqrt{Z} & \sqrt{Z} c_k \\ c_k/\sqrt{Z} & 1/\sqrt{Z} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{Z} & \sqrt{Z} c_k \\ c_k/\sqrt{Z} & \sqrt{Z} \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k &= \\ = \frac{1}{t_k^2} [\bar{U} \bar{D}]_k \begin{bmatrix} 1 - c_k^2 & 0 \\ 0 & c_k^2 - 1 \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k. \end{aligned} \quad (8.2.7)$$

Поскольку $(1-c_k^2)/t_k^2 = t'_k/t_k = Y_k/Y_{k+1}$, то можно переписать результат в виде

$$Y_{k+1} \left[U\left(\frac{1}{Z}\right)U(Z) - D\left(\frac{1}{Z}\right)D(Z) \right]_{k+1} = Y_k \left[U\left(\frac{1}{Z}\right)U(Z) - D\left(\frac{1}{Z}\right)D(Z) \right]_k. \quad (8.2.8)$$

Уравнение (8.2.8) показывает, что на любой частоте ω энергия, проходящая через k -й слой, равна энергии, проходящей через $(k+1)$ -й слой.

Эта теорема о потоке энергии приводит к некоторым общим положениям о свойствах волн, рассеянных слоистыми структурами.

На рис. 8.7 приведена геометрическая схема отражения сейсмических волн в случае слоистой среды. Возбужденный на дневной поверхности импульс распространяется вниз внутрь среды, где он отражается на границах слоев и выходит обратно к поверхности Земли в виде отраженной волны $-R(Z)$. Поскольку эта поверхность является идеально отражающей, то она посылает волну вида $R(Z)$ обратно в землю. Волна $E(Z)$, прошедшая через основание слоистой толщи, распространяется далее, внутрь Земли. Применяя теорему о потоке энергии к этой схеме, можно заключить, что поток энергии в верхнем слое равен потоку энергии в нижнем пространстве, т. е.

$$Y_1 \left\{ R\left(\frac{1}{Z}\right)R(Z) - \left[1 + R\left(\frac{1}{Z}\right) \right] \left[1 + R(Z) \right] \right\} = -Y_k E\left(\frac{1}{Z}\right)E(Z)$$

или

$$1 + R\left(\frac{1}{Z}\right) + R(Z) = \frac{Y_k}{Y_1} E\left(\frac{1}{Z}\right)E(Z). \quad (8.2.9)$$

Этот весьма примечательный результат говорит о том, что, зная наблюдаемую выходящую волну $E(Z)$, можно определить с помощью функции автокорреляции форму колебаний, регистрируемых на поверхности. Позднее увидим, что волна $E(Z)$ минимально фазовая и, следовательно, она может быть получена из функции R с помощью факторизации спектра.

Рассмотрим теперь геометрическую схему землетрясения (рис. 8.8). Волновой импульс I падает на поверхность и отражается обратно вниз. Волна $X(Z)$ выходит на свободную поверхность и отражается обратно вниз. Волна $P(Z)$ возвращается обратно в землю и там рассеивается. Применяя теорему о потоке энергии к этой схеме, получим

$$Y_1 \left[X\left(\frac{1}{Z}\right)X(Z) - X\left(\frac{1}{Z}\right)X(Z) \right] = Y_k \left[1 - P(Z)P\left(\frac{1}{Z}\right) \right]$$

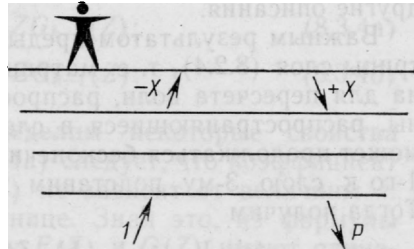
или

$$1 = P(Z)P\left(\frac{1}{Z}\right). \quad (8.2.10)$$

Полученный результат говорит о том, что рассеянная в обратном направлении из слоя волна $P(Z)$ как бы прошла через фазовый фильтр. Это можно было бы предвидеть заранее на основе общих физических представлений, поскольку в рамках принятой модели процесса вся энергия падающей на поверхность волны в конечном счете отражается вниз без затухания. Единственное, что может наблюдаться — это частотно-зависимая задержка колебаний.

РИС. 8.8. Геометрическая схема землетрясения.

Волны: I — падающая на поверхность; X — отраженная, P — рассеянная внутри Земли



В заключение рассмотрим теорему, связывающую поток энергии с функциями импеданса и полной проводимости (эти функции имеют преобразования Фурье с положительной действительной частью). Предположим, что нисходящая волна $D(Z)$ интенсивнее восходящей волны $U(Z)$ на всех частотах, т. е.

$$D\left(\frac{1}{Z}\right)D(Z) - U\left(\frac{1}{Z}\right)U(Z) > 0 \quad |Z|=1. \quad (8.2.11)$$

(Отметим, что это не означает, что $|d_t| > |U_t|$). Выражение (8.2.11) можно записать сокращенно в виде

$$\bar{D}D - \bar{U}U > 0. \quad (8.2.12)$$

На основании формулы (8.2.11) или (8.2.12) можно показать, что величина $(D-U)/(\bar{D}+U)$ имеет преобразование Фурье с положительной действительной частью. Вычислим действительную часть этой функции

$$\begin{aligned} 2 \operatorname{Re} \frac{D-U}{\bar{D}+U} &= \frac{D-U}{\bar{D}+U} + \frac{\bar{D}-\bar{U}}{D+U} = \\ &= \frac{(D-U)(\bar{D}+\bar{U}) + (\bar{D}-\bar{U})(D+U)}{(\bar{D}+U)(D+U)} = \frac{2(\bar{D}D - \bar{U}U)}{(D+U)(\bar{D}+U)}. \end{aligned} \quad (8.2.13)$$

Числитель выражения (8.2.13) согласно гипотезе (8.2.12) является положительным. Знаменатель выражения (8.2.13) также положителен, поскольку спектр функции времени $d_t + u_t$, как и любой спектр, всегда положителен. Таким образом, функция $(D-U)/(\bar{D}+U)$ является действительной положительной величиной. Интерпретация величины $(D-U)/(\bar{D}+U)$ состоит в том, что числитель ее $(D-U)$ характеризует вертикальную составляющую скорости смещения частиц среды, а знаменатель $(D+U)$ — давление при распространении волны в среде.

8.3. РАСЧЕТ ВОЛН ПО КОЭФФИЦИЕНТАМ ОТРАЖЕНИЯ ГРАНИЦ

Слоистую среду можно описать путем задания коэффициентов отражения для всех границ раздела. Альтернативные описания предусматривают задание какой-либо одной из рассеянных в слоистой среде волн $R(Z)$, $E(Z)$, $X(Z)$ или $P(Z)$. Конечной целью является настолько хорошее овладение алгеброй таких задач, чтобы, начиная с какого-либо описания слоистой среды, получить из него все другие описания.

Важным результатом предыдущего раздела было получение матрицы слоя (8.2.4), т. е. матрицы, которая может быть использована для пересчета волн, распространяющихся в одном слое, в волны, распространяющиеся в следующем слое и т. д. Этот процесс может продолжаться бесконечно. Для понимания пересчета от слоя 1-го к слою 3-му подставим в выражение (8.2.4) $k=1$ и $k=2$. Тогда получим

$$\begin{aligned} \begin{bmatrix} U \\ D \end{bmatrix}_3 &= \frac{1}{\sqrt{Z}t_2} \begin{bmatrix} 1 & Zc_2 \\ c_2 & Z \end{bmatrix} \frac{1}{\sqrt{Z}t_2} \begin{bmatrix} 1 & Zc_1 \\ c_1 & Z \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_1 = \\ &= \frac{1}{\sqrt{Z^2 t_1 t_2}} \begin{bmatrix} 1+Zc_1c_2 & Zc_1+Z^2c_2 \\ c_2+Zc_1 & Zc_1c_2+Z^2 \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_1. \end{aligned} \quad (8.3.1)$$

Исследование этого примера наводит на мысль о существовании общего выражения для произведения k матриц слоев

$$\frac{1}{\sqrt{Z^k} \prod_{j=1}^k t_j} \begin{bmatrix} F(Z) & Z^k G\left(\frac{1}{Z}\right) \\ G(Z) & Z^k F\left(\frac{1}{Z}\right) \end{bmatrix}. \quad (8.3.2)$$

Теперь проверим, что формула (8.3.2) действительно является общим выражением. Предположим, что формула (8.3.2) верна для $k-1$. Тогда, умножив ее при $k-1$ на матрицу следующего k -го слоя, мы должны получить такое же выражение, как и при прямой замене величины $k-1$ на k .

Рассмотрим произведение

$$\begin{aligned} &\frac{1}{\sqrt{Z}t_k} \begin{bmatrix} 1 & c_k Z \\ c_k & Z \end{bmatrix} \frac{1}{\sqrt{Z^{k-1}} \prod_{j=1}^{k-1} t_j} \begin{bmatrix} F(Z) & Z^{k-1} G\left(\frac{1}{Z}\right) \\ G(Z) & Z^{k-1} F\left(\frac{1}{Z}\right) \end{bmatrix} = \\ &= \frac{1}{\sqrt{Z^k} \prod_{j=1}^k t_j} \begin{bmatrix} F(Z) + c_k Z G(Z) & Z^{k-1} G\left(\frac{1}{Z}\right) + c_k Z^k F\left(\frac{1}{Z}\right) \\ c_k F(Z) + Z G(Z) & c_k Z^{k-1} G\left(\frac{1}{Z}\right) + Z^k F\left(\frac{1}{Z}\right) \end{bmatrix}. \end{aligned} \quad (8.3.3)$$

Из формулы (8.3.3) следует, что масштабный множитель в ней имеет тот же вид, что и при прямой замене величины $k-1$ на k

в формуле (8.3.2) для $k-1$. Кроме того, 22-й элемент матрицы может быть получен из 11-го элемента заменой Z на $1/Z$ и умножением на Z^k . Аналогично 21-й элемент матрицы получается из 12-го элемента. Таким образом, формула (8.3.2) действительно является общим выражением для произведения матриц слоев. Полиномы $F(Z)$ и $G(Z)$ порядка k образуются следующим образом [из первой колонки правой части выражения (8.3.3)]:

$$F_k(Z) = F_{k-1}(Z) + c_k Z G_{k-1}(Z); \quad (8.3.4a)$$

$$G_k(Z) = c_k F_{k-1}(Z) + Z G_{k-1}(Z). \quad (8.3.4b)$$

Исследуя формулы (8.3.4), определим некоторые свойства функций F и G . Так, из формулы (8.3.4a) следует, что коэффициент f_0 при первом члене многочлена $F(Z)$ не зависит от величины k . Коэффициент $(f_0)_k$ всегда равен единице. Зная это, из формулы (8.3.4b) получим $(g_0)_k = c_k$. Поскольку $F(Z)$ и $G(Z)$ имеют одинаковые степени по Z , то формула (8.3.4b) указывает на независимость коэффициента при наивысшей степени многочлена $G(Z)$, например $(g_k)_k$ от величины k , и, следовательно, он равен начальному значению c_1 . Наконец, зная это, на основании формулы (8.3.4a) можно сделать вывод о том, что коэффициент при наивысшей степени многочлена $F(Z)$ всегда будет равен произведению $c_1 c_k$. В результате получим

$$F(Z) = 1 + f_1 Z + f_2 Z^2 + \dots + c_1 c_2 Z^{k-1}; \quad (8.3.5a)$$

$$G(Z) = c_k + g_1 Z + g_2 Z^2 + \dots + c_1 Z^{k-1}. \quad (8.3.5b)$$

Используя рекуррентные соотношения (8.3.4), можно установить, что в формулах (8.3.5) коэффициенты полинома F содержат лишь четные степени коэффициентов c , а коэффициенты полинома G — только нечетные степени. Следовательно, если все коэффициенты c изменяют знак, то и полином G изменяет свой знак, а полином F сохраняет знак.

Полиномы $F(Z)$ и $G(Z)$ не являются независимыми, так как между ними существует неожиданная зависимость. С помощью формул (8.3.4) можно проверить, что

$$\left[F(Z) F\left(\frac{1}{Z}\right) - G(Z) G\left(\frac{1}{Z}\right) \right]_k = (1 - c_k^2) \left[F(Z) F\left(\frac{1}{Z}\right) - G(Z) G\left(\frac{1}{Z}\right) \right]_{k-1}. \quad (8.3.6)$$

Так как $F_1(Z) = 1$ и $G_1(Z) = c_1$, то, применяя итеративную процедуру (8.3.6), найдем:

$$\left[F(Z) F\left(\frac{1}{Z}\right) - G(Z) G\left(\frac{1}{Z}\right) \right]_k = \prod_1^k (1 - c_k^2) = \prod_1^k t't. \quad (8.3.7)$$

Уравнение (8.3.7) интересно тем, что в левой части содержит два спектра (спектры функций f_t и g_t), а его правая часть является положительной, не зависящей от частоты постоянной. Таким образом, спектр функции f_t больше спектра функции g_t . Теперь можно применить теорему о введении дополнительной информации для преобразования сейсмического импульса в минимально фазовый сигнал и получения из формулы (8.3.4а) и условия $|c_k| < 1$ вывода

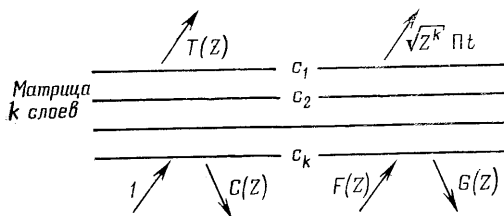


РИС. 8.9. Падающие I и $F(Z)$, отраженные $C(Z)$ и $G(Z)$ и проходящие $T(Z)$ и $\sqrt{Z^k} \Pi t$ волны в пачке слоев, залегающей между двумя полупространствами

о том, что волна $F_k(Z)$ минимально фазовая, если таковой является волна $F_{k-1}(Z)$. Следовательно, если волна $F_1(Z)$ минимально фазовая, то все волны $F_k(Z)$ также минимально фазовые. Поскольку волна $F(Z)$ является минимально фазовой, то ее можно определить по своему спектру $F(Z)F(1/Z)$ или спектру g_t (вместе с простым числом Πt). Однако нельзя получить волну G из F . Прежде чем продолжить обсуждение, рассмотрим один пример.

Предположим, что пачка слоев помещена между двумя полупространствами (рис. 8.9). Импульс падает снизу. Рассеянную в обратном направлении волну обозначим через $C(Z)$, а проходящую — через $T(Z)$.

Математически эта ситуация описывается уравнением

$$\begin{bmatrix} 1 \\ C(Z) \end{bmatrix} = \frac{1}{\sqrt{Z^k \Pi t}} \begin{bmatrix} F(Z) & Z^k G\left(\frac{1}{Z}\right) \\ G(Z) & Z^k F\left(\frac{1}{Z}\right) \end{bmatrix} \begin{bmatrix} T(Z) \\ 0 \end{bmatrix}. \quad (8.3.8)$$

Решим уравнение (8.3.8) проходящей волны

$$T(Z) = \frac{\sqrt{Z^k \Pi t}}{F(Z)}. \quad (8.3.9)$$

Подставив этот результат во второе из уравнений (8.3.8), получим рассеянную обратно волну:

$$C(Z) = \frac{G(Z) T(Z)}{\sqrt{Z^k \Pi t}} = \frac{G(Z)}{F(Z)}. \quad (8.3.10)$$

Математически волна $F(Z)$ является минимально фазовой, а физически волны $C(Z)$ и $T(Z)$ имеют конечную энергию. Поэтому зна-

менатели выражений (8.3.9) и (8.3.10) не имеют нулей внутри единичной окружности. Так как известно, что рассеянная в обратном направлении (назад) волна $C(Z)$ несет меньше энергии, чем падающая волна, а также учитывая выражение (8.2.13), зададим положительную действительную функцию выражением:

$$\frac{1 - C(Z)}{1 + C(Z)} = \frac{1 - G(Z)/F(Z)}{1 + G(Z)/F(Z)}. \quad (8.3.11)$$

Теперь рассмотрим как восстановить коэффициенты отражения c по наблюдаемой рассеянной волне $C(Z)$. В соответствии с рис. 8.9 запишем

$$\begin{bmatrix} 1 \\ C(Z) \end{bmatrix} = \frac{1}{t_k \sqrt{Z}} \begin{bmatrix} 1 & c_k Z \\ c_k & Z \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_{k-1}. \quad (8.3.12)$$

Первый коэффициент полинома $C(Z)$ есть c_k [физически это очевидно и подтверждается формулой (8.3.5)]. Следовательно, матрица слоя в выражении (8.3.12) известна. Умножив обе части выражения (8.3.12) на матрицу, обратную матрице слоя, найдем величины $U_{k-1}(Z)$ и $D_{k-1}(Z)$. Следующий коэффициент отражения c_{k-1} равен отношению d_0/u_0 . Этот процесс можно продолжить до тех пор, пока не будут определены все коэффициенты отражения c_k .

Теперь вернемся к геометрии отраженных волн. Имеем

$$\begin{bmatrix} 0 \\ E(Z) \end{bmatrix} = \frac{1}{\sqrt{Z^k} \Pi t} \begin{bmatrix} F(Z) & Z^k G\left(\frac{1}{Z}\right) \\ G(Z) & Z^k F\left(\frac{1}{Z}\right) \end{bmatrix} \begin{bmatrix} -R \\ 1+R \end{bmatrix}. \quad (8.3.13)$$

Решим первое из уравнений (8.3.13) относительно волны $R(Z)$:

$$R(Z) = \frac{Z^k G(1/Z)}{F(Z) - Z^k G(1/Z)}. \quad (8.3.14)$$

Знаменатель выражения (8.3.14) встречается настолько часто, что следует ввести для него обозначение

$$A(Z) = F(Z) - Z^k G\left(\frac{1}{Z}\right), \quad (8.3.15)$$

Функция $A(Z)$, подобно функции $F(Z)$, является минимально фазовой. Второе уравнение в (8.3.13) позволяет определить нисходящую волну

$$\begin{aligned} E(Z) &= \frac{Z^k F(1/Z) + [-G(Z) + Z^k F(1/Z)] R}{\sqrt{Z^k} \Pi t} = \\ &= \frac{Z^k F(1/Z) [F(Z) - Z^k G(1/Z)] + [-G(Z) + Z^k F(1/Z)] Z^k G(1/Z)}{A(Z) \sqrt{Z^k} \Pi t}. \end{aligned}$$

Упрощая это выражение с помощью формулы (8.3.7), получим:

$$E(Z) = \frac{\sqrt{Z^k} \Pi t'}{A(Z)}. \quad (8.3.16)$$

Рассмотрим теперь действительную положительную функцию

$$\frac{D-U}{D+U} = \frac{1+R-(-R)}{1+R-R} = 1 + 2R(Z) = \frac{W=1+2R(Z)}{P=1}, \quad (8.3.17)$$

где W — вертикальная скорость; P — давление.

Как уже указывалось выше (см. раздел 8.2), если уравнение (8.3.17) рассматривается с позиций акустики, то величина $Y = (D-U)/(D+U)$ интерпретируется как отношение вертикальной составляющей скорости к давлению и называется *полной проводимостью*, которая является обратной величиной полного сопротивления (импеданса).

На этом мы закончим анализ решения задачи вычисления волн по заданным коэффициентам отражения. В следующем разделе мы рассмотрим обратные задачи, т. е. задачи вычисления коэффициентов отражения по известным волнам различных типов.

Упражнения

1. Рассмотрите схему распространения волн, показанную на рис. 8.9, при $c_1=1/2$; $c_2=-1/2$ и $c_3=1/3$. Чему равны при этом коэффициенты полиномов $T(Z)$ и $C(Z)$?

2. Для простой границы раздела имеются известные простые соотношения $t=1+c$, $t'=1+c'$ и $c=-c'$. Какие аналогичные зависимости Вы сможете найти для обобщенной границы раздела, показанной на рис. 8.9? [Например, покажите, что выражение $1-T(Z)T'(1/Z)=C(Z)C(1/Z)$ является аналогом зависимости $1-tt'=c^2$.]

3. Покажите, что волны $T(Z)$ и $T'(Z)$ имеют одинаковую форму колебаний волны с точностью до масштабного множителя. Докажите, что пакки слоев с различной структурой могут иметь одинаковые величины $T(Z)$.

4. Пусть некоторый импульс падает на пачку слоев и при этом образуется отраженная волна $C(Z)$. Чему равен коэффициент отражения волны при встрече

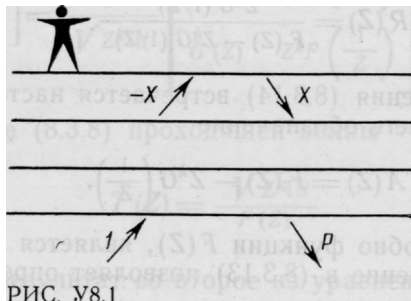


РИС. У8.1.

ее с первым слоем? Какой вид будет иметь отраженная волна как функция C в ситуации, отличающейся от рассмотренной выше тем, что удалена первая отражающая граница?

5. Рассмотрите модель Земли, представленную пачкой слоев, залегающих на полупространстве. Пусть импульс падает на нее снизу (рис. У8.1). По заданным полиномам $F(Z)$ и $G(Z)$ и элементам произведения матриц слоев, найдите вол-

ны X и P . Проверьте свой ответ, доказав, что $P(Z)\bar{P}(1/Z)=1$. Какова связь между волнами X и E ? Эта зависимость иллюстрирует принцип взаимности, согласно которому источник и приемник можно поменять местами.

6. Покажите, что $1+R(1/Z)+R(Z)$ (масштабный множитель) $X(Z)X(1/Z)$. Это доказывает, что для получения сейсмограммы отраженной волны можно использовать функцию автокорреляции сейсмограммы проходящей волны.

7. Рассмотрите рис. У8.2. Рассчитайте волну R' через волну R .

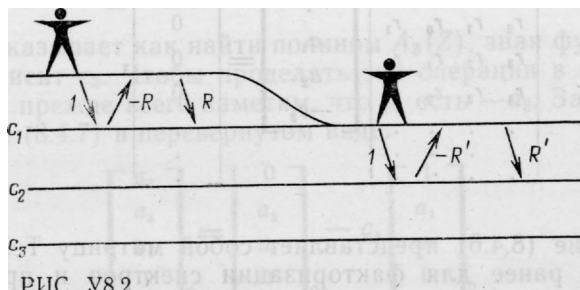


РИС. У8.2.

8.4. ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ ПО ВОЛНОВЫМ ПОЛЯМ

Для решения обратных задач наилучшим является, по-видимому, уравнение Кюнетца [31]:

$$1 + R\left(\frac{1}{Z}\right) + R(Z) = \frac{Y_k}{Y_1} E\left(\frac{1}{Z}\right) E(Z). \quad (8.4.1)$$

Нам понадобится еще выражение для нисходящей волны

$$E(Z) = \frac{\sqrt{Z^k} \Pi t'}{A(Z)}. \quad (8.4.2)$$

Необходимо также вспомнить, что $Y_k/Y_1 = \Pi t'/t'$. С учетом этих данных выражение (8.4.1) приобретает вид

$$1 + R(Z) + R\left(\frac{1}{Z}\right) = \frac{\Pi t' t}{A(Z) A(1/Z)}. \quad (8.4.3)$$

Умножив обе части последнего выражения на $A(Z)$, получим:

$$\left[1 + R(Z) + R\left(\frac{1}{Z}\right)\right] A(Z) = \frac{\Pi t' t}{A(1/Z)}. \quad (8.4.4)$$

Так как функция $A(Z)$ является минимально фазовой, то $A(Z)$ можно записать как $1/B(Z)$ или $A(1/Z) = 1/B(1/Z)$.

Тогда выражение (8.4.4) приобретает вид

$$[1 + R(Z) + R(1/Z)] A(Z) = \Pi t' t \left(1 + \frac{b_1}{Z} + \frac{b_2}{Z^2} + \dots\right). \quad (8.4.5)$$

Отождествляя коэффициенты при нулевой и положительных степенях Z из совместных уравнений, получаем систему уравнений, которая для трехслойной модели среды принимает вид (при $r_0=1$):

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \\ r_4 & r_3 & r_2 & r_1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ -c_3 \end{bmatrix} = \begin{bmatrix} \text{Пит} \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \end{bmatrix}. \quad (8.4.6)$$

Выражение (8.4.6) представляет собой матрицу Тейлица, применявшуюся ранее для факторизации спектров и предсказания временных рядов. Отметим, что множитель $-c_3$ введен в формулу (8.4.6) в качестве коэффициента при наивысшей степени полинома $A(Z)$. Это оправдано тем, что ранее функция $A(Z)$ была определена с помощью полиномов $F(Z)$ и $G(Z)$, которые, в свою очередь, выражаются через коэффициенты c_k . Рассмотрев вновь систему уравнений Тейлица (8.4.6) и метод Левинсона для решения уравнений вида (3.3.10), можно вычислить коэффициенты отражения границ по волновому полю.

Первые четыре уравнения в системе (8.4.6) являются нормальными. Задавая первые три отраженных импульса r_1 , r_2 и r_3 , можно решить уравнения относительно полинома $A(Z)$ и получить коэффициент отражения c_3 . Пятое из уравнений (8.4.6) при известном полиноме $A(Z)$ можно использовать для определения импульса r_4 . Если рассматриваемая модель действительно трехслойная, то решение будет правильным. В противном случае расхождение будет указывать на наличие нового слоя, коэффициент отражения которого c_4 определяется путем расширения уравнения (8.4.6) от 4-го порядка до 5-го. В результате, задав r_k отраженных импульсов с помощью рекурсии Левинсона, можно последовательно определять коэффициенты отражения границ c_k .

Теперь предположим, что мы наблюдаем нисходящую волну $E(Z)$. Одним из способов определения коэффициентов отражения является формирование величины $1+R(Z)+R(1/Z)$ с помощью автокорреляционной функции для волны $E(Z)$; затем, применяя рекурсию Левинсона, находим коэффициенты отражения. Единственным недостатком этого способа является то, что волна $E(Z)$ включает неограниченное число коэффициентов, поэтому на практике необходимо производить некоторое усечение коэффициентов. Чтобы избежать этого, можно воспользоваться альтернативным методом. Задав волну $E(Z)$ в виде отношения полиномов, найдем функцию $A(Z)$. Сущность рекурсии Левинсона заключается в построении полинома $A(Z)$ с помощью соотношения $A_{k+1}(Z) = A_k(Z) - cZ^k A_k(1/Z)$. В частности, из формулы (3.3.12) получим

выражение

$$\begin{bmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}_3 = \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix}_2 - c_3 \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix}_2, \quad (8.4.7)$$

которое показывает как найти полином $A_3(Z)$, зная функцию $A_2(Z)$ и коэффициент c_3 . Чтобы проделать эти операции в обратном направлении, прежде всего заметим, что c_3 есть $-a_3$. Затем запишем выражение (8.4.7) в перевернутом виде:

$$\begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ 1 \end{bmatrix}_3 = \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix}_2 - c_3 \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix}_2. \quad (8.4.8)$$

Теперь умножим выражение (8.4.7) на $1/(1-c_3^2)$ и сложим результат с уравнением (8.4.8), умноженным на $c_3/(1-c_3^2)$. Перевернутые векторы в правой части полученного равенства сокращаются, что приводит к следующему результату:

$$\frac{1}{1-c_3^2} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}_3 + \frac{c_3}{1-c_3^2} \begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ 1 \end{bmatrix}_3 = \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix}_2. \quad (8.4.9)$$

Уравнение (8.4.9) представляет собой искомый результат, позволяющий привести полином $A_{k+1}(Z)$ к полиному $A_k(Z)$ и одновременно получить коэффициент отражения c_{k+1} . Программа для реализации этого процесса на ЭВМ приведена на рис. 8.10, а для ре-

РИС. 8.10. Программа для вычисления коэффициентов отражения c_k с помощью фильтра ошибки предсказания $A(Z)$ (использование комплексных величин необязательно)

```

COMPLEX A, C, AL, BE, TOP, CONJG
C(1)=-1.; R(1)=1.; A(1)=1.; V(1)=1.
300 DO 310 I=1, N
310 C(I)=A(I)
DO 330 K=1, N
J=N-K+2
AL=1./ (1.-C(J)*CONJG(C(J)))
BE=C(J)*AL
JH=(J+1)/2
DO 320 I=1, JH
TOP=AL*C(I)-BE*CONJG(C(J-I+1))
C(J-I+1)=AL*C(J-I+1)-BE*CONJG(C(I))
320 C(I)=TOP
330 C(J)=-BE/AL

```

шения обратной задачи — на рис. 8.11. Программа вычисляет функции R и A по заданным коэффициентам отражения c .

В заключение рассмотрим решение задачи при наличии случайных источников. На рис. 8.12 показана геометрическая схема землетрясения, где широкополосный сигнал падает на границы.

В верхнем слое сумма волн стремится к нулю, обеспечивая нулевое давление на свободной поверхности. Разность восходящих и нисходящих волн дает наблюдаемую на поверхности вертикальную составляющую скорости смещения частиц среды. Для введения элемента статистики падающий импульс был свернут с «белым светом» W_t в виде ряда случайных чисел. Следовательно, все

```

COMPLEX C,R,A,BOT,CONJG
C(1)=-1.; R(1)=1.; A(1)=1.; V(1)=1.
100 DO 120 J=2,N
    A(J)=0.
    R(J)=C(J)*V(J-1)
    V(J)=V(J-1)*(1.-C(J)*CONJG(C(J)))
110 R(J)=R(J)-A(I)*R(J-I+1)
    JH=(J+1)/2
    DO 120 I=1,JH
        BOT=A(J-I+1)-C(J)*CONJG(A(I))
        A(I)=A(I)-C(J)*CONJG(A(J-I+1))
120 A(J-I+1)=BOT

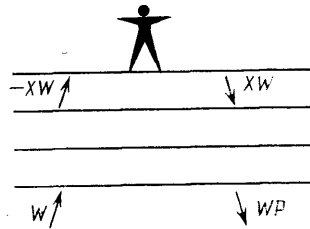
```

РИС. 8.11. Программа для вычисления величин R и A по заданным коэффициентам отражения c

внутренние волны на схеме (см. рис. 8.12) задаются сверткой шума W_t с соответствующей моделью импульса падающей волны. Теперь предположим, что заданы волны в верхнем слое $D = -U =$

РИС. 8.12. Геометрическая схема землетрясения с широкополосным сигналом W , распространяющимся снизу.

Волны: XW — отраженная от поверхности; WP — рассеянная внутрь Земли



$=XW$ и необходимо рассмотреть продолжение их вниз. Матрица слоя для рассматриваемого случая имеет вид

$$\begin{bmatrix} U \\ D \end{bmatrix}_{k+1} = \frac{1}{\sqrt{Z}t_k} \begin{bmatrix} 1 & cZ \\ c & Z \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}_k \quad (8.4.10)$$

Последнее выражение можно переписать в виде

$$\begin{bmatrix} -U \\ D \end{bmatrix}_{k+1} = \frac{1}{\sqrt{Z}t_k} \begin{bmatrix} 1 & -cZ \\ -c & Z \end{bmatrix} \begin{bmatrix} -U \\ D \end{bmatrix}_k \quad (8.4.11)$$

Схема фильтра предсказания ошибки Бурга может быть записана в следующей форме¹:

$$\begin{bmatrix} e^+ \\ e^- \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & -c \\ -c & 1 \end{bmatrix} \begin{bmatrix} e^+ \\ e^- \end{bmatrix}_k \quad (8.4.12)$$

что делает ее эквивалентной (с точностью до масштабного множителя) продолженным вниз формам волновых поверхностей. При

¹ В разд. 7.2 использованы обозначения e_- и e_+ .

этом не установлено, соответствуют ли коэффициенты отражения, вычисленные по формуле Бурга

$$\hat{c}_k = \frac{2 \sum_1^t e_k^+ e_k^-}{\sum (e_k^+ e_k^+ + e_k^- e_k^-)}, \quad (8.4.13)$$

коэффициентам отражения c_k , вычисленным согласно физической модели. Чтобы установить, как бурговский коэффициент отражения c_k связан с коэффициентом отражения границы в среде c_k , вычисленным с помощью рекурсии Левинсона, определим функции \mathbf{f}^+ и \mathbf{f}^- для $k=2$ в виде

$$[\mathbf{f}^+] = \begin{bmatrix} x_0 \\ x_1 \ x_0 \\ x_2 \ x_1 \ x_0 \\ \quad x_2 \ x_1 \\ \quad \quad x_2 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ 0 \end{bmatrix} \quad \text{и} \quad [\mathbf{f}^-] = \begin{bmatrix} x_0 \\ x_1 \ x_0 \\ x_2 \ x_1 \ x_0 \\ \quad x_2 \ x_1 \\ \quad \quad x_2 \end{bmatrix} \begin{bmatrix} 0 \\ a_1 \\ 1 \end{bmatrix}. \quad (8.4.14)$$

Затем сформируем из них скалярное произведение вида

$$\begin{aligned} (\mathbf{f}^-)^T \mathbf{f}^+ &= [0 \ a_1 \ 1] \begin{bmatrix} x_0 & x_1 & x_2 \\ & x_0 & x_1 & x_2 \\ & & x_0 & x_1 & x_2 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \ x_0 \\ x_2 \ x_1 \ x_0 \\ \quad x_2 \ x_1 \\ \quad \quad x_2 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ 0 \end{bmatrix} = \\ &= [0 \ a_1 \ 1] \begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ 0 \end{bmatrix}. \end{aligned} \quad (8.4.15)$$

Теперь используем тот факт, что ряд $(1, a_1)$ удовлетворяет системе размерностью 2×2 . С помощью рекурсии Левинсона запишем выражение (8.4.15) в виде

$$(\mathbf{f}^-)^T \mathbf{f}^+ = [0 \ a_1 \ 1] \begin{bmatrix} v \\ 0 \\ e \end{bmatrix} = e. \quad (8.4.16)$$

Аналогично можно показать, что $\mathbf{f}^+ \cdot \mathbf{f}^+ = \mathbf{f}^- \cdot \mathbf{f}^- = v$. Таким образом, коэффициент отражения по методу Левинсона можно вычислить по формуле

$$c = \frac{2(\mathbf{f}^+ \cdot \mathbf{f}^-)}{(\mathbf{f}^+ \cdot \mathbf{f}^+) + (\mathbf{f}^- \cdot \mathbf{f}^-)} = \frac{2e}{2v}. \quad (8.4.17)$$

Бург в отличие от Левинсона опускает в формулах (8.4.14) члены, связанные с краевым эффектом. Поэтому вместо равенств (8.4.14) он получает

$$[e^+] = \begin{bmatrix} x_1 \ x_0 \\ x_2 \ x_1 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \end{bmatrix}; \quad [e^-] = \begin{bmatrix} x_1 \ x_0 \\ x_2 \ x_1 \end{bmatrix} \begin{bmatrix} a_1 \\ 1 \end{bmatrix}. \quad (8.4.18)$$

Для достаточно длинной последовательности данных результаты, полученные обоими методами, очевидно, неразличимы. В то же время для выборок данных конечной длительности необходимо сделать выбор между этими методами. Метод Левинсона с учетом равенства (8.4.14) эквивалентен предположению о том, что значения данных в выборке стремятся к нулю на концах интервала. В большинстве случаев это неверно, так что применение метода Бурга обычно более предпочтительно.

Упражнения

1. Начальный импульс и первая часть сейсмограммы отраженных волн, т. е. $1+2R(z)$ имеют вид $1+2(Z/4+Z^2/16+Z^3/64+\dots)$. Определите величины первых трех коэффициентов отражения. Какой вид будет иметь следующая точка на сейсмограмме отраженных волн, если предположить, что отражающих границ в среде больше нет?

2. Сейсмограмма $X(Z)=1/(1-0,1Z+0,9Z^2)$ получена на поверхности нескольких слоев, залегающих над полупространством. Изобразите временную функцию сигнала и укажите ее резонансную частоту и время затухания. Найдите коэффициенты отражения, если сейсмограмма $X(Z)$ образуется источником импульсов неизвестной амплитуды, расположенным в полупространстве ниже слоев.

3. Источник типа b_0+b_1Z , расположенный в глубине полупространства, на котором залегает пачка слоев, образует сейсмограмму $B(Z)X(Z)=1-Z+Z^2/2-Z^3/2+Z^4/4-Z^5/4+Z^6/8-Z^7/8+\dots$. Определите структуру пачки слоев и временную функцию источника.

9. МАТЕМАТИЧЕСКАЯ ФИЗИКА СЛОИСТЫХ СРЕД

Математическое описание явлений, происходящих в слоистых средах, имеет много общих черт с описаниями таких разнообразных физических явлений, как акустические и электромагнитные волны, магнитостатика, гравитационно-упругий сферический резонанс Земли, тепловой поток, диффузия газов, электрический ток в среде, обладающей сопротивлением, сейсмические волны, волны в воде и атмосфере и т. п. В дальнейшем будем рассматривать общую теорию явлений в слоистых средах и более детально остановимся на акустических волнах.

Под слоистой средой будем понимать среду, физические свойства которой (сжимаемость, проводимость, плотность и т. п.) являются функциями только одной пространственной координаты. Обычно используют декартовы координаты, однако при рассмотрении геофизических явлений в глобальном масштабе (в масштабе земного шара) можно использовать сферические координаты.

9.1. ОТ ФИЗИКИ К МАТЕМАТИКЕ

Первый этап. Он заключается в том, чтобы записать все основные дифференциальные уравнения в частных производных классической физики, которые относятся к интересующей нас проблеме. Для упрощения задачи не следует применять уравнения со вторыми производными по пространственным координатам, которые выводятся из уравнений первого порядка. Нужно использовать уравнения с первыми производными. Следует выписывать уравнения для каждой компоненты вектора или матрицы.

Известно, что в акустике градиент давления p вызывает ускорение массы плотностью ρ . Для удобства ограничим движение плоскостью x, z . Полагая, что величины u и w являются составляющими скорости смещения частиц среды по осям x и z , запишем:

$$\rho \frac{\partial u}{\partial t} = - \frac{\partial p}{\partial x}; \quad (9.1.1)$$

$$\rho \frac{\partial w}{\partial t} = - \frac{\partial p}{\partial z}. \quad (9.1.2)$$

Другое уравнение, играющее важную роль в акустике, утверждает, что дивергенция скорости, умноженная на несжимаемость¹ K ,

¹ В отечественной литературе по акустике обычно используют обратную величину $\beta = 1/K$, называемую сжимаемостью среды. (Прим. перев.)

дает скорость изменения давления:

$$\frac{\partial p}{\partial t} = -K \left(\frac{\partial u}{\partial x} + \frac{\partial \omega}{\partial z} \right) + s(x, z, t). \quad (9.1.3)$$

В уравнение (9.1.3) включен также источник давления s . Уравнения (9.1.1)–(9.1.3) будем считать априорно заданными¹. Функция s может характеризовать, например, источник химической энергии типа взрыва; поэтому она может обращаться в нуль везде, кроме точки, в которой расположен источник. Во многих случаях представляют интерес распределенные на плоскости (в пространстве) источники, например радиоактивные породы в модели теплового потока Земли. Для большей общности можно было бы в уравнение (9.1.1) или (9.1.2) ввести источники скорости смещения, но для удовлетворительной иллюстрации основных принципов вполне достаточно рассмотреть уравнение (9.1.3).

Второй этап. Переменные, характеризующие волновое возмущение, являются неизвестными величинами, а свойства вещества — известными параметрами. Подсчитаем число уравнений и число неизвестных. Как видно, мы располагаем тремя уравнениями, содержащими три неизвестные величины u , ω и p . Величины K , ρ и s считаются известными. Отметим, что рассматриваемые уравнения линейны относительно неизвестных. Теперь сделаем предположение о слоистости среды, т. е. будем считать, что величины K и ρ являются функциями глубины z и не зависят от x . При этом линейные уравнения имеют постоянные коэффициенты по отношению к переменным x и t . Следовательно, можно ожидать, что решение будет всегда иметь синусоидальный характер по x и t . Характер решения по координате z из-за произвольной зависимости коэффициентов K и ρ от z неизвестен. Это приводит к третьему этапу.

Третий этап. Производится Фурье-преобразование по времени и пространственным координатам уравнений с постоянными коэффициентами. Иными словами, в уравнениях (9.1.1)–(9.1.3) делается следующая подстановка:

$$\begin{bmatrix} U(x, z, t) \\ W(x, z, t) \\ P(x, z, t) \\ S(x, z, t) \end{bmatrix} = \begin{bmatrix} U(k, z, \omega) \\ W(k, z, \omega) \\ P(k, z, \omega) \\ S(k, z, \omega) \end{bmatrix} e^{-i\omega t + ik_x x}. \quad (9.1.4)$$

После подстановки экспоненциальные множители сокращаются, в результате чего получаем

$$-i\omega\rho(z)U = -ik_x P; \quad (9.1.5a)$$

$$-i\omega\rho(z)W = -\frac{\partial P}{\partial z}; \quad (9.1.5б)$$

$$-i\omega P = -K(z) \left(ik_x U + \frac{\partial W}{\partial z} \right) + S. \quad (9.1.5в)$$

¹ Уравнения (9.1.1)–(9.1.3) образуют полную систему линейных общих уравнений акустики для давления и скорости частиц. (Прим. перев.)

Четвертый этап. Исключим алгебраические неизвестные. Из уравнений (9.1.5) следует, что в них имеются члены, содержащие производные $\partial P/\partial z$ и $\partial W/\partial z$, и в то же время нет члена с производной $\partial U/\partial z$. Это означает, что величина U является алгебраической переменной и может быть исключена чисто алгебраическими средствами. Это можно сделать подстановкой выражения (9.1.5а) в уравнение (9.1.5в).

Пятый этап. Перенесем члены с производными $\partial/\partial z$ в левую часть равенства, а все остальные — в правую и расположим члены в форме матрицы. Тогда будем иметь

$$\frac{\partial P}{\partial z} = i\omega\rho W;$$

$$\frac{\partial W}{\partial z} = i\left(\frac{\omega}{K} - \frac{k_x^2}{\omega\rho}\right)P + \frac{S}{K}$$

и затем

$$\frac{\partial}{\partial z} \begin{bmatrix} P \\ W \end{bmatrix} = i \begin{bmatrix} 0 & \omega\rho \\ \frac{\omega}{K} - \frac{k_x^2}{\omega\rho} & 0 \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{S}{K} \end{bmatrix}. \quad (9.1.6)$$

Шестой этап. Физическая задача, с которой мы начали, может быть описана матричным дифференциальным уравнением первого порядка:

$$\frac{\partial}{\partial z} \mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{s}, \quad (9.1.7)$$

где \mathbf{x} — вектор, характеризующий интересующие нас переменные акустического поля; \mathbf{A} — матрица, зависящая от временных и пространственных частот и свойств среды; \mathbf{s} — векторная функция источников, возможно всюду равная нулю.

Прежде чем рассматривать методы решения уравнения (9.1.7), можем непосредственно заключить, что в свободной от источников области переменные поля \mathbf{x} являются более гладкими функциями, чем свойства среды. Чтобы показать это, рассмотрим два однородных слоя, имеющие общую границу. На контакте матрица \mathbf{A} , очевидно, имеет вид разрывной ступенчатой функции. Теперь посмотрим, могут ли волновые поля иметь ступенчатые разрывы по координате x . Очевидно, не могут, так как ступенчатая разрывность по x имела бы следствием $\partial x/\partial z = \infty$, тогда как уравнение (9.1.7) утверждает, что для области, свободной от источников, $\partial \mathbf{x}/\partial z = \mathbf{A}\mathbf{x}$. Предполагается, что величины \mathbf{A} и \mathbf{x} являются конечными. Это не означает, что все переменные поля всегда являются гладкими функциями координат. В частности, алгебраические переменные, исключенные на 4-м этапе, могут быть и часто будут разрывными на границах слоев.

Упражнения

1. Какой вид примет выражение (9.1.7) в случае уравнений теплового потока? Включите в уравнения радиоактивный источник. [Указание. Посмотрите уравнения (10.1.1) и (10.1.2).]

2. Используя уравнения Максвелла $\nabla \times \mathbf{E} = -\mu \mathbf{H}$, $\nabla \times \mathbf{H} = \mathbf{J} + \epsilon \mathbf{E}$ и закон Ома $\mathbf{J} = \sigma \mathbf{E}$, где σ — проводимость, получите формулу (9.1.7) в предположении, что $\partial/\partial y = 0$.

3. Электрическое поле в ионосфере в соответствии с уравнениями электростатики можно получить из уравнения для потенциала $\nabla \phi = -\mathbf{E}$ при условии, что дивергенция электрического тока равна нулю ($\nabla \cdot \mathbf{J} = 0$), а закон Ома должен иметь добавочный член, обусловленный действием ветра (источник тока, создающий дифференциальное сопротивление перемещению ионов и электронов в магнитном поле Земли) $\mathbf{J} = \sigma \mathbf{E} + \boldsymbol{\tau} \mathbf{V}$. Считайте, что величина \mathbf{V} известна. Какой вид примет уравнение (9.1.7), если величины σ и $\boldsymbol{\tau}$ являются скалярами. Покажите, как продолжить вычисления в случае, если величины σ и $\boldsymbol{\tau}$ являются матрицами (считайте, что для любой матрицы существует обратная матрица).

4. Известно, что в магнитостатике уравнения Максвелла имеют вид $\text{rot } \mathbf{H} = \mathbf{J}$, $\text{div } \mathbf{B} = 0$ и $\mathbf{B} = \mu \mathbf{H}$. Какую форму примет уравнение (9.1.7) при заданном \mathbf{J} ?

