

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Новосибирский национальный исследовательский
государственный университет»

Факультет информационных технологий

Лихачев Алексей Валерьевич

Алгоритмы томографической реконструкции

УЧЕБНОЕ ПОСОБИЕ

Новосибирск 2013

Аннотация

Пособие включает основные положения, излагаемые в рамках курса "Методы и алгоритмы вычислительной томографии", а также содержит дополнительные сведения, способствующие в более полной мере освоению студентами знаний и компетенций, предусмотренных положениями госстандарта 230100.62. В тексте описаны перспективные методы неразрушающего исследования внутренней структуры объектов различной природы, основанные на принципах компьютерной томографии. При этом большое внимание уделяется выводу формул обращения и алгоритмов реконструкции. Математическая строгость изложения соответствует уровню подготовки студентов старших курсов бакалавриата по направлению 230100.62 "Информатика и вычислительная техника". Пособие также будет полезно для студентов старших курсов физическо-математических и технических специальностей, аспирантов и всех, кто интересуется проблемами современной томографии.

Учебное пособие подготовлено в рамках реализации
Программы развития НИУ-НГУ на 2009-2018 г.г.

Ведение.	5
-----------------	---

Глава 1

Обзор некоторых приложений.

1.1 Медицинская рентгеновская томография.	7
1.2 Методы диагностики СПЕСТ и РЕТ.	12
1.3 Исследования плазмы.	15
1.3.1 Определение локальных коэффициентов эмиссии.	15
1.3.2 Определение функции распределения.	20
1.4 Реконструкция по интерферометрическим измерениям.	22

Глава 2

Связь с уравнением переноса излучения.

2.1 Вывод лучевого уравнения томографии.	26
2.2 Уравнения для СПЕСТ и РЕТ.	32

Глава 3

Математические основы томографии.

3.1 Преобразование Радона.	35
3.2 P -Преобразование.	40
3.3 Теорема о центральном слое.	42
3.4 Томография в расходящихся лучах.	44
3.4.1 Верная схема на плоскости.	44
3.4.2 Трёхмерная задача.	47
3.5 Регуляризация некорректно поставленных задач.	49

Глава 4

Фильтрация одномерных проекций.

4.1 Общие замечания.	53
4.2 Линейная низкочастотная фильтрация.	57
4.3 Фильтр Шеппа-Логана.	61
4.4 Фильтр $1/z^2$.	64

Глава 5
**Задача трёхмерной томографии с источником излучения,
движущимся по кривой.**

5.1 Формула обращения Туя.	68
5.2. Теорема Гранжа.	73
5.3. Алгоритм Фельдкампта.	78
5.4. Спиральное движение источника.	86

Глава 6
Алгебраические методы реконструкции.

6.1 Дискретизация задачи.	90
6.2 Аддитивные алгоритмы.	92
6.3 Использование квадратичного целевого функционала.	95
6.4 Метод максимума энтропии.	98
6.5 Метод байесовской оценки.	101

Глава 7
**Задача эмиссионной томографии в частично
поглощающей среде.**

7.1 Постоянное поглощение. Формула обращения для двумерного случая.	106
7.2 Переменное поглощение. Итерационные методы решения задачи.	110
Список литературы.	115

Введение.

В настоящее время томография прочно вошла во многие области человеческой деятельности. Она находит применение в медицинской диагностике, в промышленности – для тестирования качества изделий и контроля технологических процессов, в геологии – для поиска и разведки месторождений полезных ископаемых, а также в системах безопасности, которая, к сожалению, становится всё более актуальной. Велика роль томографии при проведении научных исследований, в частности в экспериментах, связанных с изучением плазмы, пламени и потоков. Она используется в электронной и рентгеновской микроскопии – для получения структуры кристаллов и макромолекул, в астрофизике – для исследования звёзд и планетарных систем.

В связи со столь широким полем приложений, большое число методов изучения внутренней структуры объектов, различающихся на первый взгляд, включаются в понятие томография. Поэтому тут требуется уточнение. Под *томографическими методами* мы будем понимать совокупность процессов регистрации данных и их последующей математической обработки, основным этапом которой является получение локальных характеристик объекта по измеренным интегральным величинам. При этом сам процесс вычисления будет называться *реконструкцией* или *восстановлением*. Термин *томография* будет употребляться как синоним к словосочетанию "томографические методы". Результаты измерений, по которым производится реконструкция, будут называться *проекционными данными*.

При помощи томографии исследуются объекты различной природы. В соответствии с этим физические принципы, лежащие в основе систем сбора проекционных данных, могут быть весьма разнообразны. Наиболее известным применением является медицинская рентгеновская томография. Здесь измеряемой величиной является интенсивность проникающего излучения, прошедшего через тело пациента. Интенсивность излучения также регистрируется в задачах дефектоскопии, диагностики плазмы и во многих других случаях. Таким образом, томография, основанная на

регистрации интенсивности излучения, охватывает большой круг задач. Именно они являются в основном предметом рассмотрения предлагаемого учебного пособия. Однако существуют методы, оперирующие с проекционными данными совсем иной природы. Например, ЯМР-томография, электротомография, геотомография и другие.

В основе любого томографического метода лежит соотношение, связывающее проекционные данные с искомыми распределениями параметров. Особую роль играет уравнение, согласно которому индивидуальное измерение, например показание отдельного детектора, является интегралом вдоль отрезка прямой линии от функции, подлежащей определению. Его вывод из уравнения переноса излучения для различных диагностических задач приводится в главе 2. Подход к исследованию внутренней структуры объекта, основанный на такой модели сбора проекционных данных, в дальнейшем будет называться *лучевой томографией*, а соответствующее ему уравнение – *уравнением лучевой томографии*.

Уравнение лучевой томографии, с одной стороны, достаточно адекватно описывает многие задачи, с другой – позволяет свести реконструкцию к обращению известных в интегральной геометрии преобразований, что даёт возможность использовать хорошо разработанный математический аппарат. Методам решения задач лучевой томографии посвящена основная часть данного пособия, главы 3-6.

Описанное выше приближение учитывает лишь один процесс взаимодействия излучения с веществом: поглощение (в этом случае говорят о *трансмиссионной* томографии), или испускание (*эмиссионная* задача). Иногда этого оказывается недостаточно, чтобы описать реальную ситуацию. Большой интерес вызывает эмиссионная томография среды, где также присутствует поглощение. В первую очередь это связано с развитием методов медицинской диагностики СПЕСТ и РЕГ. Возникающие при этом задачи рассматриваются в главе 7.

Глава 1

Обзор некоторых приложений.

1.1 Медицинская рентгеновская томография.

Среди большого разнообразия существующих томографических методов наибольшего успеха достигла медицинская рентгеновская томография. Исторически она получила название СТ (computerized tomography), т.е. компьютерная томография. Хотя этот термин до сих пор широко используется, фактически он является устаревшим, поскольку в настоящее время любая задача реконструкции решается посредством компьютера, и, в этом смысле, медицинская рентгеновская томография является не более "компьютерной", чем любая другая.

Когда проводится обычная рентгенография пациента, изображения его внутренних органов накладываются друг на друга. Для того чтобы преодолеть этот недостаток, возникла идея проводить диагностику не по одному, а по нескольким рентгеновским снимкам, выполненным под разными ракурсами. Путём специальной математической обработки по такому набору данных можно оценить распределение коэффициента ослабления рентгеновского излучения (а по нему – распределение плотности) внутри диагностируемой части тела пациента.

Прототип медицинской томографической установки был впервые описан нейрорентгенологом Ольдендорфом (W. Oldendorf) в работе [1], вышедшей в 1961 году. Двумя годами позже математик Кормак (A. Cormack) предложил алгоритм реконструкции распределения коэффициента ослабления рентгеновского излучения по измерениям его интенсивности после прохождения через тело пациента, [2]. Кормак совместно с Хаунсфилд (G. Hounsfield) за основополагающие работы по реконструктивной томографии были удостоены Нобелевской премии по физиологии и медицине за 1979 год.

В 1972-1973 годах Хаунсфилд разработал первую томографическую систему – сканер головного мозга [3]. Сканер позволял получать изображения размером 80×80 пикселей с

разрешением 3 мм. Данные, которые требовались для получения одного изображения, измерялись примерно в течение четырёх с половиной минут. Такое высокое время регистрации, среди прочих причин, было обусловлено конструкцией первых томографов. В них имелась одна остронаправленная рентгеновская трубка и один детектор, которые синхронно передвигались вдоль рамы. Измерения обычно проводились в 160-180 положениях трубки, затем рама поворачивалась на угол в 1 градус и измерения повторялись, см. рис. 1.1,*а*. Столь значительная продолжительность сканирования накладывала ограничение на область исследования: первые медицинские томографы использовались только для диагностики головного мозга.

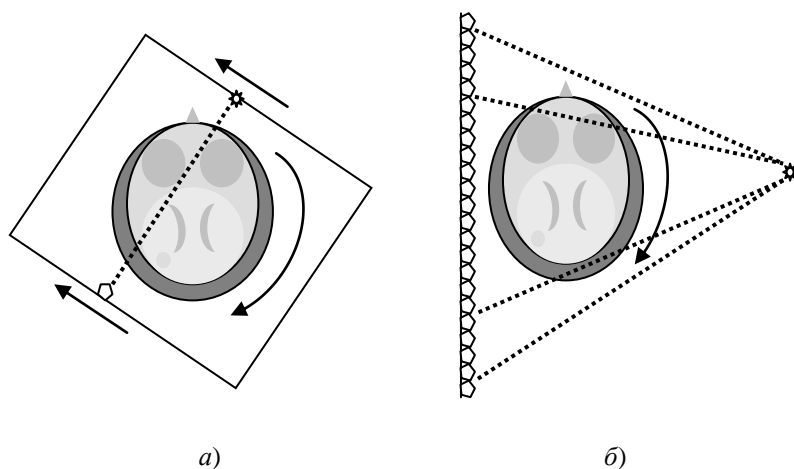


Рис. 1.1: Двумерные схемы томографического сканирования:
а) параллельная; б) веерная.

В дальнейшем томографические сканеры совершенствовались. Однако при этом измерения проводились по-прежнему в одной плоскости, в которой затем и восстанавливалось распределение плотности. Изображения различных сечений тела пациента получали путём его смещения относительно регистрирующего оборудования. Подобные схемы сбора проекционных данных в дальнейшем мы будем характеризовать как двумерные, а связанные с ними задачи реконструкции функции двух переменных *задачами*

двумерной томографии. Описание различных схем двумерного сканирования можно найти, например, в [4].

На рис. 1.1,б представлена так называемая *веерная* схема. Рентгеновская трубка излучает широкий веерный пучок, направленный на линейку детекторов. Система регистрации вращается вокруг пациента. Томографы с веерным сканированием начали появляться с середины 70-х годов прошлого столетия. В целях более удобной технической реализации, детектора в них располагались по дуге. Угловой размер дуги подбирался таким, чтобы стягиваемый ею сектор круга с центром в месте положения источника захватывал весь исследуемый объект. Данные снимались для большого числа положений источника. Эти томографы позволили во много раз уменьшить время сбора данных, что дало возможность проводить исследования движущихся органов тела, в частности легких.

С развитием техники были созданы двумерные детектора, обеспечивающие получения высококачественных изображений. Появление такого оборудования дало возможность регистрировать одним измерением излучение, испущенное источником в широком конусе. Что, в свою очередь, привело к возникновению новых схем сканирования, которые мы будем называть *трёхмерными*. В трёхмерных схемах сканирования источник движется относительно объекта по некоторой кривой в пространстве. Двумерный детектор может двигаться согласованно с источником, но может и оставаться неподвижным. При этом с практической точки зрения является важным выполнение условия, чтобы при любом положении источника всё излучение, прошедшее через тело пациента, попало на регистрирующую поверхность детектора. Во-первых, это нужно для правильной работы алгоритмов реконструкции. Во-вторых, пациент подвергается вредному воздействию рентгеновского излучения, которое, проходя мимо детектора, не даёт никакого вклада в диагностическую информацию.

Трёхмерные схемы сканирования будут рассмотрены в главе 5. Здесь мы упомянем лишь *винтовую* или *спиральную* геометрию сбора данных, которая широко применяется в медицинской рентгеновской томографии. Она стала возможной благодаря появлению конструкции сканера с кольцом скольжения, позволяющим трубке и детекторам вращаться непрерывно. В

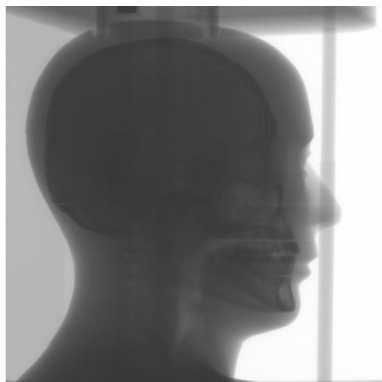
спиральном томографе осуществляется вращение регистрирующей системы одновременно с движением пациента через раму, на которой она располагается. В результате чего источник описывает вокруг пациента спираль. Её шаг можно регулировать, меняя поступательную скорость движения стола или вращательную скорость кольца. Фотография современного спирального томографа представлена на рис. 1.2.



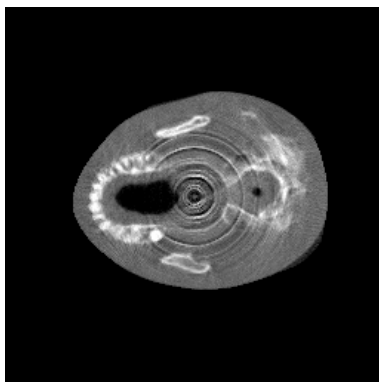
Рис. 1.2: Современный томограф со спиральной геометрией регистрации данных.

Рисунок 1.3 иллюстрирует возможности рентгеновской томографии. На рис. 1.3,а показано одно из распределений

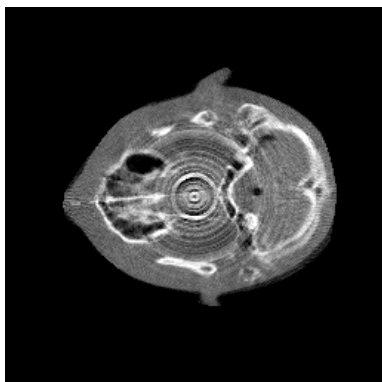
интенсивности, зарегистрированное двумерным детектором при сканировании муляжа головы человека. На рис. 1.3,б,в,г представлены различные сечения, полученные в результате томографической реконструкции. Количество ракурсов наблюдения (т.е. поворотов регистрирующей системы относительно объекта) составляло 325.



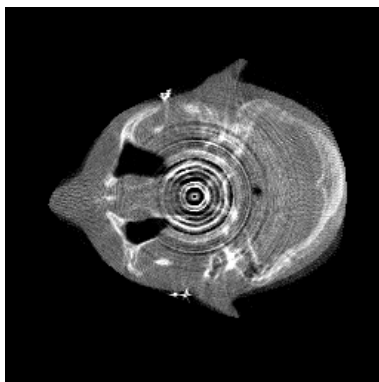
a)



б)



в)



г)

Рис. 1.3: *a)* распределение интенсивности, зарегистрированное двумерным детектором; *б), в), г)* сечения объекта, полученные в результате томографической реконструкции.

В настоящее время медицинская рентгеновская томография находится в стадии бурного развития. Диагностика почти каждого органа фактически стала её отдельным разделом, что нашло своё отражение в большом количестве публикаций. Отметим лишь несколько работ. В [5] проводится сравнение рентгеновской томографии с другими методами определения опухолей головного мозга. В [6] обобщены исследования, посвящённые выявлению посредством томографии поражений головного мозга, образовавшихся в результате кровоизлияния. В [7] дан обзор томографии лёгких. В [8] приводится сравнительный анализ рентгеновской, ультразвуковой и ЯМР-томографии как средств определения патологий печени. Рентгеновская томография традиционно занимается поиском повреждений костей и суставов. Относительно недавно она стала применяться также в стоматологии.

1.2 Методы диагностики SPECT и PET.

В последние годы большую популярность приобрели методы медицинской диагностики SPECT (Single Photon Emission Computerized Tomography) и PET (Positron Emission Tomography). В отличие от рентгеновской томографии, позволяющей получать только изображения анатомических структур и изменений в них, эти методы дают возможность проводить количественный анализ биохимических и физиологических процессов, происходящих в организме. Эта информация помогает выявить функциональные изменения, вызванные заболеванием, задолго до появления каких-либо морфологических изменений. В частности SPECT и PET позволяют диагностировать опухоли на стадии их зарождения.

В основе SPECT и PET лежит один и тот же принцип. Разница между ними заключается в его реализации. Обе методики являются комбинацией двух изобретений, представленных к Нобелевской премии – *радиоактивного индикатора* и методов томографии. Радиоактивный индикатор – это нестабильный изотоп какого-либо химического элемента, обладающий подходящими для диагностических целей свойствами. При проведении исследования путём SPECT или PET он вводится в организм человека, где и происходит его распад. Образовавшиеся в результате фотоны

регистрируются детекторами. По этим данным производится томографическая реконструкция распределения изотопа в теле. Анализируя это распределение, можно получить информацию о состоянии внутренних органов.

Изотопы, используемые в СПЕКТ, распадаясь, непосредственно выделяют фотон. В методе PET применяются, изотопы испускающие позитрон, который при аннигиляции с электроном вещества производит два γ -кванта, распространяющихся в противоположных направлениях. Радиоизотопы для PET, производятся при помощи циклотрона. Все они имеют короткое время полураспада, от нескольких минут до нескольких часов. Основные радионуклиды, используемые в клинических исследованиях методом PET, это углерод ^{11}C , азот ^{13}N , кислород ^{15}O и фтор ^{18}F . Эти химические элементы есть почти во всех соединениях в организме человека. Наиболее часто применяется изотоп ^{18}F , который используется для маркировки глюкозы. Полученный в результате препарат, распадаясь, позволяет следить за метаболизмом глюкозы в организме. В частности, злокачественные клетки используют её больше, чем здоровые ткани, поэтому такие исследования особенно эффективны для обнаружения рака и определения его стадии.

Различие в реакции распада изотопов, применяемых в СПЕКТ и PET, в конечном итоге приводит к тому, что в математической основе этих методов оказываются разные уравнения томографии, см. раздел 2.2. Помимо этого, условия, налагаемые на схему регистрации проекционных данных, для PET являются более жёсткими, поскольку для него необходима одновременная регистрация двух фотонов, распространяющихся в противоположных направлениях.

Первый прототип PET сканера для исследования головного мозга был введён в эксплуатацию в 1952 году в Массачусетском госпитале. Он имел два детектора, расположенных друг напротив друга. Установка могла обнаружить опухоль и определить её положение относительно срединной линии мозга. В начале 1960-ых были созданы PET сканеры с кольцом из 32 датчиков. По зарегистрированным на них данным можно было получить изображение сечения тела с разрешением порядка 2 см. Появившиеся в 1968 году PET сканеры следующего поколения

имели несколько колец, что позволило одновременно получать несколько срезов исследуемого объекта.

Начиная с конца 1970-ых, PET сканеры стали широко использоваться для проведения клинических исследований. Были спроектированы новые установки, в которых за счёт увеличения числа датчиков были повышены разрешающая способность и чувствительность. Новым этапом развития оборудования для СПЕКТ и PET стали сканеры, оснащённые подвижными тонкими вольфрамовыми кольцами – септами. Они способны работать в двух режимах: послойном, когда выпущенные септы выделяют набор исследуемых слоёв, и в трёхмерном, при втянутых септах, когда регистрируются пары фотонов, распространяющихся в любых направлениях. Таким образом, на этом оборудовании стало возможным собирать проекционные данные, как для двумерной, так и для трёхмерной реконструкции. Современные PET сканеры имеют до 16 колец, при этом пространственное разрешение составляет около 5 мм. Подробности можно найти, например, в [9]. На рис. 1.4 приведена фотография современного PET сканера.



Рис. 1.4: Современный PET сканер.

Развитие медицинских технологий привело к появлению новых диагностических устройств, получивших название ПЕТ/СТ сканеры. При их использовании можно одновременно получать, посредством ПЕТ, информацию о функционировании исследуемого органа, и реконструировать его структуру методом рентгеновской томографии. Это даёт возможность повысить эффективность диагностики. Кроме того, рентгеновские томограммы могут использоваться для коррекции аттенуации для ПЕТ, что существенно сокращает время исследования.

SPECT и ПЕТ применяются в различных областях медицины. Большинство клинических исследований этими методами проводится в области онкологии [10]. Они позволяют различать злокачественные и доброкачественные опухоли и определять степень их распространения. Полное сканирование всего тела пациента, занимающее около часа, позволяет выявить метастазы любой локализации. SPECT и ПЕТ используются при ранней оценке результатов химиотерапии. В кардиологии с их помощью можно получить информацию о кровоснабжении миокарда, оценить скорость происходящих в нём метаболических процессов. В фармакологии с применением этих методов проводят исследования лекарств, в частности, изучают их распределение в организме.

Методами SPECT и ПЕТ можно получать сведения о кровообращении мозга, скорости усваивания им кислорода и глюкозы, а также о других процессах. В связи с этим они используются при изучении физиологических механизмов, на которых основываются психологические функции сознания. При исследованиях реконструируется распределение тока крови, маркируемой обычно радиоизотопом ^{15}O , в объёме мозга [11]. Предполагается, что через возбуждённые участки проходит большее количество крови. Таким образом можно определить зоны активации мозга во время тех или иных действий, совершаемых человеком.

1.3 Исследования плазмы.

1.3.1 Определение локальных коэффициентов эмиссии.

Диагностика плазмы – одна из областей, где томорафические методы уже давно и прочно заняли своё место. По этому вопросу

имеется обширная литература, список которой постоянно пополняется. Её подробный обзор можно найти в монографии [12]. Здесь же мы лишь опишем основные экспериментальные постановки, приводящие к задачам томографии.

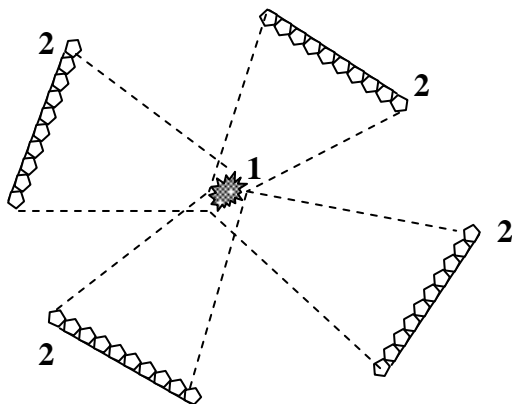


Рис. 1.5: Схема регистрации данных для двумерной томографии плазмы: 1 – исследуемый объект (сечение плазмы); 2 – линейки детекторов.

Определение распределения локальных коэффициентов эмиссии в плазме по результатам измерения суммарной интенсивности излучения теперь стало классическим примером применения томографии. Зная это распределение, и принимая некоторые предположения относительно свойств плазмы, в частности, состояние локального термодинамического равновесия, можно найти распределения концентрации и температуры различных компонентов плазмы, см. [12]. Описание данного метода диагностики можно найти практически в любом обзоре по экспериментальным исследованиям плазмы.

Рассмотрим следующую схему измерений, изображённую на рис. 1.5. Излучение плазмы регистрируется посредством нескольких линеек детекторов, ориентированных к ней под различными углами. Предполагается, что за время измерения плазма не успевает претерпеть значительного изменения, т.е. процесс является квазистационарным. Пусть каждый из детекторов измеряет интенсивность излучения, попавшего на него из очень узкого

конуса, в пределе переходящего в прямую. Помимо этого будем пренебрегать эффектами поглощения, рассеяния и других процессов взаимодействия излучения с веществом. В таких условиях, эмиссионные и трансмиссионные задачи будут описываться одним и тем же математическим уравнением, см. главу 2. Поэтому для реконструкции распределения локальных коэффициентов эмиссии плазмы могут быть использованы те же методы, что и для медицинской рентгеновской томографии.

После того, как посредством томографических методов получено распределение локальных коэффициентов эмиссии плазмы, могут быть рассчитаны другие её параметры. Для этой цели в случае наличия состояния локального термодинамического равновесия используется соотношение, [12]:

$$\varepsilon_{\lambda}(\vec{r}) = 1.06 \times 10^{-5} \left(\frac{g_1 f_{21}}{U(T) \lambda^3} \right) N_0(T, \vec{r}) \exp\left(-\frac{E_2}{T(\vec{r})}\right). \quad (1.1)$$

В (1.1) рассматривается спектральная линия с длиной волны λ , соответствующая переходу атома (иона) из возбуждённого состояния 2, в состояние 1 с меньшей энергией (в частном случае это может быть основное состояние). Через $\varepsilon_{\lambda}(\vec{r})$ обозначена мощность, излучаемая в спектральной линии в точке \vec{r} однородным прозрачным слоем плазмы толщиной в 1 см с площади 1 см^2 в единице телесного угла (т.е. коэффициент эмиссии); $N_0(T, \vec{r})$ – плотность атомов в основном состоянии; g_1 – статистический вес нижнего уровня перехода; E_2 – энергия возбуждённого уровня; f_{21} – сила осциллятора, соответствующая переходу $2 \rightarrow 1$; $U(T)$ – статистическая сумма атома (иона); $T(\vec{r})$ – температура.

Большое количество работ, связанных с томографией, посвящено исследованиям плазмы в токамаках. Токамак (от Тороидальная Камера с Магнитными Катушками) – это установка, предназначенная для магнитного удержания плазмы. Он представляет собой вакуумную камеру в форме тора, снабжённую катушками для создания магнитного поля, которое называют *тороидальным*. Вакуумную камеру заполняют смесью дейтерия и

третия. Затем в ней создаётся вихревое электрическое поле, которое вызывает протекание тока и зажигание в камере плазмы.

Протекающий через плазму ток, во-первых, производит её омический нагрев, во-вторых, создает магнитное поле, называемое *полоидальным*. Магнитное поле сжимает протекающий через плазму ток. В результате образуется конфигурация, в которой винтовые магнитные силовые линии обвивают плазменный шнур. При этом шаг при вращении в тороидальном направлении не совпадает с шагом в полоидальном направлении. Магнитные линии оказываются незамкнутыми, они бесконечно много раз закручиваются вокруг тора, образуя магнитные поверхности тороидальной формы. Наличие полоидального поля необходимо для стабильного удержания плазмы. Принципиальная схема токамака изображена на рис. 1.6.

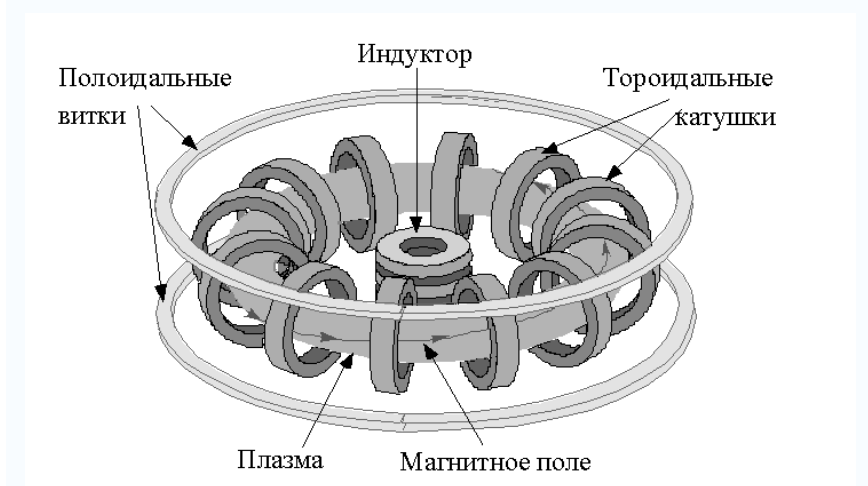


Рис. 1.6: Принципиальная схема токамака.

Томографическая реконструкция сечения плазменного шнура в токамаке обычно производится для трёх-пяти (иногда для десяти) ракурсов наблюдения. Излучение регистрируется детекторами, помещёнными в камеры-обскуры. По результатам реконструкции делаются выводы о распределениях температуры и плотности плазмы, а также об условиях устойчивости и разрушения плазменного шнура.

В последнее время повышенный интерес проявляется к изучению ближайшего к Земле космического пространства, в том числе структуры ионосферы. Широкие возможности для таких исследований предоставляет метод искусственных облаков, сущность которого состоит в том, что по интенсивности их свечения проводится реконструкция распределения концентрации ионов внутри облака, находящегося в неоднородном магнитном поле Земли, а это, в свою очередь, даёт возможность установить картину его силовых линий.

Впервые этот метод был предложен в 1950 году для изучения диффузионных процессов в верхней атмосфере. Впоследствии была разработана томографическая методика реконструкции трёхмерного распределения коэффициента в объёме облака. Она была опробована на объекте, содержащем излучающие ионы бария, который создавался с борта метеорологической ракеты и наблюдался из трех пунктов, расположенных на Земле. По оценкам, усреднение по времени за счёт рассогласования моментов начала съёмки с различных ракурсов и продолжительной экспозиции происходило в интервале порядка двух секунд. Характерная скорость движения облака на этой стадии менее 100 м/с. При таких скоростях указанному времени усреднения соответствует разрешение около 200 м.

Наиболее масштабными среди естественных ионосферных образований являются полярные сияния, рис. 1.7. По современным представлениям они тесно связаны с глобальными процессами, влияющими на всю планету. В начале девяностых годов прошлого века разрабатывался международный проект ALIS (Aurora Large Imaging System) по исследованию, в том числе томографическому, полярных сияний при помощи сети наземных станций, которые планировалось размещать в Швеции, Норвегии, Финляндии и на севере европейской части России [13]. Выбор места для размещения станций нуждается в тщательном анализе, который проводился несколькими исследовательскими группами путём численного моделирования. Расчёты производились с применением реальной географической карты. В пунктах предполагаемого размещения станций моделировались проекции полярного сияния. При этом считалось, что оно находится в месте своего наиболее частого появления.



а)



б)

Рис. 1.7: Полярные сияния различной формы.

1.3.2 Оценка функции распределения.

Другой задачей исследования плазмы, решаемой при помощи томографических методов, является оценка функций распределения

её компонентов (ионов и электронов) по скоростям. Эта диагностика основана на явлении рассеяния лазерного излучения плазмой. Поясним её на примере работы [14], где изучалась бариевая плазма. Ионы бария возбуждались при помощи лазера с перестраиваемой частотой, после чего регистрировалось излучение, соответствующее переходу на основной уровень. Плазма была помещена в однородное магнитное поле, причём предполагалось, что благодаря вращению ионов по ларморовским орбитам, их распределение по скоростям симметрично относительно оси, направленной вдоль магнитного поля.

Пусть ν_0 – частота рассматриваемого перехода, ν_L , \vec{k}_L – соответственно частота и волновой вектор центра лазерной линии, \vec{V} – скорость иона относительно источника лазерного излучения. Имеет место эффект Доплера: $\nu_L - \vec{k}_L \vec{V} = \nu_0$. Пусть ось Z выбрана вдоль направления магнитного поля, а направление лазерного пучка составляет угол φ с осью X и угол θ с осью Z . Флюоресценция регистрируется из точки \vec{r}_0 . При изменении лазерной частоты ω_L меняется интенсивность зарегистрированной флюоресценции, которая, согласно эффекту Доплера, пропорциональна плотности ионов, имеющих определенную величину компоненты скорости вдоль направления лазерного пучка. Предполагая линейную зависимость интенсивности излучения от числа излучающих ионов и стационарность процесса, можно записать:

$$I_{\theta,\varphi}(\nu_L, \vec{r}_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\vec{v}, \vec{r}_0) d\mathbf{v}_{\perp}. \quad (1.2)$$

Здесь $I_{\theta,\varphi}(\nu_L, \vec{r}_0)$ – интенсивность флюоресценции из точки \vec{r}_0 при лазерной частоте ν_L и направлении лазерного пучка, определяемом углами θ , φ ; $g(\vec{v}, \vec{r}_0)$ – нормированная функция распределения ионов по скоростям. Интегрирование в (1.2) ведется в пространстве скоростей по плоскости перпендикулярной направлению лазерного пучка. Выражение (1.2) является трёхмерным преобразованием Радона функции $g(\vec{v}, \vec{r}_0)$, см. раздел 3.1. В данном случае оно служит уравнением томографии,

связывающим измеряемую флюоресценцию с искомой функцией распределения.

1.4 Реконструкция по интерферометрическим измерениям.

Интерферометрические исследования оптически прозрачных объектов, имеющих внутри заметные вариации коэффициента преломления (иногда такие объекты называются фазовыми), известны достаточно давно. Однако только в последние десятилетия применение томографических методов позволило более полно использовать богатую информацию, содержащуюся в интерферограммах. Стало возможным восстановление локальных значений коэффициента преломления в некотором сечении или даже во всем объёме изучаемого фазового объекта. Некоторые аспекты применения томографии в этой области освещены в монографии [15].

Объясним принцип интерферометрии на примере исследования газовых потоков, поскольку в этой области она находит широкое применение. Пусть имеется однородный поток, с коэффициентом преломления n_0 . После воздействия на него (например, погружения в поток твёрдого тела) возникает неоднородное возмущение, характеризуемое распределением коэффициента преломления $n(\vec{r})$. Обычно предполагается, что луч света (т.е. направление нормали в каждой точке волнового фронта) не изменяет своего направления, а фронт волны запаздывает в среде, имеющей более высокий коэффициент преломления. При этом разность хода, выраженная в долях длины световой волны, между лучами, прошедшими через возмущённый и невозмущённый потоки имеет вид

$$\Delta N = \frac{1}{\lambda} \int_{-\infty}^{\infty} (n(p) - n_0) dp. \quad (1.3)$$

Здесь λ – длина световой волны, p – координата вдоль луча. Пределы интегрирования в (1.3) распространены до бесконечности в предположении, что разность $(n(p) - n_0)$ отлична от нуля только в ограниченной области.

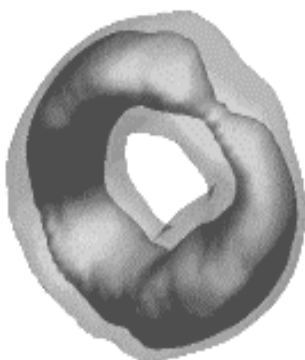
Разность хода ΔN для каждой пары лучей определяется путём расшифровки интерференционной картины. По полученным значениям ΔN из (1.3) в каждой точке определяется изменение показателя преломления, вызванное появлением неоднородности. Зная его, можно, исходя из определённых предположений, вычислить распределения плотности и температуры. Уравнение (1.3) связывает искомую функцию с результатами измерений через интеграл вдоль прямой. По сути, оно совпадает с уравнением, которое используется в медицинской рентгеновской томографии и эмиссионной томографии плазмы, см. раздел 2.1.

До недавнего времени при исследовании потоков задача томографии решалась, как правило, в предположении осевой симметрии. Однако развитие голографической интерферометрии [16] позволило относительно простыми способами получать данные (1.3) одновременно для нескольких различных направлений просвечивания. Это дало возможность восстанавливать структуры, не обладающие осевой симметрией. С точки зрения разработки методики таких исследований представляет интерес работа [17], где с большой точностью было восстановлено поле температур в сечении конвективного потока нагретого воздуха по изменению его коэффициента преломления. При реконструкции использовались данные, полученные на 19-и направлениях. Для оценки точности томографического метода проводились независимые измерения температуры при помощи специальных датчиков. Максимальная ошибка в местах, где были установлены датчики, составила менее 10-и процентов.

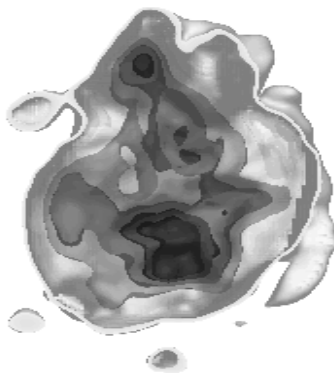
Появившееся впоследствии более сложное оборудование позволяет производить измерения для всего трёхмерного объекта, в результате чего могут быть получены двумерные проекции, что, в свою очередь, позволяет производить реконструкцию методами трёхмерной томографии.