5. Это упражнение иллюстрирует линеаризацию нелинейных задач. Для акустических волн, распространяющихся в слоистой ветреной атмосфере, используем известное решение.

Преобразуйте дифференциальные уравнения в частных производных к обыкновенному матричному дифференциальному уравнению. Указание: член, характеризующий горизонтальное ускорение, имеет вид

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial t} + \frac{\partial u}{\partial x} u + \frac{\partial u}{\partial z} w.$$

Похожее выражение получается и для вертикального ускорения. Опустите члены второго порядка в выражениях для P , U и W .

6. Тепловой поток описывается двумя уравнениями. Одно из них является следствием равенства компонент потока (H_x , H_z) теплопроводности среды σ , умноженной на градиент температуры, взятый с обратным знаком (∂_x , $\partial_z T$). Равенство производной температуры по времени, умноженной на теплоемкость c дивергенции теплового потока, взятой с обратным знаком $\partial_x H_x + \partial_z H_z$, дает второе уравнение. Используйте известные решения

$$\begin{bmatrix} T \\ H_x \\ H_z \end{bmatrix} = \begin{bmatrix} \bar{T}(z) \\ 0 \\ \bar{H}_z(z) \end{bmatrix} + \begin{bmatrix} \tilde{T}(z) \\ \tilde{H}_x(z) \\ \tilde{H}_z(z) \end{bmatrix} e^{-i\omega t + ik_x x}.$$

а. Сначала выведите стационарные уравнения для величин \bar{T} и \bar{H} , предполагая, что \tilde{T} и \tilde{H} стремятся к нулю.

б. Найдите уравнения для величин \tilde{T} и \tilde{H} , предполагая, что \bar{T} и \bar{H} удовлетворяют условиям пункта а.

в. Повторите вычисления по пунктам а и б, предполагая линейную зависимость теплоемкости и проводимости от температуры, т. е.

$$\sigma = \sigma_0(z) + \sigma_1(z)T; \quad c = c_0(z) + c_1(z)T.$$

Опустите квадратичные члены в выражениях для \tilde{T} и \tilde{H} .

7. Рассмотрите сжимаемую жидкую сферу, пульсирующую радиально под действием собственного гравитационного поля. Какой вид принимает при этом выражение (9.1.6)? Указания:

$$\rho \ddot{v} = \nabla p - \rho g \text{ — импульс силы;}$$

$$\dot{\rho} + \bar{\rho} v = 0 \text{ — масса;}$$

$$\dot{p} + K \nabla v = 0 \text{ — состояние;}$$

$$\nabla g = 4\pi \gamma \rho \text{ — гравитация (сила тяжести).}$$

9.2. ЧИСЛОВЫЕ МАТРИЦАНТЫ

Дифференциальное уравнение, как известно, связывает переменные поля в некоторой точке с переменными поля в соседних точках. Матрицант связывает переменные поля в слоистой среде на одной глубине с теми же переменными на некоторой другой глубине. Матрицант можно рассматривать так же, как и интеграл (решение) матричного дифференциального уравнения вида (9.1.7). Сначала покажем, как получить матрицант уравнения (9.1.7) численным методом при произвольном изменении плотности и сжимаемости среды с глубиной. Затем получим аналитические решения для специального случая среды с постоянными свойствами. На основании выражения (9.1.7) имеем

$$\frac{\partial \mathbf{x}}{\partial z} \approx \frac{\mathbf{x}(z + \Delta z) - \mathbf{x}(z)}{\Delta z} \approx \mathbf{A}(z) \mathbf{X}(z) + \mathbf{S}(z)$$

или

$$\mathbf{X}(z + \Delta z) = (\mathbf{I} + \mathbf{A}\Delta z) \mathbf{X}(z) + \mathbf{S}(z) \Delta z. \quad (9.2.1)$$

Очевидно, что, задав \mathbf{X} для некоторого значения z , можно использовать выражение (9.2.1) рекурсивно с целью получения \mathbf{X} для любого z . Для простоты можно принять, что $\Delta z = 1$ и использовать нижние индексы для указания z координаты. Обозначим величину $[\mathbf{I} + \mathbf{A}(z)\Delta z]$ через $\mathbf{Q}(z)$. Тогда выражение (9.2.1) примет вид

$$\mathbf{X}_{k+1} = \mathbf{Q}_k \mathbf{X}_k + \mathbf{S}_k.$$

Откуда

$$\mathbf{X}_1 = \mathbf{Q}_0 \mathbf{X}_0 + \mathbf{S}_0.$$

Следовательно,

$$\mathbf{X}_2 = \mathbf{Q}_1 (\mathbf{Q}_0 \mathbf{X}_0 + \mathbf{S}_0) + \mathbf{S}_1 = \mathbf{Q}_1 \mathbf{Q}_0 \mathbf{X}_0 + \mathbf{Q}_1 \mathbf{S}_0 + \mathbf{S}_1.$$

Соответственно

$$\begin{aligned} \mathbf{X}_3 &= \mathbf{Q}_2 (\mathbf{Q}_1 \mathbf{Q}_0 \mathbf{X}_0 + \mathbf{Q}_1 \mathbf{S}_0 + \mathbf{S}_1) + \mathbf{S}_2 = \\ &= \mathbf{Q}_2 \mathbf{Q}_1 \mathbf{Q}_0 \mathbf{X}_0 + \mathbf{Q}_2 (\mathbf{Q}_1 \mathbf{S}_0 + \mathbf{S}_1) + \mathbf{S}_2. \end{aligned}$$

Подобным же образом

$$\mathbf{X}_4 = \mathbf{M} \mathbf{X}_0 + \mathbf{T}. \quad (9.2.2)$$

Итак, имеем в общем случае численно определяемую матрицу \mathbf{M} (называемую матрицантом) и вектор \mathbf{T} , который связывает переменные поля в верхнем из слоев с теми же переменными в нижнем слое с помощью выражения

$$\mathbf{X}_{zB} = \mathbf{M} \mathbf{X}_{zH} + \mathbf{T}. \quad (9.2.3)$$

Матрицу \mathbf{M} называют также интегральной матрицей. Физические задачи формулируются разными способами с различными граничными условиями. Для акустической задачи, обсуждавшейся ранее,

величина \mathbf{X} является двухкомпонентным вектором, составляющие которого — давление и вертикальное смещение. Обе величины являются вначале неизвестными как для верхнего, так и для нижнего слоя среды. Таким образом, выражение (9.2.3) представляет собой два уравнения с четырьмя неизвестными. Решение этой задачи возможно только при введении дополнительно двух граничных условий. Если рассматривать звуковые волны в океане, то граничные условия (упрощенные) можно задать в виде нулевого давления на поверхности и нулевого вертикального смещения на дне моря. Тогда рассматриваемые граничные условия вместе с уравнением (9.2.3) дадут два уравнения с двумя неизвестными и, следовательно, могут быть решены для возмущения на поверхности и давления на дне. С их помощью давление и смещение могут быть определены всюду. Правильное определение граничных условий часто является самой трудной задачей, требующей большого искусства; мы еще вернемся к этому вопросу при решении некоторых других задач в последующих разделах.

Если части среды имеют постоянные свойства и не содержат источников, то можно получить аналитическое выражение для матрицанта. Матрицант, принимающий единичное значение для такого слоя с постоянными свойствами, называется достаточно обобщенно *матрицей слоя*. Можно проверить подстановкой, что выражение

$$\mathbf{X}_z = e^{[\mathbf{A} (z-z_0)]} \mathbf{X}_{z_0} \quad (9.2.4)$$

является решением уравнения

$$(\partial/\partial z) \mathbf{X} = \mathbf{A} \mathbf{X},$$

где

$$e^{\mathbf{A} (z-z_0)} = \mathbf{I} + \mathbf{A} (z - z_0) + \frac{\mathbf{A}^2 (z - z_0)^2}{2!} + \dots \quad (9.2.5)$$

в области пространства (\mathbf{A} является постоянной и не зависит от z) Таким образом, величина $e^{\mathbf{A} (z-z_0)}$ — это требуемый матрицант. Экспоненциальная матрица может быть вычислена с помощью выражения (9.2.2) или (9.2.5) либо с помощью теоремы Сильвестра, описанной в главе 5. В следующем разделе мы покажем как с помощью теоремы Сильвестра получают восходящие и нисходящие волны.

Упражнение

Чему равно \mathbf{Q}_k для разностного приближения?

$$\mathbf{X}(z + \Delta z) - \mathbf{X}(z) = \frac{\mathbf{A} [\mathbf{X}(z + \Delta z) + \mathbf{X}(z)]}{2}$$

9.3. ВОСХОДЯЩИЕ И НИСХОДЯЩИЕ ВОЛНЫ

Известно множество примеров, когда физические задачи для слоистых сред, свободных от источников, сводятся к уравнению вида

$$\frac{d}{dz} \mathbf{X} = \mathbf{A} \mathbf{X}, \quad (9.3.1)$$

где \mathbf{X} — вектор физических переменных; \mathbf{A} — матрица, зависящая от z , если свойства среды являются функцией глубины z . Введем новые переменные, определяемые вектором

$$\mathbf{V} = \mathbf{R} \mathbf{X}, \quad (9.3.2)$$

где \mathbf{R} — квадратная матрица собственных векторов-строк матрицы \mathbf{A} . Матрица \mathbf{R} обратна матрице \mathbf{C} собственных векторов-столбцов матрицы \mathbf{A} . Умножая уравнение (9.3.2) на матрицу \mathbf{C} слева и используя равенство $\mathbf{C} \mathbf{R} = \mathbf{I}$, получаем обратную выражению (9.3.2) зависимость, которая полезна при определении физических переменных \mathbf{X} по новым переменным:

$$\mathbf{X} = \mathbf{C} \mathbf{V}. \quad (9.3.3)$$

Подставив формулу (9.3.3) в уравнение (9.3.1), получим

$$(\mathbf{C} \mathbf{V})_z = \mathbf{A} \mathbf{C} \mathbf{V}; \quad \mathbf{C} \mathbf{V}_z = \mathbf{A} \mathbf{C} \mathbf{V} - \mathbf{C}_z \mathbf{V}.$$

Умножая последнее выражение слева на матрицу \mathbf{R} и используя равенство $\mathbf{R} \mathbf{C} = \mathbf{I}$, получаем:

$$\mathbf{V}_z = (\mathbf{R} \mathbf{A} \mathbf{C}) \mathbf{V} - \mathbf{R} \mathbf{C}_z \mathbf{V}. \quad (9.3.4)$$

Ранее мы предполагали, что матрицы \mathbf{R} и \mathbf{C} являются матрицами собственных векторов-строк и векторов-столбцов матрицы \mathbf{A} . Поэтому можно заменить произведение $\mathbf{R} \mathbf{A} \mathbf{C}$ диагональной матрицей собственных значений $\mathbf{\Lambda}$. Тогда

$$\mathbf{V}_z = \mathbf{\Lambda} \mathbf{V} - \mathbf{R} \mathbf{C}_z \mathbf{V}. \quad (9.3.5)$$

В любой области физического пространства, где среда однородна, матрица \mathbf{A} , а следовательно, и матрица \mathbf{C} не будут зависеть от z , выражение (9.3.5) преобразуется к виду

$$\frac{d}{dz} \mathbf{V} = \mathbf{\Lambda} \mathbf{V}. \quad (9.3.6)$$

Единственная матрица в уравнении (9.3.6) является диагональной, вследствие чего задача для различных переменных, образующих вектор \mathbf{V} , разбивается на отдельные задачи для каждой компоненты вектора. В волновых задачах, по-видимому, целесообразно считать компонентами вектора \mathbf{V} волновые переменные, характеризующие восходящие и нисходящие волны. Эти волны распространяются вверх и вниз в однородных областях среды без взаимодействия друг с другом. Рассмотрим пример.

В разделе 9.1 мы показали, что матричное дифференциальное уравнение первого порядка для акустической задачи в области без источников принимает вид

$$\frac{d}{dz} \begin{bmatrix} P \\ W \end{bmatrix} = \begin{bmatrix} 0 & ia^2 \\ ib^2 & 0 \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix}, \quad (9.3.7)$$

где

$$a^2 = \omega\rho; \quad (9.3.8)$$

$$b^2 = \frac{\omega}{K} - \frac{k^2 x}{\omega\rho}. \quad (9.3.9)$$

Матрица собственных векторов-столбцов \mathbf{C} и матрица собственных векторов-строк для матрицы в уравнении (9.3.7) имеют, как легко проверить, следующий вид:

$$\mathbf{C} = \begin{bmatrix} 1 & 1 \\ -\frac{b}{a} & \frac{b}{a} \end{bmatrix}; \quad (9.3.10)$$

$$\mathbf{R} = \frac{1}{2} \begin{bmatrix} 1 & -\frac{a}{b} \\ 1 & \frac{a}{b} \end{bmatrix}, \quad (9.3.11)$$

Легко также проверить, что векторы нормализованы, т. е. $\mathbf{RC} = \mathbf{CR} = \mathbf{I}$, и что

$$\Lambda = \mathbf{RAC} = \begin{bmatrix} -iab & 0 \\ 0 & iab \end{bmatrix}.$$

Нисходящая волна D связана с собственным значением iab , а восходящая волна U — с $-iab$. Определим восходящие и нисходящие волны следующим образом:

$$\begin{bmatrix} U \\ D \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -\frac{a}{b} \\ 1 & \frac{a}{b} \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix}. \quad (9.3.12a)$$

Собственный вектор-строка может содержать произвольный мультипликативный масштабный множитель, если он выделен из соответствующего собственного вектор-столбца. Это означает, что выражение (9.3.12a) не является единственным. Как оказывается, эти масштабные множители придают восходящим и нисходящим волнам физические размерности давления P . Физические переменные P и W можно найти по заданным параметрам U и D с помощью обратного соотношения

$$\begin{bmatrix} P \\ W \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -\frac{b}{a} & \frac{b}{a} \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}, \quad (9.3.12b)$$

из которого следует, что давление P характеризует сумму восходящих и нисходящих волн, а вертикальная скорость равна величине b/a , умноженной на их разность. Уравнение (9.3.5), определяющее процесс распространения волн U и D , теперь имеет вид

$$\frac{d}{dz} \begin{bmatrix} U \\ D \end{bmatrix} = \begin{bmatrix} -iab & 0 \\ 0 & iab \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix} - \frac{1}{2} \frac{(b/a)_z}{b/a} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}. \quad (9.3.13)$$

В любой области пространства, где величина b/a не является функцией z , справедливы простые несвязанные уравнения

$$\frac{d}{dz} \begin{bmatrix} U \\ D \end{bmatrix} = \begin{bmatrix} -iab & \\ & iab \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}. \quad (9.3.14)$$

Строго говоря, для обоснованного определения восходящих и нисходящих волн U и D необходимо убедиться в том, что для нисходящих волн решение принимает следующую форму:

$$d(z, t) = e^{-i\omega t + ik_z z}, \quad (9.3.15)$$

где ω и k_z должны совпадать по знаку. При этом фаза сохраняется постоянной при возрастании величин z и t . Противоположный знак должен применяться для величины U . Другими словами, параметр $k_z = ab$ должен принимать знак частоты ω . Чтобы увидеть это, извлечем квадратный корень из произведения выражений (9.3.8) и (9.3.9):

$$k_z = ab = \omega \left(\frac{\rho}{K} - \frac{k_x^2}{\omega^2} \right)^{1/2}. \quad (9.3.16)$$

Для вертикально распространяющихся волн параметр $k_x = 0$, поэтому выражение $k_z = ab$ принимает вид $k_z = \omega(\rho/K)^{1/2}$. Подставляя эту величину в формулу (9.3.15), видим, что фазовый сдвиг у экспоненты оказывается постоянным, если $z/t = (k/\rho)^{1/2}$. Из этого следует, что присущая данному веществу скорость распространения акустической волны определяется зависимостью

$$v = \left(\frac{K}{\rho} \right)^{1/2}. \quad (9.3.17)$$

На рис. 9.1 показаны лучи и фронты волн в слое. Как видно, кажущиеся длины волны λ_x и λ_z , наблюдаемые вдоль осей x и z , превышают действительную длину волны λ , наблюдаемую вдоль луча. Очевидно также, что $\lambda/\lambda_x = \sin \theta$, а $\lambda/\lambda_z = \cos \theta$. Поэтому для пространственных частот $k_x = 2\pi/\lambda_x$ и $k_z = 2\pi/\lambda_z$ удовлетворяется условие $k_x^2 + k_z^2 = (2\pi/\lambda)^2 = \omega^2/v^2$, которое является преобразованием Фурье волнового уравнения. Из приведенных формул и рис. 9.1 следует, что угол θ между вертикалью и направлением распространения луча определяется выражением

$$\sin \theta = \frac{k_x v}{\omega}. \quad (9.3.18)$$

Подставив выражения (9.3.17) и (9.3.18) в формулу (9.3.16), получим

$$k_z = ab = \frac{\omega}{v} \cos \theta. \quad (9.3.19)$$

Закон Снеллиуса гласит, что величина $\sin \theta/v$ не меняется вдоль луча при прохождении его от слоя к слою. Следствием этого является неизменность величины k_x/ω в каждом слое. Постоянство пространственной частоты вдоль луча в каждом слое существенно

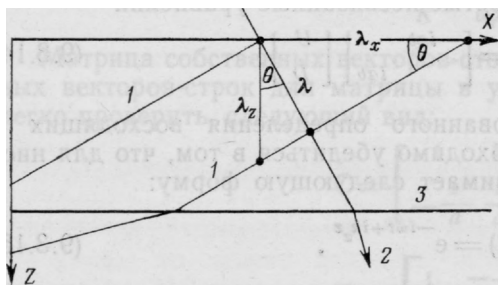


РИС. 9.1. Лучи и фронты волн в слое.

λ , λ_x , λ_z — длины волн в направленных луча и осей x и z ; θ — угол выхода луча (фронта); 1 — фронт; 2 — луч; 3 — граница слоя

для удовлетворения условий непрерывности на границах раздела слоев.

Функция времени (9.3.15) является комплексной. Чтобы получить действительную функцию времени, необходимо просуммировать или проинтегрировать выражение (9.3.15), учтя положительные и отрицательные частоты. Тогда, как мы видели в предшествующих главах, где рассматривался анализ временных рядов, должно выполняться условие $D(\omega) = \bar{D}(-\omega)$.

Величина b/a является характеристикой среды — ее волновой проводимостью Y . Извлекая квадратный корень из отношения выражений (9.3.9) и (9.3.8), получаем:

$$Y = \frac{b}{a} = \frac{(1 - v^2 k_x^2 / \omega^2)^{1/2}}{\rho v}; \quad (9.3.20)$$

$$Y = \frac{b}{a} = \frac{\cos \theta}{\rho v}; \quad (9.3.21)$$

$$I = \frac{a}{b} = \frac{\rho v}{\cos \theta}. \quad (9.3.22)$$

Сейчас мы убедимся в том, что это определение импеданса совпадает с приведенным в предыдущей главе. Чтобы сделать это, внимательно рассмотрим матрицант для пересечения слоя $\exp[\mathbf{A}(z_2 - z_1)] = \exp(\mathbf{A}\Delta z)$. С помощью теоремы Сильвестра имеем для матрицанта

$$\exp(\mathbf{A}\Delta z) = \mathbf{C} \begin{bmatrix} e^{-ik_z \Delta z} & 0 \\ 0 & e^{ik_z \Delta z} \end{bmatrix} \mathbf{R}. \quad (9.3.23)$$

Матрицант связывает волновые переменные в верхней части слоя на глубине z_1 с теми же переменными в нижней части слоя на глубине z_2 . Итак, выражение (9.3.23) дает возможность записать

$$\begin{bmatrix} P \\ W \end{bmatrix}_2 = \mathbf{C}_1 \exp(\Lambda_1 \Delta z) \mathbf{R}_1 \begin{bmatrix} P \\ W \end{bmatrix}_1. \quad (9.3.24)$$

Уравнение (9.3.24), которое кажется возникшим из теоремы Сильвестра, в действительности имеет простую интерпретацию.

Начнем с правой части. Проинтерпретируем умножение матрицы \mathbf{R}_1 на переменные P и W как преобразование для восходящих и нисходящих волн. Умножение на $\exp(\Lambda_1 \Delta z)$ означает переход через слой, а умножение на \mathbf{C}_1 снова преобразует правую часть (9.3.24) к переменным P и W , которые непрерывны при переходе через границу раздела. Умножая выражение (9.3.24) на \mathbf{R}_2 и учитывая соотношения (9.3.11) и (9.3.12), имеем:

$$\begin{bmatrix} U \\ D \end{bmatrix}_2 = \mathbf{R}_2 \mathbf{C}_1 \exp(\Lambda \Delta z) \begin{bmatrix} U \\ D \end{bmatrix}_1. \quad (9.3.25)$$

Выражение (9.3.25) определяет восходящие и нисходящие волны ниже границы раздела, как это делалось в предыдущей главе. Запишем выражение для матрицы слоя

$$\frac{1}{2} \begin{bmatrix} 1 - Y_2^{-1} \\ 1 \\ 1 \\ Y_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -Y_1 & Y_1 \end{bmatrix} \begin{bmatrix} e^{-iab\Delta z} & 0 \\ 0 & e^{iab\Delta z} \end{bmatrix}. \quad (9.3.26)$$

Определим переменную преобразования

$$Z = \exp\left(\frac{2i\omega \Delta z}{v \cos \theta}\right)$$

Зная, что время прохождения волны через слой $\Delta t = \Delta z / v \cos \theta$, приведем матрицу слоя (9.3.26) к виду

$$\frac{Y_1 + Y_2}{2Y_2} \begin{bmatrix} 1 & \frac{Y_2 - Y_1}{Y_2 + Y_1} \\ \frac{Y_2 - Y_1}{Y_2 + Y_1} & 1 \end{bmatrix} \begin{bmatrix} Z^{-1/2} & \\ & Z^{1/2} \end{bmatrix}. \quad (9.3.27)$$

Последнее выражение сопоставлено с выражением для матрицы слоя (8.2.4), имеющим вид

$$\frac{1}{t} \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \begin{bmatrix} Z^{-1/2} & 0 \\ 0 & Z^{1/2} \end{bmatrix}.$$

Легко установить, что учет зависимости $Y = b/a$ приводит к хорошо знакомому определению коэффициента отражения

$$c = \frac{Y_2 - Y_1}{Y_2 + Y_1} = \frac{I_1 - I_2}{I_1 + I_2} = \frac{\rho_1 v_1 / \cos \theta_1 - \rho_2 v_2 / \cos \theta_2}{\rho_1 v_1 / \cos \theta_1 + \rho_2 v_2 / \cos \theta_2}. \quad (9.3.28)$$

Упражнения

1. Переопределите собственные векторы так, чтобы $W=D+U$ и $P=(D-U)/Y$. Это преобразование было бы полезно, если бы мы хотели, чтобы условие $t=1+c$ относилось к нормализованным переменным вертикальной составляющей скорости вместо переменной давления, как в главе 8. Получите все уравнения этого раздела для вертикальной составляющей скорости.

2. Напишите выражение для матрицанта, пересекающего слой, через параметры a , b и мощность слоя h .

9.4. ПРИНЦИП ВЗАИМНОСТИ

Принцип взаимности гласит, что если источник и приемник (при некоторых условиях) поменять местами, то при этом будет наблюдаться одинаковое волновое поле. Этот принцип часто применяется для облегчения вычислений, а также может быть использован для упрощения сбора информации. Несколько удивительно, что этот принцип применим к Земле с ее сложной неоднородностью. Интуитивно ясно, что главной причиной действенности принципа взаимности является то, что энергия распространяется одинаково вдоль заданного луча в обоих направлениях. В любом направлении она распространяется с одной и той же скоростью и одинаковым затуханием.

Принцип взаимности хорошо выполняется для скалярных волн при скалярных источниках и приемниках (например, акустические волны давления при взрывных источниках и приемниках, чувствительных к давлению). Ситуация осложняется, когда источники или приемники обладают подвижными диафрагмами, потому что при этом важна их ориентировка. Направленные свойства источника и приемника часто характеризуют диаграммами направленности излучения. Для применения принципа взаимности в этом случае необходимо считать диаграммы направленности излучения связанными со средой, а не с источником и приемником. Таким образом, когда говорят, что источник и приемник меняются местами, то меняться местами должны только скалярные амплитуды, а диаграммы направленности источника и приемника остаются фиксированными на своих местах. Эти общие соображения в последующем будут уточняться. Станет ясным, что понятие лучей действительно оказывается неподходящим. Принцип взаимности действует также в задачах диффузии и потенциала.

Теоретические работы часто читаются довольно тяжело. Часто они начинаются с указания на то, что дифференциальный оператор вместе с соответствующими граничными условиями должен составлять самосопряженную задачу. Это означает, что для дифференциальных уравнений, записанных в конечно-разностной форме, матрица коэффициентов оказывается симметричной. Рассмотрим в качестве примера одномерную задачу для акустических волн. В соответствии с законом Ньютона произведение плотности ρ на ускорение $\partial_{tt}u$ равно сумме градиента давления с обратным зна-

на внешние возбуждающие силы:

$$\begin{bmatrix} p_0 \\ u_0 \\ p_1 \\ u_1 \\ p_2 \\ a_2 \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & A & B & \cdot & \cdot \\ \cdot & \cdot & C & D & \cdot & \cdot \\ A & C & \cdot & \cdot & \cdot & \cdot \\ B & D & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} V_0 \\ F_0 \\ V_1 \\ F_1 \\ V_2 \\ F_2 \end{bmatrix}, \quad (9.4.4)$$

где A, B, C и D указывают на симметрию матрицы. Если в первом эксперименте все внешние источники обнулить, кроме одного крайнего, являющегося объемным источником единичной интенсивности $V_0=1$, то в соответствии с выражением (9.4.4) давление в середине $p_1=A$. Если во втором эксперименте обнулить все внешние источники, кроме одного среднего объемного источника с интенсивностью $V_1=1$, то согласно выражению (9.4.4) давление p_0 на конце также будет равно A . Это и есть принцип взаимности. Отметим, что с помощью буквы D в выражении (9.4.4) подобное представление применяется к силам и смещениям. При смешанном представлении применяются буквы C и B .

В реальном эксперименте нет возможности реализовать чистый объемный источник или чистую внешнюю силу. Другими словами, внешний источник может иметь некоторый конечный нулевой импеданс. Тогда первый эксперимент следует выполнить с возбуждением в середине, имея для отклика на конце

$$\begin{bmatrix} p_0 \\ u_0 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_1 \\ F_1 \end{bmatrix}. \quad (9.4.5)$$

Меняя местами источник и приемник, получим

$$\begin{bmatrix} p_1 \\ u_1 \end{bmatrix} = \begin{bmatrix} A & C \\ B & D \end{bmatrix} \begin{bmatrix} V_0 \\ F_0 \end{bmatrix}. \quad (9.4.6)$$

Интересной особенностью выражений (9.4.5) и (9.4.6) является то, что каждая матрица в них оказывается результатом транспониро-

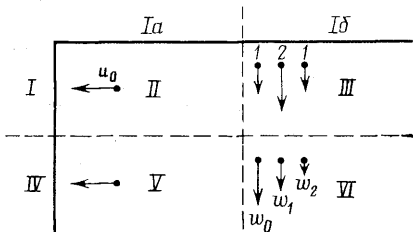


РИС. 9.2. Пример, иллюстрирующий принцип взаимности для случая $u_0 = \omega_0 + 2\omega_1 + \omega_2$.

I — первый эксперимент; Ia — положение 0; Ib — положение 1; II — наблюдаемое горизонтальное смещение; III — массив вертикальных сил; IV — второй эксперимент; V — горизонтальная единичная сила; VI — три наблюдения вертикального смещения

вания другой. Эта особенность должна учитываться, особенно при рассмотрении более сложного эксперимента, где векторы в выражениях (9.4.5) и (9.4.6) содержат больше элементов. Например, в уравнении (9.4.5) или (9.4.6) вектор мог бы содержать элементы массива физически раздельных объемных источников или датчиков (приемников) давления. Если составить симметричные алгебраи-

ческие уравнения, подобные уравнениям (9.4.3), для упругих, электромагнитных, диффузионных или потенциальных задач, то матрицы, подобные матрицам (9.4.5) и (9.4.6), всегда будут транспонированными по отношению друг к другу. Построение симметричных уравнений, подобных уравнениям (9.4.3), часто не вызывает трудностей, хотя может быть достаточно сложным делом в многомерной непрямоугольной системе координат.

В более общем случае векторы возбуждения в правой части равенства (9.4.5) или (9.4.6) можно обозначить через \mathbf{E} , а векторы ответной реакции в левой части — через \mathbf{R} . Тогда, используя обозначение \mathbf{M} для матрицы в уравнении (9.4.5) и \mathbf{M}^T для транспонированной матрицы, перепишем выражения (9.4.5) и (9.4.6) в следующем виде:

$$\mathbf{R}_0 = \mathbf{M}\mathbf{E}_1; \quad (9.4.7)$$

$$\mathbf{R}_1 = \mathbf{M}^T\mathbf{E}_0. \quad (9.4.8)$$

Теперь выведем физические следствия из выражений (9.4.7) и (9.4.8). Сначала рассмотрим скалярное произведение уравнения (9.4.7) с величиной \mathbf{E}_0^T :

$$\mathbf{E}_0^T\mathbf{R}_0 = \mathbf{E}_0^T\mathbf{M}\mathbf{E}_1.$$

Правая часть последнего выражения, являющаяся скаляром, может быть транспонирована

$$\mathbf{E}_0^T\mathbf{R}_0 = (\mathbf{E}_0^T\mathbf{M}\mathbf{E}_1)^T = \mathbf{E}_1^T\mathbf{M}^T\mathbf{E}_0.$$

Подставляя сюда величину $\mathbf{M}^T\mathbf{E}_0$ из уравнения (9.4.8), найдем:

$$\mathbf{E}_0^T\mathbf{R}_0 = \mathbf{E}_1^T\mathbf{R}_1. \quad (9.4.9)$$

Уравнение (9.4.9) представляет собой основную формулировку принципа взаимности; скалярное произведение вектора возбужде-

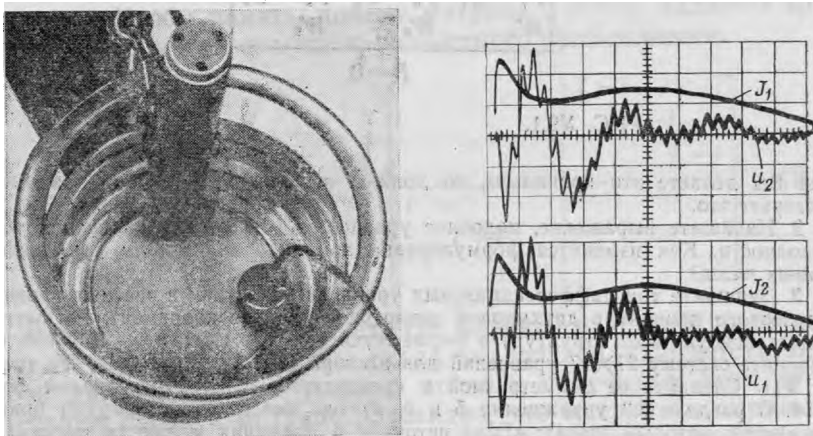


РИС. 9.3. Выполнение принципа взаимности для упругих колебаний, возбуждаемых электромагнитной системой [32]

ния и вектора ответной реакции в положении 0 равняется их скалярному произведению в положении 1. Отметим, что скалярные произведения вычисляются для векторов, наблюдаемых в различных экспериментах.

Пример упругой системы, в которой направления вектора смещения и векторов силы совпадают, изображен на рис. 9.2. Лабораторный пример Уайта [32], в котором упругие колебания возбуждаются электромагнитными силами в твердом теле, жидкости и газе, показан на рис. 9.3. Геофон представляет собой подвешенный на пружине маятник, взаимодействующий с индукционной катушкой. Первый геофон смонтирован на трубке, которая покоится на дне стеклянного эксикатора. Второй геофон прикреплен к стеклу с помощью куска глины ниже уровня воды. Верхняя пара трасс (рис. 9.3) показывает силу тока источника в первом геофоне и разность потенциалов на втором (при разомкнутой цепи). Нижние кривые показывают силу тока во втором геофоне и напряжение в первом.

Упражнения

1. Рассмотрите уравнения Пуассона $\partial_{xx}R = -E$ на сетке из пяти точек, для которых граничные условия заключаются в том, что крайние точки являются нулями. Единичное возбуждение в третьей точке сетки дает решение $(0, 1/2, 1, 1/2, 0)$. Найдите решение для единичного возбуждения во второй точке сетки.

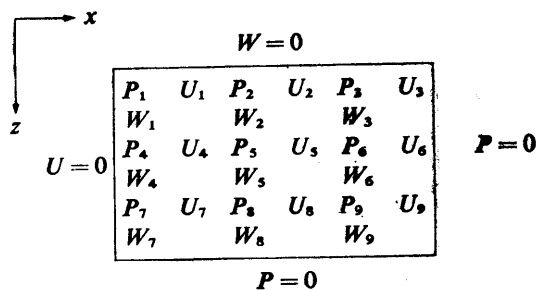


РИС. У9.1.

Если Вы делаете это правильно, то должен выполняться принцип взаимности. Проверьте это.

2. Напишите выражение, подобное уравнению (9.4.3), для уравнения теплопроводности. Как изменится формулировка принципа взаимности при введении мнимых чисел?

3. Запишите три дифференциальных уравнения акустики в частных производных первого порядка в двухмерной декартовой системе координат. Рассмотрите представленную ниже структуру в форме сетки (рис. У9.1). Напишите систему 27×27 уравнений для вектора $(U_1, P_1, W_1, U_2, P_2, W_2, \dots, U_9, P_9, W_9)$. Сделайте ее симметричной и представьте в явной регулярной форме.

4. В разделе 8.3 упражнения 5 и 6, взятые вместе, иллюстрируют принцип взаимности, который гласит: «Если источник и приемник меняются местами, то будет наблюдаться одинаковая форма волны». Решите эту задачу для источника на поверхности и приемника в середине слоя, а затем такую же задачу, поменяв местами источник и приемник с целью проверки принципа взаимности.

9.5. ПРИНЦИПЫ СОХРАНЕНИЯ И ОРТОГОНАЛЬНЫЕ МОДЫ КОЛЕБАНИИ

Ранее было показано, что задачи распространения акустических колебаний в слоистых средах приводятся к матричному дифференциальному уравнению первого порядка:

$$\frac{\partial}{\partial z} \mathbf{y} = \mathbf{A} \mathbf{y} + \mathbf{s}. \tag{9.5.1}$$

Многие задачи вида (9.5.1) можно записать в так называемой форме Аткинсона:

$$\mathbf{J} \frac{\partial}{\partial z} \mathbf{y} = [\mathbf{G}(z) + \lambda \mathbf{H}(z)] \mathbf{y}, \tag{9.5.2}$$

где \mathbf{J} — косая (несимметричная) эрмитова матрица ($\mathbf{J}^* = -\mathbf{J}$), не зависящая от z ; $\mathbf{G}(z)$ и $\mathbf{H}(z)$ — эрмитовы матрицы ($\mathbf{H}^* = \mathbf{H}$); λ — скаляр, играющий роль собственного значения. Например, в акустическом случае имеем

$$\frac{\partial}{\partial z} \begin{bmatrix} P \\ W \end{bmatrix} = i \left[\begin{pmatrix} 0 & \omega\rho \\ +\frac{\omega}{K} - \frac{k^2 x}{\omega\rho} & 0 \end{pmatrix} \begin{bmatrix} P \\ W \end{bmatrix} + i \begin{bmatrix} s_p \\ s_w \end{bmatrix} \right]. \tag{9.5.3}$$

Умножив последнее выражение слева на косую эрмитову матрицу, получим:

$$\begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} \frac{\partial}{\partial z} \begin{bmatrix} P \\ W \end{bmatrix} = \left[\begin{pmatrix} +\frac{\omega}{K} - \frac{k^2 x}{\omega\rho} & 0 \\ 0 & \omega\rho \end{pmatrix} \begin{bmatrix} P \\ W \end{bmatrix} + \begin{bmatrix} s_w \\ s_p \end{bmatrix} \right]. \tag{9.5.4}$$

В выражении (9.5.4) операторы являются самосопряженными. Это означает, что матрица в правой части и оператор в левой части эрмитовы. Чтобы понять, почему матрица $\mathbf{J}(\partial/\partial z)$ является эрмитовой, запишем ее с помощью разностного приближения:

$$\mathbf{J} \frac{\partial}{\partial z} = \frac{i}{\Delta z} \begin{bmatrix} 0 & \delta_z \\ \delta_z & 0 \end{bmatrix} = \frac{1}{\Delta z} \left[\begin{array}{c|c} i & -i \\ \hline i & -i \\ \hline i & -i \\ \hline i & -i \end{array} \right] \tag{9.5.5}$$

Из выражения (9.5.5) следует, что матрице в правой части недостает двух строк, чтобы она стала квадратной. Соответствующий выбор двух граничных условий подобен получению двух дополнительных строк в матрице. Очевидно, что матрица в выражении (9.5.5) настолько близка к эрмитовой, что достаточно выбрать еще две строки, чтобы сделать ее эрмитовой. Например, можно поме-

стить между верхней и нижней половинами матрицы в выражении (9.5.5) следующие две строки:

$$\begin{bmatrix} \dots & \dots & i \\ -i & \dots & \dots \end{bmatrix}.$$

Поскольку оператор в формуле (9.5.5) можно сделать эрмитовым с помощью выбора соответствующих граничных условий, а другие операторы в выражении (9.5.4) уже являются эрмитовыми, то, по-видимому, формула Аткинсона применима к физическим задачам, в которых выполняется принцип взаимности. Принцип взаимности применим к большинству задач геофизической разведки. Одной из простых физических задач, к которой принцип взаимности неприменим, является распространение звуковых волн в атмосфере при ветре. С точки зрения физики это объясняется тем, что волны распространяются в направлении потока ветра быстрее, чем против потока. С позиций математики в этом случае не существует матрицы \mathbf{J} , преобразовывающей выражение (9.5.1) к виду (9.5.2). Преобразование уравнения (9.5.1) в (9.5.2) возможно только в области, где отсутствуют источники. Если необходимо, чтобы величина ω играла роль собственного значения, то, считая функцию источника всюду равной нулю, можно привести равенство (9.5.4) к виду

$$\begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} \frac{\partial}{\partial z} \begin{bmatrix} P \\ W \end{bmatrix} = +\omega \begin{bmatrix} \left(\frac{1}{K} - \frac{k_x^2}{\omega^2 \rho}\right) & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix} \quad (9.5.6)$$

Здесь матрица $\mathbf{G}(z)$ обратилась в нуль, а величина k_x^2/ω^2 , пропорциональная синусу угла падения, должна рассматриваться как постоянная для различных величин собственного значения ω . Аналогично можно было бы в качестве собственного значения выбрать величину $-k_x^2$. Тогда выражение (9.5.4) имело бы вид

$$\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \frac{\partial}{\partial z} \begin{bmatrix} P \\ W \end{bmatrix} = \begin{bmatrix} \frac{\omega}{K} & 0 \\ 0 & \omega \rho \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix} - k_x^2 \begin{bmatrix} \frac{1}{\omega \rho} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix}. \quad (9.5.7)$$

Очевидно, что существует еще одна возможность выбора в качестве собственного значения угловой переменной $-k_x^2/\omega^2$ для фиксированной величины ω . Формула Аткинсона (9.5.2) непосредственно приводит к различным вариантам принципа сохранения. Вычислим вертикальную производную от квадратичной формы $\mathbf{y}^* \mathbf{J} \mathbf{y}$:

$$\begin{aligned} \frac{\partial}{\partial z} \mathbf{y}^* \mathbf{J} \mathbf{y} &= \mathbf{y}_z^* \mathbf{J} \mathbf{y} + \mathbf{y}^* \mathbf{J} \mathbf{y}_z = -\mathbf{y}_z^* \mathbf{J}^* \mathbf{y} + \mathbf{y}^* \mathbf{J} \mathbf{y}_z = -(\mathbf{J} \mathbf{y}_z)^* \mathbf{y} + \mathbf{y}^* (\mathbf{J} \mathbf{y}_z) = \\ &= -(\mathbf{G} \mathbf{y} + \lambda \mathbf{N} \mathbf{y})^* \mathbf{y} + \mathbf{y}^* (\mathbf{G} \mathbf{y} + \lambda \mathbf{N} \mathbf{y}) = (\lambda - \lambda^*) \mathbf{y}^* \mathbf{N} \mathbf{y}. \end{aligned} \quad (9.5.8)$$

Очень часто собственные значения ω , $-k_x^2$ или $-k_x^2/\omega^2$ являются действительными величинами. В таких случаях выполняется равенство $\lambda - \lambda^* = 0$ и выражение (9.5.8) показывает, что величина

y^*Jy является квадратичной функцией волновых переменных, инвариантной относительно z . В приведенном ранее примере для акустических волн рассматриваемый квадратичный инвариант пропорционален потоку энергии, а именно:

$$y^*Jy = -i [P^*W^*] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix} = i(P^*W + W^*P) = -2i \operatorname{Re}(P^*W). \quad (9.5.9)$$

Рассмотрим комплексную частоту $\omega = \omega_r + i\omega_i$. Тогда уравнение (9.5.8) для примера из акустики, описываемого уравнением (9.5.6), приобретает вид

$$-\frac{\partial}{\partial z} \operatorname{Re}(P^*W) = \omega_i \left[\left(\frac{1}{K} - \frac{k_x^2}{\omega^2 \rho} \right) P^*P + \rho W^*W \right]. \quad (9.5.10)$$

Если величины P и W имеют временную зависимость вида

$$\exp[-i(\omega_r + i\omega_i)t] = \exp(-i\omega_r t + \omega_i t),$$

то квадратичные формы, подобные P^*P и W^*W , имеют временную зависимость вида $e^{2\omega_i t}$. Тогда множитель $2\omega_i$ можно рассматривать как результат дифференцирования по времени. Следовательно, выражение (9.5.10) можно представить в виде

$$-\frac{\partial}{\partial z} \operatorname{Re}(P^*W) = + \frac{\partial}{\partial t} \left\{ \frac{1}{2} \left[\left(\frac{1}{K} - \frac{K_x^2}{\omega^2 \rho} \right) P^*P + \rho W^*W \right] \right\} = \frac{\partial}{\partial t} E. \quad (9.5.11)$$

Уравнение (9.5.11) интерпретируется следующим образом: производная по времени от плотности энергии E в некоторой точке пропорциональна дивергенции потока энергии в той же точке с обратным знаком. В других задачах квадратичные формы могут соответствовать не энергии, а, например, количеству движения и т. п.

Хорошо известная теорема из теории матриц гласит, что эрмитовы матрицы имеют действительные собственные значения. Почему же тогда мы рассматривали комплексное собственное значение в выражении (9.5.8)? Это делали для того, чтобы не было необходимости выбирать матрицу оператора конечной разности и граничные условия, обеспечивающие эрмитовость операторов. В частности, чтобы величина $\partial E / \partial t \neq 0$, энергия должна поступать к границе или исходить от нее.

Теперь предположим, что граничные условия выбраны таким образом, чтобы матрица $J\partial/\partial z$ была симметричной, а ее собственные значения были действительными. Пусть $y_n(z)$ является решением уравнения (9.5.2) с собственным значением λ_n , и пусть $y_m(z)$ также будет решением этого уравнения с другим собствен-

ным значением λ_m . Воспользовавшись тем же путем, что и для получения формулы (9.5.8), найдем

$$\frac{\partial}{\partial t} (\mathbf{y}_m^* \mathbf{J} \mathbf{y}_n) = (\lambda_n - \lambda_m) \mathbf{y}_m^* \mathbf{H} \mathbf{y}_n. \quad (9.5.12)$$

Интегрируя последнее выражение по z в пределах от z_a до z_b , получаем

$$\mathbf{y}_m^* \mathbf{J} \mathbf{y}_n \Big|_{z_a}^{z_b} = (\lambda_n - \lambda_m) \int_a^b \mathbf{y}_m^*(z, \lambda_m) \mathbf{H}(z) \mathbf{y}(z, \lambda_n) dz. \quad (9.5.13)$$

Если граничные условия были выбраны так, чтобы энергия не поступала к границам в точки z_a и z_b или не исходила от них, то левая часть должна равняться нулю. Поскольку по предположению $\lambda_n \neq \lambda_m$, то интеграл в правой части выражения (9.5.13) должен равняться нулю. Это означает, что рассматриваемые два решения (называемые модами) ортогональны. Ортогональность этих функций часто полезна при теоретических исследованиях и вычислениях. Более подробно этот вопрос, включая наиболее общую форму граничных условий для принципа сохранения энергии, освещен в работе [14].

Упражнение

Покажите, что подстановка выражения (9.5.8) в уравнение (9.5.7) приводит к определению горизонтального потока энергии. Возможно, Вам понадобится принять $k_x = k_r + ik_i$ и одновременно предположить, что $|k_r| \gg |k_i|$.

9.6. УПРУГИЕ ВОЛНЫ

В дальнейшем предполагается, что читатель знаком с классической теорией упругости. Лишь в немногих учебниках рассматриваются специальные вопросы о слоистых средах, которые имеют важное значение в сейсмологии. Много статей по этому вопросу можно найти в Бюллетене Американского сейсмологического общества (BSSA). Для читателей, которые не знакомы с BSSA, приведем результаты применения методов данной главы к уравнениям теории упругости для однородной среды.