Биологические клетки также являются фазовыми объектами. В настоящее время уровень экспериментальной техники вырос настолько, что стало возможным получение интерферограмм микрообъектов, в частности клеток *in vivo*. Томографическая реконструкция по этим данным внутренней структуры клеток открывает новые широчайшие возможности в биологических исследованиях. Начальные шаги в этом направлении были сделаны

в работе [18]. Там была получена первая двумерная томограмма распределения коэффициента преломления в поперечном сечении эритроцита человеческой крови. В дальнейшем метод был усовершенствован. В последующей работе этого же коллектива авторов была впервые проведена полная трехмерная реконструкция человеческого лимфоцита.



a)



б)

Рис. 1.8: Томографическая реконструкция клеток крови человека; *a)* эритроцит; *б)* лимфоцит.

Уже первые работы показали важность численного моделирования в процессе постановки эксперимента и обработки полученных данных. Только вычислительный эксперимент позволяет количественно судить о потенциальной точности восстановления, а также о пределах разрешения, которого можно достигнуть. Кроме того, с его помощью можно выбрать оптимальное расположение направлений, с которых проводятся измерения. В статье [19] были обобщены результаты численного моделирования и экспериментальной работы, а также намечены перспективы дальнейших исследований.

Пример результатов реконструкции приведён на рис. 1.8. На рис. 1.8,*а* изображены две изоповерхности распределения показателя преломления внутри эритроцита. На рис. 1.8,*б* представлено трёхмерное изображение лимфоцита в разрезе.

Глава 2

Связь с уравнением переноса излучения.

2.1 Вывод уравнения лучевой томографии.

Основой для многих уравнений томографии является уравнение переноса излучения. В приближении геометрической оптики оно, согласно [20], записывается следующим образом

$$\begin{aligned} \frac{\partial}{\partial s} I(\vec{r}, \vec{n}, \nu, t) + (\eta_a(\vec{r}, \vec{n}, \nu, t) + \eta_s(\vec{r}, \vec{n}, \nu, t)) I(\vec{r}, \vec{n}, \nu, t) = \\ = \frac{2}{n_r} \frac{\partial n_r}{\partial s} I(\vec{r}, \vec{n}, \nu, t) + \tilde{\mathbf{O}} I(\vec{r}, \vec{n}, \nu, t) + \varepsilon(\vec{r}, \vec{n}, \nu, t) \end{aligned} \quad (2.1)$$

Соотношение (2.1) связывает для рассматриваемого луча интенсивность излучения $I(\vec{r}, \vec{n}, \nu, t)$ со свойствами среды и распределением источников, создающих эмиссию $\varepsilon(\vec{r}, \vec{n}, \nu, t)$.

В общем случае дифференцирование вдоль луча имеет вид

$$\frac{\partial}{\partial s} = \vec{n} \nabla + \frac{\partial \vec{n}}{\partial s} \nabla_{\vec{n}} + \frac{1}{|\vec{v}_g|} \left(\frac{\partial}{\partial t} + \frac{\partial \nu}{\partial t} \frac{\partial}{\partial \nu} \right). \quad (2.2)$$

Здесь \vec{n} – единичный вектор, направленный по касательной к лучу; ∇ – градиент; \vec{r} – радиус-вектор точки; \vec{v}_g – групповая скорость волнового пакета; $\nabla_{\vec{n}}$ – градиент в пространстве вектора \vec{n} ; t – время; ν – частота излучения.

Первый член в правой части (2.1) связан с преломлением. Параметр n_r , определяемый как

$$n_r = \frac{c}{v} \sqrt{|\vec{v}_g| \left| d^3k/dvd\Omega \right|}, \quad (2.3)$$

называется лучевым показателем преломления. В (2.3) c – скорость света в вакууме, через $\left| d^3k/dvd\Omega \right|$ обозначен якобиан замены переменных при переходе от интегрирования по частоте v и телесному углу Ω к интегрированию по компонентам волнового вектора \vec{k} , который связан с частотой дисперсионным уравнением $v = v(\vec{k}, \vec{r})$, зависящим от свойств среды.

Сумма коэффициентов поглощения η_a и рассеяния η_s определяет ослабление интенсивности вдоль луча. Коэффициент η_s связан с переходом энергии излучения за счёт взаимодействия со средой на другие пространственные направления. При этом рассеяние может происходить как с изменением, так и без изменения частоты. Обратным является процесс возвращения энергии на направление \vec{n} , который описывается оператором $\tilde{\sigma}$ в уравнении (2.1),

$$\tilde{\sigma} I(\vec{r}, \vec{n}, v, t) = \int \sigma(\vec{r}, t | v, \vec{n} \leftarrow v', \vec{n}') I(\vec{r}, \vec{n}', v', t) dv' d\Omega_{\vec{n}'}. \quad (2.4)$$

При этом

$$\eta_s(\vec{r}, \vec{n}, v, t) = \int \sigma(\vec{r}, t | v', \vec{n}' \leftarrow v, \vec{n}) dv' d\Omega_{\vec{n}'}. \quad (2.5)$$

В выражениях (2.4) и (2.5) $\sigma(\vec{r}, \vec{n}, \vec{n}', v, v', t)$ означает сечение рассеяния единицы объёма для соответствующего процесса.

В томографии, обычно предполагается, что все процессы стационарны. Вообще говоря, это верно не всегда. В частности, при диагностике грудной клетки среду нельзя считать стационарной из-за дыхания, а также, из-за движения сердца. Ещё в меньшей степени условиям стационарности удовлетворяют задачи исследования плазмы. В этом случае можно говорить лишь о распределениях параметров, усреднённых по времени регистрации проекционных

данных. Другой причиной, приводящей к временной зависимости, является нестабильность источников зондирующего излучения, например, рентгеновских трубок. Несмотря на это, будем считать, что условия стационарности выполнены.

Перейдём к выводу уравнения медицинской рентгеновской томографии. При распространении рентгеновского излучения в теле человека преломление практически отсутствует. Поэтому можно считать, что направление вектора \vec{n} постоянно. Следовательно, производная $\partial\vec{n}/\partial s$ обращается в ноль, т.е. луч не искривляется. В силу предположения стационарности, производные по времени в (2.2) также обращаются в ноль. Таким образом, дифференцирование вдоль луча, в данном случае, совпадает с дифференцированием в направлении вектора \vec{n} .

Далее, будем предполагать, что среда является изотропной и дисперсия в ней отсутствует (последнее означает, что групповая скорость не зависит от частоты излучения). В этих условиях, лучевой показатель преломления совпадает с обычным показателем преломления, который, в свою очередь, равен квадратному корню из диэлектрической проницаемости. В диапазоне рентгеновских частот для большинства сред его можно считать постоянным, следовательно, будет иметь место $\partial n_r / \partial s = 0$. Учитывая принятые выше предположения, перепишем (2.1) в виде

$$\begin{aligned} \nabla I(\vec{r}, \vec{n}, \nu) + (\eta_a(\vec{r}, \vec{n}, \nu) + \eta_s(\vec{r}, \vec{n}, \nu))I(\vec{r}, \vec{n}, \nu) = \\ = \tilde{\sigma}I(\vec{r}, \vec{n}, \nu) + \varepsilon(\vec{r}, \vec{n}, \nu) \end{aligned} \quad (2.6)$$

Вид уравнения томографии существенно зависит от свойств регистрирующего оборудования. Во многих приложениях используется следующая модель. Считается, что детектор измеряет интенсивность излучения, и имеет бесконечно малую угловую апертуру и бесконечно малую площадь. Это означает, что регистрация производится в одной точке и в одном направлении. В рассматриваемой задаче существуют только внешние источники, создаваемые рентгеновской трубкой. С учётом предполагаемых свойств детектора можно считать, что для каждого луча существует

лежащий на нём точечный источник, излучающий в одном направлении (вдоль луча в сторону детектора).

Кроме прямого излучения на детектор также попадает рассеянное излучение, в частности от других источников, если таковые имеются. Даже если источник единственный, точечный и однонаправленный, т.е. $I_0(\vec{r}, \vec{n}, \nu) = I_0(\nu) \delta(\vec{r} - \vec{r}_0) \delta(\vec{n} - \vec{n}_0)$ (чего практически невозможно добиться), то при многократном рассеянии часть излучения, ушедшая с луча, вновь вернётся на него. Таким образом, процесс рассеяния не может быть полностью исключён из рассмотрения в задаче трансмиссионной томографии и проблема засветки детектора рассеянными фотонами может встать весьма остро. Однако в теле человека рентгеновское излучение рассеивается достаточно слабо, что позволяет в медицинской томографии пренебрегать членом $\tilde{\sigma} I(\vec{r}, \vec{n}, \nu)$ в уравнении (2.6). Исходя из вышесказанного, перепишем (2.6) в виде обыкновенного линейного дифференциального уравнения:

$$\frac{\partial I(p, \nu)}{\partial p} + \eta(p, \nu) I(p, \nu) = 0. \quad (2.7)$$

В (2.7) p – координата вдоль луча, $\eta \equiv \eta_a + \eta_s$ – суммарный коэффициент ослабления. Частота излучения ν рассматривается как параметр. Интенсивность источника $I_0(\nu)$, расположенного в точке $p = a$ является граничным условием для уравнения (2.7). Интегрируя его с этим условием, получаем:

$$I(p, \nu) = I_0(\nu) \exp\left(-\int_a^p \eta(p', \nu) dp'\right). \quad (2.8)$$

Пусть детектор находится на луче в точке $p = b$. Выражение (2.8), взятое при $p = b$, представляет собой интенсивность излучения на частоте ν в месте расположения детектора. Любой физический прибор измеряет данные в пределах некоторой полосы частот, причём его чувствительность в зависимости от частоты, как правило, непостоянна. С другой стороны, источники

рентгеновского излучения, в частности, рентгеновские трубки, не являются монохроматическими. Поэтому в процессе регистрации происходит интегрирование интенсивности по частотам, и показание детектора f выражается следующим образом:

$$f \equiv \int_{J(\nu) \neq 0} J(\nu) I_0(\nu) \exp\left(-\int_a^b \eta(p', \nu) dp'\right) d\nu. \quad (2.9)$$

Здесь $J(\nu)$ – чувствительность детектора, интегрирование производится по области, где $J(\nu)$ отлична от нуля.

Как правило, коэффициент ослабления тем выше, чем меньше энергия зондирующего излучения. В случае если источник не монохроматический, это приводит к тому, что спектральный состав излучения при прохождении через объект изменяется: оно становится более жёстким. Этот эффект в томографии называется *hardness*. Явление *hardness* может существенно исказить результаты реконструкции, особенно тогда, когда в диагностируемый участок тела включаются ткани, сильно различающиеся по своим свойствам, например, кости и жир.

Предположим, что источник является монохроматическим: $I_0(\nu) = I_0 \delta(\nu - \nu_0)$, а чувствительность детектора на частоте ν_0 равна единице. Обозначим $\eta(p, \nu_0)$ через $\eta(p)$. Тогда из (2.9) получаем:

$$I = I_0 \exp\left(-\int_a^b \eta(p) dp\right). \quad (2.10)$$

Уравнение (2.10) является основной математической моделью для медицинской рентгеновской томографии, а также для ряда других томографических методов. Поделив (2.10) на I_0 и прологарифмировав, имеем

$$-\ln\left(\frac{I}{I_0}\right) \equiv f = \int_a^b \eta(p) dp. \quad (2.11)$$

Левая часть (2.11) предполагается известной, таким образом, мы приходим к уравнению, связывающему проекционные данные f с искомой функцией через интеграл вдоль отрезка прямой линии. На значимость этого уравнения для томографии уже было указано во введении.

Вернёмся к стационарному уравнению переноса (2.6). По-прежнему будем пренебрегать преломлением, а также считать среду изотропной и недиспергирующей. Предположим, что внешние источники отсутствуют, но излучает сама среда. Распределение коэффициента эмиссии в ней задаётся функцией $\varepsilon(\vec{r}, \vec{n}, \nu)$. Будем предполагать, что среда не поглощает и не рассеивает собственного излучения. Тогда для точечного детектора с бесконечно узкой апертурой вместо (2.7) приходим к уравнению

$$\frac{\partial I(p, \nu)}{\partial p} = \varepsilon(p, \nu). \quad (2.12)$$

Проинтегрируем (2.12) с граничным условием $I(p, \nu)|_{p=-\infty} = 0$. Физически это означает, что детектор регистрирует лишь излучение, распространяющееся вдоль луча в положительном направлении. Получаем:

$$I(p, \nu) = \int_{-\infty}^p \varepsilon(p', \nu) dp'. \quad (2.13)$$

Если в (2.13) не рассматривать зависимость от частоты, верхний предел интегрирования отождествить с положением детектора, а в качестве нижнего предела взять координату точки пересечения луча с границей излучающей среды, наиболее удалённой от детектора, то, с точностью до обозначений, мы получим уравнение (2.11). Таким образом, мы видим, что в рассматриваемом приближении эмиссионная и трансмиссионная задачи томографии совпадают.

2.2 Уравнения для SPECT и PET.

Выведем уравнение эмиссионной томографии в среде, которая частично поглощает собственное излучение. Такая постановка задачи возникает в методах медицинской диагностики SPECT и PET (поскольку излучение изотопа поглощается тканями тела, см. раздел 1.2), в исследованиях оптически плотной плазмы, а также в ряде других приложений.

Сначала проведём вывод, основанный на уравнении трансмиссионной томографии (2.10). Можно полагать, что в каждой точке \vec{r}_0 , находится точечный источник интенсивностью $\varepsilon(\vec{r}_0)$. Вклад от него в показание детектора определяется согласно (2.10):

$$f(\vec{r}_0) = \varepsilon(\vec{r}_0) \exp\left(-\int_{\vec{r}_0}^b \eta(p') dp'\right). \quad (2.14)$$

Интегрирование в (2.14) ведётся по отрезку, соединяющему \vec{r}_0 и точку, в которой расположен детектор. Для того чтобы получить полный сигнал, надо проинтегрировать (2.14) по лучу, заканчивающемуся на детекторе:

$$f = \int_{-\infty}^b \varepsilon(p) \exp\left(-\int_p^b \eta(p') dp'\right) dp. \quad (2.15)$$

Отметим, что, поскольку все функции, с которыми приходится сталкиваться на практике, имеют финитный носитель, в качестве нижнего предела внешнего интегрирования в (2.15) достаточно взять ту точку пересечения рассматриваемой прямой с границей излучающей среды, которая наиболее удалена от детектора.

Соотношение (2.15) можно также вывести непосредственно из уравнения переноса излучения. Учитывая, что теперь среда как излучает, так и поглощает, вместо (2.7) запишем:

$$\frac{\partial I(p, \nu)}{\partial p} + \eta(p, \nu) I(p, \nu) = \varepsilon(p, \nu). \quad (2.16)$$

Проинтегрировав (2.16) с таким же граничным условием, что и (2.12), получаем для точки $p = b$, в которой находится детектор:

$$I(b, \nu) = \left(\int_{-\infty}^b \varepsilon(p, \nu) \exp \left(\int_{-\infty}^p \eta(p', \nu) dp' \right) dp \right) \exp \left(- \int_{-\infty}^b \eta(p', \nu) dp' \right). \quad (2.17)$$

Вносим множитель под знак интеграла и проводим преобразование:

$$\begin{aligned} & \exp \left(\int_{-\infty}^p \eta(p', \nu) dp' \right) \exp \left(- \int_{-\infty}^b \eta(p', \nu) dp' \right) = \\ & = \exp \left(- \left(\int_p^{-\infty} \eta(p', \nu) dp' + \int_{-\infty}^b \eta(p', \nu) dp' \right) \right) = \exp \left(- \int_p^b \eta(p', \nu) dp' \right). \end{aligned}$$

Таким образом получаем:

$$I(b, \nu) = \int_{-\infty}^b \varepsilon(p, \nu) \exp \left(- \int_p^b \eta(p', \nu) dp' \right) dp. \quad (2.18)$$

Уравнение (2.18) переходит в (2.15) если не принимать во внимание эффекты, связанные с немонахроматичностью излучения.

Уравнение томографии (2.15) соответствует СПЕСТ. Для РЕТ оно может быть приведено к более простому виду. Характерной особенностью метода РЕТ является то, что реконструкция изображения в нём базируется на регистрации пары γ -квантов, полученных при аннигиляции позитрона. Для этого два детектора, расположенные на противоположных сторонах сканера, должны с небольшой разницей во времени (порядка 10^{-8} с) зарегистрировать γ -квант. При этом каждый детектор может работать в режиме регистрации совпадения событий с множеством расположенных напротив детекторов.

Пусть детекторы, регистрирующие парные γ -кванты, имеют на соединяющей их прямой координаты a и b , при этом $a < b$.

Рассмотрим точку с координатой p , лежащую на этой прямой между детекторами, $a < p < b$. Поскольку позитрон в веществе проходит до аннигиляции незначительное расстояние, можно считать, что коэффициент эмиссии $\varepsilon(p)$ в рассматриваемой точке определяется концентрацией изотопа в ней. По аналогии с (2.15) сигналы, регистрируемые детекторами, будут, соответственно,

$$f_a = \int_a^b \varepsilon(p) \exp\left(-\int_a^p \eta(p') dp'\right) dp \quad (2.19)$$

и

$$f_b = \int_a^b \varepsilon(p) \exp\left(-\int_p^b \eta(p') dp'\right) dp. \quad (2.20)$$

В (2.19), (2.20) бесконечные пределы интегрирования заменены координатами детекторов, поскольку $\varepsilon(p) = 0$ при $p \leq a$, $p \geq b$. Кроме того, в (2.19) учтено, что интегрирование ведётся в сторону убывания переменных p и p' . Суммируя (2.19) и (2.20), получаем

$$f \equiv f_a + f_b = \int_a^b g(p) \exp\left(-\int_a^b \eta(p') dp'\right) dp = \eta_{ab} \int_a^b g(p) dp. \quad (2.21)$$

Выражение (2.21) является уравнением томографии для ПЕТ.

Уравнение (2.21) значительно проще, чем (2.15). Во-первых, для ПЕТ диагностики не нужно знать распределения коэффициента поглощения тканями, а лишь интегралы от него, которые можно получить в процессе измерения без дополнительной реконструкции. Во-вторых, если функция η известна, то (2.21) можно решать теми же методами, что и (2.11), разделив обе его части на коэффициенты η_{ab} . Решение же уравнения (2.15) и в этом случае вызывает затруднения. О методах обращения (2.15) изложено ниже в главе 7.

Глава 3

Математические основы томографии.

3.1 Преобразование Радона.

Уравнения, подобные (2.11), являются предметом рассмотрения интегральной геометрии. Основы этой дисциплины были заложены в начале двадцатого века работами Функа (P. Funk) и Радона (J. Radon). В частности, Радоном решалась задача о восстановлении функции двух переменных по её интегралам вдоль прямых [21]. Как было показано выше, именно к ней сводится двумерная рентгеновская томография. Поэтому становится понятным, насколько велико значение интегральной геометрии для развития методов томографической реконструкции.

Результаты Радона были обобщены на случай m -мерных пространств ($m > 1$). Впоследствии преобразование, переводящее быстро убывающую функцию $g(\vec{r})$, $\vec{r} \in \mathbb{R}^m$, в набор её интегралов по всевозможным $(m-1)$ -мерным гиперплоскостям в \mathbb{R}^m , было названо его именем. Преобразование Радона определяется согласно формуле, см. [22]:

$$f(\vec{n}, l) \equiv (\mathbf{R}g)(\vec{n}, l) = \int_{\langle \vec{r}, \vec{n} \rangle = l} g(\vec{r}) d\vec{r}, \quad (3.1)$$

где $\vec{n} \in S^{m-1}$ (S^{m-1} – единичная сфера в \mathbb{R}^m), $l \in \mathbb{R}^1$. Здесь и далее через $\langle \circ, \circ \rangle$ обозначено скалярное произведение. В частности, для двумерного пространства преобразование Радона принимает вид

$$f(\varphi, l) \equiv (\mathbf{R}g)(\varphi, l) = \int_{-\infty}^{\infty} g\left(\sqrt{l^2 + p^2}, \varphi + \arctan(p/l)\right) dp, \quad (3.2)$$

а для трёхмерного

$$f(\varphi, \theta, l) \equiv (\mathbf{R}g)(\varphi, \theta, l) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(l \cos \varphi \sin \theta - u \sin \varphi - v \cos \varphi \cos \theta, \\ l \sin \varphi \sin \theta - u \cos \varphi - v \sin \varphi \cos \theta, l \cos \theta + v \sin \theta) dudv. \quad (3.3)$$

В выражениях (3.2), (3.3) $p, u, v \in \mathbf{R}^1$, θ – полярный угол (угол между осью Z и вектором \vec{n}), φ – азимутальный угол (угол между осью X и проекцией вектора \vec{n} на плоскость $z=0$).

Преобразование Радона является функцией, определенной на единичном цилиндре $S^{m-1} \times \mathbf{R}^1$ в \mathbf{R}^{m+1} . Имеет место следующее свойство чётности $(\mathbf{R}g)(-\vec{n}, -l) = (\mathbf{R}g)(\vec{n}, l)$. Другие свойства преобразования Радона изложены, например, в монографии [22].

Зафиксируем вектор \vec{n} . Функцию $f_{\vec{n}}(l) \equiv (\mathbf{R}g)(\vec{n}, l)|_{\vec{n} = const}$, зависящую от одной переменной, обычно называют *одномерной параллельной проекцией* функции $g(\vec{r})$. Ей можно дать следующую физическую интерпретацию. Пусть в постановке эмиссионной томографии объект исследуется при помощи линеек детекторов. При этом оптическая система фокусирует на каждую из линеек участок плоского волнового фронта. Тогда распределение интенсивности зарегистрированного излучения вдоль линейки будет представлять собой одномерную проекцию от распределения коэффициента эмиссии в рассматриваемом сечении объекта.

Введём прямое и обратное многомерное преобразование Фурье согласно формулам

$$\hat{f}(\vec{\omega}) = \int_{\mathbf{R}^m} f(\vec{r}) \exp(-i\langle \vec{r}, \vec{\omega} \rangle) d\vec{r}, \quad \vec{r}, \vec{\omega} \in \mathbf{R}^m, \\ f(\vec{r}) = (2\pi)^{-m} \int_{\mathbf{R}^m} \hat{f}(\vec{\omega}) \exp(i\langle \vec{r}, \vec{\omega} \rangle) d\vec{\omega}. \quad (3.4)$$

Рассмотрим пространство бесконечно дифференцируемых функций, ограниченных по модулю вместе со всеми своими производными. На них определим линейный оператор V^γ , ($\gamma < m$),

называемый потенциалом Рисса, действие которого задаётся соотношением

$$(\mathbf{I} f)(\vec{r}) = (2\pi)^{-m} \int_{\mathbf{R}^m} |\vec{\omega}|^{-\gamma} \hat{f}(\vec{\omega}) \exp(i\langle \vec{r}, \vec{\omega} \rangle) d\vec{\omega}, \quad (3.5)$$

Преобразование Радона имеет формулу обращения:

$$g = \frac{1}{2} (2\pi)^{1-m} \mathbf{I}^{-\beta} \mathbf{R}^{\#} \mathbf{I}^{\beta-m+1} f, \quad (3.6)$$

Здесь потенциал Рисса $\mathbf{I}^{\beta-m+1}$ действует на одномерные проекции, а потенциал Рисса $\mathbf{I}^{-\beta}$ действует на функции, заданные на \mathbf{R}^m . Оператор $\mathbf{R}^{\#}$ называется обратным проецированием. Он представляет собой интегрирование по единичной сфере, т.е. по направлению вектора \vec{n} :

$$(\mathbf{R}^{\#} f)(\vec{r}) = \int_{S^{m-1}} f(\vec{n}, \langle \vec{r}, \vec{n} \rangle) d\vec{n}. \quad (3.7)$$

Доказательство формулы (3.6) приведено в [22].

Пусть $m = 2$. Тогда вектор \vec{n} определяется лишь одним углом φ . Введём на плоскости, где задана функция g , полярные координаты (ψ, r) : ψ – угол между положительным направлением оси X и вектором \vec{r} , $r = |\vec{r}|$. В этом случае оператор обратного проецирования $\mathbf{R}^{\#}$ будет иметь вид

$$\begin{aligned} (\mathbf{R}^{\#} f)(\psi, r) &= \int_0^{2\pi} f(\varphi, r \cos \psi \cos \varphi + r \sin \psi \sin \varphi) d\varphi = \\ &= \int_0^{2\pi} f(\varphi, r \cos(\varphi - \psi)) d\varphi \end{aligned} \quad (3.8)$$

Положив в (3.6) $\beta = 0$, получим

$$g = (4\pi)^{-1} \mathbf{R}^{\#} \mathbf{I}^{-1} f, \quad (3.9)$$

поскольку \mathbf{I}^0 есть тождественный оператор. Учитывая (3.5) и (3.8), запишем формулу обращения (3.9) для двумерного случая:

$$g(\psi, r) = \frac{1}{8\pi^2} \int_0^{2\pi} \left(\int_{-\infty}^{\infty} |\omega| \hat{f}_{\varphi}(\omega) \exp(i\omega r \cos(\varphi - \psi)) d\omega \right) d\varphi. \quad (3.10)$$

В (3.10) через $\hat{f}_{\varphi}(\omega)$ обозначено преобразование Фурье одномерной проекции функции $g(\vec{r})$ снятой под углом φ .

Используя сведения из теории обобщённых функций выражение (3.9) можно переписать в другом виде. Обозначим через $(\mathbf{I}^{-1} f_{\varphi})^{\wedge}(\omega)$ Фурье-образ функции полученной при воздействии потенциала Рисса на одномерную проекцию. Согласно (3.5) имеем

$$(\mathbf{I}^{-1} f_{\varphi})^{\wedge}(\omega) = |\omega| \hat{f}_{\varphi}(\omega) = \operatorname{sgn}(\omega) \omega \hat{f}_{\varphi}(\omega) = \frac{\operatorname{sgn}(\omega)}{i} \hat{f}'_{\varphi}(\omega). \quad (3.11)$$

В (3.11) через sgn обозначена функция, равная 1 при положительном значении аргумента и -1 – при отрицательном; через $\hat{f}'_{\varphi}(\omega)$ – Фурье-образ производной от проекции. Известно, что обратным преобразованием Фурье от $\operatorname{sgn}(\omega)/i$ является обобщённая функция, называемая преобразованием Гильберта:

$$(\mathbf{H}f)(l) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(p)}{l-p} dp. \quad (3.12)$$

В (3.12) интеграл понимается в смысле главного значения Коши:

$$\int_{-\infty}^{\infty} \frac{f(p)}{p} dp = \lim_{\varepsilon \rightarrow 0} \left(\int_{-\infty}^{-\varepsilon} \frac{f(p)}{p} dp + \int_{\varepsilon}^{\infty} \frac{f(p)}{p} dp \right). \quad (3.13)$$

Применяя к (3.11) теорему о свёртке и используя (3.8), получаем из (3.9) следующее выражение для формулы обращения двумерного преобразования Радона:

$$g(\psi, r) = \frac{1}{4\pi^2} \int_0^{2\pi} \left(\int_{-\infty}^{\infty} \frac{1}{r \cos(\varphi - \psi) - l} \frac{\partial}{\partial l} f_{\varphi}(l) dl \right) d\varphi. \quad (3.14)$$

Формула (3.10) (или её эквивалент (3.14)) является одной из основных для двумерной томографии. Она может быть интерпретирована следующим образом. Каждая одномерная проекция фильтруется посредством фильтра высоких частот, который в Фурье-пространстве представляется умножением на модуль частоты. В современной литературе эта операция также называется *грат-фильтрацией*. Вслед за этим производится интегрирование по вектору направлений – обратное проецирование. Поэтому алгоритмы, непосредственно реализующие формулу (3.10) или (3.14), ниже будут называться двумерными алгоритмами *обратного проецирования с фильтрацией*. Таких алгоритмов существует большое количество. Основное их отличие друг от друга – это способ построения фильтра. Некоторые из этих способов описаны в главе 4.

Согласно определению, трёхмерное преобразование Радона ставит в соответствие функции трёх переменных набор интегралов от неё по всевозможным плоскостям в трёхмерном пространстве. Выше было показано, что для широкого класса методов томографической диагностики измеряемой величиной является интеграл от искомой функции вдоль отрезка прямой. (Лишь в разделе 1.3.2 приведён пример, в котором проекционными данными являются интегралы по плоскостям.) В связи с этим трёхмерное преобразование Радона не играет в томографии столь важной роли, как двумерное. Чаще всего оно используется при построении алгоритмов реконструкции пространственного распределения коэффициента аттенюации, при движении источника излучения по некоторой кривой вокруг исследуемого объекта, см. главу 5.

Один из возможных вариантов формулы обращения трёхмерного преобразования Радона имеет вид (см., например [22]):

$$g(\vec{r}) = -\frac{1}{8\pi^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{2\pi} \frac{\partial^2 f(\vec{n}(\theta, \varphi), l)}{\partial l^2} \Big|_{l=\langle \vec{r}, \vec{n}(\theta, \varphi) \rangle} |\sin \theta| d\varphi d\theta, \quad (3.15)$$

где углы θ и φ определены таким же образом, как и для (3.3), вектор \vec{n} имеет координаты $(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$. Нетрудно видеть, что (3.15) получается из общей формулы (3.6) при $m=3$, $\beta=0$. Действительно, при этих значениях параметров имеем: $g = (8\pi^2)^{-1} \mathbf{R} \# \mathbf{I}^{-2} f$. Потенциал Рисса в степени -2 производит умножение Фурье-образа функции на ω^2 , что эквивалентно действию на саму функцию оператора $-\partial^2/\partial l^2$. Отсюда, учитывая (3.7), приходим к (3.15).

Существенным отличием выражений (3.14) и (3.15) является то, что второе из них, в отличие от первого, является локальным, т.е. для реконструкции функции в некоторой точке достаточно знать интегралы по плоскостям, имеющим пересечение с малой окрестностью этой точки. В то же время, для использования формулы (3.14) или (3.10), нужно иметь всю совокупность проекционных данных. Такая ситуация имеет общий характер: формулы обращения преобразования Радона в нечётномерных пространствах локальны, а чётномерных – не локальны.

3.2 P-Преобразование.

Наряду с преобразованием Радона были также введены преобразования, ставящие в соответствие быстро убывающей функции m переменных ($m > 2$) набор её интегралов по прямым. Одним из них является **P**-преобразование, рассмотренное в [22]:

$$f(\vec{n}, \vec{r}) \equiv (\mathbf{P}g)(\vec{n}, \vec{r}) = \int_{-\infty}^{\infty} g(\vec{r} + p\vec{n}) dp \quad (3.16)$$

Здесь $\vec{r} \in \mathbf{R}^m$, $\vec{n} \in \mathbf{S}^{m-1}$, $p \in \mathbf{R}^1$. При $m=3$ **P**-преобразование имеет вид:

$$f(\varphi, \theta, \vec{r}) \equiv (\mathbf{P}g)(\varphi, \theta, \vec{r}) = \int_{-\infty}^{\infty} g(x + p \cos \varphi \sin \theta, y + p \cos \varphi \sin \theta, z + p \cos \theta) dp. \quad (3.17)$$

В формуле (3.17) θ , φ – углы, определяемые как и ранее, задают направление прямой, по которой ведётся интегрирование; $\vec{r} = (x, y, z)$ – некоторая принадлежащая ей точка. Очевидно, что величина функции $f(\varphi, \theta, \vec{r})$ не зависит от положения точки \vec{r} на прямой интегрирования.

Зафиксировав в (3.17) углы θ и φ , мы получим математическое выражение для *двумерной параллельной проекции*. В частности, двумерной параллельной проекцией приближённо можно считать распределение интенсивности излучения по поверхности плоского детектора, когда излучающий объект находится от него на достаточно большом расстоянии.

\mathbf{P} -преобразование имеет формулу обращения, аналогичную (3.6). Согласно [22] она имеет следующий вид

$$g = \frac{1}{|S^{m-2}|} (2\pi)^{-1} \mathbf{I}^{-\beta} \mathbf{P}^{\#} \mathbf{I}^{\beta-1} f. \quad (3.18)$$

В (3.18) $|S^{m-2}|$ – площадь единичной сферы в пространстве \mathbf{R}^{m-1} , $\mathbf{P}^{\#}$ – оператор обратного проецирования для \mathbf{P} -преобразования:

$$(\mathbf{P}^{\#} f)(\vec{r}) = \int_{S^{m-1}} f(\vec{n}, E_{\vec{n}} \vec{r}) d\vec{n}, \quad (3.19)$$

где $E_{\vec{n}} \vec{r}$ – ортогональная проекция вектора $\vec{r} \in \mathbf{R}^m$ на подпространство, перпендикулярное вектору \vec{n} . В (3.18) потенциал Рисса $\mathbf{I}^{\beta-1}$ действует на функции, определённые на пространстве \mathbf{R}^{m-1} , т.е. на параллельные многомерные проекции, потенциал Рисса $\mathbf{I}^{-\beta}$ действует на функцию, полученную в результате обратного проецирования (3.19), зависящую от m переменных.

Реализации общей формулы (3.18) для трёхмерного случая даны, например, в [23, 24]. В этих работах полагается $\beta = 0$, т.е. в них осуществляется следующий алгоритм. Проекция фильтруется двумерным центрально симметричным гагр-фильтром, после чего производится обратное проецирование согласно (3.19).

3.3 Теорема о центральном слое.

Преобразования \mathbf{R} и \mathbf{P} , рассмотренные в двух предыдущих разделах связаны с преобразованием Фурье. Имеет место *теорема о центральном слое*, иногда называемая *центральной проекционной теоремой*. Она выражается следующими соотношениями, [22]:

$$(\mathbf{R}_{\vec{n}}g)^{\wedge}(\omega) = \hat{g}(\vec{n}\omega), \quad \vec{n} \in S^{m-1}, \quad \omega \in \mathbf{R}^1, \quad (3.20)$$

$$(\mathbf{P}_{\vec{n}}g)^{\wedge}(\vec{\omega}) = \hat{g}(\vec{\omega}_1), \quad \vec{n} \in S^{m-1}, \quad \vec{\omega} \in \mathbf{R}^{m-1}, \quad \vec{\omega}_1 \in \vec{n}_{\perp}. \quad (3.21)$$

Здесь через $\mathbf{R}_{\vec{n}}g$ и $\mathbf{P}_{\vec{n}}g$ обозначены соответственно одномерная и $m-1$ -мерная параллельные проекции функции g .

Приведём наглядное доказательство формулы (3.20) для $m = 2$. При этом будем считать, что все записанные ниже интегралы существуют. Пусть вектор \vec{n}_y параллелен оси Y , т.е. имеет координаты $(0,1)$. Очевидно, что в данном случае одномерная проекция функции $g(x, y)$ получается путём её интегрирования по прямым, параллельным оси X .

$$f_{\vec{n}_y}(y) \equiv (\mathbf{R}_{\vec{n}_y}g)(y) = \int_{-\infty}^{\infty} g(x, y) dx. \quad (3.22)$$

Возьмём от (3.22) преобразование Фурье по переменной y :

$$\hat{f}_{\vec{n}_y}(\omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \exp(-iy\omega_y) dx dy. \quad (3.23)$$

Умножим правую часть (3.23) на выражение $\exp(-ix\omega_x)$. При $\omega_x = 0$ оно равно единице. Таким образом имеем

$$\hat{f}_{\vec{n}_y}(\omega_y) = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} g(x, y) \exp(-i(x\omega_x + y\omega_y)) \Big|_{\omega_x=0} dx dy. \quad (3.24)$$

Левая часть (3.24) есть ни что иное, как сечение двумерного Фурье-образа $\hat{g}(\omega_x, \omega_y)$, взятое вдоль прямой $\omega_x = 0$. Для произвольного вектора \vec{n} можно осуществить такой поворот координатных осей, чтобы в новой системе он был направлен вдоль оси Y . Тогда в этой системе будет иметь место формула (3.24). Следовательно, для рассматриваемого частного случая мы пришли к утверждению теоремы о центральном слое.