В теории упругости для однородной среды обычно используют следующие величины и их обозначения: смещения (u , w) в направлениях x и z , матрицу напряжений τ , коэффициенты Ламе λ и μ и плотность ρ . Закон Гука и закон Ньютона с учетом зависимости от времени $e^{-i\omega t}$ можно представить так:

$$\frac{\partial}{\partial z} \begin{bmatrix} U \\ \tau_{zz} \\ W \\ \tau_{zx} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -\partial_x \frac{1}{\mu} \\ 0 & 0 & -\rho\omega^2 & -\partial_x \\ \frac{-\lambda}{(\lambda + 2\mu) \partial_x} & \frac{1}{\lambda + 2\mu} & 0 & 0 \\ -\rho\omega^2 - \gamma & \frac{-\lambda \partial_x}{\lambda + 2\mu} & 0 & 0 \end{bmatrix} \begin{bmatrix} U \\ \tau_{zz} \\ W \\ \tau_{zx} \end{bmatrix}, \quad (9.6.1)$$

где

$$\gamma = \frac{4\mu(\lambda + \mu)}{(\lambda + 2\mu)} \partial_x^2. \quad (9.6.2)$$

Введем обозначения:

$$\begin{aligned} \alpha^2 &= \frac{\lambda + 2\mu}{\rho}; & \beta^2 &= \frac{\mu}{\rho}; & m^2 &= \frac{-\omega^2}{\alpha^2} - \partial_{xx}; \\ n^2 &= \frac{-\omega^2}{\beta^2} - \partial_{xx}; & l^2 &= \frac{-\omega^2}{\beta^2} - 2\partial_{xx}. \end{aligned} \quad (9.6.3)$$

Если свойства среды не изменяются в направлении x , то преобразование вектор-строки собственных значений матрицы \mathbf{R} к переменным восходящих и нисходящих волн имеет вид

$$\begin{bmatrix} p^+ \\ s^+ \\ p^- \\ s^- \end{bmatrix} = \frac{\Lambda^{-1}}{2\omega^2\rho} \begin{bmatrix} 2\mu m \partial_x & m & \mu l^2 & \partial_x \\ \mu l^2 & -\partial_x & -2\mu n \partial_x & n \\ -2\mu m \partial_x & -m & \mu l^2 & \partial_x \\ -\mu l^2 & \partial_x & -2\mu n \partial_x & n \end{bmatrix} \begin{bmatrix} u \\ \tau_{zz} \\ w \\ \tau_{zx} \end{bmatrix}. \quad (9.6.4)$$

где

$$\Lambda = \begin{bmatrix} m & & & \\ & n & & \\ & & -m & \\ & & & -n \end{bmatrix}.$$

Соответственно преобразование вектор-столбца собственных значений обращенной матрицы \mathbf{C} записывается в следующем виде:

$$\begin{bmatrix} u \\ \tau_{zz} \\ w \\ \tau_{zx} \end{bmatrix} = \begin{bmatrix} -\partial_x & -n & -\partial_x & -n \\ -\mu l^2 & 2\mu n \partial_x & -\mu l^2 & 2\mu n \partial_x \\ -m & \partial_x & m & -\partial_x \\ -2\mu m \partial_x & -\mu l^2 & 2\mu m \partial_x & \mu l^2 \end{bmatrix} \begin{bmatrix} p^+ \\ s^+ \\ p^- \\ s^- \end{bmatrix}, \quad (9.6.5)$$

Рассматриваемые матрицы хорошо разбиваются на блоки размерности 2×2 . Читатель может проверить, что $\mathbf{CR} = \mathbf{RC} = \mathbf{I}$ и $\mathbf{C}\Lambda\mathbf{R} = \mathbf{A}$.

10. ДВУХМЕРНЫЕ И ТРЕХМЕРНЫЕ КРАЕВЫЕ ЗАДАЧИ

Решению краевых задач путем аппроксимации дифференциальных уравнений уравнениями в конечных разностях¹ посвящен ряд учебников, например [33, 34]. В главе кратко рассмотрим основные представления этого направления. Общая идея решения двухмерных задач состоит в том, что строят одну или несколько двухмерных сеток, на которых переменные поля представляются как функции двух пространственных координат, а затем вводят начальные условия. Решение задачи обычно осуществляют на ЭВМ. Известны многочисленные обширные исследования, посвященные уравнению диффузии. Гораздо меньше исследований посвящено волновому уравнению. При моделировании волнового уравнения возникают трудности, связанные с тем, что для получения хороших результатов, вероятно, недостаточно даже десяти точек на длину волны. При этом практически невозможно выполнить указанное условие для очень многих длин волн с использованием приемлемой сетки. В рассматриваемом случае энергия быстро распространяется к краям сетки, где она отражается обратно, независимо от вашего желания. Один из способов ослабления этих трудностей — использование системы координат, перемещающейся вместе с волнами. Это упрощает рассмотрение отражения волн от поверхности земли, где они наблюдаются, внутрь земли. Такой подход позволяет обрабатывать сейсмические данные методом отраженных волн.

10.1. КЛАССИЧЕСКИЕ НЕСТАЦИОНАРНЫЕ КРАЕВЫЕ ЗАДАЧИ

Для облегчения понимания основных проблем, возникающих при решении интересующих нас задач, рассмотрим одномерное, достаточно простое уравнение теплового потока, получаемое из двух очевидных уравнений. Первое из них гласит, что поток тепла H является следствием градиента температуры и пропорционален коэффициенту теплопроводности среды σ :

$$H = -\sigma \frac{\partial T}{\partial x}. \quad (10.1.1)$$

Второе гласит, что изменение температуры во времени пропорционально дивергенции теплового потока H и обратно пропорционально удельной теплоемкости C среды:

$$\frac{\partial T}{\partial t} = -\frac{1}{C} \frac{\partial H}{\partial x}. \quad (10.1.2)$$

¹ Широко применяются также термины: «разностные уравнения» и «разностные схемы». (Прим. перев.)

Подставив выражение (10.1.1) в уравнение (10.1.2), получим (пренебрегая производной от σ) уравнение теплопроводности:

$$\frac{\partial T}{\partial t} = \frac{\sigma}{C} \frac{\partial^2 T}{\partial x^2}. \quad (10.1.3)$$

Обычно в теории разностных уравнений полагают, что температуру $T(x, t) = T(k\Delta x, n\Delta t)$ в уравнении (10.1.3) можно представить полиномом вида T^n_k , где верхний индекс обозначает время. Введя обозначение $b = \sigma\Delta t/2C\Delta x^2$, запишем уравнение теплопроводности (10.1.3) в конечно-разностной форме:

$$T_k^{n+1} - T_k^n = 2b(T_{k+1}^n - 2T_k^n + T_{k-1}^n). \quad (10.1.4)$$

Если температура T^n_k известна во всех точках k пространства для некоторого определенного значения времени n , то уравнение (10.1.4) позволяет рассчитать температуру для любого момента времени. Читатель может заметить, что производная по времени относится (центрируется) к точке $T_k^{n1/2}$, тогда как пространственная производная — к точке T_k^n . Это может привести к определенным трудностям. Уравнение теплопроводности медленно сглаживает пространственные волны большой длины, а короткие волны — быстро. Разностное уравнение теплопроводности делает то же самое, только на очень коротких волнах сказывается различие в центровке производных по времени и пространственным переменным. В результате очень короткие волны не будут ослабляться достаточно быстро, а могут даже усиливаться. Действительно, при уменьшении величины Δx создаются условия для существования на сетке все более коротких волн. При этом всегда наблюдается усиление, что приводит к разрушению решения. Эта ситуация, называемая неустойчивостью решения, достаточно подробно рассматривается во всех книгах по конечно-разностным методам. Может показаться, что центрирование временной разности с помощью приближенной замены производной $\partial T/\partial t$ разностью $(T_k^{n+1} - T_k^{n-1})/(2\Delta t)$ позволяет избежать неустойчивости. Однако это оказывается даже хуже, так как оно порождает неустойчивость для любого Δx . Причина заключается в том, что уравнение теплопроводности имеет первый порядок относительно времени. В то же время временные разности, взятые через два шага, образуют разностное уравнение, которое имеет второй порядок относительно времени. Уравнение второго порядка всегда имеет два решения, причем одно из них удовлетворяет уравнению теплопроводности, а второе решение имеет колебательный характер и экспоненциально возрастает как $(1, -2, 4, -8, \dots)$. Оно быстро разрушает решение уравнения теплопроводности.

Рассматриваемые трудности можно исключить с помощью схемы Кранка — Николсона, которая всегда гарантирует стабильность решения для любого Δx и может также применяться к волновым уравнениям для акустических, электромагнитных и упругих коле-

баний. По схеме Кранка — Николсона относят пространственную разность к точке $T_k^{n+1/2}$ следующим образом:

$$T_k^{n+1} - T_k^n = b(T_{k+1}^n - 2T_k^n + T_{k-1}^n) + b(T_{k+1}^{n+1} - 2T_{k+1}^n + T_{k-1}^{n+1}). \quad (10.1.5)$$

Очевидная задача в схеме Кранка — Николсона заключается в том, чтобы способ получения $(n+1)$ -го временного уровня из n -го уровня был не длиннее другого, известного. Переносим все члены с показателем $n+1$ в левую часть выражения (10.1.5), а члены с показателем n — в правую часть, получаем:

$$-bT_{k+1}^{n+1} + (1+2b)T_k^{n+1} - bT_{k-1}^{n+1} = D_k^n. \quad (10.1.6)$$

В правой части последнего выражения величина D_k^n является известной функцией от T_n . Выражение (10.1.6) представляет собой систему совместных уравнений для T^{n+1} . Записав ее в матричной форме, получим систему уравнений, называемую тридиагональной:

$$\begin{bmatrix} (1+2b) & -b & \text{Нули} \\ -b & (1+2b) & -b \\ & & (1+2b) \\ \text{Нули} & & -b \end{bmatrix} \begin{bmatrix} T_0 \\ T_1 \\ \vdots \\ T_N \end{bmatrix}^{n+1} = \begin{bmatrix} D_0 \\ D_1 \\ \vdots \\ D_N \end{bmatrix}. \quad (10.1.7)$$

Систему уравнений типа (10.1.7) можно очень легко решить. Ниже будет показано, что решение это только немного сложнее решения уравнения (10.1.4). Схему Кранка — Николсона можно рекомендовать для решения дифференциальных уравнений в частных производных численными методами, так как оно позволяет использовать любые Δx и Δt без потери устойчивости.

Теперь рассмотрим двумерную задачу распространения теплового потока. Уравнение теплопроводности в этом случае принимает вид

$$\frac{\partial T}{\partial t} = \frac{\sigma}{C} \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right). \quad (10.1.8)$$

Просто и эффективно решить это уравнение можно методом расщепления, в котором используются два различных уравнения для чередующихся временных шагов. Эти уравнения имеют вид

$$\frac{\partial T}{\partial t} = \frac{2\sigma}{C} \frac{\partial^2 T}{\partial x^2} \quad \text{для всех } y; \quad (10.1.9a)$$

$$\frac{\partial T}{\partial t} = \frac{2\sigma}{C} \frac{\partial^2 T}{\partial y^2} \quad \text{для всех } x, \quad (10.1.9б)$$

Каждое из уравнений (10.1.9a) и (10.1.9б) можно решить по схеме Кранка — Николсона. Существует много более известных методов решения рассматриваемой задачи, чем метод расщепления, но их погрешности при отбрасывании членов (асимптотическая разность между разностным и дифференциальным уравнениями) не стремятся к нулю сколько-нибудь быстрее, чем такая же погрешность для метода расщепления.

Попробуем сформулировать задачу для акустических колебаний по схеме Кранка — Николсона. Пусть величины u и w обозначают скорости смещения частиц в направлениях x и z , а величины P , ρ и K — соответственно давление, плотность и несжимаемость. Приравнявая ускорение к градиенту давления, получаем

$$\begin{aligned}\rho \frac{\partial u}{\partial t} &= -\frac{\partial P}{\partial x}; \\ \rho \frac{\partial w}{\partial t} &= -\frac{\partial P}{\partial z}.\end{aligned}$$

Кроме того, изменение давления в сочетании с дивергенцией скорости дает

$$\frac{\partial P}{\partial t} = -K \left(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right)$$

Переходя к матричной форме и полагая, что ∂_x обозначает $\partial/\partial x$ и т. д., имеем

$$\frac{\partial}{\partial t} \begin{bmatrix} P \\ U \\ W \end{bmatrix} = \begin{bmatrix} 0 & -K\partial_x & -K\partial_z \\ -\rho^{-1}\partial_x & 0 & 0 \\ -\rho^{-1}\partial_z & 0 & 0 \end{bmatrix} \begin{bmatrix} P \\ U \\ W \end{bmatrix}. \quad (10.1.10)$$

Приведение выражения (10.1.10) к схеме Кранка — Николсона аналогично приведению к этой схеме уравнения (10.1.8). Принципиальное различие между рассматриваемыми выражениями заключается в том, что в формуле (10.1.10) имеем векторы и матрицы, а в формуле (10.1.8) — только скаляры. Применяя метод расщепления к уравнению (10.1.10), получаем для чередующихся временных шагов:

$$\frac{\partial}{\partial t} \begin{bmatrix} P \\ U \end{bmatrix} = 2 \begin{bmatrix} 0 & -K\partial_x \\ -\rho^{-1}\partial_x & 0 \end{bmatrix} \begin{bmatrix} P \\ U \end{bmatrix} \quad (10.1.11a)$$

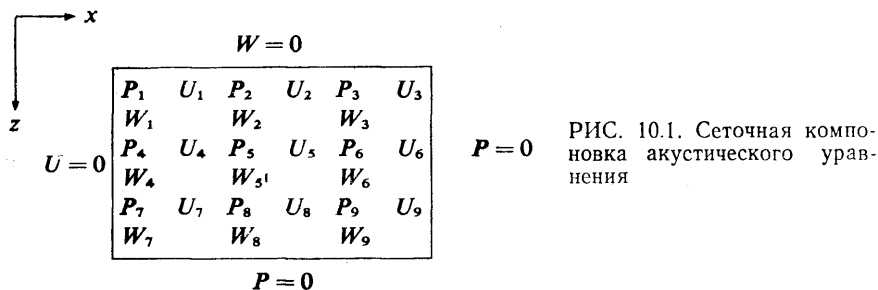
$$\frac{\partial}{\partial t} \begin{bmatrix} P \\ W \end{bmatrix} = 2 \begin{bmatrix} 0 & -K\partial_z \\ -\rho^{-1}\partial_z & 0 \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix}. \quad (10.1.11b)$$

При формулировании граничных условий для уравнения (10.1.11a) удобно определять P и U на чередующихся квадратах шахматной доски (см. рис. 10.1). Такое представление исключает необходимость задания производных ∂_x и ∂_z более чем на одном интервале.

Последний вопрос, имеющий большое практическое значение, связан с быстрым методом решения тридиагональной системы уравнений, подобной системе (10.1.6) или (10.1.7). Более общая система уравнений описывается выражением

$$A_k T_{k+1} + B_k T_k + C_k T_{k-1} = D_k. \quad (10.1.12)$$

Для уравнения теплопроводности все элементы в выражении (10.1.12) являются скалярными величинами. В других физических задачах величины A , B и C следует считать матрицами размерности 2×2 , величину T — вектором размерности 2×1 для каждого b на $(n+1)$ -м временном уровне, а величину D_k — вектором функции переменных поля с размерностью 2×1 , известной на n -м



временном уровне. Переходя к решению поставленной задачи, запишем другое уравнение с неизвестными пока величинами E_k , F_k и тем же решением T_k , что и для уравнения (10.1.12):

$$T_k = E_k T_{k+1} + F_k. \quad (10.1.13)$$

Перепишем уравнение (10.1.13) со смещенными на единицу индексами

$$T_{k-1} = E_{k-1} T_k + F_{k-1}. \quad (10.1.14)$$

Подставив найденное выражение в уравнение (10.1.12), получим

$$A_k T_{k+1} + B_k T_k + C_k (E_{k-1} T_k + F_{k-1}) = D_k. \quad (10.1.15)$$

Преобразуем уравнение (10.1.15) таким образом, чтобы привести его к виду, подобному выражению (10.1.13):

$$T_k = -(B_k + C_k E_{k-1})^{-1} A_k T_{k+1} + (B_k + C_k E_{k-1})^{-1} (D_k - C_k F_{k-1}). \quad (10.1.16)$$

Сопоставляя выражения (10.1.16) и (10.1.13), легко получить следующие рекурсивные выражения для определения величин E_k и F_k :

$$E_k = -(B_k + C_k E_{k-1})^{-1} A_k; \quad (10.1.17a)$$

$$F_k = (B_k + C_k E_{k-1})^{-1} (D_k - C_k F_{k-1}). \quad (10.1.17b)$$

При вычислениях на ЭВМ для упрощения решения задачи следует всегда стремиться использовать матрицы, содержащие нули, как, например, в уравнениях (10.1.11).

Перейдем к рассмотрению граничных условий. Предположим, что величина T_0 задана. Тогда равенство (10.1.13) удовлетворяется при условиях $E_0 = 0$, $F_0 = T$ и можно последовательно вычислить все значения E_k и F_k . Если задана величина T_N , то с помощью выражения (10.1.13) можно последовательно рассчитать величины T_{N-1} , T_{N-2} , ..., T_0 . Другим удобным видом граничных условий является задание отношений $r_1 = T_0/T_1$ и $r_2 = T_N/T_{N-1}$. Начнем

с выбора величин $E_0=r_1$, $F_0=0$ и вычисления значений E_k и F_k . Решив выражение для r_2 относительно величины T_N и используя равенство (10.1.14), найдем

$$T_{N-1} = E_{N-1}T_N + F_{N-1};$$

$$T_N/r_2 = E_{N-1}T_N + F_{N-1};$$

$$T_N = \left(\frac{1}{r_2} - E_{N-1}\right)^{-1} F_{N-1}.$$

Затем вычислим величины T_{N-1} , T_{N-2} так же, как прежде.

Как указывалось ранее, имеется множество деталей, связанных с численным решением дифференциальных уравнений в частных производных. В этой главе описаны только самые важные приемы

РИС. 10.2. Программа для решения тридиагональной системы уравнений

```

SUBROUTINE TRI(A,B,C,N,T,D,E,F)
DIMENSION T(N),D(N),F(N),E(N)
N1=N-1
E(1)=1.0
F(1)=0.
DO 10 I=2,N1
DEN=B+C*E(I-1)
E(I)=-A/DEN
10 F(I)=(D(I)-C*F(I-1))/DEN
T(N)=F(N1)/(1.0-E(N1))
DO 20 J=1,N1
I=N-J
20 T(I)=E(I)*T(I+1)+F(I)
RETURN
END

```

решения краевых задач. Программа для решения тридиагональной системы уравнений представлена на рис. 10.2. Предполагается, что величины A , B и C не зависят от k ; кроме того, используются граничные условия с нулевым наклоном.

Упражнения

1. Рассмотрите решение двумерного уравнения (10.1.8) по схеме Кранка — Николсона на сетке 4×4 . При этом система уравнений для неизвестной величины $T_{j,k}^{n+1}$ имеет размерность 16×16 . Найдите схему размещения нулей в матрице размерностью 16×16 . Учитывая трудности решения этой системы в явном виде, следует использовать метод расщепления.

2. Разностная аппроксимация дифференциального уравнения теплопроводности в частных производных имеет вид

$$P_j^{n+1} - P_j^n = \frac{a\Delta t (P_{j+1}^n - 2P_j^n + P_{j-1}^n)}{\Delta x^2} + s_j^n.$$

Используя решение $P_j = Q_n e^{ikj\Delta x}$, найденное методом подбора, приведите это уравнение к одномерному разностному уравнению. Запишите полученное уравнение с помощью Z -преобразования. Соответствует ли это уравнение затухающему процессу при любых действительных значениях a и при любых мнимых значениях a ? (Используйте разложение в ряд Фурье по s .)

3. Измените приведенную на рис. 10.2 программу для вычислительной машины таким образом, чтобы решить систему (10.1.7) без задания граничных условий с нулевым наклоном.

4. Напишите программу для решения уравнения (10.1.6) на вычислительной машине при $b=0,5$ с начальными условиями $T(1), \dots, T(20)=0,0$ и $T(21), \dots, T(30)=1,0$. Используйте подпрограмму пересылки (TR1).

10.2. ЭКСТРАПОЛЯЦИЯ ВОЛН В ОПТИКЕ¹

Геофизические измерения обычно выполняются вдоль линии профиля (оси x) на поверхности Земли. По данным этих наблюдений получают представление о свойствах Земли ниже ее поверхности. Первым шагом часто является экстраполяция данных наблюдений, полученных на поверхности Земли, в нижнее полупространство.

Прежде чем рассматривать численные методы продолжения волновых полей в нижнее полупространство, приведем краткий обзор методов, используемых в оптике для восстановления волновых фронтов в микроскопах и телескопах. Более полное объяснение можно найти в работе [35].

Предположим, что волновое возмущение $p(x, z, t)$ задано в точке z_0 в двумерной декартовой системе координат. Покажем, как оно экстраполируется (восстанавливается) ниже этой точки на оптической оси. При проектировании пучка световых лучей вдоль оптической оси возникают три достаточно общие задачи. Первая задача возникает при проектировании пучка лучей через апертуру или фотографический транспарант. Распределение интенсивности светового пучка является функцией прозрачности апертуры или транспаранта, которая изменяется от 0 до 1. Полагая, что оптическая ось совпадает с осью z , и ограничиваясь двумерной задачей, рассмотрим поглощающий свет транспарант $T(x)$, расположенный в точке $z_0 + dz/2$. Тогда

$$p(t, x, z_0 + dz) = T(x)p(t, x, z_0). \quad (10.2.1)$$

Вторая задача возникает при проектировании светового пучка через линзу, которая обычно приближенно рассматривается как «тонкая линза». При этом необходимо определить дифференциальную функцию задержки $\tau(x)$, которая описывает временную задержку световых лучей при прохождении их через линзу из точек, лежащих на оси x , параллельно оси z . Если линза расположена в точке $z_0 + dz/2$, то свертка волнового поля с задержанным дельта-импульсом имеет вид

$$\begin{aligned} p(t, x, z_0 + dz) &= \int p(t-s, x, z_0) \delta[s - \tau(x)] ds = \\ &= p[t - \tau(x), x, z_0]. \end{aligned} \quad (10.2.2a)$$

Рассматриваемый временной сдвиг просто выражается в частотной области, где свертка (10.2.2a) заменяется произведением.

¹ В отечественной литературе обычно используются термины: «продолжение», «аналитическое продолжение», «восстановление» (*Прим. перев.*)

Тогда

$$P(\omega, x, z_0 + dz) = P(\omega, x, z_0) e^{i\omega z(x)}. \quad (10.2.26)$$

Третья общая задача в оптике возникает при распространении волн через область свободного пространства. Удивительно, что это наиболее трудная из трех задач. Вспомним волновое уравнение

$$\left(\partial_{xx} + \partial_{zz} - \frac{1}{v^2} \partial_{tt}\right) p(t, x, z) = 0. \quad (10.2.3)$$

Принимая скорость v постоянной во времени и в пространстве, можно с помощью найденного подбором решения $p(t, x, z) = P(\omega, k_x, z) e^{-i\omega t + ik_x x}$ свести уравнение (10.2.3) к обыкновенному дифференциальному уравнению:

$$\frac{d}{dz^2} P = \left(-\frac{\omega^2}{v^2} + k_x^2\right) P. \quad (10.2.4)$$

Последнее уравнение имеет два решения $e^{ik_z z}$ и $e^{-ik_z z}$, где

$$k_z = \left(\frac{\omega^2}{v^2} - k_x^2\right)^{1/2}. \quad (10.2.5)$$

Одним из этих решений является волна, распространяющаяся вниз вдоль оси z , а другим — волна, распространяющаяся вверх вдоль той же оси. Начальные условия (и приближение, предполагающее отсутствие обратного рассеяния от линз и апертур) позволяют отбросить одно из этих решений, оставив следующее:

$$P(\omega, k_x, z) = P(\omega, k_x, z_0) e^{ik_z(z-z_0)} = P(\omega, k_x, z_0) e^{i(\omega^2/v^2 - k_x^2)^{1/2}(z-z_0)}. \quad (10.2.6)$$

Правая часть последнего выражения является произведением двух функций, зависящих от k_x . Она является также произведением двух функций частоты ω . Это означает, что с помощью стандартного преобразования Фурье можно представить выражение (10.2.6) в виде свертки во временной области, либо в области пространства x , либо сразу в обеих областях. Преобразовывая функцию передачи «фильтра»

$$\exp i \left(\frac{\omega^2}{v^2} - k_x^2\right)^{1/2} (z - z_0) \quad (10.2.7)$$

в пространственную область, получим «импульсную реакцию (отклик)», которая в рассматриваемом случае имеет физический смысл волнового поля, проходящего через точечную апертуру. Лучи, выходящие из точечного отверстия, ведут себя отчасти подобно лучу от точечного источника. На рис. 10.3 приведен энергетический спектр изотропного распределения лучей для точечного

источника. Как видно, в окрестности углов $\theta = \pm 90^\circ$ отмечается сгущение лучей при $k_x = \pm \omega/v$. Энергия, как функция параметра k_x , пропорциональна величине $d\theta / (dk_x/d\theta) = d\theta / (d \sin \theta / d\theta) = d\theta / \cos \theta = [1 - (vk_x/\omega)^2]^{-1/2} d\theta$.

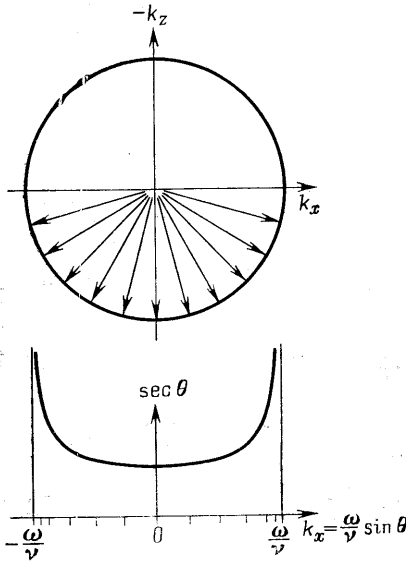


РИС. 10.3. Энергетический спектр изотропного распределения лучей точечного источника как функция волнового числа $k_x = \frac{\omega}{v} \sin \theta$

Этот результат может быть сопоставлен с функцией передачи (10.2.7). Различие между этими двумя случаями заключается в том, что функция передачи (10.2.7) имеет единичную амплитуду, не зависящую от величины k в интервале $-\omega/v < k < +\omega/v$, тогда как спектральная амплитуда точечного источника в соответствии с рис. 10.3 имеет меньшую величину в окрестности точки $k_x = 0$ и неограниченно возрастает при $k_x = \pm \omega/v$. Это означает, что апертурная (аппаратная) функция не обладает сферической изотропией, как у точечного источника, а имеет большую энергию в области $k_x = 0$, т. е. дает направленное излучение вдоль оси z . По-видимому, не существует простой аналитической процедуры для Фурье-преобразования функции (10.2.7) во временной и пространственной об-

ластях. Один из моих коллег, Филипп Шульц, осуществил численное преобразование Фурье для ряда функций передач, изображенные действительные части которых показаны на рис. 10.4—10.6. Дискретность выборки данных, подвергнутых преобразованию

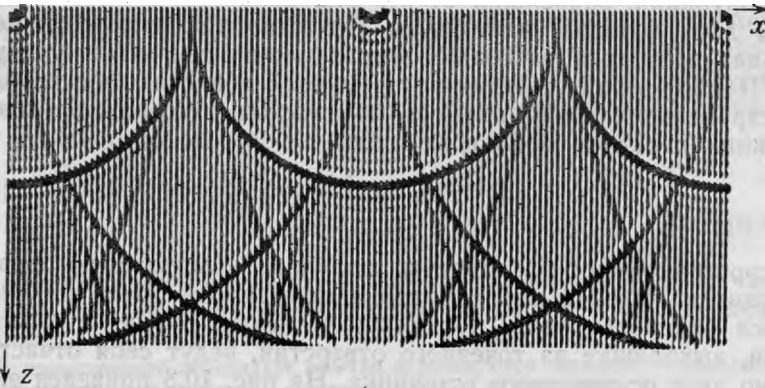


РИС. 10.4. Изображение функции передачи волнового уравнения

Фурье, вызывает периодичность результатов по всем координатам преобразования. Указанная периодичность является совершенно очевидной и хорошо видна на всех рисунках. На рис. 10.4 представлено изображение в плоскости (x, z) для фиксированного момента времени t решения функции $\exp[i(\omega^2/v^2 - k_x^2)^{1/2}z]$ в виде

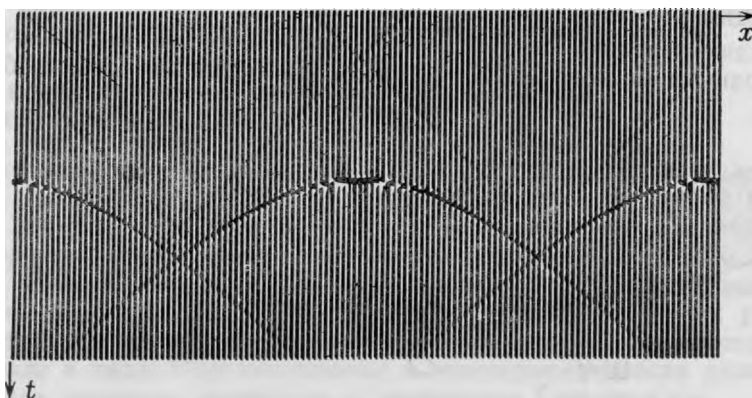


РИС. 10.5. Изображения функции передачи волнового уравнения на сейсмическом профиле

двойного ряда Фурье по параметрам k_x и ω . Фронты волн имеют вид полуокружностей, причем в области, где волны распространяются вдоль оси z , амплитуды их максимальны. На рис. 10.5 представлено изображение для функции передачи (10.2.7), полученное в виде двойного ряда Фурье по параметрам k_x и ω . Изображение показано в плоскости (x, t) при фиксированном значении глубины z_0 . Времена вступлений волны, измеренные на различных расстояниях от точечного отверстия, в начале координат на экране (x, z_0) образуют гиперболу. Лучевая теория легко

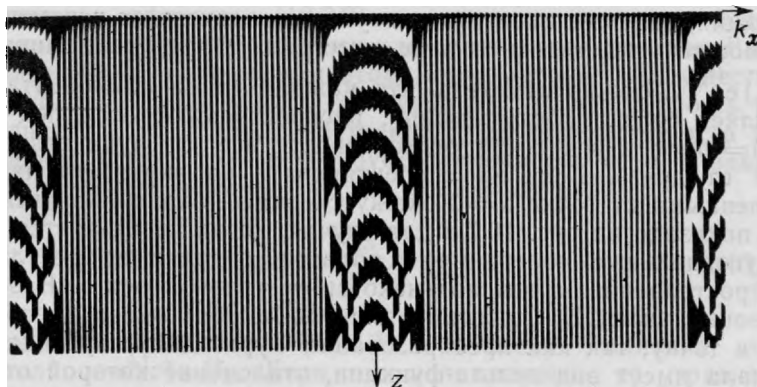


РИС. 10.6. Действительная часть точной функции передачи (10.2.7), построенная в координатах k_x и z при постоянной частоте ω

объясняет эту зависимость для времени распространения волны. Однако медленное затухание амплитуды волны вдоль гиперболы в зависимости от угла выхода луча, связанное с явлением дифракции, рассчитать аналитически не так просто, особенно на больших удалениях от оси. Действие функции наклона луча не следует путать с влиянием помех, возникающих при изображении дельта-функции на сетке. На рис. 10.6 показана действительная часть функции передачи $\exp[i\sqrt{\omega^2/v^2 - k_x^2}z]$ в координатах k_x и z при постоянной частоте ω . Как видно, эта функция скачкообразно

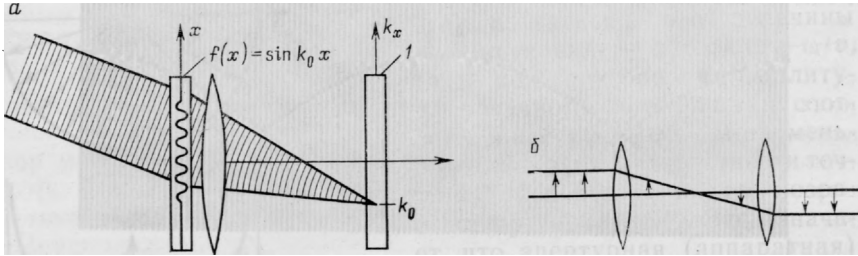


РИС. 10.7. Реализация преобразования Фурье с помощью линзы (а) и двух последовательных линз, расположенных на двойном фокусном расстоянии друг от друга (б).

f — Фурье-преобразование $f(x) = \delta(k - k_0)$

изменяется при условии $\omega^2/v^2 = k_x^2$, т. е. на границе между областями распространения и исчезновения волны.

Хорошо известно, что линзу можно использовать для получения преобразования Фурье. Фактически преобразование Фурье осуществимо лишь на бесконечности. Линза же служит для того, чтобы перенести изображение спектра из бесконечности в удобный для нас интервал расстояний. Предположим, что монохроматическое оптическое возмущение $P(x, z, \omega)$ наблюдается в точке z_0 . Функция, описывающая это возмущение, зависит от переменной x . Ее можно представить в виде интеграла Фурье по компонентам $A(k_x) e^{ik_x x}$. Важно отметить, что каждая такая компонента представляет собой плоскую волну, распространяющуюся под углом $\sin \theta = vk_x/\omega$ к оси z . На бесконечности все эти лучи отделяются друг от друга. Если их спроектировать на экран, то наиболее удаленными от точки $x=0$ окажутся самые большие значения k_x . Это показано на рис. 10.7,а для случая, когда используется линза. Синоидальные колебания в области x являются результатом распространения волны под некоторым углом к оси z . В области волновых чисел k_x (пространственных частот) линза преобразует луч в точку, так как преобразование Фурье для синусоидального сигнала имеет вид дельта-функции, отклонение которой от оптической оси пропорционально частоте колебаний синусоидального сигнала. Две последовательно установленные на двойном фокус-

ном расстоянии друг от друга линзы переворачивают изображение (рис. 10.7,б). Это является физическим проявлением того математического факта, что преобразование Фурье не является обратным к самому себе преобразованием. Обратное преобразование имеет противоположный знак в показателе экспоненты. Можно легко проверить, что произведя преобразование Фурье дважды, можно точно восстановить первоначальную форму волны.

10.3. ЭКСТРАПОЛЯЦИЯ МОНОХРОМАТИЧЕСКИХ ВОЛН ЧИСЛЕННЫМИ МЕТОДАМИ

Оптический метод экстраполяции волн неприменим для сред, скорость распространения волн в которых является функцией координат пространства, т. е. $v = v(x, z)$. В этом случае комплексная экспоненциальная функция не является решением волнового уравнения. Поэтому попытаемся найти численную процедуру экстраполяции волновых полей, не зависящую от аналитических решений или любого частного распределения скоростей распространения волн в среде. Предположение о монохроматичности решения приводит волновое уравнение к уравнению Гельмгольца:

$$P_{xx} + P_{zz} = \frac{-\omega^2}{v^2(x, z)} P. \quad (10.3.1)$$

Теперь рассмотрим возможность использования уравнения (10.3.1) для экстраполяции возмущения $P(x, z_0)$ в направлении оси z . Предположим, что значения функции P на уровне z_0 заданы для всех x . Тогда можно найти величину P_{zz} с помощью перестановки членов в уравнении (10.3.1):

$$P_{zz} = -\frac{\omega^2}{v^2} P - P_{xx}. \quad (10.3.2)$$

При заданных величинах P , P_z для всех x и некотором определенном значении z формула (10.3.2) теоретически позволяет использовать конечно-разностную схему для определения величин P и P_z на уровне $z + \Delta z$. Однако при этом возникает весьма существенная трудность. Чтобы понять ее, предположим, что скорость распространения волн в среде постоянна и не зависит от x , а преобразование Фурье трансформирует зависимость от x в зависимость от k_x . Тогда формула (10.3.2) принимает вид

$$P_{zz} = \left(-\frac{\omega^2}{v^2} + k_x^2 \right) P. \quad (10.3.3)$$

На решение уравнения (10.3.3) оказывает влияние знак множителя $-\omega^2/v^2 + k_x^2$. Если он положителен, то уравнение Гельмгольца будет иметь возрастающие и затухающие экспоненциальные решения. При отрицательном знаке решение уравнения (10.3.3) имеет синусоидальный характер. Возрастающие экспонен-

циальные решения как раз и составляют проблему. Однако их можно удержать вне поля зрения, если задать возрастающую экспоненциальную функцию с нулевой начальной амплитудой. Это можно сделать, задав определенное соотношение между величинами P и P_z . Действительно, в геофизике обычно измеряют каким-либо способом только величину P . Поэтому можно попытаться определить теоретически величину P_z по значению величины P таким образом, чтобы исключить возрастающее по экспоненте решение. Кроме того, в оптике экстраполяция возмущения $P(x, z_1)$ по $P(x, z_2)$ не требует знания производной функции $P_z(x, z_1)$. Волновое уравнение является уравнением второго порядка по z и, следовательно, имеет два решения (для восходящих и нисходящих волн). Таким образом, необходимы два граничных условия. В обычных краевых задачах физики требуется найти решение в средней области между точками z_0 и z_N , а соответствующие граничные условия должны задавать значения P в точках z_0 и z_N . Каким образом в оптике удастся избежать необходимости знать либо P_z в точке z_0 , либо P в точке z_N ? Это достигается тем, что одно из двух решений отбрасывается при определении величины k_z за счет выбора только одного из двух возможных квадратных корней. Поскольку одно решение было отброшено, то достаточно только одного граничного условия вместо двух. Отбрасывание одного из решений равносильно введению определенного предположения физического характера, которое может оказаться как справедливым, так и ошибочным. Справедливость этого предположения всегда определяется конкретными условиями задачи распространения волн в среде. Наша цель — преобразование уравнения (10.3.2) к виду, позволяющему использовать обычное для оптики предположение о том, что в среде распространяются волны, компоненты которых направлены только вдоль оси $+z$, не вводя при этом обычного в оптике предположения об однородности среды. Вместо уравнения второго порядка (10.3.2), описывающего волны, распространяющиеся как в положительном, так и в отрицательном направлении вдоль оси z , хотелось бы иметь уравнение первого порядка, описывающее волны, распространяющиеся только в направлении оси $+z$. Поскольку величину P_z мы экспериментально не наблюдаем, то ценным свойством такого уравнения первого порядка было бы требование в качестве начального условия только одной величины $P(x)$, а не двух величин P и P_z . В геофизических задачах часто можно использовать предположение о наличии только нисходящей волны (например, при описании волновых полей, излучаемых с помощью специального оборудования). С другой стороны, предположение о наличии только восходящей волны часто используется для описания последующих наблюдений. Естественно, что применимость этих предположений должна проверяться в каждом конкретном случае. Для плоской волны, распространяющейся в положительном направлении оси z , можно записать

$$P(x, z) = Q_0 e^{i(\omega/v)z},$$

где Q_0 — неизвестная постоянная, т. е. волна имеет неизвестные амплитуду и фазу. Поэтому

$$P(x, z) = Q(x, z) e^{i(\omega/v)z}.$$

В этой формулировке утверждение, что величина $Q(x, z)$ — приблизительно постоянная функция от x и z — является довольно неопределенной и требует уточнения. Ограничивая выбор величины $Q(x, z)$ медленно изменяющимися функциями, мы тем самым ограничим решение волнового уравнения $P(x, z)$ почти плоскими волнами, распространяющимися в направлении оси z . В действительности решение P может представлять собой плоскую волну, распространяющуюся под небольшим углом к оси z , а также малый участок фронта сферической волны. Это решение может описывать наблюдаемые отраженные волны или после поворота системы координат на 90° поверхностные волны.

Отношение ω/v называется пространственной частотой волны:

$$m = \frac{\omega}{v(x, z)}. \quad (10.3.4)$$

Пространственное среднее значение \bar{m}

$$\bar{m} = \frac{\omega}{v}. \quad (10.3.5)$$

В однородной среде величина \bar{m} равна пространственной частоте m . Тогда волновое возмущение имеет вид

$$P(x, z) = Q(x, z) e^{i\bar{m}z}. \quad (10.3.6)$$

Дополнительное условие медленности изменения величины $Q(x, z)$ в зависимости от z заключается в том, чтобы функция $m(x, z)$ была относительно близка к значению \bar{m} . Вычислив некоторые частные производные выражения (10.3.6), получим

$$P_x = Q_x e^{i\bar{m}z}; \quad (10.3.7a)$$

$$P_{xx} = Q_{xx} e^{i\bar{m}z}; \quad (10.3.7b)$$

$$P_z = (Q_z + i\bar{m}Q) e^{i\bar{m}z}; \quad (10.3.7в)$$

$$P_{zz} = (Q_{zz} + 2i\bar{m}Q_z - m^2Q) e^{i\bar{m}z}. \quad (10.3.7г)$$

Подставив выражения (10.3.7б) и (10.3.7г) в уравнение (10.3.1) и сократив экспоненциальный множитель, найдем:

$$Q_{xx} + Q_{zz} + 2i\bar{m}Q_z + (m^2 - \bar{m}^2)Q = 0. \quad (10.3.8)$$

Покажем, что во многих случаях функция Q изменяется настолько медленно, что в уравнении (10.3.8) можно пренебречь членом Q_{zz} по сравнению с величиной $2i\bar{m}Q_z$. При этом получаем так называемое параболическое или параксиальное (приосевое) приближение. Тогда искомое уравнение первого порядка по z с одним начальным значением имеет вид

$$Q_{xx} + 2i\bar{m}Q_z + (m^2 - \bar{m}^2)Q = 0. \quad (10.3.9)$$

Для однородной среды уравнение (10.3.9) приводится к виду

$$Q_{xx} + 2i\bar{m}Q_z = 0. \quad (10.3.10)$$

Как видно, уравнение (10.3.10) имеет ту же самую форму, что и уравнение теплопроводности, если z ассоциируется со временем, а удельная теплопроводность считается мнимой величиной. Это уравнение известно так же, как уравнение Шредингера. Его мож-

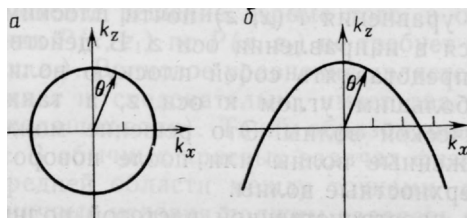


РИС. 10.8. Графики допустимых волновых чисел для волнового уравнения (а) и одностороннего волнового уравнения (б)

но решить численным методом, описанным для уравнения теплопроводности в разделе 10.1. Уравнение (10.3.10) можно применять для различных целей. Посмотрим, что мы потеряли, отбросив член Q_{zz} . Чтобы облегчить сравнение уравнения (10.3.10) с волновым уравнением, перейдем вновь от переменной Q к переменной P . Определим функцию Q и ее производные из уравнений (10.3.6) и (10.3.7)

$$Q = Pe^{-i\bar{m}z}; \quad (10.3.11a)$$

$$Q_{xx} = P_{xx}e^{-i\bar{m}z}; \quad (10.3.11б)$$

$$Q_z = (P_z - i\bar{m}P)e^{-i\bar{m}z}. \quad (10.3.11в)$$

Подставив выражения (10.3.11б) и (10.3.11в) в уравнение (10.3.10) и сократив экспоненту, получим так называемое одностороннее волновое уравнение¹:

$$P_{xx} + 2i\bar{m}(P_z - i\bar{m}P) = 0;$$

$$P_{xx} + 2i\bar{m}P_z - 2i\bar{m}^2P = 0. \quad (10.3.12)$$

Одним из способов решения любого дифференциального уравнения в частных производных с постоянными коэффициентами в декартовой системе координат, является введение комплексной экспоненты вида $e^{(ik_x x + ik_z z)}$. Если волновые числа k_x и k_z оказываются действительными, то это пробное решение можно интерпретировать как плоскую волну, распространяющуюся в направлении $k = (k_x, k_z)$. Подставляя рассматриваемую экспоненциальную функцию в уравнения (10.3.1) и (10.3.12) и сокращая

¹ Это уравнение является также параболическим. (Прим. перев.)