Выражение (3.24) означает, что Фурье-образ одномерной проекции функции $g(x, y)$ равен сечению Фурье-образа $\hat{g}(\omega_x, \omega_y)$ прямой, проходящей через начало координат в Фурье-пространстве в том же направлении, что и прямая, вдоль которой вычисляется проекция. В применении к трёхмерному \mathbf{P} -преобразованию теорема о центральном слое (3.21) формулируется аналогичным образом. Двумерное преобразование Фурье параллельной проекции $f_{\vec{n}}$ от функции трёх переменных $g(x, y, z)$ совпадает с сечением трёхмерного Фурье-образа $\hat{g}(\omega_x, \omega_y, \omega_z)$ плоскостью в пространстве Фурье, проходящей через начало координат и имеющей вектор нормали \vec{n} .

Используя теорему о центральном слое, можно получить формулу обращения преобразования Радона (3.10). Выразим искомую функцию $g(x, y)$ через её преобразование Фурье и перейдём к полярным координатам:

$$g(x, y) = \frac{1}{4\pi^2} \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} \hat{g}(\omega_x, \omega_y) \exp(i(x\omega_x + y\omega_y)) d\omega_x d\omega_y =$$

$$= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{\infty} \hat{g}(\omega, \varphi) \exp(i(x\omega \cos \varphi + y\omega \sin \varphi)) \omega d\omega d\varphi. \quad (3.25)$$

Согласно теореме о центральном слое, при каждом фиксированном φ имеем $\hat{g}(\omega, \varphi) = \hat{f}_\varphi(\omega)$. Тогда, учитывая свойства симметрии преобразования Радона, уравнение (3.25) можно переписать в виде

$$g(x, y) = \frac{1}{8\pi^2} \int_0^{2\pi} \int_{-\infty}^{\infty} |\omega| \hat{f}(\omega, \varphi) \exp(i(x\omega \cos \varphi + y\omega \sin \varphi)) d\omega d\varphi. \quad (3.26)$$

Выражая в (3.26) x и y через полярные координаты r и ψ , приходим к (3.10).

Теорема о центральном слое служит основой для алгоритмов *Фурье-синтеза*. Их идея состоит в следующем. От параллельных проекций берётся преобразование Фурье и по полученным значениям оценивается Фурье-образ искомой функции. Например, в случае двух переменных из Фурье-образов проекций составляется Фурье-образ реконструируемой функции, заданный в полярных координатах. Затем производится его интерполяция на сетку в декартовых координатах, после чего производится двумерное обратное преобразование Фурье, см. например [4].

3.4 Томография в расходящихся лучах.

3.4.1 Веерная схема на плоскости.

Обратимся к рисунку 1.1. Схема сканирования, изображённая на рис. 1.1,а, характеризуется тем, что для каждого поворота рамы, на которой крепятся источник и детектор, регистрируются данные, связанные с интегралами вдоль семейства параллельных прямых, из которых может быть получена одномерная параллельная проекция. Это позволяет применять для реконструкции формулу обращения двумерного преобразования Радона. В веерной схеме сканирования, см. рис. 1.1,б, данные, регистрируемые при фиксированном положении системы источник – детектор, не образуют параллельной проекции. Часто их называют *веерной* проекцией,

подчёркивая тем самым различие между двумя рассматриваемыми случаями.

В монографии [22] было введено **D**-преобразование:

$$f(\vec{s}, \vec{n}) \equiv (\mathbf{D}g)(\vec{s}, \vec{n}) = \int_0^{\infty} g(\vec{s} + p\vec{n}) dp, \quad (3.27)$$

где $\vec{s} \in \mathbf{R}^m$, $\vec{n} \in \mathbf{S}^{m-1}$, $p \in \mathbf{R}^1$. Параметр \vec{s} в (3.27) пробегает некоторое множество точек в \mathbf{R}^m . Для $m = 2$ при фиксированном \vec{s} уравнение (3.27) является математическим выражением веерной проекции с источником в точке \vec{s} . Когда \vec{s} пробегает окружность радиуса r_s , выражение (3.27) называется *веерным* преобразованием функции $g(\vec{r})$.

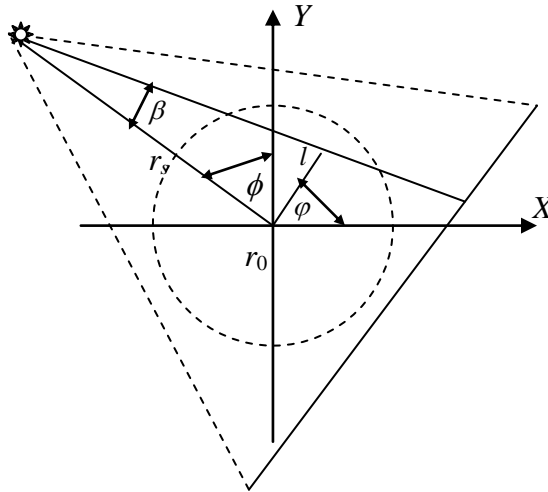


Рис. 3.1: К формуле обращения веерного преобразования.

Для реконструкции по веерным проекционным данным, не может быть непосредственно использована формула обращения (3.10). Обозначим через ϕ угол между вектором \vec{s} и осью Y , через β – угол между вектором \vec{s} и лучом, связанным с текущим

значением проекции, см. рис. 3.1. Для параллельных данных прямые, по которым ведётся интегрирование, параметризуются посредством угла φ между нормалью и ось X , и расстояния до начала координат l , см. рис. 3.1. Нетрудно видеть, что параметры ϕ , β связаны с параметрами φ , l посредством соотношений

$$l = r_s \sin \beta, \quad \varphi = \phi + \beta. \quad (3.28)$$

Таким образом, имея верные проекционные данные, можно по формулам (3.28) перевести их в параллельные данные, после чего использовать для реконструкции алгоритмы обращения двумерного преобразования Радона.

На основании формулы обращения двумерного преобразования Радона (3.14) может быть выведена формула обращения верного преобразования. Для этого в (3.14) проводится замена переменных (3.28). Опуская детали, приведём конечный результат, см. [4]:

$$g(\psi, r) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_{-\infty}^{\infty} \frac{1}{\beta' - \beta} G(\beta, \phi, \beta') d\beta d\phi. \quad (3.29)$$

где

$$G(\beta, \phi, \beta') = \begin{cases} \frac{\beta' - \beta}{\sin(\beta' - \beta)} W \left(\frac{\partial f(\phi, \beta)}{\partial \phi} - \frac{\partial f(\phi, \beta)}{\partial \beta} \right), & |\beta| \leq \arcsin\left(\frac{r_0}{r_s}\right), \\ 0, & |\beta| > \arcsin\left(\frac{r_0}{r_s}\right) \end{cases},$$

$$\beta' = \arctan\left(\frac{r \cos(\phi - \psi)}{r_s + r \sin(\phi - \psi)}\right), \quad -\frac{\pi}{2} < \beta' < \frac{\pi}{2},$$

$$W = \sqrt{(r \cos(\phi - \psi))^2 + (r_s + r \sin(\phi - \psi))^2}, \quad W > 0.$$

Здесь r_0 – радиус круга, в котором содержится носитель функции $g(\psi, r)$, предполагается, что $\arcsin(r_0/r_s) < \pi/2$.

Из сравнения формул обращения двумерного преобразования Радона (3.14) и веерного преобразования (3.29) следует важный, с точки зрения алгоритмической реализации вывод. При построении алгоритмов по формуле (3.14) либо (3.10) проекции можно сначала отфильтровать и по ним проводить обратное проецирование. В случае обращения веерного преобразования, перед фильтрацией каждая проекция должна быть умножена на значение W^{-1} , зависящее от координат точки, в которой производится реконструкция. Таким образом, для различных точек проекции должны фильтроваться разными фильтрами. Это приводит к снижению вычислительной эффективности алгоритмов, основанных на формуле (3.29).

3.4.2 Трёхмерная задача.

В настоящее время при решении задач трёхмерной томографии формула обращения **P**-преобразования используется сравнительно редко. В первую очередь это связано с конструкцией оборудования, производящего сбор проекционных данных. Во многих практически важных приложениях, в частности в медицинской рентгеновской томографии, регистрируются двумерные проекции, полученные в расходящихся лучах, а не в параллельных. Таким образом, в данном случае для построения алгоритмов реконструкции более удобным является **D**-преобразование (3.27).

Для $\vec{s} \in \mathbb{R}^3$ преобразование (3.27) иногда называют *конусным* или просто *лучевым*. Наиболее часто оно используется в следующей форме. Пусть источник зондирующего излучения движется по некоторой кривой K в трёхмерном пространстве, см. рис. 3.2. Определим лучевое преобразование на множестве $K \times S^2$. Пусть носитель функции $g(\vec{r})$ содержится в ограниченной области $D \subset \mathbb{R}^3$, а кривая задается непрерывной функцией $C: \Lambda \rightarrow \mathbb{R}^3$, т.е. $x = C_x(\lambda)$, $y = C_y(\lambda)$, $z = C_z(\lambda)$, $\lambda \in \Lambda$, Λ – интервал из \mathbb{R}^1 . Пусть кривая K удовлетворяет следующим условиям, см. [25].

- а) Кривая K проходит вне области D .
- б) Кривая K ограничена, непрерывна и почти всюду дифференцируема.

в) Для любых $\vec{n} \in S^2$, $\vec{r} \in D$ плоскость, перпендикулярная \vec{n} и проходящая через точку \vec{r} , пересекает кривую K в точке $\vec{C}(\lambda)$, для которой скалярное произведение $\langle \vec{C}'(\lambda), \vec{n} \rangle$ не равно нулю, $\vec{C}'(\lambda) = (dC_x(\lambda)/d\lambda, dC_y(\lambda)/d\lambda, dC_z(\lambda)/d\lambda)$.

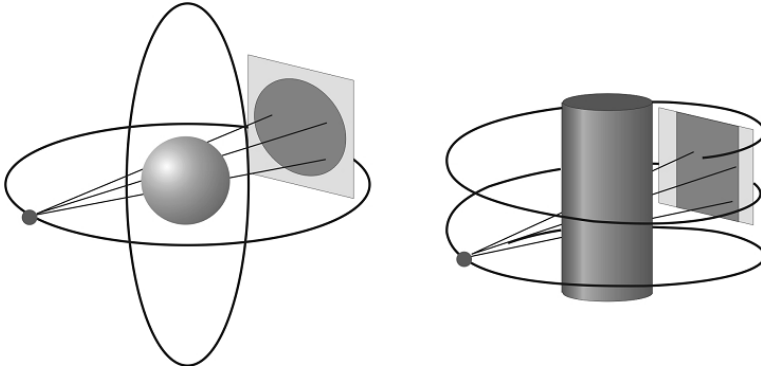


Рис. 3.2: Схемы регистрации проекционных данных с источником, движущимся по различным криволинейным траекториям в трёхмерном пространстве.

При выполнении условий а) – в) существует формула обращения

$$g(\vec{r}) = - \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\sin \theta}{2\pi i \langle \vec{C}'(\lambda), \vec{n} \rangle} \frac{\partial F(\lambda, \vec{n})}{\partial \lambda} d\theta d\varphi, \quad (3.30)$$

где

$$F(\lambda, \vec{n}) = (2\pi)^{\frac{3}{2}} \int_{\mathbb{R}^3} f_1(\lambda, \vec{\beta}) \exp(-2\pi i \langle \vec{\beta}, \vec{v} \rangle) d\vec{\beta}, \quad (3.31)$$

$$f_1(\lambda, \vec{\beta}) = \frac{1}{|\vec{\beta}|} \int_0^\infty g \left(\vec{C}(\lambda) + \frac{\vec{\beta}}{|\vec{\beta}|} p \right) dp,$$

\vec{r} , λ и \vec{n} связаны между собой соотношением:

$$\langle \vec{r}, \vec{n} \rangle = \langle \vec{C}(\lambda), \vec{n} \rangle, \quad \langle \vec{C}'(\lambda), \vec{n} \rangle \neq 0, \quad (3.32)$$

Здесь $\vec{\beta}, \vec{v} \in \mathbb{R}^3$, единичный вектор \vec{n} равен $\vec{v}/|\vec{v}|$ и имеет координаты $(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$.

На рис. 3.2 изображены различные траектории источника: две окружности, лежащие во взаимно перпендикулярных плоскостях, и спираль. Первая из них удовлетворяет условиям а) – в). Вторая – спираль – будет удовлетворять им в случае, если её размер вдоль оси будет больше либо равен размеру объекта.

Отметим, что приведённые выше условия носят достаточный характер. Если они выполняются, то реконструкцию можно проводить по формуле (3.30). Если же траектория источника им не удовлетворяет, то это не означает, что не существует другой формулы, дающей точное решение задачи в пределе бесконечного количества проекционных данных. Такие формулы обращения были действительно получены, в частности, при изменённом условии в).

Более детально формула (3.30) будет рассматриваться в главе 5, посвящённой алгоритмам трёхмерной томографической реконструкции для случая, когда источник движется по кривой в пространстве.

3.5. Регуляризация некорректно поставленных задач.

В начале 20-ого века Адамар (J. Hadamard) ввёл понятие корректной постановки задач математической физики. Пусть даны функции $g \in U_g$, $f \in U_f$ и оператор $\mathbf{A}: U_g \rightarrow U_f$; U_g , U_f – метрические пространства, с расстояниями $\rho_g(g_1, g_2)$, $\rho_f(f_1, f_2)$, соответственно. Далее, пусть определено понятие решения уравнения $\mathbf{A}g = f$, в том смысле, что определен некоторый оператор $\mathbf{T}: U_f \rightarrow U_g$, такой что каждому элементу $f \in U_f$ отвечает единственный элемент $g = \mathbf{T}f \in U_g$. Задача определения решения $g = \mathbf{T}f$ называется *устойчивой* на пространствах U_g ,

U_f , если для любого числа $\varepsilon > 0$ можно указать такое число $\delta(\varepsilon) > 0$, что из неравенства $\rho_f(f_1, f_2) \leq \delta(\varepsilon)$ следует $\rho_g(g_1, g_2) \leq \varepsilon$, где $g_1 = \mathbf{T}f_1$, $g_2 = \mathbf{T}f_2$, $g_1, g_2 \in U_g$, $f_1, f_2 \in U_f$. Согласно Адамару, см. [26], задача определения g по f называется *корректно поставленной* на паре метрических пространств U_g , U_f если удовлетворяются требования:

- а) для всякого $f \in U_f$ существует решение $g \in U_g$;
- б) решение определяется однозначно;
- в) задача устойчива на пространствах U_g , U_f .

Задачи, не удовлетворяющие хотя бы одному из требований а) – в) называются *некорректно поставленными*. При выполнении условий а) – б) оператор \mathbf{A} имеет обратный, а если выполняется и условие в), то обратный оператор непрерывен из U_f в U_g .

Решение вопроса о корректности конкретной томографической задачи определяется тем, какое уравнение связи между измерениями и искомыми распределениями лежит в её основе. При этом важную роль играет то, для какой пары пространств U_g , U_f ведётся рассмотрение. Вообще говоря, путём соответствующего выбора метрик $\rho_g(g_1, g_2)$, и $\rho_f(f_1, f_2)$ задачу можно сделать устойчивой. Однако на практике такой возможности нет. Функция f , известна из измерений с некоторой погрешностью, и способ определения этой погрешности задаёт метрику пространства U_f .

В большинстве случаев использования преобразований \mathbf{R} , \mathbf{P} или \mathbf{D} , которые действуют из \mathfrak{F} в \mathfrak{F} , где \mathfrak{F} – пространство быстро убывающих бесконечно дифференцируемых функций (пространство Шварца), обратные операторы \mathbf{R}^{-1} , \mathbf{P}^{-1} и \mathbf{D}^{-1} существуют, являются однозначными, но не являются непрерывными, см., например, [27]. Таким образом, задачи лучевой томографии некорректно поставлены. Вследствие этого, при неточном задании функции f получаемое решение g может существенно отличаться от истинного. В измеренных проекционных данных всегда присутствуют случайные ошибки.

Поэтому при решении таких задач томографии, необходимо применять *методы регуляризации*.

Тихоновым было введено следующее понятие корректности[26]. Задача решения уравнения $\mathbf{A}g = f$ называется *корректно поставленной по Тихонову*, если выполнены следующие условия.

а) Известно, что решение задачи существует и принадлежит некоторому множеству U_{gK} пространства U_g , т.е. $g \in U_{gK} \subset U_g$.

б) Решение единственно на множестве U_{gK} , т.е. для любой правой части $f \in U_{fK}$ существует единственный элемент $g \in U_{gK}$. Множество U_{fK} состоит из элементов $\mathbf{A}g$, где $g \in U_{gK}$.

в) Если вариации правой части не выводят её за пределы множества U_{fK} (следовательно, соответствующие g принадлежат U_{gK}), то существует непрерывная зависимость решения от правой части, а также существует обратный оператор \mathbf{A}^{-1} , являющийся непрерывным и, следовательно, ограниченным.

Множество U_{gK} , на образе которого U_{fK} оператор \mathbf{A}^{-1} существует и непрерывен, называется *множеством корректности*. Из сравнения условий корректности по Адамару и Тихонову видно, что корректность по Тихонову может быть достигнута за счёт сужения пространства U_g до множества корректности U_{gK} . Поэтому задачу, корректную по Тихонову (которая, возможно, некорректна по Адамару), называют *условно-корректной задачей*.

Известные подходы к нахождению множества корректности опираются на лемму, доказанную Тихоновым в 1943 году.

Пусть U_{gC} является компактным множеством пространства U_g , а оператор \mathbf{A} осуществляет непрерывное взаимно однозначное отображение U_{gC} на компактное множество $U_{fC} \subset U_f$. Тогда обратное отображение $U_{fC} \rightarrow U_{gC}$ является непрерывным.

Следовательно, решения для некорректно поставленной задачи достаточно выбирать из компактного множества U_{gC} . Для его построения на практике используют *стабилизирующий функционал*.

Неотрицательный функционал $\Omega(g)$, определенный на элементах пространства U_g , называется стабилизирующим если:

а) точное решение задачи принадлежит области определения функционала $\Omega(g)$;

б) для любого числа C_Ω множество $U_{g\Omega}$ элементов g , для которых выполняется неравенство $\Omega(g) \leq C_\Omega$ компактно в U_g .

Множество $U_{g\Omega}$ является множеством корректности для задачи $\mathbf{A}g = f$. Одним из способов решения условно-корректных задач состоит в нахождении экстремума стабилизирующего функционала.

Тихоновым было также введено понятие *регуляризующего семейства* линейных непрерывных операторов \mathbf{T}_α . Они обладают следующим свойством: для любой функции f , принадлежащей области определения оператора \mathbf{A}^{-1} , $\mathbf{T}_\alpha f \rightarrow g$ при $\alpha \rightarrow 0$, где g решение уравнения $\mathbf{A}g = f$; α – положительная функция от погрешности δ , с которой задана функция f . Если α задаётся в виде последовательности $1/n$, $n \in N$, то семейство $\mathbf{T}_{1/n}$ определяет регуляризующий итерационный процесс.

В [28] сформулированы общие принципы построения операторов \mathbf{T}_α . А именно, \mathbf{T}_α определяется по формуле

$$\mathbf{T}_\alpha = \overline{\mathbf{A}^* \mathbf{A}, \alpha} \mathbf{A}^*, \quad \alpha > 0. \quad (3.33)$$

В (3.33) черта сверху означает замыкание, \mathbf{A}^* – оператор, сопряженный к \mathbf{A} , $\mathcal{G}(\mathbf{A}^* \mathbf{A}, \alpha)$ – операторная функция. В [28] приведены условия, налагаемые на \mathcal{G} , для того чтобы семейство \mathbf{T}_α было регуляризующим.

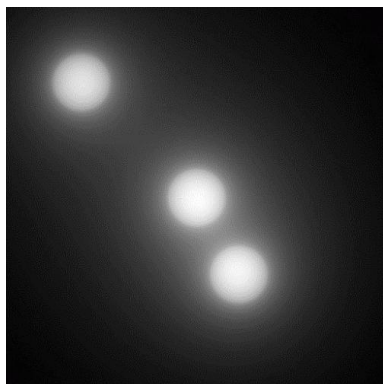
Помимо регуляризации Тихонова существуют другие методы получения устойчивых решений некорректно поставленных задач. За ними также закрепился термин "регуляризация". В дальнейшем изложении будут приведены примеры регуляризующих алгоритмов томографической реконструкции.

Глава 4

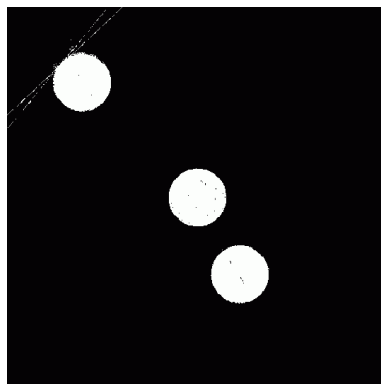
Фильтрация одномерных проекций.

4.1 Общие замечания.

В настоящее время большое количество алгоритмов двумерной томографии строится на основе формулы (3.10), или эквивалентной ей формулы (3.14). Как уже упоминалось в разделе 3.1, эти алгоритмы имеют следующую структуру. Каждая одномерная проекция фильтруется гаусс-фильтром, после чего производится обратное проецирование. Последняя процедура определяется формулой (3.8) и представляет собой интегрирование по углу, задающему направления сбора проекционных данных. Её алгоритмическая и программная реализация, как правило, не представляет серьёзных трудностей. Проблемы возникают при реализации гаусс-фильтра.



a)



б)

Рис. 4.1: Иллюстрация роли гаусс-фильтрации для томографической реконструкции: *a)* обратное проецирование по исходным проекциям; *б)* обратное проецирование после фильтрации.

Отметим, что если применять процедуру обратного проецирования непосредственно к проекциям, то полученное в результате изображение оказывается размытым, как это видно из рис. 4.1,*а*. Выявление высокочастотных структур на томограмме, в частности, разрывов, обеспечивается именно благодаря гапр-фильтрации, см. рис. 4.1,*б*.

Рассмотрим формулу обращения (3.14). В ней фильтрация производится посредством дифференцирования и свёртки проекции с сингулярным ядром $1/l$. Когда знаменатель в подынтегральном выражении становится равным нулю, интеграл расходится, в связи с этим он понимается в смысле главного значения, см. выражение (3.13). Чтобы построить алгоритм реконструкции, надо заменить сингулярное ядро некоторой функцией, которая бы допускала интегрирование в (3.14) и, в то же время, обладала бы свойствами, близкими к свойствам обобщённой функции $1/l$. Кроме того, надо численно реализовать процедуру дифференцирования.

На практике дифференцирование проекции и преобразование Гильберта часто осуществляют одной операцией – свёрткой с функцией, аппроксимирующей гапр-фильтр. Выбор такой функции можно производить, используя представление гапр-фильтра в пространстве Фурье, где он имеет вид модуля частоты. В частности, отсюда следует, что аппроксимирующую функцию следует выбирать такой, чтобы проекция после свёртки с ней имела нулевой момент (т.е. интеграл по области определения) равный нулю. Другие требования к аппроксимирующей функции будут приведены ниже.

В случае аппроксимации гапр-фильтра, исходя из его частотных характеристик, процесс разработки алгоритма на основании формулы (3.14) включает следующие этапы. Строится Фурье-образ, обладающий подходящими свойствами, после чего производится обратное преобразование Фурье. Полученная в результате функция используется для свёртки с одномерными проекциями. Однако построенный Фурье-образ может не иметь аналитического выражения для своего обратного преобразования. В этом случае удобнее использовать формулу обращения (3.10), т.е. производить фильтрацию непосредственно в Фурье области.

При построении функции, реализующей гапр-фильтр, возникает проблема, определения её спектральной ширины, поскольку спектр

проекция, вообще говоря, неограничен. Рассмотрим этот вопрос более подробно.

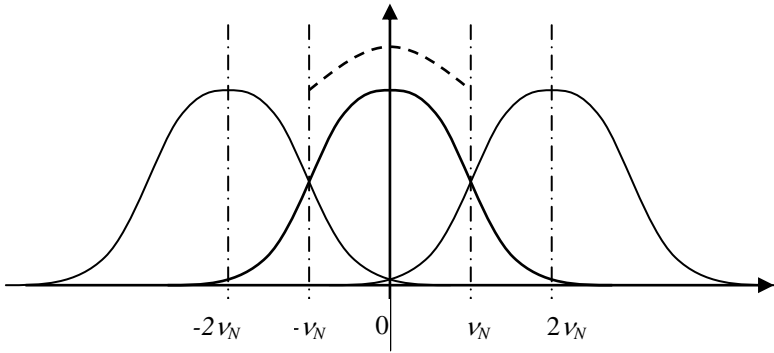


Рис. 4.2: К определению спектра дискретизованной проекции.

В данной главе мы будем использовать обычную частоту ν , которая связана с циклической частотой ω посредством соотношения $\omega=2\pi\nu$. При этом прямое и обратное преобразование Фурье будут иметь вид:

$$\begin{aligned} \hat{f}(\vec{\nu}) &= \int_{\mathbb{R}^m} f(\vec{r}) \exp(-2\pi i \langle \vec{r}, \vec{\nu} \rangle) d\vec{r}, & \vec{r}, \vec{\nu} \in \mathbb{R}^m, \\ f(\vec{r}) &= \int_{\mathbb{R}^m} \hat{f}(\vec{\nu}) \exp(2\pi i \langle \vec{r}, \vec{\nu} \rangle) d\vec{\nu}. \end{aligned} \quad (4.1)$$

Пусть проекции дискретизованы с шагом $h = 2/(N-1)$, N – число узлов, область дискретизации – $[-1; 1]$. Как известно, Фурье-образ дискретизованной функции является непрерывной периодической функцией с периодом $2\nu_N$, где ν_N – частота Найквиста, $\nu_N = 1/2h = (N-1)/4$. Обозначим через $\hat{f}_\varphi(\nu)$ истинный Фурье-образ проекции, зарегистрированной под углом φ (на рис. 4.2 он изображён жирной линией). Его связь с Фурье-образом дискретной проекции $\hat{f}_\varphi^h(\nu)$ выражается соотношением

$$\hat{f}_\varphi^h(\nu) = 2\nu_N \sum_{n=-\infty}^{\infty} \hat{f}_\varphi(\nu - 2\nu_N n) \quad (4.2)$$

Будем рассматривать $\hat{f}_\varphi^h(\nu)$ в пределах центрального периода $[-\nu_N; \nu_N]$, который также называется *главным частотным диапазоном*. Введём функцию $f_\varphi^1(l)$, Фурье-образ которой равен $\hat{f}_\varphi^h(\nu)/2\nu_N$, если $\nu \in [-\nu_N; \nu_N]$, если же $\nu \notin [-\nu_N; \nu_N]$, то $\hat{f}_\varphi^1(\nu)$ равен нулю. На рис. 4.2 $\hat{f}_\varphi^1(\nu)$ изображён штриховой линией.

Отличие Фурье-образов $\hat{f}_\varphi(\nu)$ и $\hat{f}_\varphi^1(\nu)$ состоит, во-первых, в отсутствии $\hat{f}_\varphi^1(\nu)$ на частотах, бóльших по модулю частоты Найквиста. Во-вторых, в пределах главного частотного диапазона $\hat{f}_\varphi^1(\nu)$ является суммой $\hat{f}_\varphi(\nu)$ с высокочастотными компонентами соседних периодов, см. рис. 4.2. Будем считать, что Фурье-образ проекции убывает достаточно быстро с ростом модуля частоты, и, рассматривать влияние только двух периодов (слева и справа от центрального). В таком случае имеет место

$$\hat{f}_\varphi^1(\nu) = \hat{f}_\varphi(\nu) + \hat{f}_\varphi(\nu - 2\nu_N) + \hat{f}_\varphi(\nu + 2\nu_N) \quad (4.3)$$

Проекция фильтруется посредством умножения на $|\nu|$. В результате фильтрации будем иметь $|\nu|\hat{f}_\varphi^1(\nu)$. Введём обозначение $\hat{\chi}(\nu) \equiv |\nu|(\hat{f}_\varphi^1(\nu) - \hat{f}_\varphi(\nu))$, учитывая (4.3), получим

$$\hat{\chi}(\nu) = \begin{cases} |\nu|(\hat{f}_\varphi(\nu - 2\nu_N) + \hat{f}_\varphi(\nu + 2\nu_N)), & |\nu| \leq \nu_N \\ -|\nu|\hat{f}_\varphi(\nu), & |\nu| > \nu_N \end{cases} \quad (4.4)$$

Пусть $f_\varphi^{Fl}(l)$ – функция, получающаяся в результате точной процедуры фильтрации проекции $f_\varphi(l)$. После её алгоритмической

реализации вместо $f_\varphi^{Flt}(l)$ будем иметь $f_\varphi^{Flt}(l) + \chi_\varphi(l)$, где $\chi_\varphi(l)$ – обратное преобразование Фурье от (4.4).

Оценим $\chi_\varphi(l)$. В случае, когда проекции дискретизованы, обратное преобразование Фурье производится путём суммирования по частотам, модуль которых меньше либо равен частоте Найквиста, поэтому имеем:

$$\chi_\varphi(l) \approx \int_{-v_N}^{v_N} |v| \left(\hat{f}_\varphi(v - 2v_N) + \hat{f}_\varphi(v + 2v_N) \right) \exp(2\pi i l v) dv. \quad (4.5)$$

Вычисления по формуле (4.5) можно производить если известна оценка Фурье-образа проекций при $|v| > v_N$.

4.2 Линейная низкочастотная фильтрация.

Из сказанного в предыдущем разделе следует, что для того, чтобы избежать перекрытия спектра проекции в соседних периодах, фильтр $|v|$ следует использовать в комбинации с фильтром низких частот с полушириной пропускания не превосходящей частоты Найквиста. Также низкочастотная фильтрация применяется для подавления случайного шума, который всегда присутствует в проекционных данных. Поэтому при построении алгоритмов реконструкции вместо фильтра $|v|$ используют произведение $|v|\hat{\Phi}(v)$, где $\hat{\Phi}(v)$ – чётная функция частоты, равная нулю при $|v| > v_{\max}$. В связи с этим мы будем записывать её в виде $\hat{\Phi}(|v|/v_{\max})$. Чётность требуется для того, чтобы после фильтрации проекция оставалась вещественной.

Рассмотрим следующее семейство фильтров

$$\hat{\Phi}\left(\frac{|v|}{v_{\max}}\right) = \begin{cases} 1 - \varepsilon \frac{|v|}{v_{\max}}, & |v|/v_{\max} \leq 1 \\ 0, & |v|/v_{\max} > 1 \end{cases}, \quad (4.6)$$

где $\varepsilon \in [0; 1]$. На рис. 4.3 изображены фильтры семейства (4.6) при различных значениях параметра ε .

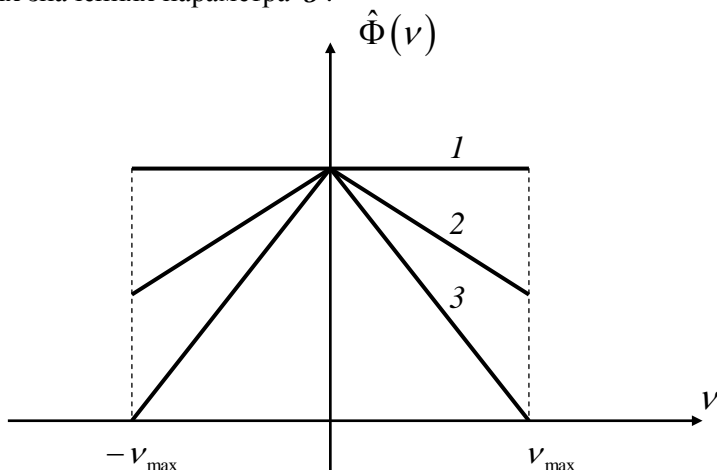


Рис. 4.3: Фильтры семейства (4.6) в частотной области;
 $1 - \varepsilon = 0; 2 - \varepsilon = 0.5; 3 - \varepsilon = 1.$

В пространственной области фильтрация проекций представляет собой свёртку, с некоторой функцией $K(l)$. Вычислим сворачивающее ядро для семейства (4.6):

$$\begin{aligned}
 K(l) &= \int_{-v_{\max}}^{v_{\max}} |v| \left(1 - \varepsilon \frac{|v|}{v_{\max}} \right) \exp(2\pi i v l) dv = \\
 &= \int_{-v_{\max}}^{v_{\max}} |v| \left(1 - \varepsilon \frac{|v|}{v_{\max}} \right) \cos(2\pi v l) dv + \\
 &+ i \int_{-v_{\max}}^{v_{\max}} |v| \left(1 - \varepsilon \frac{|v|}{v_{\max}} \right) \sin(2\pi v l) dv. \quad (4.7)
 \end{aligned}$$

Мнимая часть (4.7) обращается в ноль, а подынтегральное выражение в оставшемся члене – чётная функция. Тогда

$$K(l) = 2 \int_0^{v_{\max}} v \left(1 - \varepsilon \frac{v}{v_{\max}} \right) \cos(2\pi vl) dv. \quad (4.8)$$

Интегрируем каждое из слагаемых в (4.8) по частям:

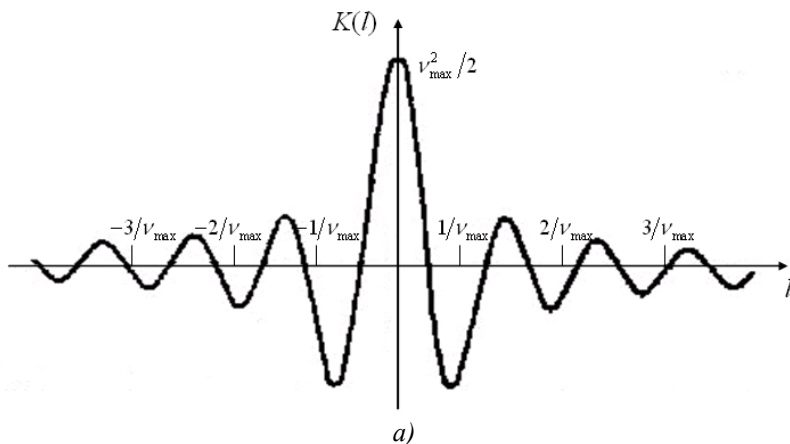
$$K_1(l) = \frac{\cos(2\pi l v_{\max}) - 1}{(2\pi)^2} + \frac{v_{\max} \sin(2\pi l v_{\max})}{2\pi}, \quad (4.9)$$

$$K_2(l) = \frac{2v_{\max} \cos(2\pi l v_{\max})}{(2\pi)^2} + \left(v_{\max}^2 - \frac{2}{(2\pi)^2} \right) \frac{\sin(2\pi l v_{\max})}{2\pi}. \quad (4.10)$$

В итоге получаем:

$$K(l) = 2 \left(K_1(l) - \frac{\varepsilon}{v_{\max}} K_2(l) \right), \quad (4.11)$$

Отметим, что функция $K(l)$ не обращается в бесконечность при $l = 0$. Действительно, воспользовавшись правилом Лопиталья, вычислим пределы выражений (4.9) и (4.10) при $l \rightarrow 0$. Получим соответственно $v_{\max}^2/2$ и $v_{\max}^3/3$.