экспоненты, получим два алгебраических уравнения, называемых дисперсионными соотношениями:

$$-k_x^2 - k_z^2 + m^2 = 0; \quad (10.3.13)$$

$$-k_x^2 - 2\bar{m}k_z + 2\bar{m}^2 = 0. \quad (10.3.14)$$

Графики для этих двух уравнений приведены на рис. 10.8. Для волнового уравнения график является окружностью; это гово-

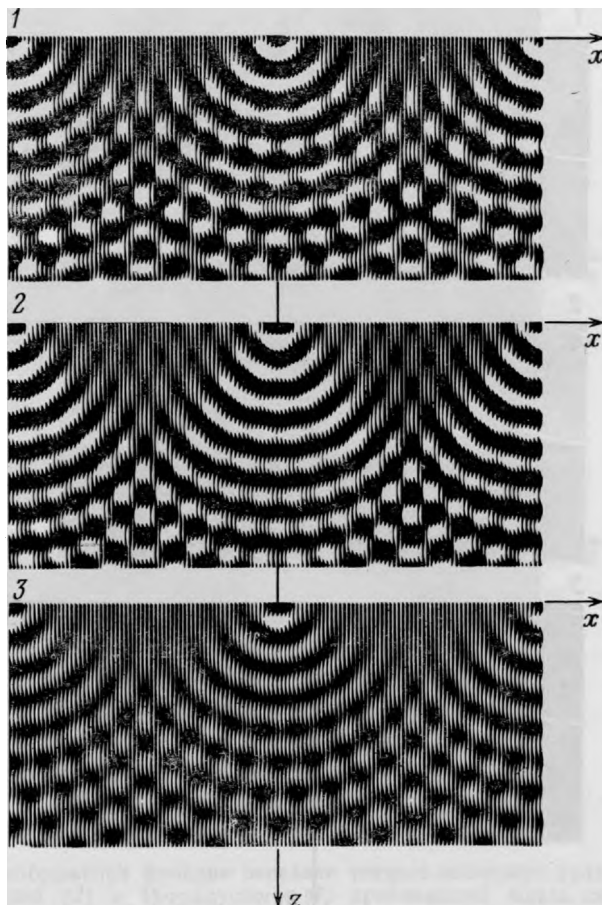


РИС. 10.9. Изображения функции передачи для точного (1) волнового уравнения и его 15-градусного (2) и 45-градусного (3) приближений (для монохроматической волны)

рит о том, что волновое число в произвольно выбранном направлении, т. е. $k = (k_x^2 + k_z^2)^{1/2}$ равно постоянной ω/v . Однако для одностороннего волнового уравнения это неверно. В этом случае имеем лишь приближенное равенство $k_x^2 + k_z^2 \approx \omega^2/v^2$ при малых углах θ . Рис. 10.8,а иллюстрирует тот факт, что уравнение

(10.3.14) есть начальное условие относительно z , так как каждому значению величины k_x отвечают два значения величины k_z . В то же время в соответствии с рис. 10.8,б имеется только одно значение величины k_z для каждого значения k_x .

На рис. 10.9 — 10.12 показаны результаты преобразования приближенной функцией $e^{ik_z z}$ с помощью волнового уравнения функции

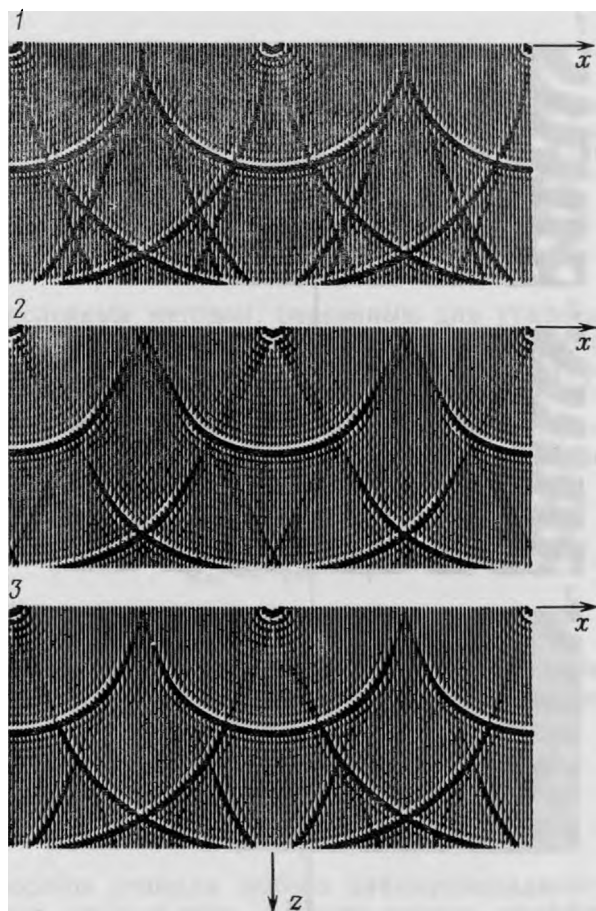


РИС. 10.10. Изображения функции передачи точного волнового уравнения (1) и его 15-градусного (2) и 45-градусного (3) приближений

$e^{ik_z z}$ и их преобразования Фурье. На рис. 10.9 на кадре 1 в плоскости (x, z) изображен результат преобразования представленной рядом Фурье по параметру k_x функции передачи точного волнового уравнения $\exp[i(1 - k_x^2 v^2 / \omega^2)^{1/2} \omega z / v]$ при фиксированной частоте ω_0 . На кадре 2 изображено то же самое для функции передачи 15-градусного приближения $\exp[i(1 - k_x^2 v^2 / 2\omega^2) \omega z / v]$. На нижнем кадре —

то же самое для функции передачи 45-градусного приближения $\exp \left[i \frac{\omega}{v} \frac{4\omega^2 - 3k_x^2 v^2}{4\omega^2 - k_x^2 v^2} z \right]$ из упражнения 2 этого раздела. Физически картина представляет собой результат прохождения волн через маленькие отверстия, размещенные периодически вдоль оси x . На

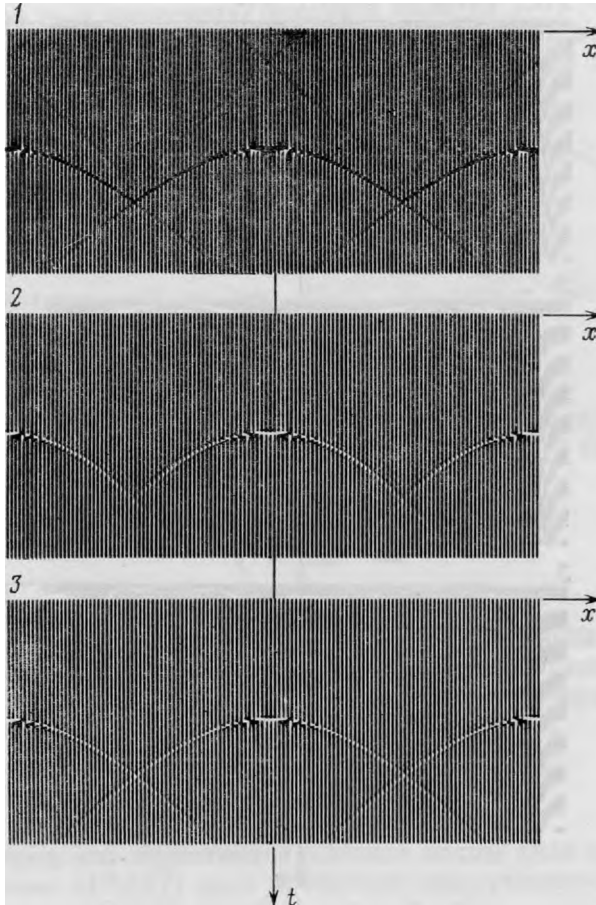


РИС. 10.11. Изображения функции передачи точного волнового уравнения (1) и его 15-градусного (2) и 45-градусного (3) приближений вдоль сейсмического профиля

рис. 10.10 приведены результаты преобразования тех же функций передачи волнового уравнения, что и на рис. 10.9, представленных двойным рядом Фурье по параметрам k_x и ω в плоскости (x, t) для фиксированного момента времени (для волнового уравнения с зависящими от времени коэффициентами).

На рис. 10.11 приведены результаты преобразования представленных двойным рядом Фурье по параметрам k_x и ω функций

передачи для различных приближений волнового уравнения в плоскости сейсмического профиля (x, t) при фиксированном z . Точное решение волнового уравнения представляет собой дельта-функцию, заданную вдоль гиперболы; 15-градусное приближение является параболой. Как видно, приближенные решения ухуд-

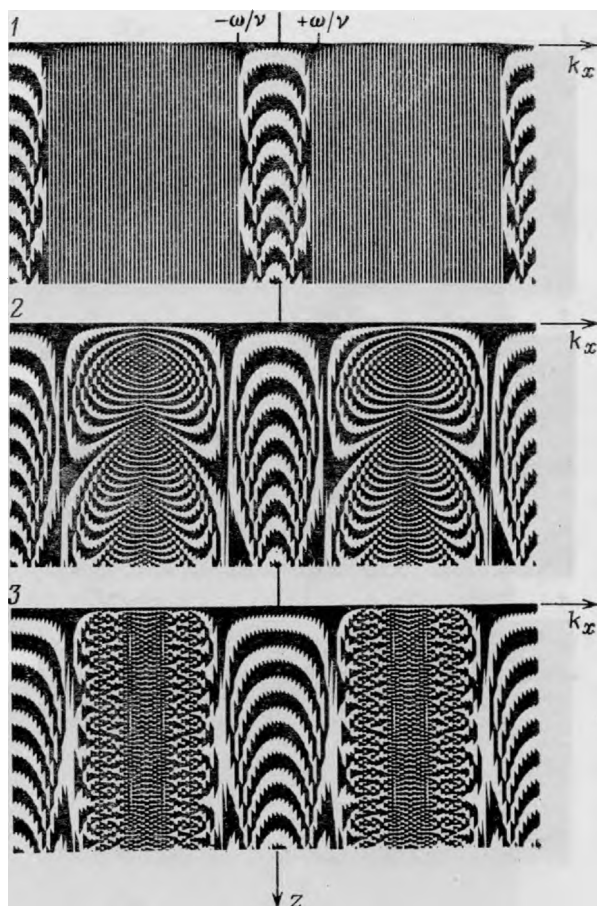


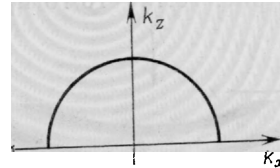
РИС. 10.12. Изображения функции передачи точного волнового уравнения (1) и его 15-градусного (2) и 45-градусного (3) приближений для монохроматических колебаний в плоскости (k_x, z) . Показана только действительная часть [36]

шаются с увеличением угла быстрее точного решения. На рис. 10.12 показаны действительные части функции передачи различных приближений волнового уравнения для монохроматических колебаний в плоскости (k_x, z) . Точная функция передачи изменяется скачкообразно вплоть до полного исчезновения сигнала при $|k_x v / \omega| = \sin 90^\circ = 1$. Экспоненциальное затухание для $k_x > \omega / v$ заметно вблизи $z=0$; 15-градусное и 45-градусное приближения являются фазовыми фильтрами и они восстанавливают

бесконечно малые зоны интересующей нас картины. Чтобы устранить помехи в виде коротких горизонтальных размытых волн в пространственной области на двух предыдущих рисунках, эти бесконечно малые зоны были сдвинуты с помощью ступенчатой функции. Из приведенных результатов следует возможность заметного улучшения приближенных функций передачи, в некоторых случаях за счет применения веерного фильтра [36].

Хорошие результаты дает одностороннее волновое уравнение, обладающее дисперсионным соотношением в виде полуокружно-

РИС. 10.13. Дисперсионное соотношение для идеального одностороннего волнового уравнения является полуокружностью



сти, представленной на рис. 10.13. Уравнение указанной полуокружности имеет вид

$$k_z = \sqrt{m^2 - k_x^2}. \quad (10.3.15)$$

Выражение (10.3.15) является основным соотношением, используемым для экстраполяции волн в оптике. Используя биномиальное разложение, запишем для выражения (10.3.15):

$$k_z = m \left(1 - \frac{k_x^2}{2m^2} - \frac{k_x^4}{8m^4} - \dots \right). \quad (10.3.16)$$

Это выражение сходится для всех $0 < k_x < m$.

Теперь интуитивно можно записать дифференциальное уравнение в частных производных с дисперсионным соотношением в виде указанной полуокружности. На основании выражения (10.3.16) запишем

$$\partial_z P = im \left(1 + \frac{\partial_{xx}}{2m^2} - \frac{\partial_{xxxx}}{8m^4} + \dots \right) P. \quad (10.3.17)$$

Естественно, что подстановка плоской волны вида $e^{i(k_x x + ik_z z)}$ в уравнение (10.3.17) дает требуемое дисперсионное соотношение (10.3.16), имеющее вид полуокружности. Очевидно, что чем выше требуемая угловая точность, тем больше членов уравнения (10.3.17) требуется для вычислений. Для сокращения записи уравнение (10.3.17) можно представить следующим образом:

$$\partial_z P = i(m^2 + \partial_{xx})^{1/2} P. \quad (10.3.18)$$

Это ничего не дает, но, как оказывается, уравнение (10.3.18) представляет собой релятивистское уравнение Шредингера.

Из выражения (10.3.18) легко получить волновое уравнение, дифференцируя его по z :

$$\partial_{zz} P = i \partial_z (m^2 + \partial_{xx})^{1/2} P.$$

Полагая пространственную частоту m не зависящей от z , можно изменить порядок дифференцирования в последнем выражении. Тогда

$$\partial_{zz}P = i(m^2 + \partial_{xx})^{1/2} \partial_z P.$$

Подставляя сюда выражение (10.3.18), найдем соотношение:

$$\partial_{zz}P = -(m^2 + \partial_{xx})^{1/2} (m^2 + \partial_{xx})^{1/2} P = -(m^2 + \partial_{xx}) P,$$

являющееся волновым уравнением.

На рис. 10.14—10.16 приведены решения в конечных разностях

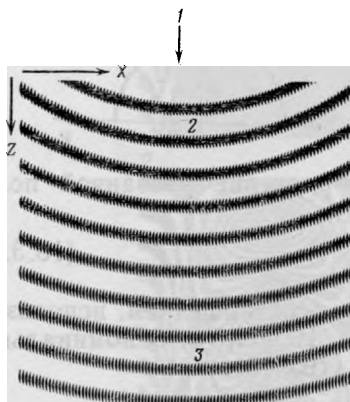


РИС. 10.14. Расширяющаяся монохроматическая цилиндрическая волна [3] (по оси x 100 точек, по оси z 30 точек).

1 — источник; 2 — сильная волна; 3 — слабая волна

для волнового уравнения в параболическом приближении для однородной среды. На рис. 10.14 и 10.15 наблюдаются расширяющиеся в плоскости (x, z) монохроматические цилиндрические

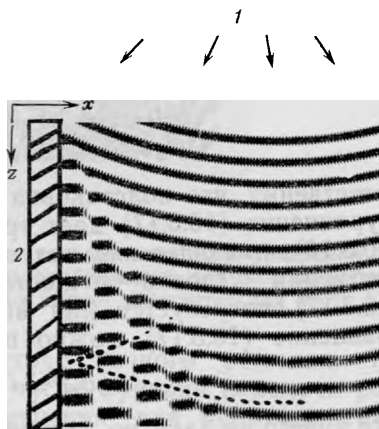


РИС. 10.15. Та же модель, что на рис. 10.14, но левой границей является недеформируемая стенка.

1 — источник; 2 — жесткая стенка

волны, фронты которых являются концентрическими окружностями с убывающей по z амплитудой. Расчет волнового поля начинается разностным методом решения для верхней части изображения на линейной сетке из 100 точек. Затем, используя разност-

ные уравнения, мы перемещаемся по сетке вниз; 30 шагов охватывают все изображение. На каждую точку требуется около шести комплексных перемножений, что занимает около 5 с машинного времени на нашей ЭВМ. Для вывода изображения использован многоканальный плоттер [3]. Рис. 10.15 отличается от рис. 10.14 тем, что слева помещена недеформируемая стена, которая отражает волны обратно в среду. Фронт отраженной волны показан с помощью более короткой волны из двух пунктирных линий [3].

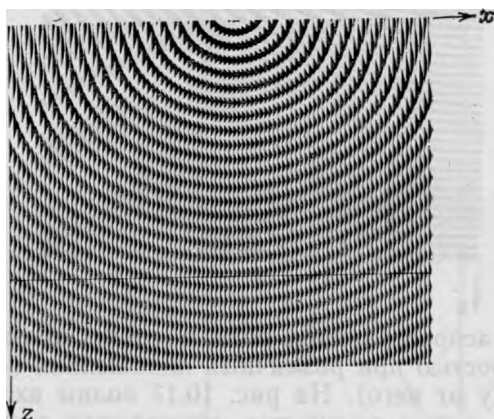


РИС. 10.16. Расходящаяся цилиндрическая волна (экстраполяция) [5]

На рис. 10.16 показан результат экстраполяции в нижнее полупространство с помощью теоретического решения уравнений из упражнения 2 этого раздела, использованного в качестве граничного условия. Как видно, фронты волн являются недостаточно круговыми, что препятствует использованию уравнения (10.3.18). Отметим также, что на изображении не обнаруживается затухание амплитуды (пропорционально $r^{-1/2}$) для волн теоретических, распространяющихся под углом около 60° к вертикали. Такие волны затухают медленнее, так как кривая фазы для угла в 60° более пологая, чем окружность [5].

Обратимся теперь к вопросу об использовании параболического приближения при наличии пространственных изменений скорости распространения волн в среде. Некоторые важные детали этого вопроса рассматриваются в упражнениях, но можно легко сделать некоторые усовершенствования уравнения (10.3.9). Основная идея заключается в том, чтобы аппроксимировать окружность параболой. При этом фактический радиус окружности не должен изменяться. Это приводит к предположению о том, что выражение (10.3.12) или (10.3.14) может быть использовано с заменой величины \bar{m} на m , как в уравнении (10.3.17). Тогда уравнение (10.3.12) будет иметь вид

$$P_{xx} + 2im(x, z)P_z + 2m^2(x, z)P = 0. \quad (10.3.19)$$

Уравнение (10.3.19) не требует введения предположения о том, что $m \approx \bar{m}$, ввиду чего можно рассматривать задачи с изменяющимися

в широком диапазоне скоростями. На самом деле, как будет показано в упражнениях, справедливость уравнения (10.3.19) также зависит от приближения, учитывающего насколько малы изменения скорости волн в среде на интервале, равном длине волны.

Рис. 10.17—10.19 иллюстрируют распространение волн в неоднородных средах. На рис. 10.17 и 10.18 показаны процессы

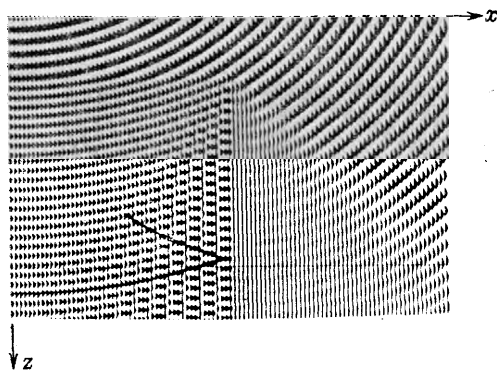


РИС. 10.17. Набегание волны сверху на погруженный в жидкость блок из вещества с более низкой чем в жидкости скоростью [5]

распространения волн в среде, включающей блок с низкой скоростью при различных положениях источника (над блоком и сбоку от него). На рис. 10.17 волны входят в верхнюю часть блока, а затем полностью отражаются внутрь блока от его боковой стенки (явление полного внутреннего отражения). Это дает тень

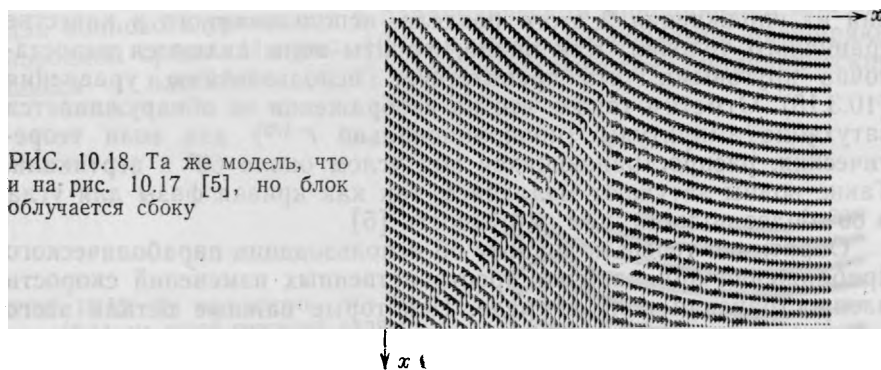


РИС. 10.18. Та же модель, что и на рис. 10.17 [5], но блок облучается сбоку

на внешней стороне блока [5]. При облучении того же блока сбоку (рис. 10.18) наблюдается частичное отражение волн от боковой стенки блока и интерференция между волнами, входящими в блок через различные грани [5]. На рис. 10.19,а показано поле плоских волн, падающих сверху на призму, а на рис. 10.19,б — схема их распространения в модели. Внутри призмы видно укорочение длины волны. При выходе из призмы волны отклоняются вправо, а их амплитуда увеличивается за счет сжатия пучка. Справа внизу они интерферируют с волнами, распространяющи-

мися вдоль боковой грани призмы. Это приводит к амплитудной модуляции волн. Изогнутые волновые фронты — следствие дифракции волн на углах призмы. Особенно интересна дифракция от верхнего угла призмы, которая четко выделяется в виде подобной углу фигуры около правой грани призмы. Энергия для этой дифракции переносится волной, распространяющейся вдоль пра-

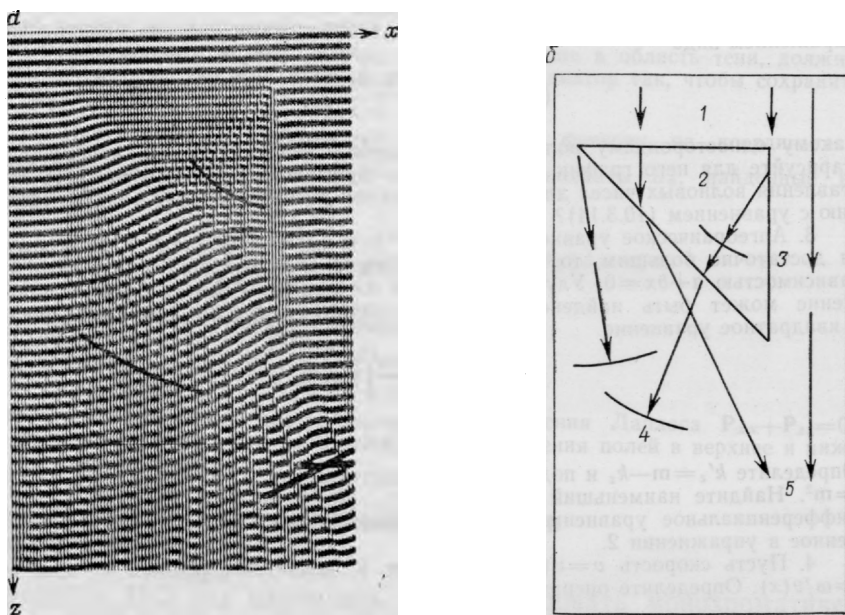


РИС. 10.19. Плоские волны, падающие сверху на правильную 45-градусную призму (а), и схема их распространения (б) [5].

1 — падающая плоская волна; 2 — призма; 3 — затухание энергии; 4 — дифракция от угла; 5 — интерференция между классическими лучами

вого вертикального края призмы. Расчет этого поля требует 10с машинного времени на ЭВМ IBM 360-67, установленной в Стенфорде [5].

Очевидно, что рассматриваемое приближение лучше на высоких частотах (при коротких длинах волн). Это приближение хорошо известно в волновой теории. Иногда оно называется приближением геометрической оптики. Однако читатель не должен думать, что рассматриваемая волновая теория сводится к геометрической оптике. В действительности все явления физической оптики (например, интерференция, дифракция и фокусировка волн в области конечного размера) полностью сохраняются. Практически при этом не накладываются ограничения на физическую оптику. В некоторых упражнениях ниже приведены примеры, включающие градиенты скорости, выраженные с помощью членов низких частот. Вопрос о значимости каждого из этих членов на практике решается в каждом конкретном приложении отдельно.

Упражнения

1. Переменная Q имеет практическое преимущество над P , так как она более медленно изменяется по координате z . Поэтому по ней можно производить менее частую выборку и тем самым уменьшить затраты времени на вычисления. Преобразуйте уравнение (10.3.19) в уравнение относительно переменной Q , используя при этом выражение (10.3.11). Сравните полученный результат с уравнением (10.3.9). Какое из двух уравнений, полученное Вами или уравнение (10.3.9) точнее? Почему?

2. Хорошее приближение квадратного корня обеспечивается рациональным выражением вида

$$(1+x)^{1/2} \approx \frac{1+3x/4}{1+x/4}.$$

Какому одностороннему волновому уравнению соответствует это приближение? Нарисуйте для него график дисперсионного соотношения. Какова точность представления волновых чисел для выбранных углов распространения волны по сравнению с уравнением (10.3.14)?

3. Алгебраическое уравнение $a+bx+cx^2=0$ имеет два корня. Если b является достаточно большим, то можно аппроксимировать меньший корень линейной зависимостью $a+bx=0$. Улучшенное и все же линейное относительно x приближение может быть найдено путем подстановки величины $x=-a/b$ обратно в квадратное уравнение

$$a+bx+c\left(\frac{-a}{b}\right)x=0;$$

$$ab+(b^2-ac)x=0.$$

Определите $k'_z = m - k_z$ и подставьте величину $k_z = m - k'_z$ в уравнение $k^2_x + k^2_z = m^2$. Найдите наименьший корень для k'_z . Покажите, что это дает такое же дифференциальное уравнение в частных производных, как и уравнение, приведенное в упражнении 2.

4. Пусть скорость $v = v(x) \neq v(x, z)$, т. е. является функцией от x , а $m = \omega/v(x)$. Определите оператор

$$\text{Op} = m + \frac{1}{2m} \partial_{xx} - \frac{m_x}{2m^2} \partial_x.$$

Отметим, что

$$\partial_z P = i \text{Op} P;$$

$$\partial_{zz} P = i \partial_z \text{Op} P = i \text{Op} \partial_z P = -\text{Op}^2 P$$

$(\partial_{zz} + \text{Op}^2) P = 0 =$ волновое уравнение + погрешность.

Исследуйте каждый член погрешности и решите, является ли он важным на высоких частотах (выберите члены, пропорциональные n -й степени длины волны) и при малых углах относительно оси z .

5. Просмотрите раздел, посвященный теореме Сильвестра для матриц. В какой степени квадратный корень матрицы аналогичен квадратному корню оператора?

6. Выведите волновое уравнение для выходящей волны в цилиндрической системе координат.

7. Выведите волновое уравнение для выходящей волны в сферической системе координат.

8. Полученные в упражнении 3 формулы дают хорошее широкоугольное приближение для волнового уравнения при постоянном m , а формулы, приведенные в упражнении 4, справедливы при $m = m(x)$. Используйте метод, примененный в упражнении 3, в случае, когда $m = m(x)$. Необходимо отметить, что, хотя $bx - xb = 0$, равенство $(m \partial_x - \partial_x m) P = 0$ не выполняется, если $m \neq m(x)$. Примените рассмотренный в упражнении 3 метод, избегая по возможности использования принципа коммутативности.

9. Рассмотрите поверхностную волну, распространяющуюся на поверхности несовершенной сферы. Выведите уравнение первого порядка относительно функ-

ции Φ (координаты долготы) и второго порядка относительно θ (координаты широты) для волны, возбужденных вдоль экватора. Предположите, что все величины не зависят от радиальной координатной оси.

10. Модифицируйте указанную в упражнении 4 к разделу 10.1 программу таким образом, чтобы решить уравнение (10.3.10). Вам будет необходимо просмотреть правила работы транслятора для комплексной арифметики. После расчета величины $Q(x, z)$ умножьте ее на экспоненту e^{imz} для получения функции $P(x, z)$. Напечатайте только действительную часть функции $P(x, z)$. Физической интерпретацией этого результата является свечение, возникающее позади края непрозрачного экрана. Волны, дифрагировавшие в область тени, должны иметь полукруглые фронты, если вы разместили индикатор так, чтобы сохранить $\Delta z = \Delta x$ на выходе.

11. Пусть $Z = e^{ik_x \Delta x}$ обозначает дискретность функции по координате x . Определите величину $A(Z) = \sum a_n Z^n$ через коэффициенты a_n , найденные из условия

$$a_0 + \sum_{n=1}^{\infty} a_n \left(Z^n + \frac{1}{Z^n} \right) = |k_x| \Delta x \quad \text{для } |k_x| \Delta x \leq \pi.$$

Докажите, что любое решение уравнения

$$\frac{\partial P(Z)}{\partial z} = \pm A(Z) P(Z)$$

является решением для дифференциального уравнения Лапласа $P_{xx} + P_{zz} = 0$. Эти решения могут быть использованы для продолжения полей в верхнее и нижнее полупространства.

10.4. ЭКСТРАПОЛЯЦИЯ ВОЛНОВЫХ ИМПУЛЬСОВ В ПРОСТРАНСТВЕ

В разделе 10.3 мы научились экстраполировать монохроматические волны в пространстве. Для восстановления волновых импульсов в пространстве можно предварительно разложить их с помощью преобразования Фурье в монохроматические волны, затем экстраполировать, как в предыдущем разделе, и, наконец, преобразовать снова во временную область, используя преобразование Фурье. Несмотря на то что в данном и предыдущем разделах рассматриваются сходные в спектральном представлении задачи, метод прямого восстановления волновых импульсов во временной области часто оказывается предпочтительнее по практическим соображениям. Хотя исследование во временной области сложнее, чем в частотной (все временные точки должны рассматриваться вместе, а каждое значение частоты является изолированной точкой), они позволяют достичь гораздо лучшего понимания, особенно с позиций принципа причинности. Мы обнаружим, что процедуры экстраполяции волн подобны фильтрам (в действительности они являются особым видом многомерных фазовых фильтров) и что элементы обратной связи этих фильтров должны быть минимально фазовыми. Существуют два независимых способа получения решения во временной области.

Первый способ начинается с преобразования скалярного волнового уравнения

$$0 = P_{xx} + P_{zz} - v^{-2} P_{tt} \quad (10.4.1)$$

в системе координат, перемещающейся вдоль оси z со скоростью \bar{v} , которую обычно принимают равной величине v или превышающей ее. Независимо от того, каким образом энергия распространяется в фиксированной системе отсчета, при рассмотрении в движущейся системе координат она будет оставаться постоянной или уменьшаться в обратном направлении. Преобразование координат

$$x' = x; \quad (10.4.2a)$$

$$z' = \bar{v}t - z; \quad (10.4.2б)$$

$$t' = t, \quad (10.4.2в)$$

изображено на рис. 10.20 для $\bar{v} = v$. В системе координат со штрихом все волны имеют составляющую скорости, направленную

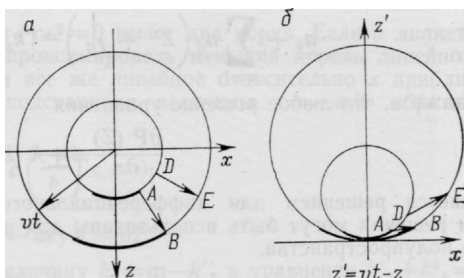


РИС. 10.20. Расходящаяся сферическая волна в фиксированных координатах (а) и в координатах, движущихся в направлении оси z со скоростью волны (б)

в положительную сторону оси z' . Знание волнового поля P для настоящего и прошедшего времен при всех x' и фиксированном z' должно быть достаточным для определения волнового поля P для настоящих и прошедших значений времени при $(x', z' + \Delta z')$, потому что прежде, чем что-либо случится на уровне $z' + \Delta z'$, что-то должно было случиться на уровне z' . Таким образом, вследствие ограничения $\bar{v} \geq v$ линейные операторы, которые мы выведем для экстраполяции волнового поля P в положительном направлении оси z' , должны удовлетворять принципу причинности. Пусть величина P' обозначает возмущение в движущейся системе отсчета. Тогда

$$P(x, z, t) = P'(x', z', t'). \quad (10.4.3)$$

Для частных производных напомним

$$P_x = P'_{x'}; \quad (10.4.4)$$

$$P_{xx} = P'_{x'x'}.$$

Аналогично

$$P_z = P'_{x'}x'_z + P'_{z'}z'_z + P'_{t'}t'_z = -P'_{z'}.$$

Поэтому

$$P_{zz} = P'_{z'z'}, \quad (10.4.5)$$

$$P_t = P'_{x'}x'_t + P'_{z'}z'_t + P'_{t'}t'_t = \bar{v}P'_{z'} + P'_{t'}.$$

Следовательно,

$$\begin{aligned}
 P'_{tt} &= \bar{v} (\bar{v} P'_{z'z'} + P'_{t'z'} + \bar{v} P'_{z't'} + P'_{t't'}) = \\
 &= \bar{v}^2 P'_{z'z'} + 2\bar{v} P'_{z't'} + P'_{t't'}
 \end{aligned}
 \tag{10.4.6}$$

Подставив выражения (10.4.4), (10.4.5) и (10.4.6) в уравнение (10.4.1), получим

$$P'_{x'x'} + \left[1 - \left(\frac{\bar{v}}{v} \right)^2 \right] P'_{z'z'} - 2 \frac{\bar{v}}{v^2} P'_{z't'} - \frac{1}{v^2} P'_{t't'} = 0. \tag{10.4.7}$$

При постоянной скорости можно принять, что $v(x, z) = \bar{v}$. Случай, когда $v \neq \bar{v}$, оставим для упражнений. Наибольший интерес для нас в уравнении (10.4.7) представляют те волны, которые распространяются со скоростью, приблизительно равной скорости в новой системе координат. В движущейся системе отсчета такие волны имеют близкое к нулю доплеровское смещение. Это предполагает отсутствие в выражении (10.4.7) члена с $P'_{t't'}$. Таким образом уравнение (10.4.7) принимает вид

$$P'_{t'z'} = \frac{v}{2} P'_{x'x'}. \tag{10.4.8}$$

Если применить преобразование Фурье по времени, то уравнение (10.4.8) примет вид $-i\omega P'_{z'} = (v/2) P'_{x'}$, т. е. станет идентичным уравнению для монохроматических волн:

$$Q_{xx} + 2imQ_z = 0,$$

полученному в предыдущей главе. Таким образом, отбрасывание члена $P'_{t't'}$ является хорошо знакомой нам аппроксимацией окружности параболой.

При решении уравнения (10.4.8) на ЭВМ можно принять любой из двух подходов. Первый из них заключается в том, что поле P' задается вначале на сетке по координатам x' и z' , а затем уравнение (10.4.8) используется для экстраполяции волн по оси t' . Во втором подходе уравнение (10.4.8) используется для экстраполяции волн по оси z' .

Прежде чем рассматривать численный метод решения уравнения (10.4.8), попытаемся получить его с помощью совершенно различных преобразований координат. Примем новую систему координат, фиксированную в пространстве относительно старой системы. Однако в каждой точке пространства в новой системе отсчета будем использовать различные часы. Все часы идут с одинаковой скоростью, но они запущены таким образом, что плоская волна, проходящая в положительном направлении оси z , будет иметь одно и то же время вступления, измеренное всеми часами. Это отчасти похоже на движение реактивного самолета в запад-

ном направлении. Уравнения преобразования имеют вид

$$x'' = x; \quad (10.4.9a)$$

$$z'' = z; \quad (10.4.9б)$$

$$t'' = t - \frac{z}{v}. \quad (10.4.9в)$$

Возмущение, возникшее при $(x, z, t) = 0$, изображено на рис. 10.21. Гиперболы слева (рис. 10.21,а) показывают времена вступления волн для $z=0, \Delta z, 2\Delta z$. Если время является функцией положения точки приема, заданного зависимостью $t'' = t - z/v$, то

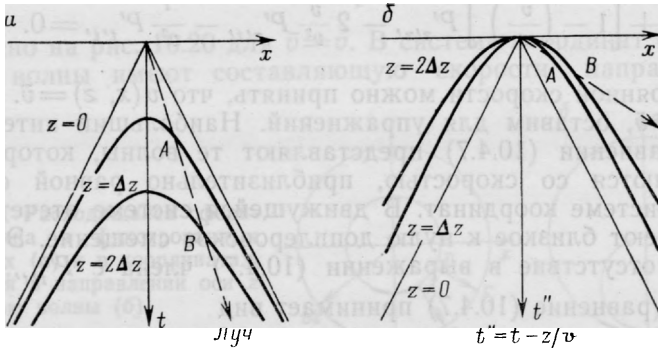


РИС. 10.21. Времена вступления волн от точечного источника при $x=0, z=0, t=0$ в обычной (а) и движущейся (б) системах координат

времена вступлений волн будут такими, как показано справа на рис. 10.21,б. Энергия перемещается в направлении оси t'' , так как на волновом фронте $z=vt \cos \theta$ имеем $t'' = t - z/v = t(1 - \cos \theta)$. Большой интерес для вычислений представляет возможность отсчета времени относительно времени прихода самого раннего из возможных лучей. Определим возмущение в новой системе отсчета через функцию P'' следующим образом:

$$P(x, z, t) = P''(x'', z'', t''). \quad (10.4.10)$$

Поступая, как и прежде, найдем

$$P_{xx} = P''_{x''x''}; \quad (10.4.11)$$

$$P_{zz} = P''_{z''z''} - 2v^{-1}P''_{t''z''} + v^{-2}P''_{t''t''}; \quad (10.4.12)$$

$$P_{tt} = P''_{t''t''}. \quad (10.4.13)$$

Подставляя эти соотношения в волновое уравнение (10.4.1), получим:

$$P''_{t''z''} = \frac{v}{2} (P''_{x''x''} + P''_{z''z''}). \quad (10.4.14)$$

Последний член в уравнении (10.4.14) является малой величиной более высокого порядка для волн, распространяющихся в направлениях, составляющих небольшие углы с осью z . Следова-

тельно, решение волнового уравнения для волн, распространяющихся в положительном направлении оси z , является произвольной функцией вида $f(t-z/v) = f''(t'')$. Таким образом, производная $\partial f''/\partial z''$ обращается в нуль для волны, идущей вдоль оси z'' . Пренебрегая в выражении (10.4.14) членом $P''_{z''z''}$, получаем уравнение:

$$P''_{t''z''} = -\frac{v}{2} P''_{x''x''}, \quad (10.4.15)$$

которое совпадает с уравнением (10.4.8). Введение временной зависимости $e^{-i\omega t}$ в выражение (10.4.8) или (10.4.15) дает уравнение (10.3.10), которое было получено ранее для экстраполяции монохроматических волн.

Другой подход заключается в том, что нестационарное волновое уравнение можно получить простой заменой величины $-i\omega$ в уравнениях для монохроматических волн величиной ∂_t .

Рассмотрим схему конечных разностей для решения уравнения (10.4.8) или (10.4.15). Опустим штрихи. Пусть приращение $j\Delta t$ относится ко времени, а приращение $n\Delta z$ — к координате z . Обозначим через δ разностный оператор, а через \mathbf{P}^n_j — вектор для всех значений n и j . С помощью этого вектора будем задавать значения давления вдоль оси x . Используя матричную алгебру, можно избавиться от необходимости написания нижнего (подстрочного) индекса для обозначения зависимости от x . Обозначим через \mathbf{T} тридиагональную матрицу с отрицательным разностным оператором второго порядка $(1, -2, 1)$ на диагонали. Учитывая все эти обозначения, преобразуем уравнение (10.4.8) или (10.4.15) к виду

$$\delta_z \delta_t \mathbf{P}^n_j = -\frac{v\Delta z \Delta t}{8\Delta x^2} \mathbf{T} \mathbf{P}^n_j. \quad (10.4.16)$$

Введем обозначение $a = v\Delta z \Delta t / 8\Delta x^2$. Теперь мы должны решить точнее, какие приближения первой разности следует использовать в выражении (10.4.16). Воспользуемся схемой Кранка — Николсона, эквивалентную билинейному преобразованию. Сначала образуем конечные разности по времени

$$\delta_z (\mathbf{P}^n_{j+1} - \mathbf{P}^n_j) = -a \mathbf{T} (\mathbf{P}^n_{j+1} + \mathbf{P}^n_j),$$

а затем конечные разности по пространственным координатам

$$(\mathbf{P}^{n+1}_{j+1} - \mathbf{P}^{n+1}_j) - (\mathbf{P}^n_{j+1} - \mathbf{P}^n_j) = -a \mathbf{T} (\mathbf{P}^{n+1}_{j+1} + \mathbf{P}^{n+1}_j + \mathbf{P}^n_{j+1} + \mathbf{P}^n_j), \quad (10.4.17)$$

Будем считать величину \mathbf{P}^{n+1}_{j+1} неизвестной, а все остальные величины известными. Переносим неизвестную в левую часть, а известные в правую, получаем:

$$(\mathbf{I} + a \mathbf{T}) \mathbf{P}^{n+1}_{j+1} = \mathbf{P}^{n+1}_j + \mathbf{P}^n_{j+1} - \mathbf{P}^n_j - a \mathbf{T} (\mathbf{P}^{n+1}_j + \mathbf{P}^n_{j+1} + \mathbf{P}^n_j). \quad (10.4.18)$$

Для каждого n и j правая часть сжимается в известный вектор. Левая часть является тридиагональной матрицей $(\mathbf{I} + a\mathbf{T})$, умноженной на неизвестный вектор \mathbf{P}_{j+1}^{n+1} . Решение этих уравнений является предельно простым и может быть получено тем же способом, что и для уравнения теплопроводности в разделе 10.1. Граничные условия по x содержатся на концах \mathbf{I} . Граничные условия по z и t достаточно задать для всех x в виде функции \mathbf{P}_0^n для всех n и функции \mathbf{P}_j^0 для всех j . Возможны также другие комбинации граничных условий.

Очень важна устойчивость решения. Можно показать, что рекурсивное соотношение (10.4.18) является устойчивым при любом положительном значении a . Если собственные значения и собственные векторы матрицы \mathbf{T} заданы, а все векторы \mathbf{P}^n_j выражены через собственные векторы матрицы \mathbf{T} , то уравнение (10.4.18) разбивается на множество отдельных уравнений по одному для каждого из собственных значений матрицы \mathbf{T} . Собственные векторы матрицы \mathbf{T} имеют компоненты, которые являются синусоидальными функциями от x . Если существуют границы по x , то существует и набор дискретных частот. В противном случае спектр непрерывен. Чтобы убедиться в этом, заметим, что для неограниченной среды функции \mathbf{TP} представляет собой последовательность $(-1, 2, -1)$, свернутую с экспонентой $e^{ik_x m \Delta x}$;

$$(-e^{ik_x \Delta x} + 2 - e^{-ik_x \Delta x}) e^{ik_x m \Delta x}.$$

Таким образом, собственное значение матрицы равно $2 - 2 \cos k_x \times \Delta x = (2 \sin k_x \Delta x / 2)^2$. Так как любое собственное значение матрицы должно лежать в интервале между 0 и +4, то достаточно исследовать уравнение (10.4.18), где вектор \mathbf{P}^n становится скалярной P^n_j функцией от k_x , вектор \mathbf{I} заменяется единицей, а векторная матрица \mathbf{T} заменяется на T — произвольное число в интервале от 0 до 4. Можно доказать, что при граничных условиях, характеризующих сохранение энергии, собственные значения матрицы также лежат в интервале между 0 и 4. Теперь, предполагая, что функция P^n_j известна для всех j при некотором определенном значении n , исследуем устойчивость определения функции P_j^{n+1} для всех j . Перенеся неизвестные в уравнении (10.4.17) в левую часть, получим:

$$(1 + aT) P_{j+1}^{n+1} - (1 - aT) P_j^{n+1} = (1 - aT) P_{j+1}^n - (1 + aT) P_j^n. \quad (10.4.19)$$

Для устойчивости выражения (10.4.19) важно, чтобы при последовательном увеличении j коэффициент при функции P_{j+1}^{n+1} превышал коэффициент при функции P_j^{n+1} . При уменьшении j должно быть справедливо обратное соотношение. Устойчивость можно исследовать с помощью Z -преобразования, описанного в предыдущих главах. С помощью Z -преобразования уравнения (10.4.19) можно

показать, что коэффициент при Z^j в уравнении

$$[(1+aT) - Z(1-aT)]P(Z)^{n+1} = [(1-aT) - Z(1+aT)]P(Z)^n \quad (10.4.20)$$

дает выражение (10.4.19). Функция фильтра для вычисления величины $P(Z)^{n+1}$ по $P(Z)^n$ имеет вид

$$\frac{(1-aT) - Z(1+aT)}{(1+aT) - Z(1-aT)}. \quad (10.4.21)$$

Отметим, что для положительных значений величин a и для всех T , лежащих в интервале от 0 до 4, полином в знаменателе является минимально фазовым. Это означает, что рекуррентное соотношение (10.4.19) устойчиво, а поскольку выражение (10.4.21) представляет собой фазовый фильтр, то рекуррентное соотношение будет устойчивым также и для больших n .

Мы получили точное, законченное и до некоторой степени трудное доказательство устойчивости. Читатель, несомненно, заметит, что в его собственных приложениях встречаются незначительно отличающиеся уравнения, возможно с зависимостью $v = v(x, z)$ или с увеличенной при больших углах точностью. Какие рекомендации можно дать относительно формулирования задач, чтобы они были устойчивыми при экстраполяции волнового поля? Во-первых, следует учитывать физическое представление о характере протекания волнового процесса. Во-вторых, если возникает неустойчивость, то она, вероятно, возникает при непредвиденных значениях величин ω , k_x , k_z или их отношений. Заметим, что уравнение (10.4.15) в области преобразования Фурье имеет вид

$$\omega k_z = -\frac{v}{2} k_x^2. \quad (10.4.22)$$

Если мы предполагаем экстраполировать волны в направлении оси z'' , то мы будем, по существу, формировать функции $\exp(ik_z z'')$ либо $\exp(-ik_x^2 z''/\omega)$. Читатель должен вспомнить все важные свойства фазовых фильтров и факторизации спектров. Если процесс распространения волн в среде моделируется фазовыми фильтрами, являющимися по предположению физически осуществимыми или удовлетворяющими принципу причинности, то производная по фазе или групповое время задержки сигнала должны быть положительной величиной для всех частот. В этом случае для производной по фазе получим выражение

$$\frac{d}{d\omega}(k_z z) = -z \frac{v}{2} k_x^2 \frac{d}{d\omega} \left(\frac{1}{\omega} \right) = z \frac{v}{2} \frac{k_x^2}{\omega^2}, \quad (10.4.23)$$

являющееся, как и требуется, положительным для всех частот и всех ω . Положительность только самих величин ω и k_x практического интереса не представляет. Если для какого-либо значения ω

или k_x групповое время задержки окажется отрицательным, то уравнения экстраполяции волн во временной области будут иметь разрыв.

В заключение рассмотрим пример, изображенный на рис. 10.22. В первом кадре показан плоский фронт волны, деформирующейся как при прохождении через область, в которой скорость периодически изменяется в направлении x . С точки зрения оптики первый кадр на рис. 10.22 характеризует импульсную плоскую волну сразу

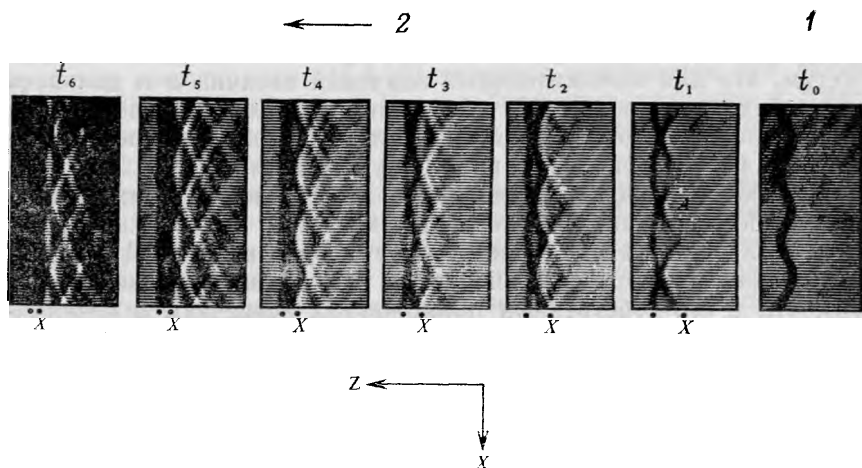


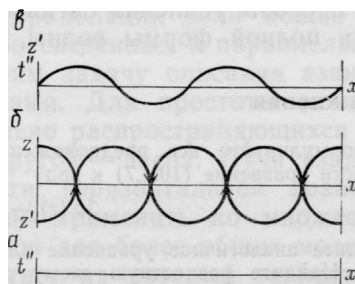
РИС. 10.22. Возмущенная плоская волна, распространяющаяся в однородной среде.

1 — начало; 2 — направление распространения волны

после выхода ее из фазовой решетки. С позиций атмосферной акустики искажения волнового фронта могли возникнуть, например, при прохождении плоской волны через элементы атмосферы с периодической циркуляцией, которые изображены на рис. 10.23. Последующие кадры на рис. 10.22 изображают процесс распространения фронта волны. В работах по оптике [35] решение для монохроматических волн обычно получают на бесконечности. Наиболее очевидный вывод из этого состоит в том, что энергия распространяется, перемещаясь от одного кадра к следующему. Простой импульс на первом кадре превратился в растянутый колебательный процесс в последнем кадре. С течением времени энергия первого импульса уменьшается, а энергия последующих колебаний увеличивается. Другой, весьма примечательной особенностью рассматриваемого процесса распространения волн является то, что после некоторого достаточно длительного промежутка времени первые вступления волн стремятся снова выстроиться в ряд так, что возмущения волнового фронта со временем восстанавливаются. В противоположность этому волновая кода (шлейф волны) все больше усложняется по мере распространения волны и превращается в пространственно некогерентную волну. (Это ими-

тирует поведение большинства волн при геофизических наблюдениях). Можно также отметить несколько других, менее заметных особенностей рис. 10.22. Для шлейфа волны характерно движение энергии в обратном направлении, т. е. от первого вступления волны к последующим. В то же время фазовые фронты в шлейфе волны, обозначенные через x , движутся вперед по направлению к вступлению волны. Одновременно наклон фронта волны к оси z (или видимое направление ее распространения) стремится увели-

РИС. 10.23. Возможный способ формирования возмущенной волны.
 a — плоская волна; b — колеблющиеся воздушные ячейки; $в$ — возмущенная волна



читься по мере приближения к нижнему кадру. Это является подтверждением того факта, что последующие вступления волны распространялись по более длинным лучевым траекториям. Кроме того, на втором кадре в области точки A наблюдается двумерная фокусировка (каустика) и сдвиг фазы сигнала на $\pi/2$, являющийся причиной образования дублетов. Хотя данный расчет был проделан, начиная с кадра t_0 и кончая кадром t_6 , он мог бы быть осуществлен и в обратном порядке, т. е. начиная с кадра t_6 и кончая кадром t_0 . После перестройки во времени мы собрали бы на кадре t_0 всю энергию сигнала.

Задача представления возмущения неограниченной протяженности по x на конечной расчетной сетке решена с помощью периодического возмущения, имеющего нулевой наклон на боковых границах. Граничные условия с нулевым наклоном при этом эквивалентны бесконечному периодическому продолжению по x . Величина $v\Delta t\Delta z/\Delta x^2 = 1/4$ выбрана таким образом, чтобы получить необходимые изменения в последовательности кадров на рис. 10.22, представляющих собой пять вычислительных итераций процесса распространения волн. Масштаб представления решения может быть изменен несколькими способами благодаря взаимосвязи параметров $v\Delta t$, Δx и Δz .

Возможно, имеет смысл рассмотреть различные процессы усиления данных в соответствии с рис. 10.22. В процессе, называемом «управление лучом», результаты наблюдений, подобные представленным на рис. 10.22, можно было бы просуммировать по координате x с целью усиления сигнала и подавления шума. Ясно, что в рассматриваемом нами случае (см. рис. 10.22) усилятся первые вступления волны и в то же время ослабится случайный шум. Суммирование по x будет также стремиться уничтожить энергию сигнала, сосредоточенную в осциллирующих шлейфах.

волны. При необходимости реального увеличения отношения сигнал/шум едва ли возникнет необходимость в процедурах обработки, ослабляющих энергию сигнала. Несмотря на то что величины z' или t'' растут, положение осложняется, так как энергия сигнала перетекает от начального импульса к осциллирующим волновым шлейфам. То, что часто считают «шумом, генерируемым сигналом», может оказаться полезным сигналом в потенциально ценной форме. Можно действительно ожидать поразительных результатов, если способы усиления сигналов будут основываться на использовании полной формы волны, а не только начального импульса.

Упражнения

1. Сформулируйте все предположения, которые необходимы для того, чтобы привести уравнение (10.4.7) к виду

$$[v(z) - \bar{v}] P'_{z'} = P'_{t'}.$$

Выведите аналогичное уравнение для координат, помеченных двумя штрихами.

2. Найдите разностную схему для уравнения из упражнения 1, которая экстраполирует волновое поле с уровня z' на уровень $(z' + \Delta z')$. Покажите, что необходимо использовать прошедшее время, если $\bar{v} > v$, и будущее время, если $\bar{v} < v$.

3. Пусть координатные преобразования заданы следующим образом:

$$x' = x;$$

$$z' = z;$$

$$t' = t - \int_0^z v^{-1}(z) dz.$$

Запишите скалярное волновое уравнение в этих координатах.

4. Покажите, что если при трансформации скорость \bar{v} в формулах (10.4.9a), (10.4.9б) и (10.4.9в) принимает любое значение, меньшее v в волновом уравнении, то в результате будут получены устойчивые разностные уравнения.

5. Рассмотрите разностное уравнение $(1 + \delta_{xx}/12)\delta_{zt}P = b\delta_{xx}P$. При каком значении b это уравнение приводится к явной форме? Является ли временная рекурсия устойчивой для этого значения b ?

10.5. ВЗАИМОДЕЙСТВИЕ ЛУЧЕЙ

Большую часть наших знаний о внутреннем строении Земли дают границы раздела, превращающие волны, распространяющиеся вниз, в волны, выходящие на поверхность. Математически строгое разложение возмущений в слоистой среде на нисходящие $[\exp(ik_z z)]$ и восходящие волны $[\exp(-ik_z z)]$ возможно, но в настоящее время такое разложение не разработано для двух- и трехмерных неоднородных сред. То, что у нас имеется в этой области, представляет собой набор специально разработанных для расчета волн методов, строгое обоснование которых предполагает полное отсутствие горизонтального распространения энергии¹ или его пренебрежимо малость. На практике не представ-

¹ Вследствие дифракции наблюдается распространение энергии вдоль фронта волны. (Прим. перев.)

ляет интереса точное разложение волн на восходящие и нисходящие компоненты. Нам больше интересуют описание взаимодействия между приближенно коллимированными (сведенными в параллельный пучок) лучами. В голографии к ним относятся *падающие* (или опорные) и *рассеянные* лучи. В сейсмологии ими могут быть падающий луч продольной волны и луч рассеянной поперечной волны. Они не требуют какой-либо определенной ориентации относительно друг друга или относительно вертикали.

Рассмотренные ранее методы экстраполяции волн можно использовать для описания лучей, грубо сведенных в параллельный пучок вдоль оси z . Теперь рассмотрим задачу описания взаимодействия между двумя такими лучами. Для простоты вначале изучим два приблизительно вертикально распространяющихся луча, один из которых идет вниз, а другой — вверх. Лучи взаимодействуют между собой в плоскости горизонтальной границы раздела. Развиваемый подход можно применить ко множеству простых моделей. Точность результатов для более общих моделей является практическим вопросом, ответ на который зависит от конкретных условий задачи. Недостаточная точность является следствием многих причин, среди которых можно отметить следующие:

1. Угловую зависимость скорости распространения волн в коллимированном пучке, которая возникает вследствие приближений, подобных приближению Френеля.

2. Пренебрежение дифракцией.

3. Возможную неспособность уравнения для двух коллимированных лучей описать все существенные для данной задачи лучи из числа возникающих на сложной границе раздела.

4. Приближенность описания упругих продольных волн скалярным волновым уравнением.

Влияние недостаточной точности должно оцениваться с учетом точности экспериментальных работ, требуемой точности, точности и стоимости конкурирующих методов. Такие оценки выходят за рамки настоящей книги.

В этом разделе мы будем описывать только первичную (однократно) отраженную сейсмическую энергию, используемую в методе отраженных волн. Волны большой амплитуды возбуждаются на поверхности Земли с помощью взрывчатых веществ или других источников, обладающих достаточной энергией. Эти волны проникают в землю, где малая доля их энергии отражается слабыми границами в среде и передается обратно к чувствительным приемникам, расположенным на поверхности. Встречающиеся на практике ситуации, когда заметное количество энергии отражается вверх и вниз по несколько раз (многократные отражения), будут обсуждены в следующей главе. Для горизонтально-слоистой среды можно использовать уравнение (9.3.13).

$$\frac{d}{dz} \begin{bmatrix} U \\ D \end{bmatrix} = \begin{bmatrix} -iab & \\ & iab \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix} - \frac{1}{2} \frac{Y_z}{Y} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}. \quad (10.5.1)$$

Поскольку практическая ситуация, которую мы пытаемся описать, удовлетворяет неравенству $U \ll D$, то можно аппроксимировать нижнюю строку в выражении (10.5.1) следующим образом:

$$D_z = iab D - \frac{Y_z}{2Y} D. \quad (10.5.2)$$

Чтобы физически объяснить соотношение (10.5.2), применимое также в условиях, когда величина a , b и Y зависят от z , отметим, что решение уравнения (10.5.2) имеет следующий вид:

$$D = D_0 Y^{-1/2} \exp\left(i \int_0^z ab dz\right). \quad (10.5.3)$$

Это легко проверить прямой подстановкой. Из выражения (10.5.2) следует, что произведение iab определяет фазу (или скорость) волны, а отношение Y_z/Y — изменение ее амплитуды. Таким образом, можно интерпретировать член Y_z/Y как соответствующий физическому эффекту, связанному с коэффициентом прозрачности. Часто информация о скорости в произведении ab приближительна, а положения границ раздела в Земле, задаваемые скачками функции Y_z/Y , вообще неизвестны. Это означает, что нет необходимости отказываться от поиска решения D , если погрешность амплитуды решения, определяемая неизвестными коэффициентами прозрачности, допустима.

В разделе 10.3 описана возможность рассмотрения неплоских волн при условии, что величина iab является квадратным корнем дифференциального оператора — $(\omega^2/v^2 + \partial_{xx})$. Для пучка лучей, коллимированных в направлении вниз вдоль оси z , первое приближение квадратного корня имеет вид $i\omega/v [1 + v^2(\partial_{xx}/2\omega^2)]$. При использовании предположения о коллимированности пучка ($\partial_{zz}0$) и равенстве нулю неизвестного градиента полной проводимости нисходящую волну D можно рассчитать с помощью уравнения

$$D_z = \frac{i\omega}{v} D + \frac{iv}{2\omega} D_{xx}. \quad (10.5.4)$$

Это выражение имело бы еще большее сходство с полученными ранее результатами, если предположить, что среда однородна (скорость $v = \bar{v}$). Воспользуемся преобразованием вида $D = D' e^{imz}$, где $m = \omega/v$. При этих условиях уравнение (10.5.4) преобразуется к виду

$$D'_z = \frac{iv}{2\omega} D'_{xx}. \quad (10.5.5)$$

Чтобы решить уравнение (10.5.4) или (10.5.5) в области, расположенной внутри Земли, достаточно знать функцию D на поверхности Земли (для всех x при $z=0$). В сейсмической разведке методом отраженных волн величина D обычно может аппроксимироваться дельта-функцией в месте взрыва.

Теперь обратимся к расчету восходящей волны U . Из верхней строки выражения (10.5.1) имеем

$$U_z = -iabU - \frac{Y_z}{2Y}(U - D). \quad (10.5.6)$$

Если пренебречь влиянием коэффициента прозрачности на величину U , обозначая в то же время взаимодействие между волнами U и D через коэффициент отражения, то последнее выражение примет вид

$$U_z = -iabU + \frac{Y_z}{2Y}D. \quad (10.5.7)$$

Поскольку коэффициент отражения

$$c = \frac{Y_2 - Y_1}{Y_2 + Y_1} = -c',$$

то [для дифференцируемой функции $Y(z)$] уравнение (10.5.7) можно записать в виде

$$U_z = -iabU - c'(z)D. \quad (10.5.8)$$

Как и для нисходящих волн, можно обобщить задачу для плоских волн в случае пучков лучей и использовать приближение квадратного корня. Тогда

$$U_z = -\frac{i\omega}{v}U - \frac{iv}{2\omega}U_{xx} - c'(x, z)D. \quad (10.5.9)$$

Замена переменных $U = U''e^{-imz}$ и $D = D'e^{imz}$ в предположении о постоянстве скорости в среде и непостоянстве ее проводимости позволяет преобразовать уравнение (10.5.9) к виду

$$U''_z = -\frac{\bar{v}/2}{-i\omega}U''_{xx} - c'(x, z)D'e^{2imz}. \quad (10.5.10)$$

Важно понять, как можно получить решение уравнения (10.5.10). Прежде всего, нужно вычислить функцию D' с помощью некоторого другого уравнения. При решении уравнения (10.5.10) будем считать член $c(x, z)D'$, соответствующим источнику, генерирующему возмущение U' . Существуют две важные задачи: *прямая* и *обратная*. Прямая задача — это синтез данных. Обратная задача заключается в том, чтобы определить функции $c(z)$ и $U''(z)$ на основании выборки данных U'' , заданной на поверхности Земли (при $z=0$). Более полно обратная задача рассматривается в следующей главе, посвященной обработке сейсмических данных. Здесь мы ограничимся прямой задачей. Граничное условие для U'' , которое дает возможность использовать уравнение (10.5.9) для нахождения функции U'' всюду, заключается в том, чтобы вели-

чина U'' равнялась нулю для всех x внутри Земли на глубине z_N , ниже всех обнаруживаемых отражающих границ. Тогда уравнение (10.5.10) пересчитывается на более высокие уровни от z_N к z_{N-1} , z_{N-2} и т. д. При этом функция U'' остается равной нулю до тех пор, пока мы не дойдем до первого облученного отражателя, т. е. до самого глубокого места, где величины $c(x, z)$ и D' одновремен-

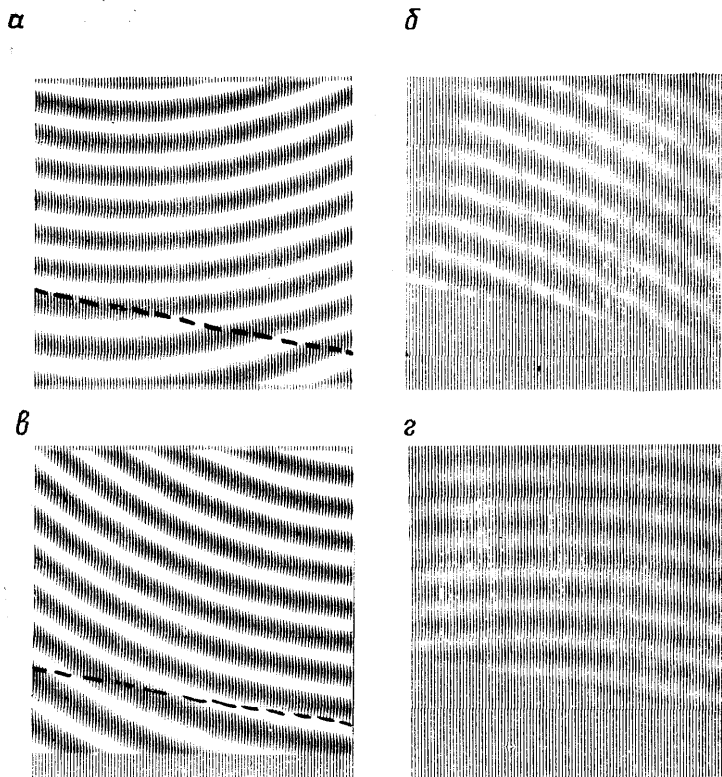


РИС. 10.24. Нисходящие (*a*, *в*) и восходящие (*б*, *г*) волны при двух положениях источника (*a*, *в*)

но не равны нулю. В этой точке «включается» источник, т. е. соответствующий ему член в уравнении (10.5.10) становится для восходящей волны ненулевым. Одновременно и функция $U'' \neq 0$ в направлении вверх. Этот расчет иллюстрируется рис. 10.24. На рис. 10.24,*a*, *в* показаны нисходящие сферические волны от двух источников, расположенных в разных местах. Эти волны «освещают» погруженную границу раздела, которой соответствует скачок как импеданса, так и скорости. Ниже границы длина волны увеличивается. На рис. 10.24,*б*, *г* показаны восходящие (отраженные от границы) волны, которых нет ниже границы раздела. На границе раздела фазы нисходящей и восходящей волн одинаковы.

Вычисления могут быть проделаны также во временной области. Запишем преобразования координат для нисходящей волны

$$x' = x; \quad (10.5.11a)$$

$$z' = z; \quad (10.5.11b)$$

$$t' = t - \frac{z}{v} \quad (10.5.11b)$$

и преобразования координат для восходящей волны

$$x'' = x; \quad (10.5.12a)$$

$$z'' = z; \quad (10.5.12b)$$

$$t'' = t + \frac{z}{v}. \quad (10.5.12b)$$

Теперь можно выразить величины U и D в обеих системах координат (10.5.11) и (10.5.12):

$$U(x, z, t) = U'(x', z', t') = U''(x'', z'', t''); \quad (10.5.13a)$$

$$D(x, z, t) = D'(x', z', t') = D''(x'', z'', t''). \quad (10.5.13b)$$

Правило дифференцирования сложной функции дает

$$\partial_x D = \partial_{x'} D'; \quad (10.5.14a)$$

$$\partial_z D = \left(\partial_{z'} - \frac{1}{v} \partial_{t'} \right) D'; \quad (10.5.14b)$$

$$\partial_t D = \partial_{t'} D' = \partial_{t''} D''; \quad (10.5.14b)$$

$$\partial_x U = \partial_{x'} U'; \quad (10.5.15a)$$

$$\partial_z U = \left(\partial_{z'} + \frac{1}{v} \partial_{t'} \right) U'; \quad (10.5.15b)$$

$$\partial_t U = \partial_{t'} U'. \quad (10.5.15b)$$

Рассматривая среду с постоянной скоростью $v = \bar{v}$ и умножая обе части выражений (10.5.4) и (10.5.9) на $-i\omega$, а затем, отождествляя множитель $-i\omega$ с производной по времени, получаем

$$D_{zt} = -\frac{1}{v} D_{tt} + \frac{\bar{v}}{2} D_{xx}; \quad (10.5.16a)$$

$$U_{zt} = \frac{1}{v} U_{tt} - \frac{\bar{v}}{2} U_{xx} - c'(x, z) D_t. \quad (10.5.16b)$$

Уравнения (10.5.16а) и (10.5.16б) с помощью выражений (10.5.14) и (10.5.15) легко преобразуются к виду

$$D'_{z't'} = \frac{\bar{v}}{2} D'_{x'x'}, \quad (10.5.17а)$$

$$U''_{z''t''} = -\frac{\bar{v}}{2} U''_{x''x''} - c'(x'', z'') D''_{t''}. \quad (10.5.17б)$$

Теперь формулу (10.5.17а) можно использовать для вычисления величины D' , а формулу (10.5.17б) — для определения D'' . Вычитая (10.5.12в) из (10.5.11в), получаем:

$$t' = t'' - \frac{2z''}{\bar{v}}.$$

Следовательно, с помощью формулы (10.5.13б) зависимость (10.5.17б) можно выразить через величину D' следующим образом:

$$U''_{z''t''} = -\frac{\bar{v}}{2} U''_{x''x''} - c'(x'', z'') \partial_{t''} D' \left(x'', z'', t'' - \frac{2z''}{\bar{v}} \right). \quad (10.5.18)$$

Этот результат во временной области является преобразованным выражением (10.5.10).

Упражнения

1. Покажите, что величина $1/2[(d/dz) \ln \Gamma]$ является коэффициентом отражения волны c' при распространении ее над границей раздела.

2. Вспомните из выражения (9.3.20), что определение импеданса Y содержит величину k_x . Эта зависимость была устранена при выводе формулы (10.5.10). Измените формулу (10.5.10) таким образом, чтобы включить в нее члены, содержащие $\partial D''/\partial x$. Это изменение позволяет выразить коэффициент отражения в функции угла падения волны на границу.

10.6. ЧИСЛЕННАЯ ВЯЗКОСТЬ¹

При положительной численной вязкости отклонение поведения коротких волн при решении разностного уравнения от ожидаемого их поведения, соответствующего решению дифференциального уравнения, является таким, что короткие волны затухают с глубиной. Численная вязкость может быть также отрицательной, вызывая усиление коротких волн вместо их ослабления. Независимо от того, имеется ли хорошее обоснование решений или нет для изучения численной вязкости, ученые часто неохотно занимаются ею по нескольким причинам. Во-первых, даже если дифференциальные уравнения не нарушают принцип причинности, то возможна неустойчивость, связанная с отрицательной вязкостью в разностных уравнениях. Во-вторых, из-за проблемы экономии машинного вре-

¹ Этим термином автор определяет искусственно вводимое затухание высокочастотных компонент волнового поля. Это обеспечивает повышение устойчивости решения конечно-разностного уравнения. (Прим. перев.)

мени [особенно в многомерной задаче, такой как $P_{zt} = (v/2)P_{xx}$] может потребоваться, чтобы форма волны была задана с помощью немногих реальных точек. В-третьих, при обработке данных наблюдений, когда поле $P(x, t)$ экстраполируется с уровня z_1 на уровень z_2 , может оказаться, что эти данные несовместимы с определенными допущениями, на которых основаны уравнения экстраполяции волнового поля.

Например, предположим, что поле $P(x, t)$ имеет преобразование Фурье $P'(k_x, \omega)$. Поскольку $k_x^2 + k_z^2 = \omega^2/v^2$, то свободно распространяющиеся волны характеризуются соотношением $|k_x| < \omega/v$. Следовательно, спектр $P'(k_x, \omega)$ должен обращаться в нуль, если не выполняется условие $|k_x| < \omega/v$. При выводе уравнения $P_{zt} = (v/2)P_{xx}$ дополнительно предполагалось, что волны имеют малые углы распространения. Следовательно, неравенство становится более сильным при $|k_x| \ll \omega/v$. Данные наблюдений, несомненно, не будут точно удовлетворять этим условиям. Поэтому имеются две возможности. Во-первых, можно пренебречь искажениями в части пространства преобразования (k_x, ω) , где сигналы имеют небольшую энергию, а разностное уравнение не усиливает их до неприемлемого уровня. Во-вторых, можно преобразовать наше разностное или дифференциальное уравнение таким образом, чтобы в рассматриваемой части пространства преобразования (k_x, ω) существовала регулируемая положительная численная вязкость. Этот подход иногда называют веерной фильтрацией из-за клинообразной формы области пространства (ω, k_x) , в которой ослабляется решение.

Оператор ∂_{xx} имеет преобразование Фурье $-k_x^2$. Оператор δ_{xx} равнозначен свертке на оси x с коэффициентами $(1, -2, 1)/\Delta x^2$; следовательно, ее преобразование Фурье имеет вид $[\exp(-ik_x\Delta x) - 2 + \exp(ik_x\Delta x)]/\Delta x^2$. Запишем это следующим образом:

$$\text{ФП}(\partial_{xx}) = -k_x^2; \quad (10.6.1a)$$

$$\text{ФП}\left(\frac{\delta_{xx}}{\Delta x^2}\right) = -\widehat{k}_x^2 = -\frac{2(1 - \cos k_x \Delta x)}{\Delta x^2} = -\frac{2^2}{\Delta x^2} \sin^2\left(\frac{k_x \Delta x}{2}\right). \quad (10.6.1b)$$

Приближение величины \widehat{k}_x с помощью величины k_x задается выражением

$$\widehat{k}_x \approx \frac{2}{\Delta x} \sin k_x \frac{\Delta x}{2}. \quad (10.6.2)$$

Погрешность приближения $\widehat{k}_x \approx k_x$ табулирована. В табл. 2 приведены относительные погрешности вычислений для коротких волн, часто ассоциируемые с выражением дифференциальных уравнений в конечно-разностной форме.

Схеме Кранка — Николсона соответствует другое приближение. Здесь оператор $\partial/\partial t$, имеющий преобразование $-i\omega$, аппроксимируется билинейным преобразованием. Величина $\widehat{\omega}$ выражается

через частоту ω :

$$-i\hat{\omega}\Delta t = \frac{2(1 - e^{i\omega\Delta t})}{1 + e^{i\omega\Delta t}}.$$

Умножая числитель и знаменатель правой части последнего выражения на экспоненту $e^{-i\omega\Delta t/2}$ получаем:

$$-i\hat{\omega}\Delta t = 2 \frac{e^{-i\omega\Delta t/2} - e^{i\omega\Delta t/2}}{e^{-i\omega\Delta t/2} + e^{i\omega\Delta t/2}} = -2i \operatorname{tg} \frac{\omega\Delta t}{2};$$

$$\hat{\omega} = \frac{2}{\Delta t} \operatorname{tg} \frac{\omega\Delta t}{2}. \quad (10.6.3)$$

Это приближение также табулировано (см. табл. 2).

Для образования разностных приближений более высокого порядка решим уравнение (10.6.2) относительно величины ik_x . Из формулы (10.6.2) имеем

Таблица 2

Число точек на длину волны, $\frac{\pi}{2\pi/\omega\Delta t}$	$\omega\Delta t$ или $k_x\Delta x$, рад	Относительная погрешность величины $2 \operatorname{tg} \frac{\omega\Delta t}{2}$	Относительная погрешность величины $2 \sin k_x\Delta x/2$	Относительная погрешность уравнения (10.6.3)
$\pi \times 10^{12}$	2×10^{-12}	$10^{-2n}/3$	$10^{-2n}/6$	$0(10^{-4n})$
20 000 000	0,314 159	0,008 272	-0,004 116	-0,000 021
16 000 000	0,392 699	0,012 968	-0,006 434	-0,000 051
12 000 000	0,523 599	0,023 218	-0,011 449	-0,000 159
10 000 000	0,628 318	0,033 675	-0,016 504	-0,000 330
8 000 000	0,785 398	0,053 325	-0,025 834	-0,000 812
6 000 000	1,047 197	0,097 645	-0,046 109	-0,002 613
4 000 000	1,570 796	0,240 396	-0,104 913	-0,013 849
3 000 000	2,094 395	0,492 833	-0,189 390	-0,046 111
2 100 000	2,991 992	1,596 763	-0,400 123	-0,203 548

$$ik_x = \frac{2}{\Delta x} \operatorname{Arsh} \left(\frac{i\hat{k}_x\Delta x}{2} \right). \quad (10.6.4)$$

Вспомним степенной ряд для обратной гиперболической функции арка-синуса:

$$\operatorname{Arsh} u = u - \frac{1}{2} \frac{u^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \frac{u^5}{5} - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \frac{u^7}{7} + \dots \quad (10.6.5)$$

Обратное преобразование Фурье для выражения (10.6.4) с учетом разложения (10.6.5) позволяет определить разложение в степенной ряд функции ∂_x по степеням величины δ_x .

В настоящее время сейсмические данные метода отраженных волн часто не обеспечивают полноценной выборки по горизонтальной координате x . Следовательно, имеет смысл предложить для величины ∂_{xx} более точное приближение к величине δ_{xx} . Возво-

дя в квадрат уравнение (10.6.4) и оставляя только первые два члена в разложении функции Arsh , найдем:

$$-k_x^2 \approx \frac{4}{\Delta x^2} \left(u^2 - \frac{u^4}{3} \right), \quad (10.6.6)$$

где $u = -i\hat{k}_x \Delta x / 2$. Применяв обратное преобразование, получим:

$$\partial_{xx} \approx \frac{\delta_{xx}(1 - \delta_{xx}/12)}{\Delta x^2}. \quad (10.6.7)$$

Чаще всего удобно использовать это выражение в форме рациональной дроби:

$$\partial_{xx} \approx \frac{\delta_{xx}/\Delta x^2}{1 + \delta_{xx}/12}. \quad (10.6.8)$$

С помощью особого приема выражение (10.6.8) можно использовать для решения задач без перехода к разностным операторам более высокого порядка. Отметим, что использование соотношения (10.6.8) в дифференциальном уравнении типа $P_t = P_{xx}$ приводит к формуле

$$\left(1 + \frac{\delta_{xx}}{12} \right) \delta_t P = \frac{\Delta t}{\Delta x^2} \delta_{xx} P. \quad (10.6.9)$$

Новый член δ_{xxt} подходит к старой схеме вычислений и, таким образом, равнозначен изменению коэффициентов; т. е. едва ли приведет к какому-либо увеличению затрат времени на ЭВМ. Данные, приведенные в табл. 2, свидетельствуют об увеличении точности решения. На основе табл. 2 и допустимой погрешности для некоторого частного применения, например на уровне 3%, можно определить минимальное допустимое число точек на длину волны, например по 10 точек на длину волны по осям z и t и $3\frac{1}{2}$ точки на

длину волны по оси x . Тогда полезная ширина полосы частот составит $-2\pi/10 < \omega \Delta t < +2\pi/10$, (где 2π является интервалом периодичности выборки сейсмических данных). Это значительно меньше полной полосы частот сигнала. В этом случае отношение полезной полосы частот к полной равно 1/5. Чтобы использовать большую часть имеющейся полосы частот, необходимо допустить большую погрешность или разработать более сложные разностные приближения для дифференциальных операторов. Волновое поле можно рассматривать как имеющее энергию всюду в пространстве (ω, k_x) . Однако разностные уравнения должным образом моделируют волновое уравнение лишь в ограниченных областях, заштрихованных на рис. 10.25. Энергия составляющих волнового поля с $|k_x| > |\omega/v|$ соответствует поверхностным волнам, погрешностям сбора данных (погрешности в статических поправках, случайные шумы или несглаженные флуктуации сигнала от трассы к трассе) и другим помехам. Эти составляющие не могут ничего внести

в процесс миграции и, следовательно, они должны быть вырезаны с помощью веерного фильтра [36] или, как сделано в данной книге, с помощью численной вязкости. По практическим соображениям часто оказывается необходимым вырезать лучи с углами падения, лежащими вне определенного диапазона. Такую возможность дает широкодиапазонный веерный фильтр, вырезающий область $|k_x| > |\omega/v \sin \theta|$ (θ — угол наклона). Несмотря на то что

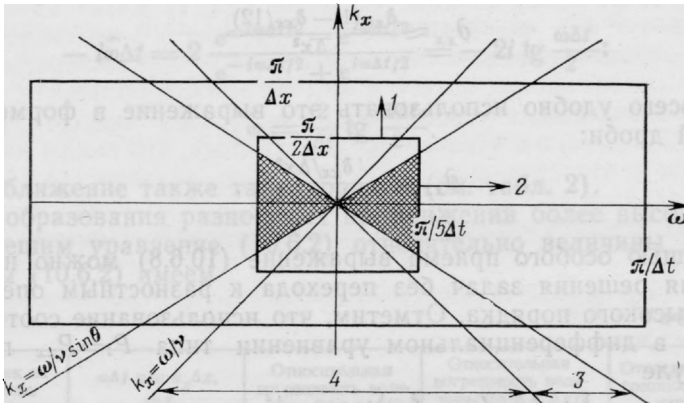


РИС. 10.25. Волновое поле в плоскости (ω, k_x) .

1 — k_x очень большие; 2 — ω очень большие; 3 — лучи, слишком близкие к горизонтальной плоскости; 4 — область непропускания в направлении оси z

информация может содержать $1/2$ частоты кодирования (частоту Найквиста) как по ω , так и по k_x , на практике применяются операторы конечной длины с узкой полосой пропускания. Использование простых время-разностных операторов дает в результате полосу пропускания около $1/4$ частоты Найквиста. В принципе частота кодирования не представляет трудностей, так как данные могут быть интерполированы перед обработкой. Кроме того, можно применять более сложные (т. е. более длинные) разностные операторы.

Управление численной вязкостью решения рассмотрим на примере дифференциального уравнения вида $P_{zt} = (v/2)P_{xx}$. Для ослабления решения вне используемой полосы частот, где $|k_x \Delta x| > \pi/5$, добавим новый член к правой части, т. е. изменим выражение

$$\partial_z P = \frac{v}{-2i\omega} \partial_{xx} P \quad (10.6.10)$$

путем соответствующего выбора дополнительного члена. Тогда

$$\partial_z P = \frac{v}{-2i\omega} \partial_{xx} P + a \partial_{xx} P. \quad (10.6.11)$$

Чтобы определить значение константы a , преобразуем координату x в уравнении (10.6.11):

$$\partial_z P = \left(\frac{vk_x^2}{2i\omega} - ak_x^2 \right) P. \quad (10.6.12)$$

Уравнение (10.6.12) имеет следующее решение:

$$P(z) = P(z_0) \exp \left[\left(\frac{vk_x^2}{2i\omega} - ak_x^2 \right) (z - z_0) \right]. \quad (10.6.13)$$

Мнимая часть экспоненты дает только фазовый угол, который можно не учитывать, так как нас интересует лишь амплитуда. Пусть $z - z_0 = d$. Тогда выражение (10.6.13) принимает вид

$$\left| \frac{P(z)}{P(z_0)} \right| = \exp(-ak_x^2 d). \quad (10.6.14)$$

Теперь нужно решить, какое ослабление нам необходимо. Предположим, мы хотим, чтобы при $k_x \Delta x = \pi/4$ выражение (10.6.14) давало ослабление в e раз. Тогда из уравнения (10.6.14) имеем

$$\begin{aligned} -ak_x^2 d &= -1; & a \left(\frac{\pi}{4} \right)^2 \frac{d}{\Delta x^2} &= 1; \\ \alpha_x (k_x \Delta x)^2 \frac{d}{\Delta x^2} &= 1; & a &= \frac{16\Delta x^2}{d\pi^2}. \end{aligned} \quad (10.6.15)$$

Таким образом, для получения выражения (10.6.11) в формулу (10.6.10) был добавлен член с коэффициентом a , который стремится к нулю, если квадратная сетка имеет промежутки Δx^2 . Включение этого члена дает гауссову функцию затухания пространственной частоты для выражения (10.6.14). Введение члена с вязкостью, по-видимому, существенно не увеличит затрат времени на ЭВМ.

Теперь модифицируем уравнение экстраполяции волн таким образом, чтобы ослабить чрезмерные наклоны $[\sin \varphi = kv/\omega]$. Строго говоря, это не чистая численная вязкость, так как мы изменили основное дифференциальное уравнение. Это похоже на численную вязкость тем, что используется специально разработанная модификация уравнения, предназначенная для исправления определенного недостатка. Модифицируем дифференциальное уравнение (10.6.10) для изучения выражения

$$\partial_z P = \frac{v \partial_{xx}}{2(-i\omega + \omega_0)} P. \quad (10.6.16)$$

Чтобы подобрать численное значение величины ω_0 , освободимся от иррациональности в знаменателе. Тогда

$$\partial_z P = \frac{v}{2} \frac{i\omega + \omega_0}{(\omega^2 + \omega_0^2)} \partial_{xx} P. \quad (10.6.17)$$

Теперь можно пренебречь мнимой частью в правой части уравнения (10.6.17), поскольку она влияет только на фазу решения P . Применяв преобразование Фурье к выражению (10.6.17) по координате x , получим:

$$\partial_z P = - \frac{v\omega_0 k_x^2}{2(\omega^2 + \omega_0^2)} P. \quad (10.6.18)$$

Здесь возможны два случая. Выберем величину ω_0 очень маленькой с тем, чтобы в интервале частот, не представляющем интереса ($\omega < \omega_0$), уравнение (10.6.18) обеспечивало бы поглощение пространственных частот, а в рабочем диапазоне частот ($\omega > \omega_0$) оно было бы равносильно выражению

$$\partial_z P = \frac{-v\omega_0}{2} \frac{k_x^2}{\omega^2} P. \quad (10.6.19)$$

Очевидно, что выражение (10.6.19) описывает затухание, являющееся гауссовой функцией угла наклона. Определение численного значения частоты ω_0 оставим для упражнений.

Упражнения

1. Какая частота ω_0 в уравнении (10.6.16) может ослабить в e раз исходную амплитуду волны, распространяющейся с уровня z_1 до уровня z_2 под углом 30° к оси z ? Таким образом, можно считать, что частота ω_0 должна быть малой величиной. Необходимо сравнить ее с величиной, имеющей физическую размерность обращенного времени. Приведите примеры ситуаций, когда частота ω_0 является малой величиной, и ситуацией, когда это условие не наблюдается.

2. Покажите, что параметр b в выражении $P = iv/2\omega(\partial_{xx} + b\partial_{xxx})P$ может применяться для осуществления затухания компонент, разрушающих решение в приближении вида $\exp[-bk^2z(z-z_0)]$. Это может быть полезно, когда величина Δz выбрана слишком большой.

3. Рассмотрите экстраполяцию волн на одном шаге в направлении оси z с помощью уравнения $P_z = -a\omega^2 P$. Примените билинейное преобразование $-i\omega = 2(1-Z)/(1+Z)$ и установите, что рассматриваемое уравнение не может быть использовано, так как результирующий многочлен обладает неминимально фазовым делителем.

4. Покажите, что уравнение $P_z(-\omega^2 \Delta t^2/2 + i\omega \Delta t)P$, подобное уравнению, использованному в упражнении 3, приводит во временной области к фильтру, удовлетворяющему принципу причинности. (Осуществите экстраполяцию поля по z с помощью схемы Кранка — Николсона, т. е. методом билинейного преобразования).

5. Заданная выборка данных $P(x, t)$ должна удовлетворять уравнению $P_{zt} = P_{xx}$. Замечено, что преобразованные данные $Q(x, t)$, где $Q(x, t) = P(x, t)e^{\alpha t}$, соответствуют умеренно малому ряду чисел, ввиду чего ряд Q может демонстрировать применение арифметики целых чисел. Какому дифференциальному уравнению должна удовлетворять величина Q ?

11. ОБРАБОТКА СЕЙСМИЧЕСКИХ ДАННЫХ НА ОСНОВЕ ВОЛНОВОГО УРАВНЕНИЯ

Система координат, применяемая теоретиками для описания процессов распространения сейсмических волн, не совпадает с системой, используемой на практике для описания схемы наблюдений. Теоретики обычно рассматривают процесс во времени для отдельного положения источника (или взрыва), а экспериментаторы наблюдают волны, возбужденные множеством отдельных источников. Наша задача заключается прежде всего в том, чтобы представить волновое уравнение в удобной для геофизиков-разведчиков системе координат.

11.1. ПРОДОЛЖЕНИЕ В НИЖНЕЕ ПОЛУПРОСТРАНСТВО СЕЙСМИЧЕСКИХ ДАННЫХ И РАЗРЕЗОВ

В морской сейсморазведке исследования проводятся на корабле, который оснащен источником импульсных колебаний периодического действия и буксирует за собой косу со звуковыми приемниками длиной от 2 до 3 км. В идеале курс корабля является прямой линией, которую можно принять в качестве оси x . В благоприятных условиях можно полагать, что все сейсмические волны, представляющие для нас интерес, распространяются в вертикальной плоскости, проходящей через линию курса корабля. Эта плоскость называется плоскостью сейсмического разреза. Описание волнового процесса в трех измерениях не представляет особой сложности. Однако техника реализации конечно-разностных методов пока разработана для двух измерений, поэтому ограничим теорию двумя измерениями. Кроме того, это ограничение согласуется с современной методикой сейсмической съемки и возможностями наиболее совершенных вычислительных машин. Амплитуда волнового импульса, распространяющегося в трехмерной среде от точечного источника, затухает обратно пропорционально времени распространения волны. (Площадь сферического фронта волны возрастает пропорционально t^2 , следовательно, энергия на единицу площади уменьшается как t^2 , а расхождение амплитуды волны — пропорционально t^{-1}). Волновой фронт от линейного источника, расположенного на поверхности перпендикулярно к курсу корабля, имеет расхождение, пропорциональное $t^{-1/2}$. Таким образом, при любой попытке сжать трехмерную среду в одно- или двухмерную математическую модель надо учитывать необходимость введения масштабного множителя $t^{1/2}$ либо t . Значительно большее значение для практики, чем учет расхождения волн, имеет выделение всех сей-

смических лучей, распространяющихся от источника до приемников, лежащих на траектории движения корабля в одной плоскости. Другими словами, следует избегать регистрации сигналов, соответствующих боковым отраженным волнам. Эти сигналы часто могут быть ослаблены или уничтожены благодаря правильному выбору курса корабля. Однако если такие данные записаны, то необходимо определить, с какой стороны пришли отражения. Один из способов представления боковых отражений заключается в том, чтобы считать линию движения корабля осью цилиндрической системы координат. При этом время задержки отраженного сигнала не является мерой глубины до отражателя, а характеризует радиус цилиндра, на поверхности которого располагается отражатель. Интерпретация облегчается, если плоскость сейсмического профиля слабо наклонена к вертикали. Задача интерпретации значительно усложняется в условиях, когда структура среды во всех трех измерениях настолько сложна, что существует несколько плоскостей, в которых распространяются колебания, а данные наблюдений являются суперпозицией многих из них. Иначе говоря, в условиях, когда среда неоднородна во всех направлениях, нельзя ограничиться двумерными данными и соответствующими им вычислительными методами.

На рис. 11.1 показаны наиболее важные связи между двумя системами координат. Координаты взрывного источника звука s и приемника звука g задаются вдоль курса корабля, который является осью x . Вдоль этой оси определяется расстояние взрыв — прибор, т. е. координата смещения ($f=g-s$) и средней точки $y=(g+s)/2$ между источником и приемником. Теоретики обычно используют величины s и g как координаты волнового поля давления, а интерпретаторы — величины f и y .

Рассмотрим волны давления $P(s, g)$, для которых координаты точек взрыва и приема являются независимыми переменными. В действительности источники и приемники не распределены непрерывно вдоль оси x , но обычно они расположены достаточно близко друг к другу, в результате чего давление P для любых точек s и g может быть найдено с помощью интерполяции. Если данные наблюдений не могут быть проинтерполированы, то они не удовлетворяют требованиям, необходимым для использования их в дифференциальных уравнениях.

Другой независимой переменной является время t . Начальная точка на оси времени выбирается таким образом, чтобы время t равнялось нулю до момента взрыва. После того, как отраженные сигналы от взрыва в точке s полностью затухнут (обычно это занимает около 6 с), в точке $s+\Delta s$ производится следующий взрыв. При этом время отсчитывается от нуля. На рис. 11.1 ось t можно считать находящейся за плоскостью листа. Обе системы координат как (s, g) , так и (y, f) ортогональны. Неортогональные координаты, такие как (s, f) , используются при морских наблюдениях, но они редко применяются в процессе анализа данных, ввиду чего мы рассматривать их не будем.