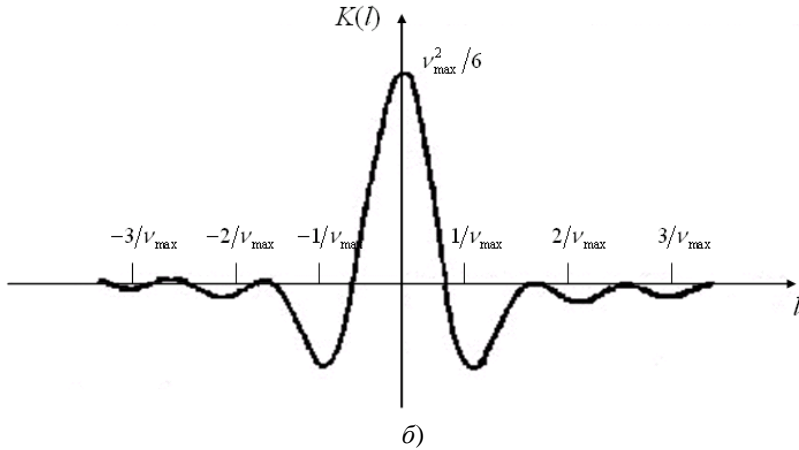


Рис. 4.4: Сворачивающее ядро (4.11): а) $\varepsilon = 0$, б) $\varepsilon = 1$

На рис. 4.4 приведены ядра свёртки (4.11) при $\varepsilon = 0$, рис. 4.4,а, и при $\varepsilon = 1$, рис. 4.4,б. Как видно из рис. 4.4, с увеличением l функция $K(l)$ при $\varepsilon = 1$ убывает по модулю быстрее, чем при $\varepsilon = 0$. Поэтому использование сворачивающего ядра (4.11) с $\varepsilon = 1$ при построении численных алгоритмов реконструкции представляется более предпочтительным.

Положим $v_{\max} = v_N$ и вычислим значения ядра свёртки $K(l)$ в узлах сетки kh , $k = \dots -2, -1, 0, 1, 2, \dots$. Заметим, что при этом в формулах (4.9), (4.10) синусы обращаются в ноль, а косинусы равны единице при чётных k и минус единице при нечётных. В результате получим следующее выражение:

$$K(kh) = \begin{cases} \frac{1}{4\pi h^2} \left(\frac{1}{2} - \frac{\varepsilon}{3} \right), & k = 0, \\ -\frac{\varepsilon}{2\pi^3 k^2 h^2}, & k \neq 0, k \text{ чётное}, \\ -\frac{1}{2\pi^3 k^2 h^2} (1 - \varepsilon), & k \text{ нечётное}. \end{cases} \quad (4.12)$$

Одна из первых программных реализаций гагр-фильтрации была проведена согласно формуле (4.12) при $\varepsilon=0$ в работе [29], опубликованной в 1971 году.

Из рис. 4.3 видно, что при ε близком к единице фильтр (4.6) сильно занижает низкочастотную часть Фурье-образа проекции. В связи с этим ряд исследователей считает, что фильтры с гладкой передаточной функцией, имеющей пологий профиль вблизи точки $\nu=0$, могут обеспечить более высокое качество реконструкции. Одним из таких фильтров является обобщённое окно Хэмминга:

$$\hat{\Phi}\left(\frac{|\nu|}{\nu_{\max}}\right) = \begin{cases} \varepsilon + (1-\varepsilon)\cos\left(\frac{2\pi|\nu|}{\nu_{\max}}\right), & |\nu|/\nu_{\max} \leq 1 \\ 0, & |\nu|/\nu_{\max} > 1 \end{cases}. \quad (4.13)$$

В (4.13) величина параметра ε может принимать значения из интервала $[0.5; 1]$. Отметим, что сворачивающее ядро для обобщённого окна Хэмминга имеет осциллирующий характер, как и для фильтра (4.6).

4.3 Фильтр Шеппа-Логана.

В работе [30] был предложен следующий фильтр низких частот:

$$\hat{\Phi}\left(\frac{|\nu|}{\nu_{\max}}\right) = \begin{cases} \text{sinc}\left(\frac{\pi|\nu|}{2\nu_{\max}}\right), & |\nu|/\nu_{\max} \leq 1 \\ 0, & |\nu|/\nu_{\max} > 1 \end{cases}, \quad (4.14)$$

где $\text{sinc}(x) = \sin(x)/x$. Вычислим для (4.14) сворачивающее ядро.

$$K(l) = \int_{-\nu_{\max}}^{\nu_{\max}} |\nu| \text{sinc}\left(\frac{\pi|\nu|}{2\nu_{\max}}\right) \exp(2\pi i \nu l) d\nu =$$

$$= \int_{-v_{\max}}^{v_{\max}} |v| \left(\frac{\pi |v|}{2v_{\max}} \right)^{-1} \sin \left(\frac{\pi |v|}{2v_{\max}} \right) (\cos(2\pi vl) + i \sin(2\pi vl)) dv. \quad (4.15)$$

Как и ранее, мнимая часть $K(l)$ равна нулю. Используя четность подынтегральной функции, перепишем (4.16) в виде

$$K(l) = \frac{4v_{\max}}{\pi} \int_0^{v_{\max}} \sin \left(\frac{\pi v}{2v_{\max}} \right) \cos(2\pi vl) dv. \quad (4.16)$$

Далее воспользуемся известной тригонометрической формулой: $\sin a \cos b = 0.5(\sin(a-b) + \sin(b+a))$. Применив её, получаем

$$\begin{aligned} K(l) &= -\frac{2v_{\max}}{\pi} \left(\frac{\cos((\pi/2v_{\max} - 2\pi l)v)}{\pi/2v_{\max} - 2\pi l} \Big|_0^{v_{\max}} + \frac{\cos((2\pi l + \pi/2v_{\max})v)}{2\pi l + \pi/2v_{\max}} \Big|_0^{v_{\max}} \right) = \\ &= -\frac{2v_{\max}}{\pi} \left(\frac{\cos(\pi/2 - 2\pi lv_{\max})}{\pi/2v_{\max} - 2\pi l} - \frac{1}{\pi/2v_{\max} - 2\pi l} + \frac{\cos(2\pi lv_{\max} + \pi/2)}{2\pi l + \pi/2v_{\max}} - \frac{1}{2\pi l + \pi/2v_{\max}} \right). \end{aligned}$$

После простых преобразований имеем

$$K(l) = \frac{4v_{\max}^2}{\pi} \left(\frac{\pi/2 - 2\pi lv_{\max} \sin(2\pi lv_{\max})}{\pi^2/4 - 4\pi^2 l^2 v_{\max}^2} \right). \quad (4.17)$$

Функция (4.17) является пространственным представлением фильтра Шеппа-Логана, названного так по фамилиям авторов работы [30]. Так же как и ядро свёртки (4.11) ядро Шеппа-Логана всюду ограничено и непрерывно. Оно представлено на рис. 4.5.

Положим $v_{\max} = v_N$ и вычислим значения ядра свёртки $K(l)$ в узлах сетки kh , $k = \dots -2, -1, 0, 1, 2, \dots$:

$$K(kh) = \frac{4}{4h^2 \pi} \left(\frac{\pi/2 - (2\pi kh/2h) \sin(2\pi kh/2h)}{\pi^2/4 - 4\pi^2 k^2 h^2/4h^2} \right) = \frac{1}{h^2 \pi} \left(\frac{\pi/2 - \pi k \sin(\pi k)}{\pi^2/4 - \pi^2 k^2} \right) =$$

$$= \frac{1}{h^2 \pi} \left(\frac{\pi/2}{\pi^2/4 - \pi^2 k^2} \right) = \frac{1}{2h^2} \left(\frac{1}{\pi^2/4 - \pi^2 k^2} \right) = \frac{2}{h^2 \pi^2} \left(\frac{1}{1 - 4k^2} \right). \quad (4.18)$$

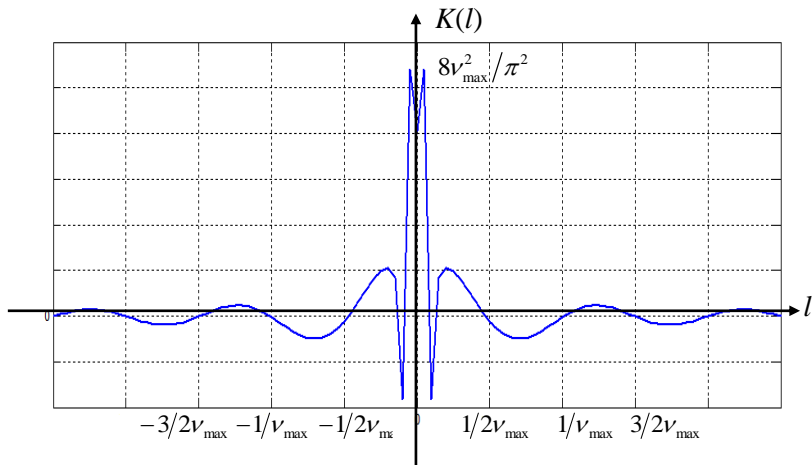


Рис. 4.5: Вид ядра Шеппа логана в пространственной области.

В дискретном виде ядро Шеппа-Логана, как и ядро (4.12), положительно при $k = 0$ и отрицательно при всех остальных k .

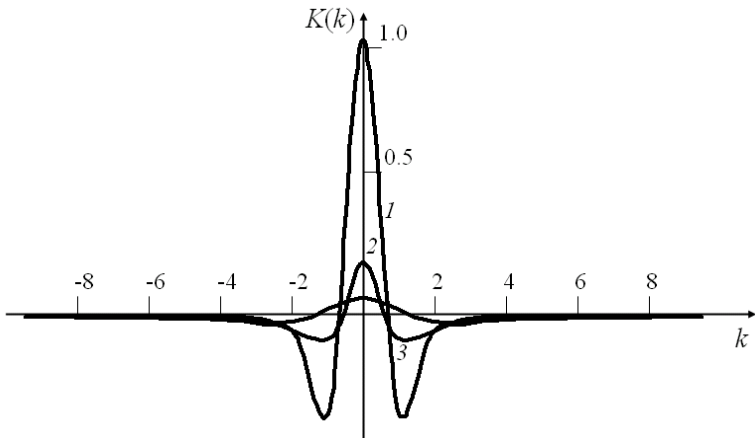


Рис. 4.6: Ядра свёртки при дискретном значении переменных:
 1 – ядро (4.18); 2 – ядро (4.11), $\varepsilon = 0$; 3 – ядро (4.11), $\varepsilon = 1$.

На рис. 4.6 представлены ядра свёртки для дискретного значения переменной: кривая 1 – фильтр Шеппа-Логана, кривые 2, 3 – фильтр (4.12) при $\varepsilon = 0$ и $\varepsilon = 1$ соответственно. Для наглядности значения в узлах соединены непрерывной линией, хотя в действительности функции от дискретной переменной между узлами не определены. Значения функций $K(k)$ даны в относительных единицах.

Исследования, проведённые рядом авторов, показывают, что алгоритмы обратного проецирования с фильтрацией при использовании фильтров (4.18) или (4.12) достаточно хорошо восстанавливают функции с эффективной шириной спектра ν_{\max} . При этом ν_{\max} должно быть достаточно большим, а также должны выполняться условия: $\nu_{\max} \leq M$ и $\nu_{\max} \leq 1/2h$, где M – количество зарегистрированных одномерных проекций.

4.4 Фильтр $1/Z^2$.

В данном разделе рассматривается другой подход к реализации гаур-фильтрации параллельных проекций, использующий аппарат обобщённых функций. Изложение ведётся в соответствии с работой [31]. Выразим реконструируемую функцию через преобразование Фурье следующим образом:

$$g(x, y) = \frac{1}{4\pi^2} \int_0^\pi \left(\int_{-\infty}^{\infty} |\omega| \hat{g}(\omega \cos \varphi, \omega \sin \varphi) \exp(i\omega(x \cos \varphi + y \sin \varphi)) d\omega \right) d\varphi. \quad (4.19)$$

Введём функцию $S(z, \varphi)$:

$$S(z, \varphi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\omega| \hat{g}(\omega \cos \varphi, \omega \sin \varphi) \exp(i\omega z) d\omega, \quad z \in \mathbb{R}^1. \quad (4.20)$$

При фиксированном φ выражение (4.20) является обратным преобразованием Фурье от произведения $|\omega| \hat{g}(\omega \cos \varphi, \omega \sin \varphi)$. По теореме о центральном слое

$$\hat{g}(\omega \cos \varphi, \omega \sin \varphi) = \int_{-\infty}^{\infty} f(z, \varphi) \exp(-iz\omega) dz \equiv \hat{f}(\omega, \varphi). \quad (4.21)$$

Подставляя (4.21) в (4.20) получаем

$$S(z, \varphi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\omega| \hat{f}(\omega, \varphi) \exp(i\omega z) d\omega. \quad (4.22)$$

Используя теорему о свёртке для обобщённых функций, имеем

$$S(z, \varphi) = \left((\mathbf{F}^{-1} |\omega|) * f \right) (z, \varphi). \quad (4.23)$$

В уравнении (4.23) через \mathbf{F}^{-1} обозначен оператор одномерного обратного преобразования Фурье; свёртка производится по первой переменной. Обратным преобразованием Фурье от $|\omega|$ является обобщённая функция $-1/(\pi z^2)$. Регуляризовав её, получаем:

$$S(z, \varphi) = \frac{1}{\pi} \int_0^{+\infty} \frac{f(z+z', \varphi) + f(z-z', \varphi) - 2f(z, \varphi)}{z'^2} dz'. \quad (4.24)$$

Подробности вывода формулы (4.24) можно найти в [31]. Из выражений (4.19), (4.20) видно, что функция $g(x, y)$ получается обратным проецированием функции $(2\pi)^{-1} S(x \cos \varphi + y \sin \varphi, \varphi)$. Тогда при фиксированном φ функция $S(x \cos \varphi + y \sin \varphi, \varphi)$ есть проекция, отфильтрованная гагр-фильтром, при этом процедура фильтрации выражается формулой (4.24).

Предположим, что проекции заданы в интервале $[-1; 1]$ на сетке с шагом $h = 2/(N-1)$. Будем аппроксимировать проекцию $f_\varphi(l)$ функцией $\tilde{f}_\varphi(l)$, обладающей следующими свойствами.

а) Функция $\tilde{f}_\varphi(l)$ имеет непрерывную производную $\tilde{f}'_\varphi(l)$.

б) В узлах сетки аппроксимирующая функция совпадает с имеющимися значениями проекции $\tilde{f}_\varphi(kh) = f_\varphi(kh)$, $1 \leq k \leq N$.

в) Равенство $\tilde{f}'_\varphi(kh) = (f_\varphi((k+1)h) - f_\varphi((k-1)h))/2h$ имеет место в узлах сетки, причём $f_\varphi(0) = f_\varphi((M+1)h) = 0$.

г) На интервалах $[kh; (k+1)h]$ функция $\tilde{f}_\varphi(l)$ является полиномом третьей степени.

Перечисленные условия позволяют в явном виде получить коэффициенты соответствующего сплайна:

$$\tilde{f}_\varphi(l) = \sum_{k=1}^M f_\varphi(kh) P(l+1+(1-k)h). \quad (4.25)$$

Полином $P(l)$ в (4.25) определяется следующим образом:

$$P(l) = \begin{cases} \frac{3}{2h^3}l^3 - \frac{5}{2h^2}l^2 + 1, & 0 \leq l \leq h \\ -\frac{1}{2h^3}l^3 + \frac{5}{2h^2}l^2 - \frac{4}{h}l + 2, & h < l \leq 2h \end{cases}. \quad (4.26)$$

При этом $P(-l) = P(l)$ и $P(l) = 0$ при $|l| > 2h$. Обратим внимание на то, коэффициенты полинома $P(l)$ не зависят от угла, под которым была снята проекция.

Функция $P(l)$ имеет разрывы второй производной, но модуль второй производной интегрируемый. Можно показать, что свёртка $P(l)$ с обобщённой функцией $-1/(\pi z^2)$ выражается формулой (4.24), в которой $f_\varphi(l)$ заменено на $P(l)$. Обозначим результат этой свёртки через $K(l)$. Из выражений (4.24), (4.25) следует, что оценкой функции $S(x \cos \varphi + y \sin \varphi, \varphi)$ при фиксированном φ является свёртка проекции $f_\varphi(l)$ с ядром $K(l)$. Таким образом,

мы получили ещё одно выражение для гаур-фильтрации. Непосредственное вычисление функции $K(l)$ приводит к следующему результату:

$$K(l) = \begin{cases} -\frac{8}{h\pi^2} \ln(0.5), & l = 0, \\ -\frac{1}{h\pi^2} (-9.5 \ln(h) + 20 \ln(2h) - 10.5 \ln(3h)), & l = \pm h, \\ -\frac{1}{h\pi^2} (8 \ln(h) - 36 \ln(2h) + 48 \ln(3h) - 20 \ln(4h)), & l = \pm 2h, \\ -\frac{1}{h^3 \pi} (2(3l+2h)(l+h) \ln|l+h| + 2(3l-2h)(l-h) \ln|l-h| - \\ - 9l^2 \ln|l| - 0.5(3l+4h)(l+2h) \ln|l+2h| - \\ - 0.5(3l-4h)(l-2h) \ln|l-2h|), & l \neq 0, \pm h, \pm 2h. \end{cases} \quad (4.27)$$

На рис. 4.7 приведено ядро (4.27) при различных значениях h .