Теоретики обычно работают в плоскости (g, t) при фиксированном s . В сейсморазведке данные в плоскости (g, t) называются профилем. Интерпретаторы сейсмических данных обычно работают с амплитудой волны как в плоскости (y, t) , так и в плоскости (f, t) . Представление данных наблюдений в плоскости (y, t) называется сейсмическим разрезом, а в плоскости (f, t) — выборками общих средних точек или общих точек отражения, которые часто

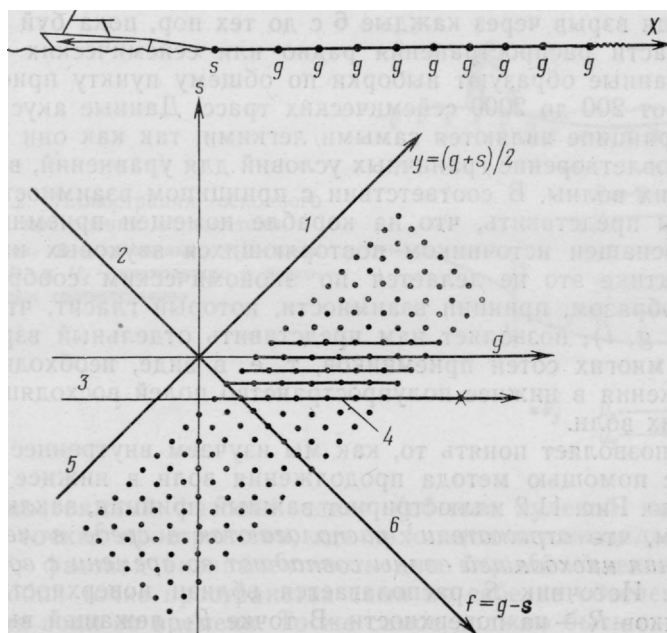


РИС. 11.1. Соотношения координат источника s , приемника g , смещения $f=g-s$ и средней точки $y = (g+s)/2$.

1 — ось средних точек; 2 — выборка общих средних точек; 3 — профиль; 4 — сейсмическая коса; 5 — разрез; 6 — ось смещений (выносов)

называются выборками общих глубинных точек. Эта терминология появилась, когда строение реальных геологических сред обычно представляли с помощью моделей горизонтально-слоистой среды, в которой точки отражения, называемые глубинными точками, располагаются непосредственно под средними точками. Такое представление неудачно из-за значительной путаницы, которая связана с осью глубин. Мы будем избегать термина «глубинная точка». Представление данных над средней точкой y при фиксированном смещении f , т. е. в плоскости (y, t) , является единственной из ортогональных плоскостей, которая может продолжаться на сотни и тысячи километров. Профили и выборки имеют длину в несколько километров, что связано с ограниченной длиной сейсмической косы и с ограниченными размерами области (что более существенно), в пределах которой могут быть зарегистрированы отражения, вызванные взрывом.

Профили и выборки в настоящее время обычно содержат около 48 сейсмических трасс. Так как полезные данные часто наблюдаются на расстояниях, превышающих размеры 3-километровой косы, то иногда применяется другой вид регистрации данных с помощью акустического буя, который представляет собой буй с одним приемником звука (называемым гидрофоном) и радиопередатчиком. Буй выбрасывается за борт корабля, который удаляется от него, производя взрыв через каждые 6 с до тех пор, пока буй окажется вне области распространения радио или сейсмических сигналов. Такие данные образуют выборки по общему пункту приема и содержат от 200 до 2000 сейсмических трасс. Данные акустического буя в принципе являются самыми легкими, так как они обеспечивают удовлетворение граничных условий для уравнений, восстанавливающих волны. В соответствии с принципом взаимности можно было бы представить, что на корабле помещен приемник звука, а буй оснащен источником повторяющихся звуковых импульсов. На практике это не делается по экономическим соображениям. Таким образом, принцип взаимности, который гласит, что $P(g, s, t) = P(s, g, t)$, позволяет нам представить отдельный взрыв с помощью многих сотен приемников, т. е. в виде, необходимом для продолжения в нижнее полупространство полей восходящих и нисходящих волн.

Это позволяет понять то, как мы изучаем внутреннее строение Земли с помощью метода продолжения волн в нижнее полупространство. Рис. 11.2 иллюстрирует важный принцип, заключающийся в том, что *отражатели*¹ *располагаются в среде в местах, где вступления нисходящей волны совпадают во времени с восходящей волной*². Источник S располагается вблизи поверхности, а ряд приемников R — на поверхности. В точке P_1 , лежащей выше отражателя, нисходящая волна D_1 наблюдается раньше восходящей волны U_1 , которая прошла от источника до отражателя, а затем вернулась вверх в точку P_1 . В точке P_2 , которая находится в непосредственной близости к отражающей границе или на самой границе произвольной формы, нисходящая и восходящая волны D_2 и U_2 перекрываются во времени, что может быть использовано для определения местоположения отражателя. В точке P_3 , расположенной ниже отражателя, в принципе нет восходящей волны. Однако практические процедуры вычисления восходящих волн U на различных глубинах большей частью равнозначны сдвигу восходящих волн, наблюдаемых в точках R , ко все более ранним моментам времени, соответствующим все большему и большему глубинам. Поскольку эти процедуры не содержат знания о положении границы раздела, то они будут прогнозировать показанную пунктиром ложную восходящую волну U_3 в точке P_3 до прихода туда нисходящей волны [5].

¹ Отражателями могут быть как протяженные гладкие отражающие границы, так и дифрагирующие объекты. (Прим. перев.)

² Этот принцип совпадения времен широко используется в способах огибающих и лучевых диаграмм, а также в методе полей времен. (Прим. перев.)

Чтобы наилучшим образом пояснить эту идею, используем монохроматические волны. Это позволяет рассчитать и представить рассматриваемый процесс в плоскости (x, z) для фиксированной частоты ω , вместо трехмерного пространства (x, z, t) . В случае монохроматических волн представление о совпадении во времени двух зависящих от времени волновых сигналов трансформируется в представление о том, что обе волны являются когерентными и

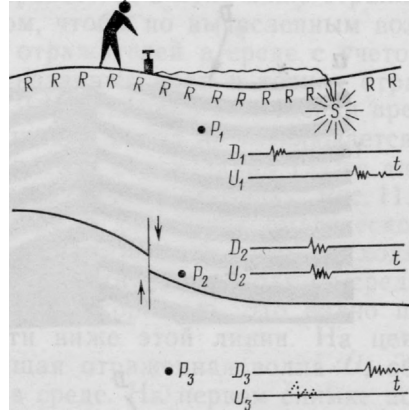


Рис. 11.2. Иллюстрация основного принципа построения отражателя. S — источник; R — приемники; P_i — точки в среде; D_i и U_i — нисходящие и восходящие волны соответственно

имеют постоянный фазовый сдвиг (обычно нулевой или равный 180°) в точках пространства, где находится отражатель. Мы увидим, что фазовая эквивалентность для каждой отдельной частоты в любой точке пространства сама по себе не обеспечивает совпадения волн во времени. То же самое должно быть справедливо и для других частот. Как правило, чем больше частот в фазовой когерентности, тем выше будет разрешение в пространстве.

Напомним, что предположения, сделанные в разделе 10.5, привели к понятию о том, что восходящие и нисходящие монохроматические волны можно рассчитать по формулам:

$$D'_z = \frac{iv}{2\omega} D'_{xx}; \quad (11.1.1)$$

$$U'_z = -\frac{iv}{2\omega} U'_{xx} - c(x, z) D' e^{2imz}. \quad (11.1.2)$$

В *прямой задаче* волновое поле рассчитывается на основе модели среды. В *обратной задаче* на основе наблюдаемого волнового поля определяется модель среды. Для постановки прямой задачи предположим, что взрыв производится в точке с координатами $z=0, x=0$. Это обеспечивает начальные условия типа дельта-функции для нисходящей волны. При заданной модели скоростной характеристики среды уравнение вида (11.1.1) позволяет продолжить нисходящую волну в нижнее полупространство на произвольную глубину. На достаточно большой глубине восходящая волна U

принимается равной нулю. Произведение коэффициента отражения c и нисходящей волны D в выражении (11.1.2) может рассматриваться как источник восходящей волны.

При решении обратной задачи необходимо определить коэффициент отражения границы $c(x, z)$ по заданным на поверхности Земли (при $z=0$ для всех x) результатам наблюдений восходящей волны U . При этом волну D можно рассчитать по приведенной выше формуле. Коэффициент отражения $c(x, z)$ в выражении

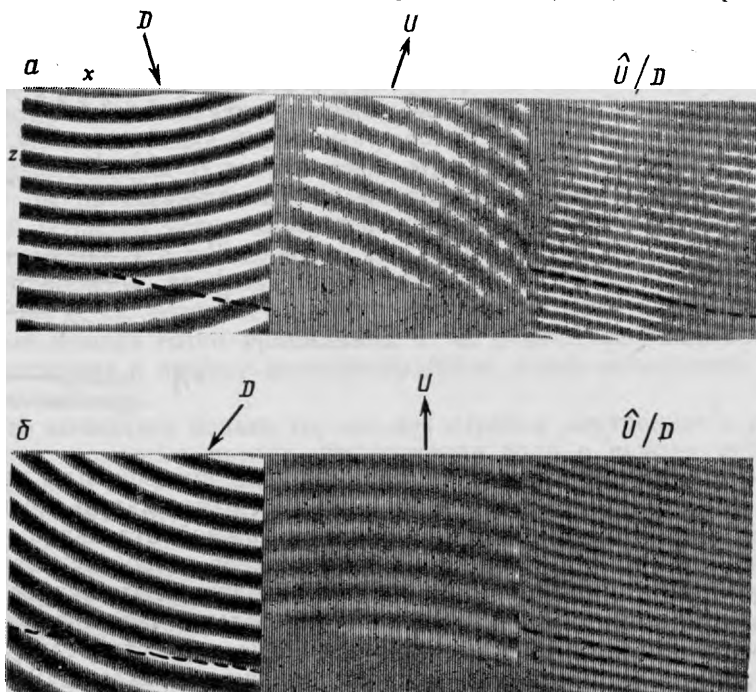


РИС. 11.3. Волновое поле одиночного источника гармонических колебаний. D — нисходящая волна; U — восходящая волна; $a, б$ — различные положения источника

(11.1.2) неизвестен. С помощью пошагового решения уравнения можно определить волну \hat{U} , являющуюся приближением к волне U

$$\hat{U}'_x = -\frac{i\nu}{2\omega} \hat{U}'_{xx} = 0, \quad (11.1.3)$$

в нижнем полупространстве, начиная с поверхности $z=0$, где величина \hat{U} задана.

Возникает вопрос, насколько величина \hat{U} отличается от U ? Между поверхностью Земли и наиболее мелко залегающим отражателем с ненулевым коэффициентом отражения c разницы между выражениями (11.1.2) и (11.1.3) не будет. В этой области величины U и \hat{U} описывают волны с идентичными скоростями,

направлениями распространения, коэффициентами расхождения и всеми другими свойствами. Если выражение (11.1.2) рассматривается для направленной вниз волны (противоположном направлению распространения восходящей волны), то члены (характеризующие источник) в уравнении (11.1.2) позволяют «отключить» восходящие волны до тех пор, пока поле U в отличие от поля \hat{U} тождественно не обратится в нуль в области ниже самого глубокого отражателя. Естественно, возникает вопрос: имеет ли практическое значение тот факт, что поле \hat{U} не обращается в нуль ниже отражателей. Наша цель заключается в том, чтобы по вычисленным волнам \hat{U} и D определить положения отражателей в среде с учетом принципа временного совпадения указанных волн в точках отражателя, как показано на рис. 11.2. Для принципа совпадения времен не имеет значения, что восходящая U волна не «отключается» ниже границы раздела. Эти представления справедливы также для многих отражателей, залегающих на различных глубинах. Рис. 11.3 иллюстрирует применение этого принципа для монохроматической волны. На левом снимке показана действительная часть нисходящей волны D . Пунктирная линия соответствует переходу от среды с низкой скоростью к среде с высокой скоростью. Это видно по увеличению длины волны в области ниже этой линии. На центральном снимке показана восходящая отраженная волна U , образующаяся при разрыве скорости в среде. На первом снимке построена действительная часть отношения \hat{U}/D . Как видно, нисходящая и восходящая волны имеют одинаковую фазу. Это дает правильный угол наклона границы. Для определения глубины одной частоты недостаточно. Вычисленная восходящая волна \hat{U} определяется по наблюдаемой на поверхности восходящей волне и скорости распространения волн в среде v ; величина Y_i/Y при этом не использовалась [3]. Необходимый для построения модели геологической среды коэффициент отражения $c(x, z)$ оценивается величиной

$$\hat{c}(x, z) = \frac{\hat{U}(x, z)}{D(x, z)}. \quad (11.1.4)$$

Для монохроматических волн совпадение фазы волн \hat{U} и D наблюдается не только на отражателе, но и на интервалах в половину длины волны выше и ниже отражателя. (Если бы мы имели волну U вместо \hat{U} , то ниже отражателя было бы нулевое поле). Теперь можно повторить все вычисления для монохроматических волн с другими частотами и просуммировать результаты. Суммирование волн будет всегда происходить синфазно на отражателе, а вне отражателя — с различными фазовыми сдвигами.

На рис. 11.4 показаны изображения, подобные тем, которые представлены на рис. 11.3, и отличающиеся тем, что в уравнениях при расчетах учтено больше членов (для лучшего представления волн, распространяющихся под большими углами к вертикали). Кроме того, использовался набор частот (для иллюстрации конструктивной интерференции волн на отражателе). На кадре 1 по-

казана волна, распространяющаяся от точечного источника, на кадре 4 — восходящая отраженная волна. Она исчезает ниже границы раздела. На кадре 5 показана восходящая волна, восстановленная по наблюдениям, проведенным на поверхности Земли. В процессе восстановления не используются данные о положении отражателей, так как они являются предметом поиска. Следовательно, в восстановленной возвращающейся волне восходящая волна имеется и в области ниже отражателя. На кадре 2 показано произведение восходящей волны на комплексно-сопряженную ве-

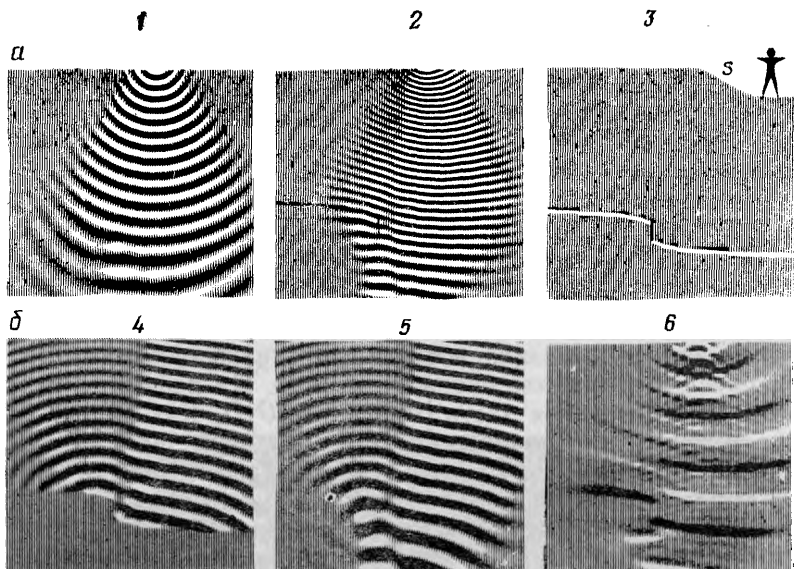


РИС. 11.4. Синтез волн от деформированного отражателя и восстановление изображения границы.

a, б — различные положения источника

1 — нисходящая волна; *2* — восстановление модели в поле монохроматической волны; *3* — модель; *4* — восходящая волна; *5* — восстановленная восходящая волна; *6* — восстановление модели по четырем частотам

личину для нисходящей волны. Отражатель располагается вдоль одной из линий нулевой фазы, но при использовании одной частоты нельзя сказать, какая из линий соответствует отражателю. Кадр 2 построен по четырем частотам, а результаты просуммированы. Полученное (кадр 6) изображение дает более точные сведения о положении отражателя. Увеличение числа частот повысило бы точность определения его положения [5]. Рис. 11.3 и 11.4 различаются еще и тем, что оценка свойств отражателя на первом из них имеет вид \hat{U}/D , т. е. определяется выражением (11.1.3), а на втором она равна $\hat{U}D^*$ (где D^* — комплексно-сопряженная величина D). Оценки \hat{U}/D и $\hat{U}D^*$ имеют одинаковые фазы, но различные амплитуды. Преимущество оценки \hat{U}/D — совпадение ее с коэффициентом отражения. Недостаток оценки \hat{U}/D — наличие узловых

точек поля D , которым теоретически должны соответствовать узлы поля U , где может возникнуть практическая проблема деления на малые числа. Вторая оценка свободна от этого недостатка и имеет требуемую характеристику с точки зрения суммирования сигналов не только по многим частотам, но и по многим положениям пункта взрыва. Величина UD^* мала там, где интенсивность нисходящей волны и «освещенность» отражателей в среде низкая и, наоборот, велика там, где «освещенность» их большая.

Рассмотрим теперь важный практический вопрос, как продолжить профиль в нижнее полупространство. Очевидный подход состоит в том, чтобы решить задачу в плоскости (g, t) отдельно для каждого пункта взрыва. Альтернативно, используя принцип взаимности, мы могли бы рассматривать отдельно задачу в плоскости (s, t) для каждой точки приема колебаний. По причинам, которые нам станут ясными позже, значительные практические преимущества представляет продолжение поля в нижнее полупространство непосредственно в координатах (y, f) . Необходимость такого подхода вызывается тем, что большие объемы сейсмических данных получаются при помощи совместно перемещаемых по поверхности Земли отдельных приемников и источников. При этом для каждого положения пункта взрыва существует лишь одна точка приема колебаний. Очевидно, что этих данных недостаточно для того, чтобы начать продолжение волнового поля в нижнее полупространство в плоскости (g, t) . Тем не менее если рассматривать всю совокупность пунктов взрыва, то оказывается, что можно продолжить волновое поле, полученное при постоянном смещении, вниз по разрезу с хорошей точностью¹.

Рассмотрим два отдельных уравнения для нисходящих и восходящих волн. Переход от координат наблюдаемого поля (s, g, e, t) к интерпретационным координатам (y, f, z, t') осуществляется по формулам

$$y = \frac{s + g}{2}; \quad (11.1.5a)$$

$$f = g - s; \quad (11.1.5б)$$

$$z = e; \quad (11.1.5в)$$

$$t' = t - \left(\pm \frac{e}{v} \right). \quad (11.1.5г)$$

Выражения (11.1.5a), (11.1.5б) описывают преобразование координат пункта взрыва s и приемника g в координаты средней точки y и в смещение. Уравнение (11.1.5в) показывает, что приемник расположен на высоте e в точке на вертикальной оси z . Выражение (11.1.5г) определяет время t' , учитывающее превышения приемника по высоте. Согласно разделу 10.4 знак плюс соответствует нисходящим волнам, распространяющимся в направлении $+z$, а знак минус — восходящим волнам. Важно отметить, что

¹ Это является синтезированием апертуры. (Прим. перев.)

в преобразовании координат (11.1.5) используется постоянная скорость \bar{v} , отличная от пространственно переменной скорости $\tilde{v}(x, z)$, присущей волновому уравнению. Несмотря на то что преобразование координат базируется на постоянной скорости, преобразованное волновое уравнение тем не менее можно применять для описания распространения волн в среде с переменной скоростью.

Теперь попытаемся описать одно и то же возмущение в новой и старой системах координат, т. е.

$$P(g, s, t, e) = Q(y, f, t', z). \quad (11.1.6)$$

Вычислим частные производные функции давления P

$$P_g = Q_y y_g + Q_f f_g + Q_{t'} t'_g + Q_z z_g; \\ P_g = (0,5\partial_y + \partial_f) Q; \quad (11.1.7a)$$

$$P_e = Q_y y_e + Q_f f_e + Q_{t'} t'_e + Q_z z_e; \\ P_e = (t'_e \partial_{t'} + \partial_z) Q; \quad (11.1.7b)$$

$$P_t = (t'_t \partial_{t'}) Q. \quad (11.1.7b)$$

Поскольку все коэффициенты при ∂_y , ∂_f , $\partial_{t'}$ и ∂_z в выражениях (11.1.7) являются постоянными, то вторые производные функции P можно найти с помощью возведения в квадрат операторов в скобках в формулах (11.1.7).

Мы хорошо знакомы с волновым уравнением

$$P_{xx} + P_{zz} = \tilde{v}^{-2} P_{tt} + s(x, z, t). \quad (11.1.8)$$

Предположим, что приемники, использованные для наблюдения величины P , размещены где-нибудь в пространстве (x, z) . Тогда величину, подобную P_{xx} , на поверхности Земли можно измерить системой приемников, расположенных в точках $g=x$. Координаты, определяющие положения приемника $(x, z) = (g, e)$ можно считать независимыми переменными. Тогда волновое уравнение примет вид

$$P_{gg} + P_{ee} = \tilde{v}^{-2} P_{tt} + \delta(g - s, e, t), \quad (11.1.9)$$

где член с дельта-функцией соответствует источнику в пространстве $(x, z, t) = (s, 0, 0)$. Если отказаться от применения этого уравнения в окрестностях источников, то можно отбросить дельта-функцию, соответствующую источнику. Вычислив вторые производные выражений (11.1.7) и подставив их в уравнение (11.1.9), получим:

$$\frac{1}{4} Q_{yy} + Q_{ff} + Q_{y_f} + (t'_e)^2 Q_{t't'} + 2t'_e Q_{t'z} + Q_{zz} = \frac{(t'_t)^2 Q_{t't'}}{\tilde{v}^2}. \quad (11.1.10)$$

Если в последнем выражении пренебречь величиной Q_{zz} , то получим приближение, подобное приближению Френеля. Более высокая точность может быть достигнута, если, как и в разд. 10.3, вычислить величину Q_{zz} . Переходя к однородной среде ($v \approx \bar{v}$) и используя условия $t'_e = -(\pm 1/\bar{v})$ и $t'_i = 1$, преобразуем уравнение к виду:

$$\pm \frac{2}{v} Q_{t'z} = -\frac{1}{4} Q_{yy} + Q_{fi} + Q_{yj}. \quad (11.1.11)$$

Уравнение (11.1.11) является дифференциальным уравнением в частных производных с четырьмя переменными и начальными условиями, заданными с помощью трех переменных. Решение задач такой сложности часто оказывается неосуществимым для вычислений, ввиду чего следует исключать одну переменную, в частности расстояние взрыв — прибор f . С этой целью проинтегрируем уравнение (11.1.11) по переменной f . Тогда

$$0 = \left(\frac{\pm 2}{v} \partial_{z't'} + \frac{1}{4} \partial_{yy} \right) \int Q df + \partial_y \int \frac{\partial Q}{\partial f} df + \int \frac{\partial^2 Q}{\partial f^2} df. \quad (11.1.12)$$

Если величины Q и Q_f обращаются в нуль при больших смещениях, которые мы принимаем в качестве пределов интегрирования, то два последних члена в правой части обращаются в нуль. Тогда вертикально накапливаемые разрезы определяются выражением

$$S = \int Q df. \quad (11.1.13)$$

Накапливание выполняется без временного сдвига, следовательно, это похоже на хорошо знакомое нам (хотя и не совсем то же самое) накапливание общих точек отражения. Таким образом, для вертикально накапливаемых разрезов имеем

$$S_{z't'} = \pm \frac{\bar{v}}{8} S_{yy}. \quad (11.1.14)$$

Выражение (11.1.14) можно получить также из уравнения (11.1.11) в предположении, что для данных наблюдений с нулевым смещением производные величины f отбрасываются. Уравнение (11.1.14) справедливо для области, расположенной в окрестности трассы разреза.

Выражение (11.1.14) не является для нас совершенно новым. Ранее мы уже рассматривали его как уравнение $P_{zi} = 0,5vP_{xx}$, описывающее распространение волнового поля. Это означает, что выражение (11.1.14) должно преобразовывать одни гиперболы в другие, однако в действительности из-за приближения, подобного приближению Френеля (которое допускает усовершенствование), оно преобразует одни параболы в другие. Иными словами, с помощью уравнения (11.1.14) осуществляются преобразования, которые применяются в сейсмической миграции¹.

¹ Под миграцией обычно понимают учет сейсмического сноса для негоризонтальных границ. (Прим. перев.)

11.2. МИГРАЦИЯ НА ОСНОВЕ ВОЛНОВОГО УРАВНЕНИЯ

Построение поперечного разреза отражающих границ внутри Земли по сейсмическому профилю называется *миграцией*. В большинстве случаев под поверхностью Земли залегают горизонтально-слоистые толщи осадочных пород. В таких условиях миграция сейсмических данных предельно проста, так как волны распространяются вертикально по отношению к отражающим границам,

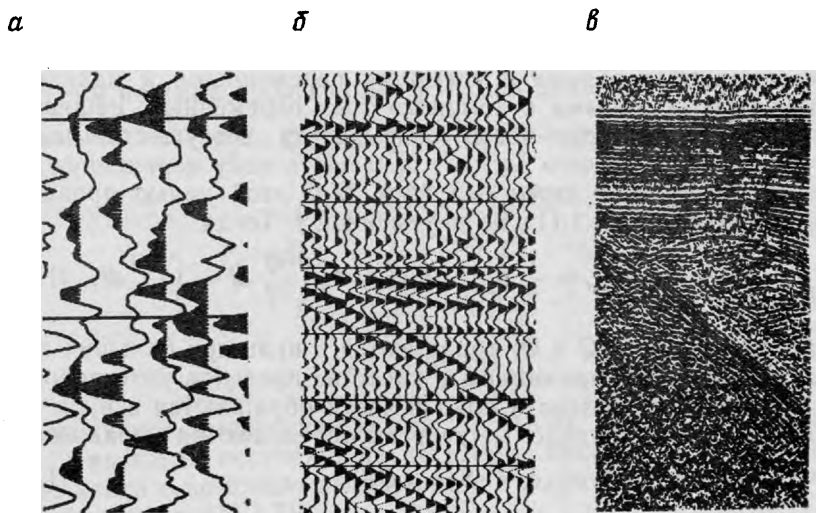


РИС. 11.5. Изображения волновых полей с различной плотностью представления информации (а, б, в). С увеличением плотности информации слева направо запись (а) становится изображением волнового поля (в)

а время их пробега туда и обратно прямо пропорционально глубине. Миграция сводится к введению коэффициента пропорциональности по отношению к оси времен. Волновая картина, представленная на рис. 11.5, показывает распределение акустического давления $P(x, t)$ при $z=0$. Если по оси времен откладывается вертикальное время распространения отраженных сигналов, то распределение акустического давления $P(x, 2z/v)$ характеризует поперечный разрез Земли. В идеальных условиях переходы между светлыми и темными областями на разрезе должны описывать границы между различными типами осадочных пород. К сожалению, вид чередующихся слоев (между черными «породами» и белыми «породами») обычно обманчив. Это связано с эффектами фильтрации в источниках и приемной аппаратуре, а также с влиянием среды. Наиболее явная и, возможно, наиболее важная информация, содержащаяся в сейсмическом разрезе, заключается в отклонении структуры среды от горизонтально-слоистых моделей. Сейсмические разрезы обычно представляются с некоторым увеличением вертикального масштаба (1—20 раз). Например, изо-

браженный на кадрах 1, 4 (рис. 11.5) вертикальный масштаб увеличен в 5 раз, ввиду чего четкий сильный отражатель, который, как кажется, наклонен под углом 45° , в действительности наклонен приблизительно под углом 9° . На основе таких изображений мы пытаемся определить внутреннее строение Земли на данном участке.

Цель миграции состоит в том, чтобы учесть тот факт, что волны распространяются не только прямо вверх и вниз, в местах, где слои внутри Земли залегают негоризонтально. При этом может

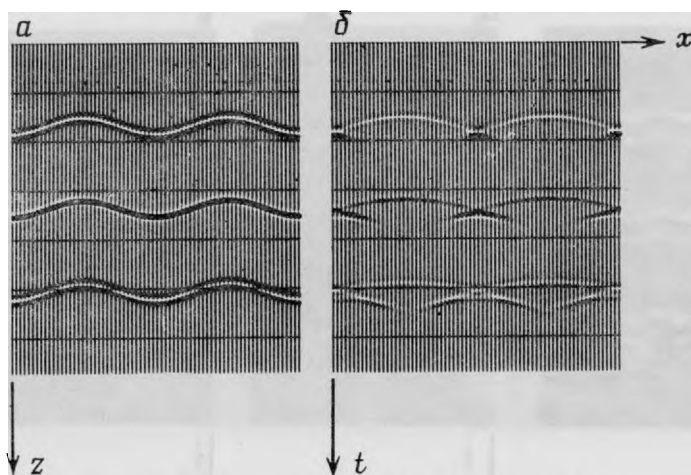


РИС. 11.6. Модель с тремя синусоидальными отражающими границами на различных глубинах (а) и рассчитанные для нее сейсмограммы для нулевого смещения (б) без вертикального увеличения масштаба

сказываться влияние бесконечного числа сложных факторов. Однако применение уравнения $U_{tz} = U_{xx}$ практически позволяет идентифицировать наблюдаемые на изображениях отклонения от горизонтального залегания с реальными обнаруженными поднятиями. Рис. 11.6 иллюстрирует применение миграционных способов, описанных ниже. Все отражающие границы на рис. 11.6 имеют одинаковую форму, ввиду чего каждая из них должна иметь одинаковые углы наклона (около 15°). В то же время отклонение временного разреза от глубинного явно увеличивается с глубиной. Значительное отклонение появляется, когда глубина границы становится сравнимой с наименьшим радиусом ее кривизны (погруженный фокус). Этот радиус приблизительно равен глубине самого мелкого отражателя, для которого сильная фокусировка очевидна [8]. Таким образом, различие между сейсмическим временным и глубинным разрезами определяется главным образом кривизной границ. Следовательно, применение миграции оказывается важным и необходимым даже в областях с относительно плоскими гладкими отражающими поверхностями.

Импульс, падающий на границу раздела, дает отраженный импульс. Однако источнику, приемнику и среде всегда присущ некоторый эффект фильтрации, который преобразует предполагаемый импульсный сигнал в короткий сейсмический импульс длиной в несколько периодов. В идеале на поперечном разрезе Земли импульсы должны отображать границы раздела, но при миграции сейсмический импульс переносится с временного разреза на глубинный. Таким образом, ограничение разрешающей способности, имеющееся на временной оси, преобразуется в ограничение разрешаю-

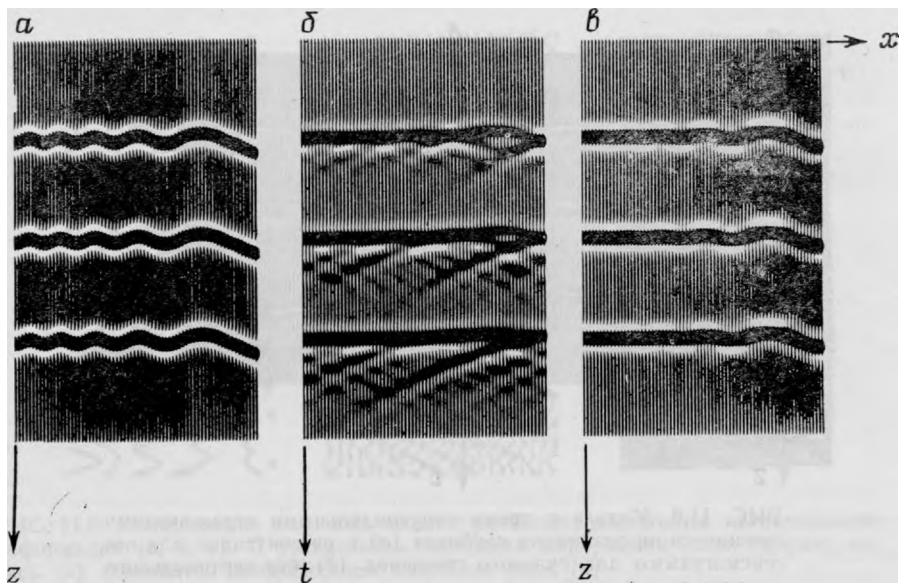


РИС. 11.7. Отражения от ундулирующих границ раздела иллюстрирующие ограничение латеральной (горизонтальной) разрешающей способности [8].

a — исходная модель; *б* — сейсмический временной разрез для нулевого смещения, построенный миграционным способом в варианте 45-градусного приближения и модифицированным путем введения численной вязкости для подавления волн с наклонами, превышающими 30°; *в* — результат восстановления исходной модели

щей способности на оси глубин. Интересный вопрос возникает относительно разрешающей способности по горизонтальной оси x . Для волн, распространяющихся под углом к поверхности, разрешающая способность сейсмического импульса преобразуется к наклонной, а не к вертикальной оси. Очевидно, что наилучшая разрешающая способность по горизонтали, которую можно было бы получить, соответствует волнам, распространяющимся по горизонтали. Для волн, идущих под углом 45°, разрешающая способность по горизонтали будет равна увеличенному в $1/\sin 45^\circ$ раз линейному разрешению по вертикали (рис. 11.7). Как видно из рис. 11.7, ограничение углов наклона регистрируемых волн приводит к снижению разрешающей способности по горизонтали: коротковолно-

вые ундуляции в левой части границы раздела разрешить не удалось (рис. 11.7,в).

Принцип действия миграции может быть пояснен в простых терминах без привлечения волнового уравнения. Рис. 11.8 и 11.9 иллюстрируют преобразование между временем и глубиной для двух специальных случаев, где одна область или другая содержит в себе только одну импульсную функцию [реальные данные содержат множество сигналов в плоскости (x, t)]. Рис. 11.8,а описывает ситуацию, в которой при совместном перемещении источ-

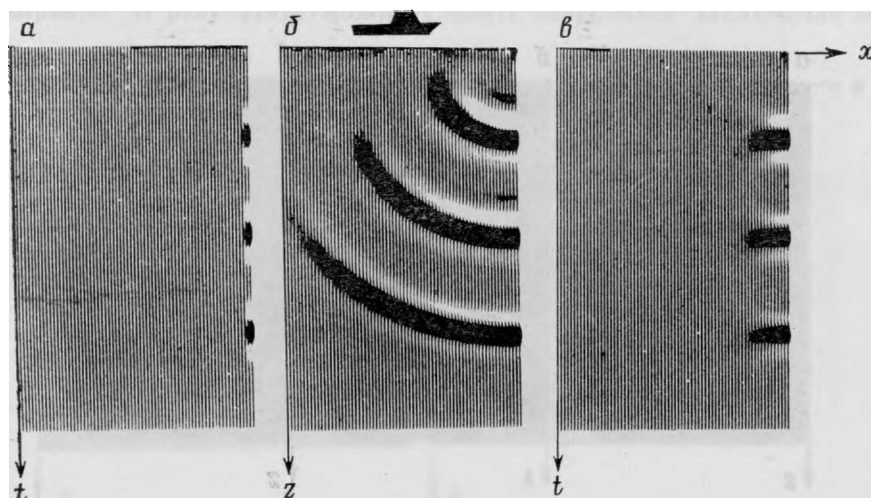


РИС. 11.8. Глубинная реакция на импульсы во временной области и восстановление таких импульсов [8].

а — модель среды; *б* — сейсмическая запись; *в* — результат восстановления модели

ника и приемника в горизонтальном направлении (при нулевом смещении) отраженные сигналы отсутствуют вплоть до правого края кадра, где наблюдаются три эхо-сигнала в последовательно возрастающие моменты времени. На основании этих наблюдаемых данных можно заключить, что структуру Земли характеризуют три границы в виде концентрических окружностей с центрами на правой грани кадра. Эти окружности показаны на рис. 11.8,б (с целью экономии места правым краем кадра является плоскость симметрии). Отметим, что нижние части окружностей темнее верхних. Это указывает на 45-градусный фазовый сдвиг двумерных волн вдали от фокуса. Волны с углами наклона, превышающими 45° , отфильтрованы путем введения численной вязкости. Потери этой энергии, а также энергии волн, распространяющихся под комплексными углами, приводят при восстановлении (рис. 11.8,в) к некоторому расширению импульсов в горизонтальном направлении. На рис. 11.9,а представлена модель среды, состоящая из трех точечных рассеивателей, и соответствующие ей сейсмические записи, на которых выделяются гиперболические кривые времени

распространения волн (рис. 11,9б). На изображении (рис. 11,9в), полученном с помощью миграции, как и на рис. 11,8,в, наблюдается уменьшение разрешения, поскольку горизонтальное разрешение в принципе не может быть лучше вертикального разрешения, обусловленного частотным спектром волн. Кроме того, на практике обычно используются лучи, выходящие под углами приблизительно только до 40° .

Рис. 11.8 иллюстрирует способ миграции, основанный на принципе линейной суперпозиции. При этом берут каждый отсчет дан-

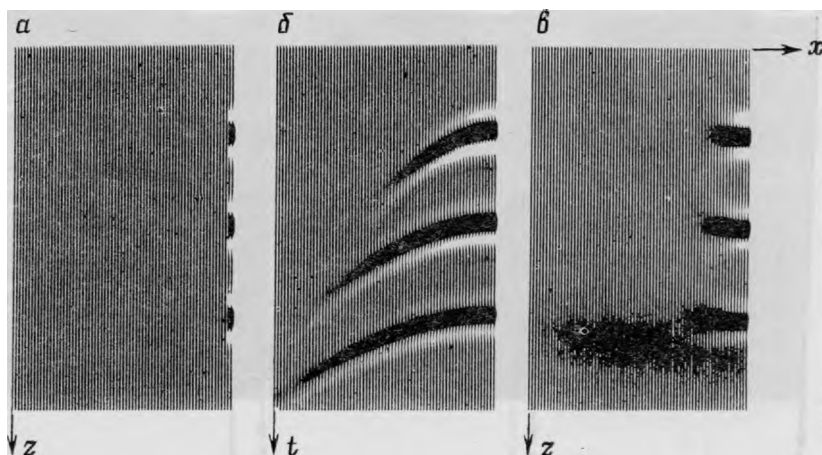


РИС. 11.9. Реакция во времени на импульсы в глубинной области и восстановление этих импульсов [8].

a — модель Земли в виде трех точечных рассеивателей, расположенных друг под другом на различных глубинах; *б* — синтетическая сейсмограмма для модели; *в* — результат миграции

ных в плоскости (x, t) и выносят его на дугу окружности в области (z, t) . Миграционный разрез является суперпозицией всех этих дуг. Можно задать естественный вопрос: зачем использовать волновое уравнение для осуществления миграции, когда ее можно реализовать с помощью дуг окружностей? Ответ на этот вопрос зависит от множества практических факторов, часть из которых определяется исходными данными. В частности, преимущество волнового уравнения заключается в том, что с его помощью неоднородность среды учитывается более просто и точно, чем с помощью построения лучей. Оба способа в настоящее время экономически доступны.

На рис. 11.10 представлен пример, иллюстрирующий происхождение термина «сейсмическая миграция». Он показывает, что глубокая граница раздела, оканчивающаяся в некоторой точке, на временном разрезе будет казаться смещенной (мигрировавшей) ниже по наклону границы из ее истинного положения. В этом

примере отклонение глубинного разреза от временного невелико и уменьшается по мере уменьшения угла наклона границы.

Однако это не означает, что сейсмические данные для районов с незначительными углами наклона границ не нуждаются в миграции (прежде всего, как мы видели, из-за существенного влияния кривизны границы). Интересный пример такого рода показан на рис. 11.11. Большое практическое значение имеет тот факт, что реальная вогнутая структура на временном разрезе кажется выпуклой. Таким образом, синклиналь может быть принята за антиклиналь. В результате можно сделать ошибочное заключение об

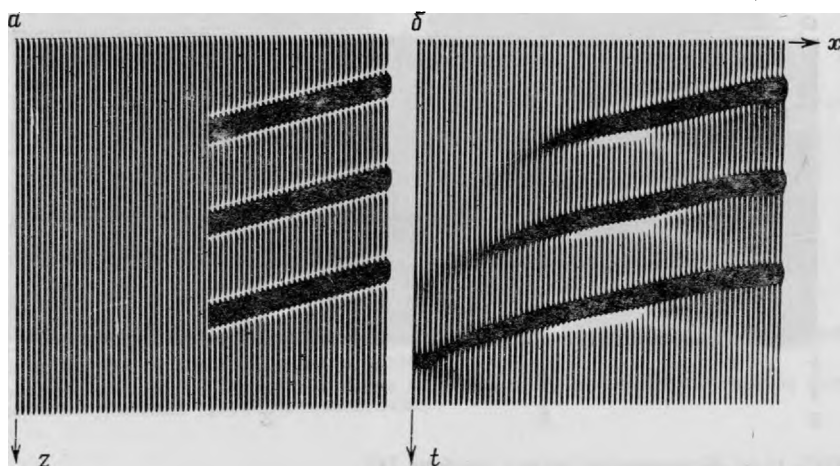


РИС. 11.10. Иллюстрация происхождения термина «сейсмическая миграция» (перенос) [8].

a — глубинный разрез, состоящий из трех ограниченных в пространстве границ раздела; *b* — синтетический временной разрез

условиях образования залежи нефти. Один из множества примеров миграции данных полевых исследований показан на рис. 11.12. На рис. 11.12, *b* видно, что на времени больше 2 с четко выделяется иррегулярная шероховатая отражающая граница. На временном разрезе (рис. 11.12, *a*) эта область нечитаема из-за сильной дифракции.

Объясним, как строить миграционные разрезы с использованием уравнений (11.1.13) и (11.1.14) (при замене t' на t):

$$S(y, z, t) = \int Q(y, f, z, t) df, \quad (11.1.13)$$

$$S_{zt} = \pm \frac{v}{8} S_{yy}. \quad (11.1.14)$$

Уравнение (11.1.13) определяет сумму всех волн по различным смещениям. Уравнение (11.1.14) показывает, как можно экстраполировать эту сумму S вниз от поверхности, на которой она известна. Напомним, что знак плюс или знак минус выбирается в зави-

симости от того, экстраполируются ли нисходящие или восходящие волны. Выражение для источника $\delta(g-s, e, t)$ в уравнении (11.1.9) в координатах средняя точка — смещение принимает вид $\delta(f)\text{const}(y)\delta(z)\delta(t)$. Так как рассматриваемая функция источника является постоянной функцией средней точки y , то можно надеяться свести уравнение (11.1.14) для нисходящей волны к виду $S_{zt}=0$. Общим решением уравнения $S_{zt}=0$ является сумма произвольных функций глубины z и времени t . Поскольку нас интересуют волновые колебания, а не теория потенциала, то мы можем отбросить произвольную функцию глубины z . Сохранение функции времени,

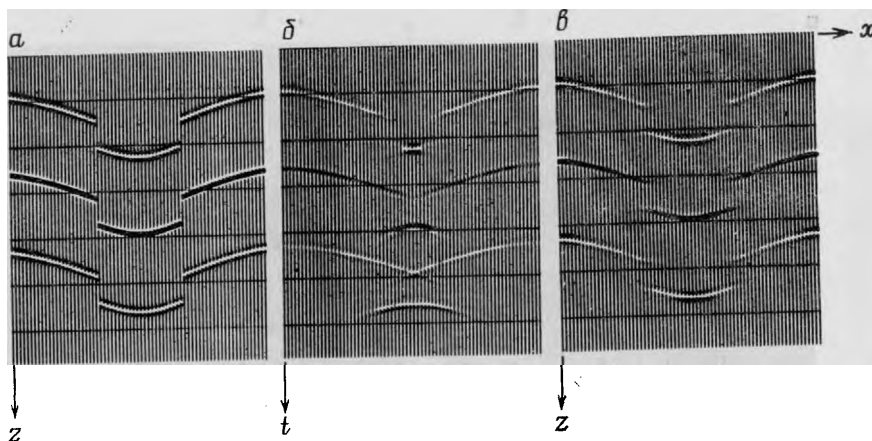


РИС. 11.11. Классическая модель грабена [8].

a — модель среды; *b* — синтетический временной разрез; *в* — результат миграции

не зависящей от глубины, соответствует зависимой от времени функции источника, т. е. дельта-функции времени. Имея теперь аналитическую функцию для сумм S нисходящих волн на всех глубинах, можно обратиться к численной экстраполяции суммы сигналов восходящих волн в нижнее полупространство. Уравнение (11.1.14) со знаком минус дает возможность производить экстраполяцию восходящей волны вниз в нижнее полупространство. Так как нисходящая волна является плоской, то восходящая волна, если смотреть вниз на отражатели, будет принимать форму отражателей. Таким образом, миграционный разрез содержит данные наблюдений на поверхности, экстраполированные вниз от поверхности до необходимой глубины.

Однако на практике исследователи сейчас не используют вертикальное суммирование, определяемое соотношением (11.1.13). Вместо этого они вводят перед суммированием временные сдвиги как функцию смещения. Главный эффект этой операции — сделать все трассы с разными смещениями возможно более похожими на трассу с нулевым смещением. Временная задержка волн, записанных на профиле с нулевым смещением, распределяется между нисходящей и восходящей траекториями движения волн поровну.

В случае, когда нисходящая волна является плоской, а восходящая волна — сложно рассеянной, эти задержки неодинаковы. На основе этих соображений можно теперь догадаться о результате применения уравнения (11.3.18), который заключается в том, что уравнение миграции для профиля с нулевым смещением или сум-

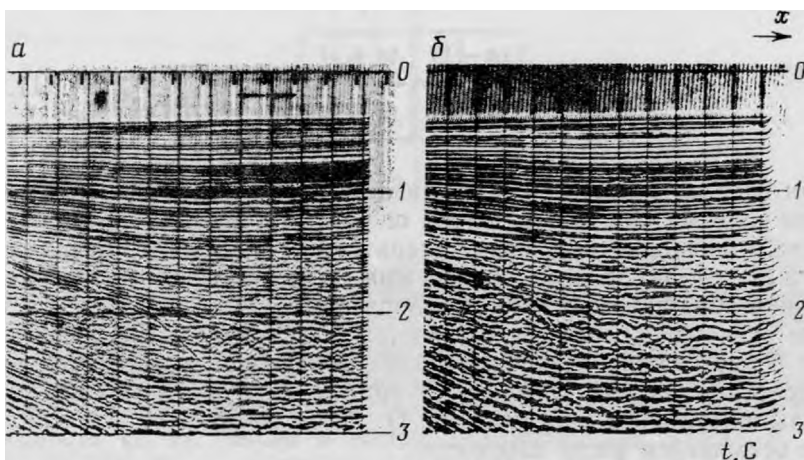


РИС. 11.12. Пример миграции сейсмических данных, полученных при наблюдениях по методу отраженных волн на море: *a* — временной разрез ОГТ (компания Ройал Датч Шелл); *b* — результат миграции на основе волнового уравнения (фирма Диджикон, Инк.)

мирование с введенными кинематическими поправками имеет сходство с уравнением (11.1.14), но содержит дополнительный множитель 2. Таким образом, начнем с выражения

$$Q_{zt} = -\frac{\bar{v}}{4} Q_{yy}. \quad (11.2.1)$$

Допущение о том, что скорость \bar{v} является переменной какой-либо из трех координат (y, z, t) на практике легко оправдать; однако, полагая скорость постоянной по y , можно упростить задачу. Произведя преобразование Фурье по y с помощью e^{iky} , приведем выражение (11.2.1) к виду:

$$0 = \left(\frac{v}{4} k_y^2 - \partial_{tz} \right) Q. \quad (11.2.2)$$

Теперь можно дискретизировать координаты z и t . Тогда функция Q в уравнении (11.2.2) может быть табулирована в плоскости (z, t), а дифференциальный оператор в выражении (11.2.2) принимает вид оператора свертки 2×2 в этой плоскости. Этот оператор имеет вид

$$\frac{vk_y^2}{16} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{\Delta z \Delta t} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (11.2.3)$$

Введем масштабный множитель

$$a = \frac{\Delta z \Delta t v k_y^2}{16}. \quad (11.2.4)$$

Тогда оператор, умноженный на $\Delta z \Delta t$, имеет вид

$$\begin{array}{|c|c|} \hline (a-1) & (a+1) \\ \hline (a+1) & (a-1) \\ \hline \end{array}. \quad (11.2.5)$$

Если оператор (11.2.5) накладывается на произвольное место таблицы величин Q и четыре числа оператора умножаются на четыре числа, лежащие под этим оператором, то смысл выражения (11.2.2) состоит в том, что сумма этих четырех произведений должна обращаться в нуль. Следовательно, если мы найдем в таблице значений Q место, где известны только три из четырех чисел для Q , то можно легко вычислить четвертое, неизвестное число. В действительности значения Q обычно заданы только вдоль нескольких границ в плоскости (z, t) . Этого достаточно для заполнения оставшейся части плоскости.

Миграцию и ее обращение представим на следующей сетке:

						$\rightarrow z$
	0					
	r_1	c_1				
	r_2		c_2			
	r_3			c_3		
	r_4				c_4	
	0	0	0	0	0	
$t \downarrow$						

(11.2.6)

На этой сетке величины r_0, \dots, r_4 представляют наблюдаемую на поверхности сейсмограмму, а величины c_0, \dots, c_4 — миграционный разрез. Нули в нижней строке отражают тот факт, что на достаточно большом времени сейсмограммы обращаются в нуль. Отметим, что при заполнении таблицы количество работы (благодаря дифракции) увеличивается при переходе от r_4 к c_4 по сравнению с переходом от r_2 к c_2 . Это происходит потому, что глубоко залегающие элементы с фиксированным наклоном мигрируют (сносятся) дальше, чем мелко залегающие. Принимая $k_y^2 = 0$, получаем $a = 0$ и описание слоистой земли. Начав с заданных на поверхности величин $r_1 - r_4$ и нижней строки нулей, можно использовать

выражение (11.2.5) для заполнения таблицы (11.2.6). Тогда она принимает вид

0					→ z
r ₁	r ₁				
r ₂	r ₂	r ₂			
r ₃	r ₃	r ₃	r ₃		
r ₄	r ₄	r ₄	r ₄	r ₄	

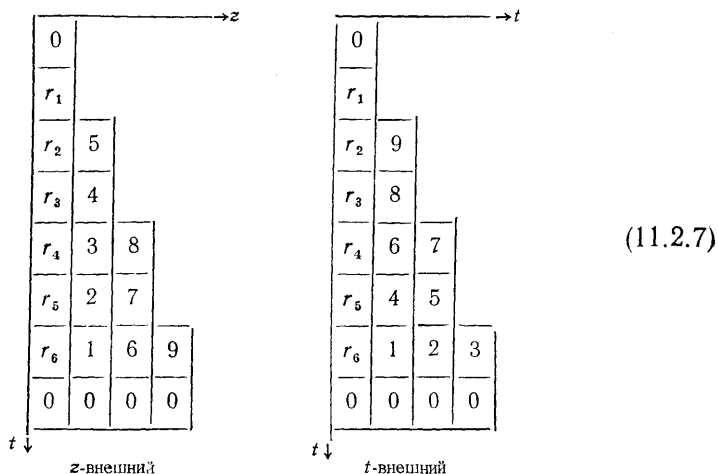
t ↓

(11.2.7)

Исследуя таблицу (11.2.7), видим, что при движении в направлении z все числа остаются неизменными. Если бы мы не приняли условия $k_y^2=0$, а вместо этого считали бы величину k_y^2 малой, то увидели бы, что числа в таблице медленно изменяются в направлении z . В случае $k_y^2 \neq 0$ должны быть сделаны некоторые предположения относительно порядка, в котором должна заполняться плоскость (z, t) . Число a всегда положительное. Очевидно, что если множитель a окажется равным $+1$, то будет невозможно найти неизвестные числа в таблице Q , если они лежат под множителем $1-a$. Оказывается, что независимо от численного (положительного) значения множителя a процесс рекурсивного поиска чисел в таблице значений Q , которые находятся под множителем $1-a$, будет также неустойчив, как деление полиномов на минимально фазовый фильтр. Одна из причин заключается в том, что рассматриваемый процесс действительно включает деление многочленов, причем многочлены в этом случае являются двухмерными. Эти осложнения целиком являются математическим следствием физического принципа причинности.

Для решения гиперболических дифференциальных уравнений самым экономным было бы получение примерно одинакового числа точек на длину волны (обычно восьми) по каждой оси координат. Это означает, что до сих пор мы использовали сильно избыточную выборку по оси z . Использование приближения Френеля с ограничением углов падения до 15° позволяет всегда выбирать величину Δz в 5 раз и более грубее, чем величину $v\Delta t$. Вопрос оптимального выбора шага сетки несколько запутан. Здесь достаточно сказать, что для некоторых данных полевых наблюдений были получены удовлетворительные результаты с помощью миграции при отношении $\Delta z/v\Delta t > 100$ и соответствующем сокращении затрат машинного времени. Для иллюстрации этого пересоставим таблицу разреза с вдвое более грубой выборкой данных по оси z . Числа в двух таблицах, представленных ниже, показывают два различных возможных порядка, удовлетворяющих принципу причинности и используемых в формуле (11.2.5) для вычислений миграции. В любом из этих двух случаев результативный

миграционный разрез интерполируется (возможно, очень грубо) по диагонали:



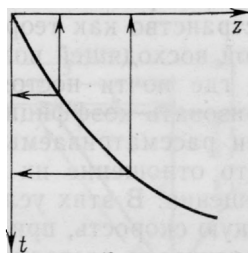
Читатель мог бы проверить в каждой из приведенных выше таблиц, что число, вычисленное на любой стадии, получено на основе трех уже известных табличных строк и что неизвестное, будучи вычислено, умножается на $(1+a)$, а не на $(1-a)$. Чтобы получить синтетические сейсмограммы для гипотетического глубинного разреза, вычисления проводят в обратном порядке.

Изучение процедур (11.2.7) показывает, что первая из них имеет существенное преимущество перед второй процедурой в смысле необходимого объема машинной памяти. На практике необходимо иметь возможность обрабатывать данные со скоростями, изменяющимися с глубиной. Этого можно достичь, используя данные c_1, c_2, \dots , при миграции не по диагонали в плоскости (z, t) , а с другой кривой в плоскости (z, t) , на которой представлена кривая двойного вертикального времени распространения отраженной волны (рис. 11.13). Ось средних точек y , по которым определяется скорость для разных точек профиля располагается перпендикулярно к плоскости рисунка. Для построения разреза данные наблюдений можно проектировать либо на ось z , либо на ось t , образуя миграционный глубинный или временной разрез. Положительным свойством миграции на основе волнового уравнения является возможность легкого изменения значений скорости без необходимости переделывать весь расчет. Если плоскость (z, t) однажды заполнена, то изменение скорости сводится к вопросу, какую кривую на ней использовать.

Наклоны кривых в плоскости (z, t) обычно малы. Это означает, что изменения скорости в модели среды, т. е. кривой на рис. 11.13, проявляются в виде более сильных изменений миграционного глубинного разреза, чем миграционного временного разреза. Поэтому обработчики часто предпочитают временной миграционный разрез.

Последний фактор, который необходимо рассмотреть, это возможность латеральных (горизонтальных) изменений скорости v . Хотя детальный анализ не проводился, можно предположить, что простое включение в уравнение (11.2.1) переменной по y скорости позволяет решать поставленную задачу (если только скорость не изменяется слишком быстро в направлении y). Какое именно изменение скорости является слишком большим, представляет вопрос, который, несомненно, зависит от исходных данных и нахо-

РИС. 11.13. Плоскость (z, t) для определения скорости при миграции в случае вертикально неоднородной среды



дится за рамками настоящей книги. Однако следует отметить, что при использовании зависящей от y скорости преобразование Фурье по оси y не сводит уравнение (11.2.1) к выражению (11.2.2). В результате этого необходимо рассматривать каждый элемент в таблице (z, t) не как скаляр, а как вектор, различные компоненты которого исходят от различных точек на оси y . Тогда мы должны считать величину $1-a$ не скалярным делителем, а тридиагональной матрицей формы $(1-\delta_{yy})$, которая должна быть обратимой. Это не создает каких-либо трудностей при практической реализации. Этот способ был использован при построении миграционных разрезов, приведенных на рисунках в этом разделе.

11.3. ВЫЧИСЛЕНИЕ СКОРОСТИ РАСПРОСТРАНЕНИЯ СЕЙСМИЧЕСКИХ ВОЛН В СРЕДЕ

Преыдушие главы были посвящены задаче определения структуры Земли. При этом скорость распространения сейсмических волн в среде мы считали известной (и для удобства постоянной) величиной, а импеданс имеющим неизвестные разрывы, соответствующие границам раздела различной формы внутри среды. Теперь рассмотрим способы вычисления скорости распространения волн в среде. Обычно это делается в предположении, что среда является горизонтально-слоистой. В этом случае скорость вычисляется на основании зависимости временного сдвига отраженной волны от величины смещения (называемого *нормальным приращением времени*, или *кинематической поправкой*), который наилучшим образом сглаживает данные общей средней точки. Покажем, как можно вычислить скорость распространения волн в среде, состоящей даже из случайных точечных неоднородностей. Это может оказаться полезным при изучении зон нарушений или об-

ластей, где отсутствуют отражения. Под последними понимают области, в которых наилучшим образом обработанный разрез характеризуется отсутствием когерентности колебаний по координате средней точки y . Это может быть следствием как плохого качества исходных данных, так и нарушения непрерывности геологических границ. Однако с теоретической точки зрения последнее обстоятельство не является препятствием для определения скорости распространения волн в среде.

В основе определения скорости лежит продолжение в нижнее полупространство как теоретической нисходящей волны, так и наблюдаемой восходящей волны. Они строятся вниз в сторону отражателей, где почти постоянное отношение их амплитуд должно характеризовать коэффициент отражения как функцию от смещения. Если рассматриваемые волны строятся с правильной скоростью, то отношение их амплитуд будет осциллирующей функцией смещения. В этих условиях задача заключается в том, чтобы найти такую скорость, при которой отношение амплитуд этих двух волн практически постоянно. При этом не важно, имеют ли отражатели какую-либо степень непрерывности по горизонтали или не имеют, так как сглаживание осуществляется для различных смещений в фиксированной средней точке на глубине отражающей границы. Если отражающие границы не обладают непрерывностью по горизонтали, то их можно назвать рассеивателями (неоднородностями). Модель Земли со случайно расположенными рассеивателями будет давать миграционные сейсмические данные, которые являются случайной функцией (исправленную за кинематическую поправку) времени и положения средней точки, но не зависят от смещения.

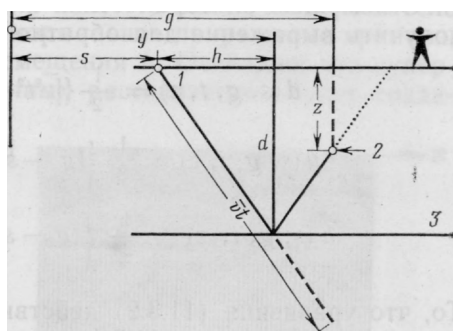
Найти хорошие методы продолжения вниз нисходящих волн достаточно просто. Из точки взрыва эти волны расходятся сферически. Для однородной среды легко записать аналитическое решение. Для слабо неоднородной среды можно воспользоваться методами, рассмотренными в предыдущих главах. Единственная проблема заключается в том, что приближение $Q_{zz} \approx 0$ ограничивает применимость этих методов углами наклона лучей, приблизительно равными 15° от вертикали. Положение легко улучшить путем перехода от декартовых координат (x, z) к полярным координатам (r, θ) . Приближение $Q_{rr} \approx 0$ в полярных координатах требует, чтобы лучи находились внутри угла $\pm 15^\circ$ от направления радиуса. Очевидно, такую систему координат для слоистых сред можно было бы разработать для обработки данных в условиях сред даже с большей скоростной неоднородностью.

Более трудной задачей является поиск хорошей системы координат для восходящих волн. Мне потребовалось два года, чтобы добиться практического решения этой задачи. Как уже указывалось, полярная система координат для нисходящей волны предпочтительнее докартовой системы. Для квазисферической волны величина $Q_{\theta\theta} \approx 0$, тогда как величина Q_{xx} быстро увеличивается, если точка отражения не находится непосредственно под источни-

ками. Поскольку мы имеем дело с уравнениями, подобными $Q_{zt} = Q_{xx}$ или $Q_{rt} = Q_{\theta\theta}/r^2$ то это означает, что величина Q_r обычно будет малой, тогда как величина Q_z мала только на оси z непосредственно под источником. Следовательно, приближение $Q_{rr} \approx 0$ намного лучше, чем $Q_{zz} \approx 0$. Преимущество координат (r, θ) состоит в том, что нисходящая волна $D(r, \theta)$ почти не зависит от горизонтальной (латеральной) координаты θ . Нам необходима система координат, в которой восходящая волна U также почти

РИС. 11.14. Коррекция сейсмических данных за нормальное приращение времени (ввод кинематических поправок) путем продолжения их в нижнее полупространство.

1 — взрыв; 2 — приемник; 3 — отражающая граница (рефлектор)



не зависит от латеральной координаты. Опытные геофизики сразу же скажут, что этому требованию удовлетворяют данные с введенными кинематическими поправками. Коррекция сейсмических записей за нормальное приращение времени (ввод кинематических поправок) приводит к сжатию временной оси на удаленных от источника трассах сейсмограммы, в результате чего волны на этих трассах приходят одновременно с вертикально падающими волнами. Благодаря этому частная производная волнового поля Q по смещению взрыв — прибор при фиксированной кинематической поправке должна быть малой величиной.

Чем ближе полученные сейсмические данные к данным для плоских горизонтальных отражений в Земле с известной скоростью, тем меньше будет производная по смещению. Волновое уравнение учитывает отклонения от такой идеализированной ситуации.

Сжатие временной оси для удаленных от источника трасс сейсмограмм при вводе кинематических поправок может рассматриваться как изменение системы координат. Обычное определение кинематических поправок не предусматривает помещение сейсμοприемников вглубь Земли. При смещении приемников вниз вдоль траекторий лучей мы можем сохранить поверхность средних точек y и поверхность полусмещений $h = f/2$ как горизонтальные (латеральные) координаты волнового поля. Латеральные производные для идеализированных данных должны обращаться в нуль.

На рис. 11.14 показана геометрия коррекции сейсмических данных за нормальное приращение времени путем продолжения их в нижнее полупространство (однородная среда с параметром \bar{v}). Переход от переменных интерпретации к переменным наблюдений

осуществляется с помощью выражений

$$s(h, y, d, z) = y - h; \quad (11.3.1a)$$

$$g(h, y, d, z) = y + \frac{(d-z)h}{d}; \quad (11.3.1b)$$

$$t(h, y, d, z) = \frac{(d^2 + h^2)^{1/2} (2d - z)}{d\bar{v}}. \quad (11.3.1в)$$

С помощью алгебраических или геометрических методов легко получить выражение для обратного перехода:

$$d(s, g, t, z) = \frac{1}{2} \{[\bar{v}^2 t^2 - (g-s)^2]^{1/2} + z\}; \quad (11.3.2a)$$

$$y(s, g, t, z) = \frac{1}{2} \left\{ (g+s) + \frac{z(g-s)}{[\bar{v}^2 t^2 - (g-s)^2]^{1/2}} \right\}; \quad (11.3.2б)$$

$$h(s, g, t, z) = \frac{1}{2} \left\{ (g-s) + \frac{z(g-s)}{[\bar{v}^2 t^2 - (g-s)^2]^{1/2}} \right\}. \quad (11.3.2в)$$

То, что уравнения (11.3.2) действительно реализуют обратный переход к уравнениям (11.3.1) легко проверить подстановкой формул (11.3.1) в выражения (11.3.2).

Для однородной среды, в которой волны распространяются со скоростью \bar{v} , можно записать решение для нисходящей волны в виде дельта-функции на расширяющейся окружности:

$$D(g, s, t, z) = \delta[(g-s)^2 + z^2 - \bar{v}^2 t^2]. \quad (11.3.3)$$

Теперь необходимо сравнить восходящую волну U , вычисляемую в переменных (h, y, d, z) , с нисходящей волной D , заданной выражением (11.3.3) в переменных (g, s, t, z) . Для этого можно преобразовать нисходящую волну D к переменным (h, y, d, z) с помощью подстановки выражений (11.3.1) в формулу (11.3.3). При этом возникает имеющая глубокий смысл возможность упрощения получаемых выражений, если предположить, что скорость распространения волн в среде \bar{v} равна скорости движения \bar{v} координатной системы. Тогда

$$D(h, y, d, z) = \delta[4d(z-d)]. \quad (11.3.4)$$

В этом случае нисходящая волна не зависит от латеральных координат h и y .

Теперь рассмотрим модель среды, которая содержит только один точечный рассеиватель, находящийся в точке с координатами (x_0, z_0) . Этот рассеиватель облучается источником типа дельта-функции, расположенным в точке $(s, 0)$. Исключая горизонтально распространяющиеся волны, имеем для восходящей волны $U(s, g, t, z)$ на малом удалении от рассеивателя:

$$U(s, g, t, z_0-0) = \delta(g-x_0) \delta[\bar{v}^2 t^2 - (s-x_0)^2 - z_0^2]. \quad (11.3.5)$$

Подставив преобразование (11.3.1) при $z=d$ в формулу (11.3.5), получим

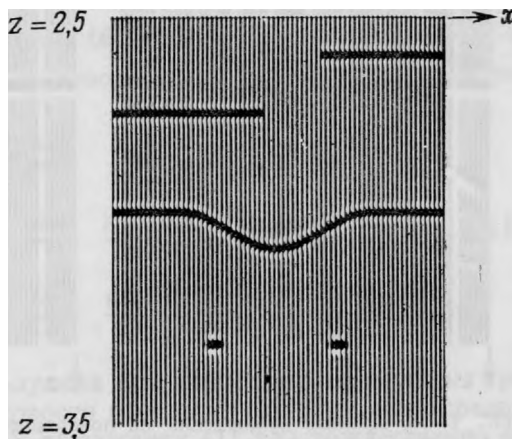
$$U(h, y, d, z=d) = \delta(y-x_0) \delta[d^2 + h^2 - (y-h-x_0)^2 - z^2_0].$$

Член $\delta(y-x_0)$ дает возможность считать $y=x_0$ во второй дельта-функции. Тогда

$$U(h, y, d, z=d) = \delta(y-x_0) \delta(d^2 - z^2_0).$$

Как видим, главная идея состоит в том, что волна на отражателе в движущихся, исправленных за нормальное приращение времени, координатах не зависит от полусмещения h . Очевидно, что суперпозиция случайного набора точечных рассеивателей будет созда-

РИС. 11.15. Модель Земли, используемая для иллюстрации скоростного анализа с помощью продолжения данных в нижнее полупространство (z — глубина)



вать при миграции волновое поле, которое является случайным по координатам y и d , но тем не менее постоянным по смещению h . Это представление будет, по-видимому, справедливым даже, если рассеиватели случайно распределены вне плоскости разреза. В трехмерном случае величину z следует рассматривать как радиальное расстояние от линии продольного профиля.

Наша цель — определить скорость распространения волн в среде. Однако скорость нужна уже для первого шага, а именно для осуществления миграции. Использование ошибочной скорости при миграции препятствует полному сжатию волны к дельта-функции на оси средних точек. Это вызвано деструктивной интерференцией между смежными средними точками. При этом в модели со случайно расположенными рассеивателями некоторая часть информации теряется. Для модели слоистой Земли это не имеет значения (даже когда сама миграция не является необходимой).

Стефен М. Догерти с целью иллюстрации этих понятий рассчитал волновое поле для модели Земли, представленной на рис. 11.15. На рис. 11.16 показаны данные наблюдений для этой модели на поверхности и результат продолжения этих данных

в нижнее полупространство. Вертикальной осью является координата d . Для горизонтальной оси координата h выбрана при шести значениях h , которые изображены группами (выборки общей средней точки). Десять таких выборок расположены (с интервалами) вдоль оси y . Внутри каждой выборки трасса со смещением $h=0$ располагается слева, а со смещением $h=h_{\max}$, соответствующую

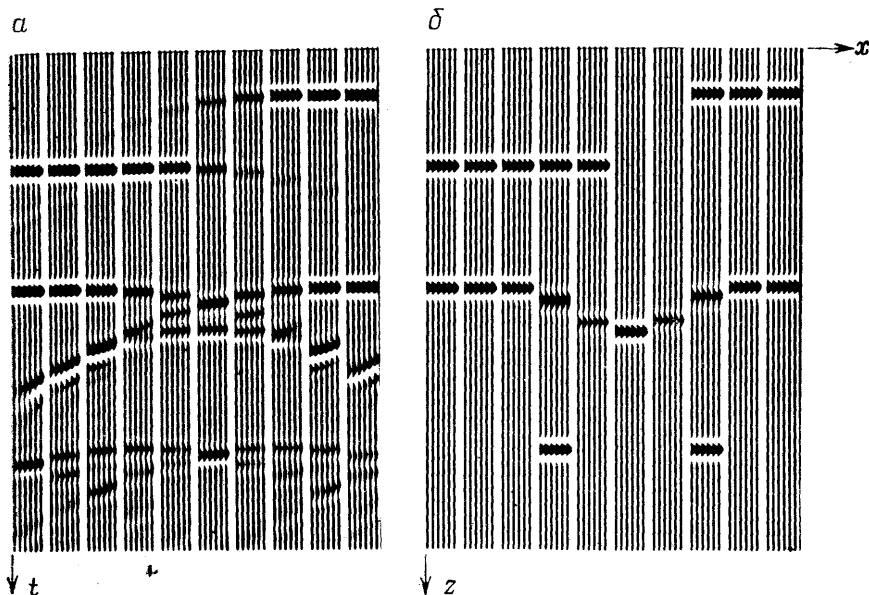


РИС. 11.16. Данные наблюдений на поверхности (а) и результат продолжения их в нижнее полупространство (б) для модели, представленной на рис. 11.15. Координаты выбраны так, чтобы изобразить трехмерные данные (y, h, d) в плоскости чертежа

щим лучам с углами выхода около 45° , — справа. На рис. 11.16,б видно, что в области отражателей волновые импульсы выравнены по горизонтали. Это свидетельствует о независимости данных в нижнем полупространстве от смещения h .

Для определения скорости несущественно, какая система координат используется для продолжения вниз наблюдаемого волнового поля. Однако это удобнее делать по сейсмограммам с введенными кинематическими поправками. Для упрощения выкладок отметим, что уравнения (11.3.2б) и (11.3.2в) подразумевают выполнение равенства

$$\frac{\partial y}{\partial(g, t, z)} = \frac{\partial h}{\partial(g, t, z)}. \quad (11.3.6)$$

Волновое уравнение

$$\left(\partial_{gg} + \partial_{zz} - \frac{\partial_{tt}}{\tilde{v}^2} \right) P = 0 \quad (11.3.7)$$

в координатах с введенными кинематическими поправками принимает вид

$$\left[(d_g \partial_d + y_g \partial_y + h_g \partial_h)^2 + (\partial_z + d_z \partial_d + y_z \partial_y + h_z \partial_h)^2 - \frac{(d_t \partial_d + y_t \partial_y + h_t \partial_h)^2}{\bar{v}^2} \right] Q = 0 \quad (11.3.8)$$

Как и ранее, при возведении в квадрат этих дифференциальных операторов в частных производных будем считать коэффициенты постоянными. Это соответствует предположению о высокочастотности сигналов. Прежде чем приступить к вычислению всех необходимых производных, введем упрощающее обозначение

$$b = [\bar{v}^2 t^2 - (g-s)^2]^{1/2}. \quad (11.3.9)$$

Вычислив требуемые производные и учитывая выражение (11.3.6), получим

$$\begin{bmatrix} d_g & d_z & d_t \\ y_g & y_z & y_t \\ h_g & h_z & h_t \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -\frac{(g-s)}{b} & 1 & \frac{\bar{v}^2 t}{b} \\ 1 + \frac{z \bar{v}^2 t^2}{b^3} & \frac{g-s}{b} & -\frac{(g-s) \bar{v}^2 t z}{b^3} \\ 1 + \frac{z \bar{v}^2 t^2}{b^3} & \frac{g-s}{b} & -\frac{(g-s) \bar{v}^2 t z}{b^3} \end{bmatrix}. \quad (11.3.10)$$

Если скорость \bar{v} , используемая для коррекции нормальных приращений времени, равна скорости распространения волны в среде \bar{v} , т. е. $\bar{v} = \tilde{v} = v$, то три члена в выражении (11.3.8) тождественно обращаются в нуль. Подстановкой можно проверить, что

$$\left(d_g^2 + d_z^2 - \frac{d_t^2}{v^2} \right) Q_{dd} = 0; \quad (11.3.11a)$$

$$2 \left(d_g y_g + d_z y_z - \frac{d_t y_t}{v^2} \right) Q_{dy} = 0; \quad (11.3.11б)$$

$$2 \left(d_g h_g + d_z h_z - \frac{d_t h_t}{v^2} \right) Q_{dh} = 0. \quad (11.3.11в)$$

Затем получаем три смешанных производных с дифференцированием по z :

$$2y_z Q_{yz} = \frac{g-s}{b Q_{yz}} = \frac{h}{d Q_{yz}}; \quad (11.3.12a)$$

$$2h_z Q_{hz} = \frac{g-s}{b Q_{hz}} = \frac{h}{d Q_{hz}}; \quad (11.3.12б)$$

$$2d_z Q_{dz} = Q_{dz}. \quad (11.3.12в)$$

Из выражения (11.3.6) следует, что коэффициенты при Q_{yy} , Q_{hh} и $2Q_{hy}$ идентичны. Выполнив необходимые преобразования, получим:

$$\left(y^2_x + y^2_z - \frac{y^2_t}{v^2}\right)(\partial_y + \partial_h)^2 Q = -\left(\frac{d}{2d-z}\right)^2 \left(1 + \frac{h^2}{d^2}\right)(\partial_y + \partial_h)^2 Q. \quad (11.3.13)$$

Как обычно, отбросив член с Q_{zz} , получим приближение, подобное приближению Френеля. В декартовой системе координат это ограничивает точную обработку лучами, лежащими внутри конуса с раскрывом около 15° от вертикали. В системе с введенными кинематическими поправками это, по-видимому, более похоже на 15° -градусное ограничение для углов наклона границ. Конечно, при необходимости всегда можно использовать методы с более высокой точностью. Объединяя вместе выражения (11.3.11) и (11.3.13), получаем основной результат:

$$\left[\partial_d + \frac{h}{d}(\partial_y + \partial_h)\right] \partial_z Q = \left(\frac{d}{2d-z}\right)^2 \left(1 + \frac{h^2}{d^2}\right)(\partial_y + \partial_h)^2 Q. \quad (11.3.14)$$

Уравнение (11.3.14) можно использовать для определения скорости путем продолжения в нижнее полупространство скорректированных за нормальное приращение времени сейсмических записей без суммирования разрезов.

По-видимому, имеет смысл рассмотреть выражение (11.3.14) для некоторых специальных случаев. На поверхности земли при нулевом смещении величина Q вследствие симметрии обращается в нуль. Для точечного рассеивателя, как было показано ранее, величина Q не зависит от h , а для идеализированных данных, полученных от слоистых отражателей, величина Q является функцией только параметра d . Учитывая широкое разнообразие практических ситуаций, логично упростить выражение (11.3.14) с помощью системы неравенств $Q_d \gg Q_y \gg Q_h$. Это приводит к выражению

$$Q_{dz} = -\left(\frac{d}{2d-z}\right)^2 \left(1 + \frac{h^2}{d^2}\right) Q_{yy}. \quad (11.3.15)$$

В отличие от ранее рассмотренных уравнений с постоянными коэффициентами, уравнение (11.3.15) содержит переменный коэффициент $d/(2d-z)$. Однако для миграции это не имеет практического значения. Введем новую переменную

$$z' = \frac{zd}{2d-z} \quad (11.3.16)$$

На поверхности Земли ($z=0$) величина $z'=0$, а на отражателях, где $z=d$, $z'=d$. Положив $Q(z, d) = Q'(z', d)$, находим

$$Q_z = z'_z \partial_{z'} Q';$$

$$Q_d = (\partial_d + z'_d \partial_{z'}) Q'.$$

С учетом этих соотношений левая часть выражения (11.3.15) принимает вид

$$Q_{dz} = (\partial_d - z'_d \partial_{z'}) z'_z \partial_z Q.$$

Подобно приближению Френеля (при отбрасывании члена с $\partial_{z'z'}$) имеем

$$Q_{dz} = z'_z Q_{dz'} = \frac{2d^2}{(2d - z)^2} Q_{dz'}.$$

Это позволяет привести выражение (11.3.15) к виду

$$Q_{dz'} = -\frac{1}{2} \left(1 + \frac{h^2}{d^2} \right) Q_{yy}. \quad (11.3.17)$$

Чтобы подтвердить необходимость введения коэффициента 2, о котором было сказано ранее в разделе 11.2, можно сделать другое преобразование от координаты d к координате t' двойного времени ни пробега волны до рефлектора и обратно:

$$t' = \frac{2d}{v}.$$

Это дает

$$Q_{t'z'} = -\frac{v}{4} \left[1 + \left(\frac{2h}{vt'} \right)^2 \right] Q_{yy}. \quad (11.3.18)$$

Очевидно, что выражение (11.3.18) должно быть проинтегрировано в интервале от $z'=0$ до $z'=t'v'/2$. Изменять масштаб по оси глубины удобно через двойное время прохождения волны

$$t'' = \frac{2z'}{v}.$$

Это приводит к уравнению с временным масштабом по оси глубин

$$Q_{t''t''} = -\frac{v^2}{8} \left[1 + \left(\frac{2h}{vt''} \right)^2 \right] Q_{yy}, \quad (11.3.19)$$

в котором величина t' является двойным временем прохождения волны, а величина t'' — осью глубины во временном масштабе, интегрируемой в интервале от поверхности $t''=0$ до отражателей на глубине $t''=t'$.

Строго говоря, уравнение преобразования (11.3.19) следует применять отдельно для каждого смещения h до того, как данные суммируются по смещению. С целью сокращения затрат машинного времени эти данные часто суммируют перед миграцией с помощью выражения (11.3.19). При таком компромиссном подходе величина h в уравнении (11.3.19) принимается равной нулю или используется некоторое среднее значение величины $2h/vt'$.

Пока мы показали, что результаты продолжения в нижнее полупространство сейсмограмм с коррекцией нормального приращения времени не зависят от величины смещения, если продолжение выполнено с правильным значением скорости. Однако мы не знаем, как оценить влияние погрешности определения скорости на результат продолжения данных в нижнее полупространство. Для этого необходимо учесть член, который ранее опускался. Этот член использовался при рассмотрении распространения волн в неоднородной среде, но мы от него отказались (для упрощения), начиная с формулы (11.3.11). Теперь необходимо до конца учитывать различие между скоростью распространения волн в среде $\bar{v}(x, z)$ и скоростью \bar{v} , при которой определялись кинематические поправки [обобщенной до зависимости $\bar{v}(z)$]. Вспомним, что для малых отклонений от слоистых моделей использовано неравенство $Q_d \gg Q_y \gg Q_h$, из которого следует, что первый из трех членов в уравнении (11.3.11) является наиболее существенным. Учет различия между рассматриваемыми двумя скоростями требует ввода в каждое из выражений (11.3.11a) данного члена:

$$\left(d^2_g + d^2_z - \frac{d^2_t}{\bar{v}^2}\right) Q_{dd} \neq 0, \quad (11.3.20)$$

$$\frac{\bar{v}^2 t^2}{\bar{v}^2 t^2 - (g-s)^2} \left(1 - \frac{\bar{v}^2}{\bar{v}^2}\right) Q_{dd} \neq 0,$$

$$\left(1 + \frac{h^2}{d^2}\right) \left(1 - \frac{\bar{v}^2}{\bar{v}^2}\right) Q_{dd} \neq 0. \quad (11.3.21)$$

Тогда выражение (11.3.15) принимает вид

$$Q_{dz} = - \left(\frac{d}{2d-z}\right)^2 \left(1 + \frac{h^2}{d^2}\right) Q_{yy} + \left(1 + \frac{h^2}{d^2}\right) \left(1 - \frac{\bar{v}^2}{\bar{v}^2}\right) Q_{dd}. \quad (11.3.22)$$

При численном решении уравнения (11.3.22) можно использовать метод расщепления. Решение строится посредством попеременного использования следующих двух уравнений:

$$Q_{dz} = - \left(\frac{d}{2d-z}\right)^2 \left(1 + \frac{h^2}{d^2}\right) Q_{yy}; \quad (11.3.23a)$$

$$Q_z = \left(1 + \frac{h^2}{d^2}\right) \left(1 - \frac{\bar{v}^2}{\bar{v}^2}\right) Q_d. \quad (11.3.23b)$$

Уравнение (11.3.23a) может быть названо «дифракционной» частью, а уравнение (11.3.23b) — частью «тонкой линзы». При построении решения для Q в направлении оси z с помощью выражения (11.3.23b) результат преобразования сейсмограммы (исправленной за нормальное приращение времени d при фиксиро-

ванном полусмещении h и средней точке y) непрерывно сдвигается по времени (d — сдвиг). Этот сдвиг увеличивается с увеличением погрешности определения скорости в соответствии с множителем $1 - \bar{v}^2/\tilde{v}^2$, а также смещения в соответствии с множителем $1 + h^2/d^2$. Таким образом, выражение (11.3.23б) изменяет кривизну данных с полусмещением h . Поскольку выражение (11.3.23а) содержит величину Q_{yy} и не имеет величин Q_h и Q_{hh} , то операции, выполняемые по формулам (11.3.23а) и (11.3.23б), переставимы (коммутативны). Следовательно, можно строить продолжение поля вниз к отражателям с помощью выражения (11.3.23а), а затем использовать формулу (11.3.23б). Важно отметить, что наиболее трудная часть работы, выполняемой по формуле (11.3.23а), зависит от скорости системы координат \bar{v} , а не от скорости распространения волн в среде \tilde{v} . Это означает, что можно строить изображение с различными скоростями распространения волн в среде \tilde{v} с небольшими затратами.

Для выбора лучшего способа вычисления скорости распространения волн в среде \tilde{v} необходимо рассмотреть вопрос взаимодействия восходящей волны U с нисходящей волной D , умноженной на коэффициент отражения c . Сопоставление этих волн можно проделать в координатах регистрируемого волнового поля, но следует отдать предпочтение системе координат с введенными кинематическими поправками. Сначала получим выражение для нисходящей волны в рассматриваемых координатах. Подставив выражение (11.3.1) в формулу (11.3.3), получим:

$$D(h, y, d, z) = \delta \left[4d(z - d) + \left(1 - \frac{\bar{v}^2}{\tilde{v}^2} \right) \left(1 + \frac{h^2}{d^2} \right) (2d - z)^2 \right]. \quad (11.3.24)$$

В данный момент мы не пытались сохранить медленные изменения амплитуды волны [сферическое расхождение опущено, начиная с формулы (11.3.3)]. Поэтому можно разделить весь аргумент дельта-функции на $-4d$. Поскольку нас интересуют малые отклонения величины \tilde{v} от \bar{v} , то дельта-функция должна обращаться в нуль в непосредственной близости к точке, в которой выполняется условие $z=d$. Поэтому, сохраняя хорошее приближение, можно подставить величину z вместо d в коэффициент при множителе $1 - \bar{v}^2/\tilde{v}^2$ в выражении (11.3.24). Учитывая предыдущие соображения, получим:

$$D(h, y, d, z) = \delta \left[d - z - \left(1 - \frac{\bar{v}^2}{\tilde{v}^2} \right) \left(1 + \frac{h^2}{z^2} \right) \frac{z}{4} \right] = \delta(d - z - s), \quad (11.3.25)$$

где функция задержки

$$s(h^2, z, \bar{v}^2/\tilde{v}^2) = \left(1 - \frac{\bar{v}^2}{\tilde{v}^2} \right) \left(1 + \frac{h^2}{z^2} \right) \frac{z}{4}. \quad (11.3.26)$$

Теперь вернемся к задаче сопоставления восходящей и нисходящей волн. Коэффициент отражения можно определить вместе с некоторой угловой зависимостью в виде степенного ряда, например,

$$c = c_0 + \frac{c_1 h}{z} + \frac{c_2 h^2}{z^2} + \dots \quad (11.3.27)$$

Для упрощения последующих результатов будем вычислять только постоянный член c_0 с помощью минимизации выражения

$$\min_c \sum_h \sum_d [U(y, z, h, d) - c(y, z) D(y, z, h, d)]^2. \quad (11.3.28)$$

Решение имеет вид

$$c(y, z) = \frac{\sum_h \sum_d UD}{\sum_h \sum_d D^2}. \quad (11.3.29)$$

Поскольку величина D почти всюду обращается в нуль, то можно использовать это для замены двойной суммы однократной, в частности, в числителе

$$\begin{aligned} \sum_h \sum_d UD &= \sum_h \sum_d U(y, z, h, d) \delta(d - z - s) = \\ &= \sum_n U[y, z, h, d = z + s(h^2, z, \bar{v}^2/\tilde{v}^2)]. \end{aligned} \quad (11.3.30)$$

Полагая, что величина N обозначает число членов в сумме смещений, получим для формулы (11.3.29):

$$c(y, z, \tilde{v}) = \frac{1}{N} \sum_h U. \quad (11.3.31)$$

В заключение рассмотрим способы определения скорости распространения волны в среде \tilde{v} , обеспечивающие наилучшую минимизацию разности $U - cD$. С этой целью можно использовать перебор по значениям скорости \tilde{v} и искать минимум выражения

$$\begin{aligned} \min_{\tilde{v}} \sum_h \sum_d (U - cD)^2 &= \min_{\tilde{v}} \sum_h (U - c)^2 = \min_{\tilde{v}} \sum_h \left(U - \frac{1}{N} \sum_h^N U \right)^2 = \\ &= \min_{\tilde{v}} \left[\left(\sum_h U^2 \right) - \frac{1}{N} \left(\sum_h U \right)^2 \right] \geq 0. \end{aligned} \quad (11.3.32)$$

Однако, как показала практика, вместо минимизации выражения (11.3.32) лучше максимизировать отрицательный логарифм или отношение подобия¹:

$$\Phi_v = \frac{(\Sigma U)^2}{N \Sigma U^2} \leq 1. \quad (11.3.33)$$

Это отношение нечувствительно к амплитуде волны U и применимо в различных условиях.

11.4. МНОГОКРАТНЫЕ ОТРАЖЕНИЯ

Точное моделирование многократных отражений дает возможность производить вычитание теоретических многократных отражений из данных полевых наблюдений и тем самым облегчает выделение первичных отражений, несущих полезную информацию. Рассмотрим сначала упрощенную модель слоистой среды для многократных отражений, а затем модифицируем ее применительно к двумерной неоднородной среде. Решим прямую и обратную задачи. В прямой задаче задается двумерное пространственное распределение коэффициентов отражения и выделяются отражения, включая дифрагированные волны с учетом многократного отражения в тонком слое. В обратной задаче определяется двумерное пространственное распределение коэффициентов отражения на основе наблюдаемого волнового поля.

Основная идея, которая использована при решении обратной задачи, заключается в том, что отражатели существуют в точках, где времена первого вступления нисходящей и восходящей волн совпадают. Следует выбирать коэффициент отражения, обеспечивающий обращение в нуль восходящей волны перед появлением в области отражателя нисходящей волны. В главе 8 установлено, что в случае слоистой среды Z -преобразование полинома для нисходящей волны $D(z)$ является минимально фазовой функцией. Это означает, что величину, обратную $D(z)$, т. е. $1/D(z)$, можно разложить по положительным степеням z , а первый коэффициент будет равен величине $1/d_0$. Следовательно, коэффициент отражения на нижней границе (подшве) n -го слоя является коэффициентом при самой низкой степени z в разложении для функции $U^n(z)/D^n(z)$. Мы хотим показать, что величина $U^n(z)/D^n(z)$ может быть измерена на поверхности при $n=0$, а величина U^{n+1}/D^{n+1} — легко вычислена на основании отношения U^n/D^n . Для решения этих задач необходимо прежде всего знать давление P и вертикальную скорость W на этой поверхности для всех моментов времени.

¹ В отечественной литературе это отношение называется коэффициентом когерентности. (Прим. перев.)

Из раздела 9.3 и уравнений (9.3.12а) и (9.3.12б) имеем

$$\begin{bmatrix} U \\ D \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -\frac{1}{Y} \\ 1 & \frac{1}{Y} \end{bmatrix} \begin{bmatrix} P \\ W \end{bmatrix}, \quad (11.4.1)$$

и

$$\begin{bmatrix} P \\ W \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -Y & Y \end{bmatrix} \begin{bmatrix} U \\ D \end{bmatrix}. \quad (11.4.2)$$

В выражениях (11.4.1) и (11.4.2) давление $P^0(z) = p_0$ на поверхности $n=0$ в момент времени $t=0$ является дельта-функцией, которая описывает мгновенное нарушение поверхностного граничного

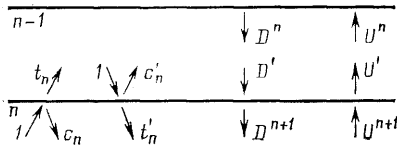


РИС. 11.17. Обозначения волн в слоистой среде, рассеянных на n -й границе раздела

условия $P^0=0$, обеспечивающее ввод энергии внутрь среды. Это является причиной обращения дельта-функции в вертикальную скорость W^0 при $t=0$ и далее, когда на отражающей поверхности величина $W^0 \neq 0$, а давление $P^0=0$. Для применения рассматриваемого метода необходимо, чтобы величина $U^0 = \frac{1}{2}(P - YW) = 0$ до прихода первого отраженного сигнала, а затем становилась характеристикой восходящей наблюдаемой формы волны W , взятой с обратным знаком, а величина $D^0 = \frac{1}{2}(P + YW)$, являющаяся дельта-функцией при $t=0$, изменялась в соответствии с формой наблюдаемой волны.

Рис. 11.17 иллюстрирует условия формирования волн на границе раздела, рассмотренные в главе 8. При распространении волны через слой наблюдается очевидное запаздывание, которое учитывается умножением на оператор половинной задержки:

$$Z^{1/2} = \exp(i\omega\Delta t/2) = \exp(i\omega\Delta z/v).$$

Тогда

$$D' = Z^{1/2}D^n; \quad (11.4.3)$$

$$U^n = Z^{1/2}U'. \quad (11.4.4)$$

При прохождении волн через слоистую среду происходит трансформация нисходящих волн в восходящие на границах раздела

$$U' = t_n U^{n+1} + c'^n D'. \quad (11.4.5)$$

Чтобы теория была достаточно общей, как в главе 8, запишем уравнение, обратное выражению (11.4.5). Согласно этому уравне-

нию нисходящие волны усиливаются вследствие обратного рассеяния восходящих волн. Чтобы упростить первое рассмотрение дифрагированных многократно-отраженных волн, будем предполагать, что нисходящие волны генерируются восходящими волнами только на поверхности $n=0$. Во всех остальных случаях будем предполагать, что нисходящие волны проходят свой путь, не получая энергии от восходящих волн. При этих допущениях

$$D^{n+1} = t'_n D'. \quad (11.4.6)$$

Практически предположение, на котором основано выражение (11.4.6), является вполне подходящим для морских данных, где свободная поверхность моря — намного более сильный отражатель, чем все остальные отражатели в разрезе. Несмотря на то что коэффициенты отражения границ, как показывает практика, изменяются в широком диапазоне, обычно коэффициент отражения около 0,1 для морского дна и 0,01 для границ, находящихся ниже. Для данных, полученных на суше, часто важно учитывать сильное поглощение и рассеивание волн в приповерхностных слоях.

В любом случае, исключая переменные D' и U' из уравнений (11.4.3) — (11.4.6), получаем

$$D^{n+1} = t' Z^{1/2} D^n; \quad (11.4.7)$$

$$U^{n+1} = \frac{U' - c'_n D'}{t_n} = \frac{Z^{-1/2} U^n + c_n Z^{1/2} D^n}{t_n}. \quad (11.4.8)$$

Разделив выражение (11.4.8) на уравнение (11.4.7), получим простую зависимость для продолжения величины U/D в нижнее полупространство. Она имеет вид

$$\frac{U^{n+1}}{D^{n+1}} = \frac{Z^{-1} U^n / D^n + c_n}{t'_n t_n}. \quad (11.4.9)$$

Чтобы понять эту простую рекуррентную зависимость, мы должны вспомнить, что коэффициент при Z^0 в выражении для U^n/D^n обращается в нуль (см. рис. 11.17), а коэффициент при Z равен $c'_n = -c_n$. Очевидно, продолжение величины U/D в нижнее полупространство является настолько простым, что результаты могут рассматриваться почти как коэффициенты отношения $U^0(Z)/D^0(Z)$ на поверхности. Единственным осложнением является делитель $t'_n t_n$, который дает монотонно возрастающий масштабный множитель $1/\prod_{k=1}^n (1 - c_k^2)$ для коэффициента Z^n в выражении для $U^0(Z)/D^0(Z)$.

Теперь подумаем, как обобщить эти представления для модели слоистой среды на модель двумерной неоднородной среды. Основное различие между этими задачами заключается в том, что продолжение волн вверх и вниз не может больше проводиться просто с помощью оператора задержки Z . Для продолжения волн вверх и вниз можно было бы использовать уравнение (10.5.16).

Однако естественнее и экономичнее использовать выражения (10.5.17а) и (10.5.18):

$$D'_{z't'} = \frac{\bar{v}}{2} D'_{x'x'}; \quad (11.4.10)$$

$$U''_{z't''} = -\frac{\bar{v}}{2} U''_{x'x''} - c'(z'') \partial_{t''} D'(x'', z'', t'' - \frac{2Z''}{\bar{v}t''}). \quad (11.4.11)$$

Уравнения (11.4.10) и (11.4.11) можно сравнить с уравнениями (11.4.7) и (11.4.8). В уравнениях (11.4.10) и (11.4.11) не учитываются коэффициенты пропускания, так как они обычно не являются важными при анализе полевых данных. Однако при необходимости они могут быть учтены с помощью более полного вывода.

Сначала рассмотрим прямую задачу. Ее легче понять при записи уравнений в конечно-разностной форме, чем в форме дифференциальных уравнений (11.4.10) и (11.4.11). Определим тридиагональную матрицу T , с помощью дифференциального оператора второго порядка δ_{xx} следующим образом:

$$T = - \left(\frac{\bar{v}\Delta z\Delta t}{8\Delta x^2} \right) \delta_{xx}. \quad (11.4.12)$$

Тогда уравнение (11.4.10) вместе с граничным условием, означающим, что нисходящая волна D равна нулю до момента $t'=0$, можно записать в табличной форме:

$$\begin{array}{|c|c|} \hline T+I & T-I \\ \hline T-I & (T+I) \\ \hline \end{array} * \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 0 \\ \hline d^0_0 & d^1_0 & d^2_0 & d^3_0 & d^4_0 \\ \hline d^0_1 & d^1_1 & d^2_1 & d^3_1 & \\ \hline d^0_2 & d^1_2 & d^2_2 & & \\ \hline d^0_3 & d^1_3 & & & \\ \hline d^0_4 & & & & \\ \hline \end{array} = 0, \quad (11.4.13)$$

где знак «*» обозначает свертку в плоскости (z, t') . Эту таблицу можно заполнить, если известны нули в верхней строке и данные наблюдений на поверхности указаны в левом столбце. Рассмотрим типичную для прямой задачи операцию. Предположим, что известны восходящая U и нисходящая D волны для всех z и $t < 4$. При $t \geq 4$ обе волны неизвестны. Типичная операция, которую мы покажем, объясняет, как получить волны U и D при $t=4$. Это означает, что известны все занесенные в D -таблицу (11.4.13) сведения, за исключением величины d^0_4 . На первом шаге берут величину D из t' координаты в табл. (11.4.13) для формирования t'' координаты. Если $\Delta z = \bar{v}\Delta t/2$, то это достигается путем сдвига вниз последующих столбцов в таблице (11.4.13). Тогда

				→z
0	0	0	0	0
d^0_0	0	0	0	0
d^0_1	d^1_0	0	0	0
d^0_2	d^1_1	d^2_0	0	0
d^0_3	d^1_2	d^2_1	d^3_0	0
d^0_4	d^1_3	d^2_2	d^3_1	d^4_0

$t'' \downarrow$

(11.4.14)

При наличии отражения все элементы данной строки, соответствующей постоянной t'' , могут вносить вклад в восходящую волну, зарегистрированную на поверхности наблюдений в момент времени t'' . Уравнение (11.4.11) можно преобразовать к виду

$$\left(\frac{\bar{v}}{2} \partial_{xx} + \partial_{zt}\right) (-U) = c'(x, z) \partial_t D, \quad (11.4.15)$$

что также позволяет выразить его в табличной форме:

<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>$T - I$</td> <td>$T + I$</td> </tr> <tr> <td>$(T + I)$</td> <td>$T - I$</td> </tr> </table>	$T - I$	$T + I$	$(T + I)$	$T - I$	*	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>d^0_1</td> <td>$-U^1_1$</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>d^0_2</td> <td>$-U^1_2$</td> <td>$-U^2_2$</td> <td>0</td> <td>0</td> </tr> <tr> <td>d^0_3</td> <td>$-U^1_3$</td> <td>$-U^2_3$</td> <td>$-U^3_3$</td> <td>0</td> </tr> </table>	0	0	0	0	0	d^0_1	$-U^1_1$	0	0	0	d^0_2	$-U^1_2$	$-U^2_2$	0	0	d^0_3	$-U^1_3$	$-U^2_3$	$-U^3_3$	0	=							
$T - I$	$T + I$																																	
$(T + I)$	$T - I$																																	
0	0	0	0	0																														
d^0_1	$-U^1_1$	0	0	0																														
d^0_2	$-U^1_2$	$-U^2_2$	0	0																														
d^0_3	$-U^1_3$	$-U^2_3$	$-U^3_3$	0																														
=	$\frac{\Delta z}{2}$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>-1</td> <td>-1</td> </tr> <tr> <td>1</td> <td>1</td> </tr> </table>	-1	-1	1	1	*	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>d^0_0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>d^0_1</td> <td>d^1_0</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>d^0_2</td> <td>d^1_1</td> <td>d^2_0</td> <td>0</td> <td>0</td> </tr> <tr> <td>d^0_3</td> <td>d^1_2</td> <td>d^2_1</td> <td>d^3_0</td> <td>0</td> </tr> <tr> <td></td> <td>d^1_3</td> <td>d^2_2</td> <td>d^3_1</td> <td>d^4_0</td> </tr> </table>	d^0_0	0	0	0	0	d^0_1	d^1_0	0	0	0	d^0_2	d^1_1	d^2_0	0	0	d^0_3	d^1_2	d^2_1	d^3_0	0		d^1_3	d^2_2	d^3_1	d^4_0	(11.4.16)
-1	-1																																	
1	1																																	
d^0_0	0	0	0	0																														
d^0_1	d^1_0	0	0	0																														
d^0_2	d^1_1	d^2_0	0	0																														
d^0_3	d^1_2	d^2_1	d^3_0	0																														
	d^1_3	d^2_2	d^3_1	d^4_0																														
			*																															
		<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td>0</td> <td>c'_1</td> <td>c'_2</td> <td>c'_3</td> <td>c'_4</td> </tr> </table>	0	c'_1	c'_2	c'_3	c'_4																											
0	c'_1	c'_2	c'_3	c'_4																														

В левой части уравнения (11.4.16) для неизвестных элементов оставлены пробелы. В правой части коэффициенты отражения должны быть свернуты в направлении вверх в нисходящую волну для того, чтобы создать источники для восходящей волны. В стро-

ке коэффициентов отражения величина c_0 принята равной нулю, так как граничное условие $d^0_t = -U^0_t$ будет задано отдельно. Таким образом, правая часть уравнения (11.4.16) полностью известна, а неизвестные в левой части могут быть заполнены справа налево путем вычисления таблицы восходящей волны при $i''=4$ для всех z . Это дает величину d^0_4 , а также позволяет вернуться назад и заполнить следующую строку в таблице нисходящей волны, которая дает возможность заполнить следующую строку в таблице восходящей волны и так до бесконечности.

Обратная задача решается подобным образом. Предположим, что величины d^0_t известны для всех t . Необходимо рекурсивно рассчитать коэффициенты отражения c'_1, c'_2 и т. д. Достаточно показать, как вычисляется коэффициент c'_1 . Предположим, что мы отбираем два левых столбца из уравнения (11.4.16). Тогда получим

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline g & e_0 \\ \hline g & e_1 \\ \hline g & e_2 \\ \hline g & e_3 \\ \hline \vdots & \vdots \\ \hline \end{array} \\
 * \\
 \begin{array}{|c|c|} \hline T-I & T+I \\ \hline T+I & T-I \\ \hline \end{array} \left\| \begin{array}{l} \uparrow \\ \text{свернуть} \\ \downarrow \\ \text{вверх} \end{array} \right. \frac{\Delta z}{2}
 \end{array}
 \begin{array}{c}
 \begin{array}{|c|c|} \hline 0 & c'_1 \\ \hline \end{array} \downarrow \begin{array}{l} \text{Умножить} \\ \text{перекрестно} \end{array} \\
 * \\
 \begin{array}{|c|c|} \hline g & 0 \\ \hline g & g \\ \hline g & g \\ \hline g & g \\ \hline g & g \\ \hline \end{array} \\
 * \\
 \begin{array}{|c|c|} \hline -1 & -1 \\ \hline 1 & 1 \\ \hline \end{array}
 \end{array}
 \quad (11.4.17)$$

В выражении (11.4.17) все ячейки слева, заполненные знаками g , заданы. Знаки g в ячейках в правой части уравнения, как и ранее легко вычисляются. Если бы коэффициент c'_1 был известен, то это была бы прямая задача для последовательного вычисления величин $\dots, e_3, e_2, e_1, e_0$. При этом вычисления необходимо начинать с приближения вида $e_N=0$ для некоторого большого N . Если использовать точное значение коэффициента c'_1 , то величина e_0 должна обращаться в нуль. Поскольку точное значение коэффициента c'_1 неизвестно, то используем значения $c'_1=+1$ и $c'_1=-1$, получив при этом величины e^+_0 и e^-_0 соответственно. Точное значение c'_1 является соответствующим образом взвешенной линейной комбинацией

$$0 = \alpha e^+_0 + \beta e^-_0, \quad (11.4.18)$$

$$1 = \alpha + \beta, \quad (11.4.19)$$

$$c'_1 = \alpha - \beta. \quad (11.4.20)$$

Решая уравнения (11.4.19) и (11.4.20) относительно α и β , получаем:

$$2\alpha = 1 + c'_1; \quad (11.4.21)$$

$$2\beta = 1 - c'_1. \quad (11.4.22)$$

Подставив выражения (11.4.21) и (11.4.22) в уравнение (11.4.18), получим:

$$0 = (1 + c'_1)e^{+0} + (1 - c'_1)e^{-0}.$$

Откуда
$$c'_1 = \frac{e_0^+ + e_0^-}{e_0^+ - e_0^-}. \quad (11.4.23)$$

Таким образом, принцип, который мы использовали для вычисления коэффициента c'_1 , состоит в том, что восходящая волна e_0 на глубине z_1 должна обращаться в нуль до появления здесь нисходящей волны.

Для практического применения представленных выше одномерного и двухмерного алгоритмов необходимо преодолеть ряд трудностей. Прежде всего возникает задача учета рассеяния энергии источника в трех измерениях. Это требует введения поправки за сферическое расхождение. К счастью, главную роль играет *отношение* восходящей волны к нисходящей волне. Поэтому истинная амплитуда, по-видимому, не столь важна. Действительно, можно показать, что расчет является инвариантным при наличии экспоненциальной регулировки коэффициента усиления. Одна из реальных проблем состоит в том, что требуемая для расчета амплитуда дельта-функции начального взрыва не может быть правильно измерена. В действительности приходится использовать *волновой импульс*, возбужденный взрывом. К счастью, волновой импульс может быть прокалиброван в глубокой воде, где дно моря является плоскостью. В таком месте возвращающаяся волна может быть приведена к нулю всюду, за исключением окрестности первичного отражения от морского дна; обозначим этот волновой импульс через P . Образуется также другой волновой импульс M , приводящий к нулю всюду, кроме окрестности отражения второй кратности. В идеальных условиях волновой импульс P , свернутый с самим собой, должен дать волновой импульс M . Отклонения от этого условия позволяют оценить амплитуду, время прихода и форму волны от взрывного источника.

Необходимо четко представлять себе, что сильно отражающая граница не может опуститься точно на один интервал дискретизации данных. Это означает, что все сигналы (4,0) (3,1) (2,2), (1,3) и (0,4) должны рассматриваться как последовательные задержки одного и того же волнового импульса. Различие спектров на высоких частотах — проблема того же типа, что и отклонение уравнений в конечных разностях от дифференциальных уравнений для коротких волн с длиной волны короче десяти интервалов дискретизации. Обычно нет необходимости точно моделировать высокие частоты, но следует осуществлять выборку данных достаточно плотно. Реальной задачей для высоких частот является исключе-

ние возможности превращения их в причину нарушения стабильности решений. Например, в расчете

$$\frac{U^0}{D^0} = -\hat{W}(Z) / [1 + \hat{W}(Z)]$$

сначала необходимо определить величину \hat{W} , которая должна быть зарегистрирована, если бы P/I для взрыва являлось идеальной дельта-функцией. Естественно, что высокие частоты в функ-

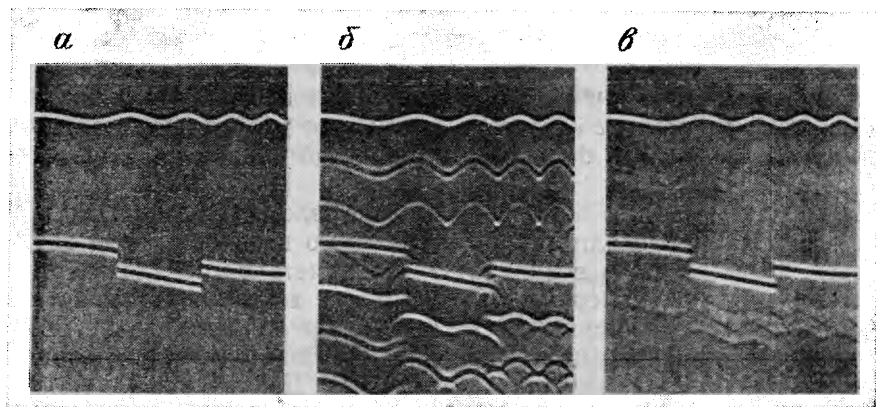


РИС. 11.18. Пример миграции для сложного разреза, волновое поле для которого содержит высокие пространственные частоты.

a — модель; *b* — синтетический временной разрез; *v* — результат миграции

ции \hat{W} являются в значительной мере бессмысленными. Единственной причиной, по которой их необходимо сохранить, состоит в том, чтобы они не обратили функцию $1 + \hat{W}$ в неминимально фазовую, что привело бы к невозможности ее деления. В этом случае высокие частоты из \hat{W} могут быть отфильтрованы. Тогда этот делитель будет стремиться к положительной действительной функции. Дон Райли [38] установил, что рассматриваемые практические трудности могут быть преодолены. На рис. 11.18 приведен один из результатов выполненных им расчетов для случая ундулирующей поверхности морского дна, под которым залегает нарушенная глубинная структура (рис. 11.18, *a*). Горизонтальная линия в нижней части разреза представляет собой метку времени, соответствующую 1 с. С целью демонстрации изображений ко всем данным была применена одинаковая регулировка усиления во времени по экспоненциальному закону, а также было использовано увеличение вертикального масштаба в 5 раз. На рис. 11.18, *b* приведен синтетический временной разрез, на котором в области ниже морского дна, наблюдаются многократные отражения от дна (придонные многократные отражения). Кроме того, ниже нарушенной глубинной структуры наблюдаются многократные отражения от глубоких границ. На рис. 11.18, *v* демонстрируется результат восстановления исходной модели.

СПИСОК ЛИТЕРАТУРЫ

1. Flinn Edward A., Enders A. Robinson and Sven Treitel (Eds.): The MIT Geophysical Analysis Group (GAG) Reports, Geophysics, vol. 32, no. 3, 1967.
2. Haskell N. A.: The Dispersion of Surface Waves on Multilayered Media, Bull. Seismol. Soc. Am., vol. 43, pp. 17—34, 1953.
3. Claerbout Jon F.: Coarse Grid Calculations of Waves in Inhomogeneous Media with Application to Delineation of Complicated Seismic Structure, Geophysics, vol. 35, no. 3, 1970.
4. Claerbout Jon F.: Numerical Holography, Acoustic Holography, vol. 3, Plenum Press, 1971.
5. Claerbout Jon F.: Toward a Unified Theory of Reflector Mapping, Geophysics, vol. 36, no. 3, 1971.
6. Claerbout Jon F. and A. G. Johnson: Extrapolation of Time Dependent Waveforms along Their Path of Propagation, Geophys. J. R. Astron. Soc., vol. 26, nos. 1—4, pp. 285—295, 1971.
7. Landers T. and Jon F. Claerbout: Numerical Calculations of Elastic Waves in Laterally Inhomogeneous Media, J. Geophys. Res., vol. 77, no. 8, pp. 1476—1482, 1972.
8. Claerbout Jon F. and S. M. Doherty: Downward Continuation of Moveout Corrected Seismograms, Geophysics, vol. 37, no. 5, pp. 741—768, 1972.
9. IEEE Trans. Audio Electrostat., Special Issue on fast Fourier transform, June, 1967.
10. Wold Herman: «Stationary Time Series», Almqvist and Wiksell, Stockholm, 1938.
11. Robinson Enders A.: «Random Wavelets and Cybernetic Systems», Griffin, London, 1962.
12. Grenander Ulf and Gabor Szego: «Toeplitz Forms and Their Applications», University of California Press, Berkeley and Los Angeles, 1958.
13. Levinson Norman: The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction, Appendix B in NORBERT WIENER: «Extrapolation, Interpolation and Smoothing of Stationary Time Series», Technology Press of the Massachusetts Institute of Technology, Cambridge, 1947.
14. Atkinson F. V.: «Discrete and Continuous Boundary Problems», Academic Press, New York, 1964.
15. Whittle P.: «Prediction and Regulation», English Universities Press, Ltd., London, 1963.
16. Kolmogorov A.: Sur l'interpolation et l'extrapolation des suites stationnaires, C. R. Acad. Sci. (Paris), vol. 208, pp. 2043—2045, 1939.
17. Gabor D.: Theory of Communication, J. IEEE, vol. 93, no. 1, pp. 429—441, 1946.
18. Mosteller F. and R. E. K. Rourke: «Sturdy Statistics», Addison Wesley Publishing Co., Inc., Reading, Mass., 1973.
19. Claerbout J. F.: Spectral Factorization of Multiple Time Series, Biometrika, vol. 53, nos. 1 and 2, pp. 264—6, 1966.
20. Seneta Eugene: «Non-Negative Matrices», John Wiley and Sons, Inc., New York, 1973.
21. Golub G.: Numerical Methods for Solving Linear Least Squares Problems, Numerische Mathematik, vol. 7, pp. 206—216, 1965.
22. Singleton R. S.: Algorithm 347 Sort, Comm. ACM, vol. 12, 1969.
23. Hoare C. A. R.: Quicksort, Comput. J., vol. 5, pp. 10—15, 1962.
24. Plackett R. L.: Studies in the History of Probability and Statistics, Chapt. XXIX, in the Discovery of the Method of Least Squares: Biometrika, vol. 59, no. 2, pp. 239—251, 1972.
25. Claerbout Jon F. and Francis Muir: Robust Modeling with Erratic Data, Geophysics, vol. 38, no. 5, pp. 826—844, 1973.

26. *Claerbout Jon F. and E. A. Robinson*: The Error in Least — Squares Inverse Filtering, *Geophysics*, vol. 29, no. 1, pp. 118—120, 1964.
27. *Burg John Parker*: The Relationship Between Maximum Entropy Spectra and Maximum Likelihood Spectra, *Geophysics*, vol. 37, no. 2, pp. 375—376, 1972.
28. *Widrow B., P. E. Mantey, J. J. Griffiths and B. B. Goode*: Adaptive Antenna Systems, *Proc. IEEE*, vol. 55, pp. 2143—2159, 1967.
29. *Levin M. J.*: Maximum Likelihood Array Processing, MIT Lincoln Laboratory Technical Report, 1964.
30. *Goupillaud P.*: An Approach to Inverse Filtering of Near Surface Layer Effects from Seismic Records, *Geophysics*, vol. 26, no 6, pp. 754—760, 1961.
31. *Kunetz G.*: Generalization des opérateurs d'antiresonance a un nombre quelconque de réflecteurs, *Geophys., Prosp.*, vol. 12, pp. 283—289, 1964.
32. *White J. E.*: Use of Reciprocity Theorem for Computation of Low — Frequency Radiation Patterns, *Geophysics*, vol. 25, no. 3, pp. 613—624, 1960.
33. *Richtmyer Robert D. and K. W. Morton*: «Difference Methods for Initial value Problems», John Wiley and Sons, Inc., New York, 1967.
34. *Mitchell A. R.*: «Computational Methods in Partial Differential Equations», John Wiley and Sons, Inc., New York, 1969.
35. *Goodman J. W.*: «Introduction to Fourier Optics», McGraw — Hill Book Company, New York, 1968.
36. *Treitel Sven, J. L. Shanks and C. W. Frasier*: Some Aspects of Fan Filtering, *Geophysics*, vol. 32, no. 5, pp. 789—800, 1967.
37. *Doherty Stephen M.*: «Structure Independent Seismic Velocity Estimation», doctoral dissertation, Geophysics Department, Stanford University, Stanford, California, 1975.
38. *Riley Don Clinton*: «Wave Equation Synthesis and Inversion of Diffracted Multiple Seismic Reflections», doctoral dissertation, Geophysics Department, Stanford University, Stanford, California, 1975.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Автокорреляция 30, 154
— случайных чисел 92
Адаптивный фильтр 158
Алгоритм Бурга 157, 158
Амплитудно-модулированный косинус 26
Ансамбль 84
Антиклиналь 271
Апертура 214
Аппроксимация луча 232
Ареа синус гиперболический 250
Арифметика целых чисел 254
Аткинсон 70
Атмосферная волна 28, 190, 204, 240
- Базисные уравнения 147
Байесов подход 84
«Белый шум» 95, 184
Бесконечно малый (исчезающе малый) 218
Биеция между двумя частотами 25
Билинейное преобразование 55, 249
Биномиальные коэффициенты 87
— в применении к функции Гаусса 100
Биномиальная функция вероятности 100
Блок Геллица 150, 162
Боковое отражение 256
Боковой лепесток (диаграммы направленности) 47
Боскович 147
Бросание монеты 84
Бург Джон 124, 154, 160, 185
Быстрое преобразование Фурье 20—24
Быстрый и нестрогий вывод метода наименьших квадратов 126
- Важнейшие классы функций 46
Ввод кинематических поправок (коррекция за нормальное приращение времени) 277
Веерная фильтрация 249, 252
Вероятность и центральная предельная теорема 97
Вертикальное накопление (суммирование) 265, 272
Вертикальная пространственная частота k_z 195, 196
Ветреная атмосфера 204, 290
Взаимная корреляция 31
Взаимный спектр 31
Взаимодействие 242—248
Взвешенные погрешности в методе наименьших квадратов 129
Взвешенная медиана 143
Взвешенное среднее 85, 86
Взрыв 188
Видров Бернард 159
Винер Норберт 64, 118
Виттль Петер 70
Воздушная волна (см. атмосферная волна) 28, 190, 204, 240
Волновое уравнение в системе координат с введенными кинематическими поправками 273
Волновое уравнение с одним направлением распространения волны 220—222
Волновой импульс 296
Волновой фронт в слое 196
— в виде полуокружности 217, 269
Восстановление (экстраполяция) 262, 271, 272, 296
Восходящая волна 168, 169, 193, 291
Временной ряд 85, 91, 154
Время нарастания 80—82
— распространения, меняющееся по гиперболическому закону 265, 275
— распространения по двойному пути 172, 285
— релаксации 77
— интегрирования с уткой 50
— совпадения 257, 258
Выборка (дискретизация) данных 12
— по общим средним точкам 257
— по общим точкам отражения 257
Выборочное среднее 84
Высокочастотное приближение 231, 283
Вычисление спектра с использованием максимума энтропии 65, 140, 154—158
— скорости 277
Вязкость 248—254
- Габор Деннис 80
Гаусс 147
Гауссова вероятность 87, 101
Геометрия землетрясения 175, 184
— отражения сейсмических волн 173
Герберт Верн 24
Гиббс 84
Гидрофон (см. сейсмоприемник) 257
Гиперболическое дифференциальное уравнение 236, 265, 275
Гиперплоскость 149
Глубинная точка 257
Голуб Гене 134
Горизонтальная пространственная частота k_x 195
Гуониллауд Пьерре 172
Гравитация 190
Граница 167
— боковая 229
— жесткая (недеформируемая) 168, 169
— раздела 168
— свободная 168, 291
— слоя 189, 227
Грубая ошибка (промах) 141
Групповая задержка 25, 53, 239
— — положительная 52
— скорость 25
- Данные акустического буя 257
— для равных смещений взрыв — прибор 257
Движущаяся система координат 234
Двойственность 21
Деление полиномов 36
Детерминант (опредетитель) 107
Деформация 38
Диаграмма направленности излучения 198
— состояний-переходов 119
Диапазон 79
Дисперсия 85
— волн в воздухе 28
— скорости по площади 277
— среднего выборки 84
Дисперсионное соотношение 223
— — для полуокружности 227
Дифракция 231, 271, 286
Дифференциальное уравнение 56, 75
Диффузия 59, 167
Длительность (продолжительность) 79
Дно моря 291, 296
Добротность 77
Доверительный интервал 101
Догерти Стефен 281
Допплеровское смещение 235
Дробный интегральный оператор 58
Дуга окружности 269
- Единичная окружность 38, 179
- Жесткая (недеформируемая) граница 168, 169
Жидкая сфера 190
- Закон Гаука 206
— Ома 190
— Снеллнуса 196

Зона тени 233

Игра в кости 98

Изображение сейсмограммы 266

Изотропная упругость 206

Изотропное распространение лучей 216

Импеданс 43, 168, 175, 196

Импульсная реакция 13

— — точечной апертуры 216

Инвариантная система координат 139

Интеграл матричного уравнения 191

Интегрирование с утечкой 50, 89

Интервал доверительный 101

Ионосфера 190

Источник 188

Квадратичная форма 45

Квадратный корень оператора 112, 113

Квадратурный фильтр 33

Квантиль 146, 147

Квантовая механика (см. также уравнение Шредингера) 79

Когерентность 93

— пространственная 240

Код (шлейф) волны 240

Колмогоров А. 72

Комплексные величины в методе наименьших квадратов 126, 127

Комплексная функция времени 28, 29

Координатная система движущаяся 234

— — для восходящих волн 263

— — для нисходящих волн 263

— — интерпретации 257

— — поля 257

— — разведочная 256

— — с введенными кинематическими поправками 277, 279

— — теоретическая 257

Координата смещения (расстояния взрыв-прибор) 256—258

Координаты интерпретации 257

Косая (несимметричная) эрмитова матрица 203

Коррекция за нормальное приращение времени (ввод кинематических поправок) 277

Коррекция сферического расхождения 255

Корреляция 28

— между случайными переменными 88, 89

Коэффициент отражения 167, 197

— — в спектральных вычислениях 156

— — как функция угла 197, 248

— — поверхности и дна моря 291

— — при взаимодействии двухмерных волн 242

Коэффициент пропускания 167, 168, 244

Краевая задача 208—254

Кули Дж. 24

Курс корабля 255

Кюнцетт Г. 181

Лаплас П. 147

Левин М. 164

Левинсон Норман 64

Линза 214, 218, 286

Линейная форма 142

Линейное приближение 13

— программирование 142, 147

Луч в слое 196

Магнитное поле 74, 190

Магнитометр полного поля 75

Мадден Теодор 113

Максимум фазы 43

Марковский процесс 118

Математическое ожидание E 84

Матрица 104—121

— вероятности 119

— косая Эрмитова 203

— интегральная 191

— неправильная 113

— несимметричная 108, 113

— перехода 119

— полиномиальная 114

— преобразованная 119

— слоя 133, 192

— спектральная 116

— треугольной формы 135, 136

— тридиагональная 210—213, 237

— Теплица (блок) 150

— для положительных действительных функций 46

— при факторизации спектра 64

Матричная теорема Сильвестра 109

Матрицант 191

Мгновенная частота 32

Медiana 102

— взвешенная 143

Метод Виттля 70, 71

— Колмогорова 72

— корней 60—62

— Кранка—Николсона 209, 210, 237, 249

— наименьших квадратов 122—149

— расщепления 210, 286

— Теплица 64, 117, 118

Миграционный временной разрез 266

Миграция на основе волнового уравнения 266

Микробарограмма 28

Минимум задержки 40

— запаздывания энергии 63—64

— фазы 39—41

Многоканальный спектр 117, 118

— фильтр 114

Многократные отражения 65, 289—296

Множители Лагранжа 132

Модель грабена 272

Момент источника 188

Моменты функции вероятности 99

Морская поверхность 170

Морские сейсмические данные метода отраженных волн 255—273

Морское дно 291, 296

Мультиканал Теплица 161, 162

Наклон 29, 129

Направляющий (управляющий) луч 241, 242

Напряжение (упругое) 38

Нарушение 262, 296

Независимые погрешности 102

— случайные числа 94

Нелинейность 14

Неминимальность фазы 41, 254, 296

Непараметрическая статистика 162

Неравенства, моделирующие норму 142, 144

— ограничивающие разрешение 80—97

Неустойчивость в дифференциальных уравнениях в частных производных 209

— в обычных дифференциальных уравнениях 35, 40

Нефть 79, 271

Низкочастотный фильтр 47

Нисходящая волна 168, 169, 193, 291

Норма линейная несимметричная 142

— квадратичная L_2 142

— модели 138

Нормальное уравнение 128

Нули фильтров 51, 54

Нутация 29

Область отсутствия отражений 278

Обнаружение слабого сигнала 164

Обработка сейсмических данных на основе волнового уравнения 255—297

Обратная задача 65, 245, 260, 294

Обратное преобразование Фурье 18

Обратное Z-преобразование 17

Обратный (инверсный) фильтр 35

Обращение во времени 60, 171

— матрицы методом окаймления (программа) 106

Огибающая 31

Ограничения 129

Односторонние функции 35

- Однородное уравнение 105
 Оператор интегрирования 58
 Определение положения очага землетрясения 128
 Оптика 214
 Ортогональные моды 203
 Осадочные породы 266
 Ослабление во времени 76, 77
 — и причинность 76
 — на поверхности 253
 — при распространении волн 76, 77
 — с глубиной по гауссовой функции 233
 Отклонение 132, 133
 Отражение от поверхности моря 65, 170
 Отрицательная частота 28
- Падающий луч 243
 Парабола 226, 265
 Параболическое (паракснальное) приближение 221, 284
 Перенос (миграция) 266—277
 Плоскость сейсмического разреза 256
 Плохо поставленная задача 104, 132—134, 138—140
 Поверхностная волна 221
 — — на несовершенной сфере 232, 233
 Повторный удар 153
 Погрешность билинейного преобразования 55—57
 — для коротких волн 248
 — определения скорости 78, 286
 — случайная 79
 — усечения (ряда) 210
 — человека 102
 Погруженный фокус 267
 Подобие 289
 Подход Гиббса 84
 Поле, координаты 256
 Полезная полоса частот 251
 Полином с матричными коэффициентами 114
 Полиномы ортогональные на единичной окружности 70
 Полная проводимость (см. полное сопротивление — импеданс) 168
 Положительная реальная функция 43—46
 Полосовой фильтр 46, 54
 Полуцелый оператор 59
 Полюс (пик) 53
 — на пьедестале 53, 54
 Поляризация 29
 Поперечные волны 243
 Поток энергии 171, 173
 Предсказующий фильтр 150, 161
 Преобразование Гильберта 31, 74
 — отражения 134
 — Фурье с помощью линзы 218, 219
 — симметричное 20
 — Хаусхолдера 134
 Преувеличение вертикального масштаба 266
 Приближение, подобное френелевскому, в геометрии временного разреза (см. параболическое приближение) 284
 — по методу наименьших квадратов (программа) 137
 Призма 231
 Принцип взаимности 198—202, 257
 Прилив (отлив) 129
 Принцип неопределенности 79, 80, 154
 — причинности и распространение волн двумерных 233—242, 275
 — одномерных 75—78
 — сохранения в форме Аткинсона 203—205
 — потока 173
 — энергии 75
 Программа быстрого преобразования Фурье 24
 — вычисления квантилей 146, 147
 — вычисления медиан 146, 147
 — вычисления спектра по методу Бурга 158
- решения трехдиагональных совместных уравнений 213
 — метода наименьших квадратов 137
 — рекурсии Левинсона 69
 — обращения матрицы методом окаймления 107
 — определения коэффициентов отражения с помощью фильтра предсказания ошибок 183
 — преобразования Гильберта 74
 — расчета сейсмограмм для слонстых сред 184
 — свертки (конволюции) 16
 — факторизации спектра 74
 Продолжение вниз данных и разрезов 255—265
 Продолжительность (во времени) 80, 81
 Производство отсчетов 12
 Пространственная частота 27, 195, 196
 Профиль 257
 Процентиль 103
- Радиус кривизны отражающей границы 267
 Разрез (сечение) 257
 — вертикального суммирования 265
 — с нулевым смещением 265, 272, 273
 Разрешение статистико-временное 84—89
 — частотно-временное 79—83
 — частотно-временное с учетом статистики 94—97
 — частотно-статистическое 90—93
 — миграции по энергии 268
 Райлы Дон 137, 206
 Распространение волн и вязкость (затухание) 76—78, 248
 Рассеяние волн в слонстой среде 171—180
 Рассеянные лучи 243
 Расчленение (матриц) на части 105, 106
 Реализуемое обращение 35—38
 Реализуемый фильтр 35
 Режекторный фильтр 53
 Резонанс 154, 171
 Рекурсия Левинсона 67—69, 163, 164
 — для фильтра, формирующего волновой импульс 164
 — многоканальная 163
 Релятивистское уравнение Шредингера 227
 Робинсон Эндриу 63
 Ряд Лорана 38, 115
- Самосопряженная задача 198
 Свертка (программа) 16
 Свободная отражающая поверхность 170
 Сейсмограммы отраженных волн 174, 186
 Сейсмодетектор 35, 170, 191, 255—258
 Сенета Евгени 121
 Сингулярный (вырожденный) 107
 Синклиналь 272
 Синус гиперболический 250
 Синусоидальный отражатель 267
 Скорость 277
 Скоростной анализ 277—288
 Слабость неравенства 80
 Слонстая среда 65, 167—175, 187—207
 Случайные числа 94
 — — после узкополосного фильтра 96, 97
 Случайный точечный рассеиватель 280
 Смещение при вычислении автокорреляции 154
 — при определении когерентности 94
 — при вычислении корреляции 89
 — взрыв — прибор 256
 Собственное значение 108, 194, 238
 Собственный вектор 108, 193, 194, 238
 Сохранение энергии 170
 Спектр выборки случайных чисел 91
 — матрицы 116
 — многоканальный 114
 — простого полюса 50
 — реального двухполюсного фильтра 50
 — уравненный 71

- флуктуаций 91, 92
- энергетический 90
- Спектры фильтров 47
- Средние величины 85, 102, 143, 144
- Статистика, центральная предельная теорема 97
- Стационарный временной ряд 94, 154
- Строгость неравенства 80
- Суммирование вертикальное 265, 272
- и миграция с учетом нормального приращения времени 284
- общих средних точек 257
- скорректированных за нормальное приращение времени записей 272
- Суперпозиция 13
- Сферические гармоники 128
- Сферические волны 273
- Теорема Кейли—Гамильтона 113
- о добавлении помех 42
- Теория потенциала 60, 129
- общего поля 74
- Теплиц 64
- Тонкая линза 214, 286
- Точность билинейного преобразования 56—58
- экстраполяции (восстановления) волн 243
- Треугольная форма матрицы 136
- Тридиагональная матрица 210, 211
- Тьюки Дж. 24
- Уайт Дж. 202
- Удвоение 22
- Узкополосный фильтр 46—50
- — в спектральных вычислениях 96
- — для случайных чисел 96
- Уолд Герман 63
- Упорядочение чисел 142, 143
- Упругие волны 206, 207
- Уравнение Гельмгольца 219
- Уравнение для коллимированных (параллельных) лучей 243
- Уравнение для монохроматических волн 235
- Лапласа 129, 233
- Максвелла 190
- Пуассона 202
- теплопроводности 190, 208
- Устойчивость 37, 38, 209, 210, 239
- Фаза для двухмерного фокуса 241
- — фазового фильтра 50—52
- — квадратурного фильтра 31, 32
- — положительной реальной функции 41, 42
- скорости 25—27
- Фазовый фильтр 50—53
- — при двухмерном распространении волн 224—226
- — при одномерном распространении волн 75, 76
- — полюса и нули 62
- Фазовая задержка 50—53
- решетка 240
- Факторизация спектра методом Виттля 70, 71
- корней 60—62
- Теплица 64—69
- Колмогорова 72—74
- спектральных матриц 114—117

- Физически нереализуемый фильтр 16
- Фильтр адаптивный 153
- , включенный параллельно 41—42
- квадратурный 31
- матричный 115—116
- многоканальный 161
- , моделирующий распространение волн 75—76
- низкочастотный 47
- предсказующий (прогнозирующий) 150—153
- рекурсивный Левинсона 163
- с ограничениями 164
- узкополосный 46, 47
- фазовый 50
- , формирующий волновой импульс 150
- , предсказующий ошибку 152, 153
- с 90° фазовым сдвигом 31
- Фильтрация с обратной связью 48, 49
- Фильтрующий эффект пласта (тонкого слоя) 167
- Флуктуация спектра 90, 91
- Функция групповой задержки 53
- импеданса 43, 44
- комплексной величины 28, 29
- минимально-фазовая 39, 40, 46
- полной проводимости 43, 44
- с монотонно возрастающей фазой 54, 77
- удовлетворяющая принципу причинности 35
- физически реализуемая 35
- пропускания (прозрачности) 214
- пропускания (прозрачности) 214
- разрывности (ступенчатая) 189
- угла наклона 218

Хор С. 143

Центральная предельная теорема 97—101

- Частичные дроби 34
- Частота мгновенная 32
- отрицательная 28
- Чендлер Уобблс 29
- Численная вязкость 248—253

Шаг сетки 274, 275

- Ширина спектра полосы пропускания 79—81
- — используемая при миграции волнового уравнения 251, 252
- Шредингера уравнение 227
- Шульц Филипп 216
- Шум, генерируемый сигналом 242

- Экспоненциальная матрица 111—112, 192
- Экспоненциальная регуловка усиления 297
- Экстраполирующий фильтр 161
- Экстраполяция волновых импульсов 233
- монохроматических волн 219
- Электрическое поле 190
- Электромагнитные волны 29
- Электростатика 190
- Энергетический спектр 90

- Явный вид 116, 242
- Ядерный взрыв 28, 164

- Z-преобразование 12
- матриц 114

ОГЛАВЛЕНИЕ

От переводчика	5
Предисловие	7
Введение	9
1. ПРЕОБРАЗОВАНИЯ	12
1.1. Дискретные данные и Z-преобразование	12
1.2. Связь Z-преобразования с Фурье-преобразованием	17
1.3. Быстрое преобразование Фурье	20
1.4. Фазовая и групповая задержки	25
1.5. Корреляция и спектры	28
1.6. Преобразование Гильберта	31
2. ОДНОСТОРОННИЕ ФУНКЦИИ	35
2.1. Инверсные фильтры	35
2.2. Минимально фазовые функции	39
2.3. Параллельное включение фильтров	41
2.4. Положительные действительные функции	43
2.5. Узкополосные фильтры	46
2.6. Фазовые фильтры	50
2.7. Режекторный фильтр и фильтр с пиком на пьедестале	53
2.8. Билинейное преобразование	55
3. ФАКТОРИЗАЦИЯ СПЕКТРА	60
3.1. Метод корней	60
3.2. Метод Робинсона о задержке энергии [11]	63
3.3. Метод Теплица	64
3.4. Экспоненциально-логарифмический метод Виттля	70
3.5. Метод Колмогорова	72
3.6. Принцип причинности и распространения волн	75
4. РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ	79
4.1. Частотно-временное разрешение	79
4.2. Статистико-временное разрешение	84
4.3. Частотно-статистическое разрешение	90
4.4. Частотно-временное разрешение с учетом статистики	94
4.5. Центральная предельная теорема	97
4.6. Доверительные интервалы	101
5. МАТРИЦЫ И МНОГОКАНАЛЬНЫЕ ВРЕМЕННЫЕ РЯДЫ	104
5.1. Обзор матриц	104
5.2. Матричная теорема Сильвестра	109
5.3. Матричные фильтры, спектры и факторизация	114
5.4. Марковские процессы	118
6. МОДЕЛИРОВАНИЕ ДАННЫХ НАБЛЮДЕНИЙ МЕТОДОМ НАИ- МЕНЬШИХ КВАДРАТОВ	122
6.1. Уравнений больше, чем неизвестных	123
6.2. Весовые функции и ограничения	129
6.3. Уравнений меньше, чем неизвестных	132
6.4. Преобразования Хаусхолдера и метод Голуба	134
6.5. Выбор нормы модели	138
6.6. Надежное моделирование	142

7. ПРИМЕНЕНИЕ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ ДЛЯ ОПРЕДЕЛЕНИЯ ФОРМЫ ВОЛНОВОГО ИМПУЛЬСА	150
7.1. Прогнозирующие и формирующие фильтры	150
7.2. Вычисление спектра методом Бурга	154
7.3. Адаптивные фильтры	158
7.4. Расчет многоканальных фильтров	161
7.5. Рекурсия Левинсона	163
7.6. Фильтры с ограничениями	164
8. ОБНАРУЖЕНИЕ СЛОЕВ С ПОМОЩЬЮ ФИЛЬТРАЦИИ РАССЕЯННЫХ ВОЛН	167
8.1. Коэффициенты отражения и прохождения волн на границах раздела	167
8.2. Поток энергии в слоистых средах	171
8.3. Расчет волн по коэффициентам отражения границ	176
8.4. Вычисление коэффициентов отражения по волновым полям	181
9. МАТЕМАТИЧЕСКАЯ ФИЗИКА СЛОИСТЫХ СРЕД	187
9.1. От физики к математике	187
9.2. Числовые матрицанты	191
9.3. Восходящие и нисходящие волны	193
9.4. Принцип взаимности	198
9.5. Принципы сохранения и ортогональные моды колебаний	203
9.6. Упругие волны	206
10. ДВУХМЕРНЫЕ И ТРЕХМЕРНЫЕ КРАЕВЫЕ ЗАДАЧИ	208
10.1. Классические нестационарные краевые задачи	208
10.2. Экстраполяция волн в оптике	214
10.3. Экстраполяция монохроматических волн численными методами	219
10.4. Экстраполяция волновых импульсов в пространстве	233
10.5. Взаимодействие лучей	242
10.6. Численная вязкость	248
11. ОБРАБОТКА СЕЙСМИЧЕСКИХ ДАННЫХ НА ОСНОВЕ ВОЛНОВОГО УРАВНЕНИЯ	255
11.1. Продолжение в нижнее полупространство сейсмических данных и разрезов	255
11.2. Миграция на основе волнового уравнения	266
11.3. Вычисление скорости распространения сейсмических волн в среде	277
11.4. Многократные отражения	289
Список литературы	297
Предметный указатель	299

ДЖОН Ф. КЛАЕРБУОТ

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ОБРАБОТКИ ГЕОФИЗИЧЕСКОЙ ИНФОРМАЦИИ С ПРИЛОЖЕНИЕМ К РАЗВЕДКЕ НЕФТИ

Редактор издательства Т. И. Борушко
 Переплет художника К. В. Голыкова
 Художественный редактор В. В. Шутько
 Технический редактор А. Е. Матвеева
 Корректор Р. Т. Баканова
 ИБ № 3410

Сдано в набор 06.01.81 Подписано в печать 23.03.81 Формат 60×90^{1/16} Бумага типографская № 1
 Гарнитура «Литературная» Печать высокая Усл. печ. л. 19,0 Усл. кр.-отг. 19,0 Уч.-изд. л. 18,79
 Тираж 2900 экз. Заказ 1017/8192-3 Цена 1 р. 70 к.

Издательство «Недра», 103633, Москва, К-12, Третьяковский проезд, 1/19
 Московская типография № 10 Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 113114, Москва, М-114, Шлюзов-
 вая наб., 10