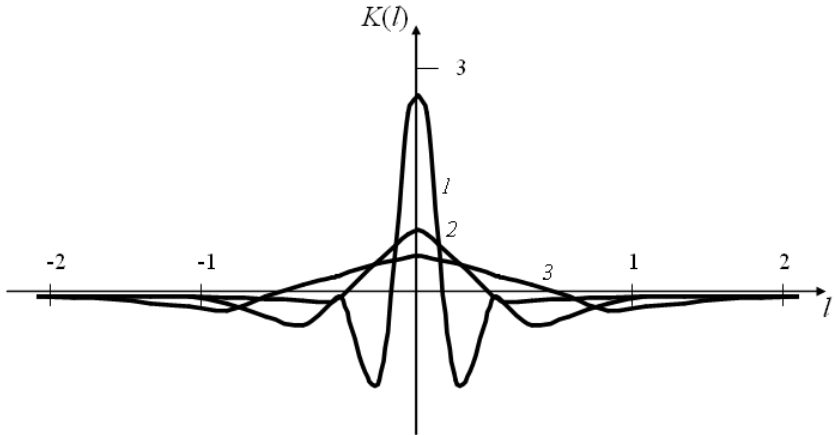


Рис.4.7: Сворачивающее ядро (4.30); кривая 1 - $h = 1$, кривая 2 - $h = 0.6$, кривая 3 - $h = 0.2$.

Ядро (4.27), в отличие от рассмотренных ранее сворачивающих ядер, как функция непрерывной переменной, не имеет периодических осцилляций. Кроме того, при изменении шага дискретизации меняется не только его нормировка, но и форма.

Глава 5

Задача трёхмерной томографии с источником излучения, движущимся по кривой.

5.1 Формула обращения Туя.

В главе 3 была приведена формула, выведенная Туем (Н. К. Туу), которая выражает функцию трёх переменных через проекционные данные, полученные при движении источника по кривой, см. (3.30). Ниже приводится её формальное доказательство согласно [25]. Перепишем преобразование (3.27) для рассматриваемого случая в виде

$$f(\lambda, \vec{n}) = \int_0^{\infty} g(\vec{C}(\lambda) + \vec{n}p) dp, \quad (5.1)$$

где $\vec{C}(\lambda)$ – уравнение траектории движения источника, $\vec{n} \in \mathbf{S}^2$. Введём функцию $f_1(\lambda, \vec{\beta})$, $\vec{\beta} \in \mathbf{R}^3$, которая при $|\vec{\beta}| = 1$ совпадает с проекционными данными (5.1):

$$f_1(\lambda, \vec{\beta}) = \int_0^{\infty} g(\vec{C}(\lambda) + \vec{\beta}p) dp. \quad (5.2)$$

Сделаем в (5.2) замену переменной интегрирования $p' = |\vec{\beta}|p$:

$$f_1(\lambda, \vec{\beta}) = \frac{1}{|\vec{\beta}|} \int_0^{\infty} g\left(\vec{C}(\lambda) + \frac{\vec{\beta}}{|\vec{\beta}|} p'\right) dp' = \frac{1}{|\vec{\beta}|} f\left(\lambda, \frac{\vec{\beta}}{|\vec{\beta}|}\right). \quad (5.3)$$

Поскольку $f(\lambda, \vec{\beta}/|\vec{\beta}|)$ – это проекционные данные, которые предполагаются известными, функция определена всюду в \mathbf{R}^3 , за исключением точки $|\vec{\beta}| = 0$.

Возьмём преобразование Фурье от $f_1(\lambda, \vec{\beta})$:

$$F(\lambda, \vec{v}) = \int_{\mathbb{R}^3} f_1(\lambda, \vec{\beta}) \exp(-2\pi i \langle \vec{\beta}, \vec{v} \rangle) d\vec{\beta}. \quad (5.4)$$

Здесь мы вновь пользуемся определением (4.1). Подчеркнём, что выражение (5.4) определяет функцию $F(\lambda, \vec{v})$ лишь формально. Действительно, как видно из (5.3), функция $f_1(\lambda, \vec{\beta})$ убывает на бесконечности как $1/|\vec{\beta}|$, и, соответственно, преобразование Фурье от неё не определено. На это указывал и сам Туй. Поэтому после формального доказательства формулы (3.30), он привёл её доказательство, исходя из теории обобщённых функций.

Докажем следующее утверждение:

$$F(\lambda, \vec{v}) = -\int_0^\infty \rho \hat{g}(\rho \vec{v}) \exp(2\pi i \rho \langle \vec{C}(\lambda), \vec{v} \rangle) d\rho. \quad (5.5)$$

Заменим в (5.4) функцию $f_1(\lambda, \vec{\beta})$ её выражением (5.2)

$$F(\lambda, \vec{v}) = \int_{\mathbb{R}^3} \left\{ \int_0^\infty g(\vec{C}(\lambda) + \vec{\beta} p) dp \right\} \exp(-2\pi i \langle \vec{\beta}, \vec{v} \rangle) d\vec{\beta}. \quad (5.6)$$

Произведём в (5.6) замену переменных интегрирования:

$$\begin{cases} \vec{r} = \vec{C}(\lambda) + \vec{\beta} p \\ \rho = 1/p \end{cases}, \quad \begin{cases} \vec{\beta} = (\vec{r} - \vec{C}(\lambda)) \rho \\ p = 1/\rho \end{cases}. \quad (5.7)$$

Якобиан, соответствующий этой замене, будет равен:

$$J \left(\frac{\partial(\vec{\beta}, p)}{\partial(\vec{r}, \rho)} \right) = \begin{pmatrix} \partial\beta_1/\partial r_1 & \partial\beta_1/\partial r_2 & \partial\beta_1/\partial r_3 & \partial\beta_1/\partial \rho \\ \partial\beta_2/\partial r_1 & \partial\beta_2/\partial r_2 & \partial\beta_2/\partial r_3 & \partial\beta_2/\partial \rho \\ \partial\beta_3/\partial r_1 & \partial\beta_3/\partial r_2 & \partial\beta_3/\partial r_3 & \partial\beta_3/\partial \rho \\ \partial p/\partial r_1 & \partial p/\partial r_2 & \partial p/\partial r_3 & \partial p/\partial \rho \end{pmatrix} =$$

$$= \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & \rho & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & -1/\rho^2 \end{pmatrix} = -\rho. \quad (5.8)$$

Таким образом получаем:

$$\begin{aligned} F(\lambda, \vec{v}) &= -\int_0^\infty \int_{\mathbb{R}^3} \rho g(\vec{r}) \exp\left(-2\pi i \rho \langle \vec{r} - \vec{C}(\lambda), \vec{v} \rangle\right) d\vec{r} d\rho = \\ &= -\int_0^\infty \rho \left\{ \int_{\mathbb{R}^3} g(\vec{r}) \exp(-2\pi i \langle \vec{r}, \rho \vec{v} \rangle) d\vec{r} \right\} \exp\left(2\pi i \rho \langle \vec{C}(\lambda), \vec{v} \rangle\right) d\rho. \end{aligned} \quad (5.9)$$

Замечаем, что выражение, стоящее в фигурных скобках в (5.9), есть преобразование Фурье функции $g(\vec{r})$ от аргумента $\rho \vec{v}$. Следовательно, утверждение (5.5) доказано.

Выразим $g(\vec{r})$ через её Фурье-образ в сферических координатах:

$$g(\vec{r}) = \int_0^{2\pi} \int_{\frac{\pi}{2}}^{\frac{\pi}{2}} \sin \theta \int_0^\infty \rho^2 \hat{g}(\rho \vec{n}) \exp(2\pi i \langle \vec{r}, \rho \vec{n} \rangle) d\rho d\theta d\varphi. \quad (5.10)$$

В формуле (5.10) $\vec{n} = \vec{v}/|\vec{v}|$, при этом координаты вектора \vec{n} : $(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$.

Далее, продифференцируем (5.5) по параметру λ :

$$\frac{\partial F(\lambda, \vec{v})}{\partial \lambda} = -2\pi i \langle \vec{C}'(\lambda), \vec{v} \rangle \int_0^\infty \rho^2 \hat{g}(\rho \vec{v}) \exp(2\pi i \rho \langle \vec{C}(\lambda), \vec{v} \rangle) d\rho, \quad (5.11)$$

где $\vec{C}'(\lambda) = (dC_x(\lambda)/d\lambda, dC_y(\lambda)/d\lambda, dC_z(\lambda)/d\lambda)$.

В случае если $\langle \vec{C}'(\lambda), \vec{v} \rangle \neq 0$, имеет место

$$\int_0^{\infty} \rho^2 \tilde{g}(\rho \vec{v}) \exp(2\pi i \rho \langle \vec{C}(\lambda), \vec{v} \rangle) d\rho = -\frac{1}{2\pi i \langle \vec{C}'(\lambda), \vec{v} \rangle} \frac{\partial F(\lambda, \vec{v})}{\partial \lambda} \quad (5.12)$$

Выражение (5.12) справедливо для любых $\vec{v} \in \mathbb{R}^3$, в том числе и тогда, когда \vec{v} является единичным вектором \vec{n} . С другой стороны (5.10) справедливо при любом \vec{r} , в частности, при таком что

$$\langle \vec{r}, \vec{n} \rangle = \langle \vec{C}(\lambda), \vec{n} \rangle. \quad (5.13)$$

С геометрической точки зрения уравнение (5.13) имеет решение, если существует плоскость, перпендикулярная вектору \vec{n} , и проходящая через точки \vec{r} и $\vec{C}(\lambda)$. Согласно условиям, которые Туй наложил на траекторию источника $\vec{C}(\lambda)$ (см. раздела 3.4.2), в случае, когда \vec{r} принадлежит носителю функции g , такая плоскость всегда существует, более того, для неё выполняется условие $\langle \vec{C}'(\lambda), \vec{n} \rangle \neq 0$. Учитывая это, а также выражение (5.12) получаем из (5.10) формулу обращения (3.30):

$$g(\vec{r}) = -\int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\sin \theta}{2\pi i \langle \vec{C}'(\lambda), \vec{n} \rangle} \frac{\partial F(\lambda, \vec{n})}{\partial \lambda} d\theta d\varphi. \quad (5.14)$$

Формула (5.14) соответствует следующему алгоритму реконструкции. Фиксируется некоторая точка \vec{r} , в которой восстанавливается функция $g(\vec{r})$. Для того, чтобы аппроксимировать двойной интеграл, с некоторыми шагами перебираются углы θ и φ , задающие направления вектора \vec{n} . Для каждого полученного \vec{n} путём решения уравнения (5.13) определяются точки $\vec{C}(\lambda_i)$, соответствующие пересечению траектории источника с плоскостью, перпендикулярной \vec{n} и проходящей через точку \vec{r} . Из всех решений (5.13) выбирается то,

для которого выполняется соотношение $\langle \vec{C}'(\lambda_i), \vec{n} \rangle \neq 0$ (согласно условиям, которые налагаются на кривую, оно всегда существует). Если таких решений несколько, то берётся любое из них. Обозначим его через $\vec{C}(\lambda_0(\vec{r}, \vec{n}))$.

После того, как положение источника $\vec{C}(\lambda_0(\vec{r}, \vec{n}))$ найдено, для него по формуле (5.4) вычисляется функция $F(\lambda_0, \vec{v})$ при $\vec{v} = \vec{n}$. При реконструкции по данным, полученным в результате измерений, конусные проекции известны только для конечного числа положений источника. Поэтому для получения функции $F(\lambda_0, \vec{n})$, вообще говоря, требуется интерполяция данных $f(\lambda, \vec{n})$ в точку λ_0 , что само-по-себе является нетривиальной задачей. Поскольку в формулу обращения (5.14) входит не сама $F(\lambda, \vec{n})$, а её производная, для оценки $\partial F(\lambda, \vec{n})/\partial \lambda|_{\lambda=\lambda_0}$ может быть использована величина

$$\left. \frac{\partial F(\lambda, \vec{n})}{\partial \lambda} \right|_{\lambda=\lambda_0} \approx \frac{F(\lambda_2, \vec{n}) - F(\lambda_1, \vec{n})}{|\lambda_2 - \lambda_1|}, \quad (5.15)$$

что позволяет избежать сложной процедуры интерполяции. Здесь $F(\lambda_1, \vec{n})$, $F(\lambda_2, \vec{n})$ соответствуют ближайшим к $\vec{C}(\lambda_0(\vec{r}, \vec{n}))$ положениям источника.

Таким образом, для реконструкции функции $g(\vec{r})$ в одной точке по формуле (5.14) требуется пять интегрирований, что приводит к большим затратам времени компьютерного счёта. Кроме того, для каждого значения вектора \vec{n} надо производить дифференцирование вдоль траектории источника. Как известно (см. [26]), процедура дифференцирования является некорректно поставленной задачей, поэтому при отсутствии процедур регуляризации она может привести к большим ошибкам. Также отметим, что, поскольку выражение (5.4) определяет $F(\lambda, \vec{v})$ только в обобщённом смысле,

то нет гарантии, что при его численной реализации будут получены значения пригодные для получения искомой функции $g(\vec{r})$.

Из сказанного выше можно заключить, что использование формулы обращения Туя на практике весьма затруднительно. Тем не менее, она имеет исключительно важное методическое значение. Во-первых, с её помощью установлены достаточные условия существования решения задачи трёхмерной томографии для случая, когда источник движется по пространственной кривой. Во-вторых, она служит основой для получения более эффективных с точки зрения алгоритмической реализации формул реконструкции для траекторий источника определённой формы.

5.2. Теорема Гранжа.

Удобный подход, с точки зрения построения численных алгоритмов трёхмерной реконструкции, был предложен Гранжа (P. Grangeat). Основная идея Гранжа, изложенная в работе [32], заключается в том, чтобы свести лучевое преобразование к трёхмерному преобразованию Радона, после чего воспользоваться формулой обращения (3.15). В данном параграфе излагается работа [32], при этом сохранены используемые там обозначения.

Пусть $\Omega = \text{supp}(g)$ – выпуклая область в R^3 , $g \in C^2(\Omega)$. Кривая, по которой движется источник, кусочно-гладкая и проходит вне области Ω . Предположим, что детектор проходит через начало координат O , а луч, соединяющий источник, находящийся в точке S , и начало координат, перпендикулярен плоскости детектора. Выберем на детекторе некоторую точку A и перепишем лучевое преобразование в виде:

$$f(S, A) = \int_0^{\infty} g \left(S + a \frac{\overrightarrow{SA}}{|\overrightarrow{SA}|} \right) da, \quad (5.16)$$

где интегрирование производится вдоль луча, соединяющего точки S и A , см. рис. 5.1.

Определим взвешенное лучевое преобразование:

$$Yg(S, A) = \frac{|\overline{SO}|}{|\overline{SA}|} f(S, A). \quad (5.17)$$

Возьмём произвольный вектор \vec{n} . Определим плоскость интегрирования: $P(\rho, \vec{n})$, где $\rho = \langle \vec{n}, \overline{OS} \rangle$. Обозначим через \vec{w} вектор $\overline{OS} / |\overline{OS}|$, предположим, что \vec{n} и \vec{w} не параллельны. Обозначим прямую пересечения детектора и плоскости $P(\rho, \vec{n})$ через $D(S, \vec{n})$. Определим базис $(\vec{u}, \vec{v}, \vec{w})$:

$$\vec{v} = \vec{w} \times \vec{n} / |\vec{w} \times \vec{n}|, \quad \vec{u} = \vec{v} \times \vec{w}, \quad (5.18)$$

где через " \times " обозначено векторное произведение. Отметим, что \vec{u} перпендикулярен прямой $D(S, \vec{n})$, а \vec{v} – параллелен ей.

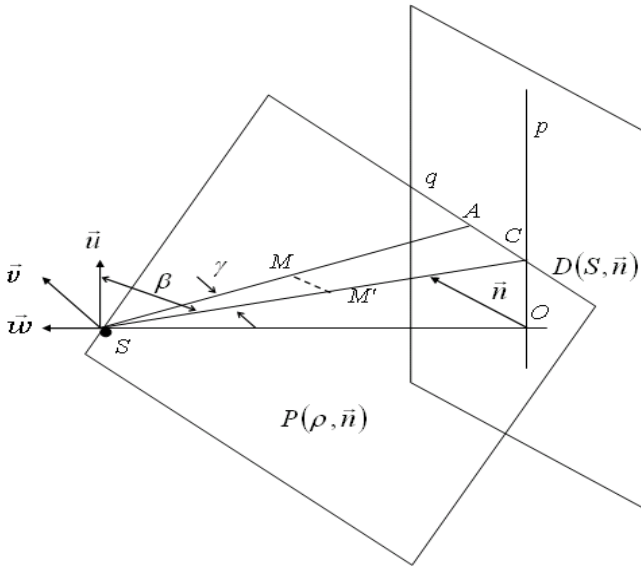


Рис 5.1: К доказательству теоремы Гранжа.

Будем задавать положение точки A на прямой $D(S, \vec{n})$ параметром q , а положение самой прямой на детекторе – параметром p , который отсчитывается в перпендикулярном ей направлении см. рис. 5.1. Положим, что пересечению детектора с рассматриваемой плоскостью $P(\rho, \vec{n})$ соответствует значение $p = 0$. Определим функцию $SYg(S, p)$:

$$SYg(S, p) = \int_{D(S, \vec{n})} Yg(S, A(q)) dq. \quad (5.19)$$

Теорема Гранжа выражается следующим соотношением:

$$\left. \frac{|\overline{OS}|^2}{|\overline{OS} \times \vec{n}|^2} \frac{\partial SYg(S, p)}{\partial p} \right|_{p=0} = \left. \frac{\partial (\mathbf{R}g)(\vec{n}, \rho)}{\partial \rho} \right|_{\rho = \langle \vec{n}, \overline{OS} \rangle}, \quad (5.20)$$

где через $(\mathbf{R}g)(\vec{n}, \rho)$ обозначено трёхмерное преобразование Радона функции $g(\vec{r})$ от аргументов \vec{n} и ρ , т.е. интеграл от $g(\vec{r})$ по плоскости, перпендикулярной вектору \vec{n} и находящейся на расстоянии ρ от начала координат. Выше эта плоскость обозначена как $P(\rho, \vec{n})$.

Доказательство.

Определим две новые координатные системы, параметризующие точку M , принадлежащую области определения функции $g(\vec{r})$, см. рис. 5.1. Одна из них – декартова система координат с центром в точке положения источника S и осями X, Y, Z , направленными вдоль векторов $\vec{i}, \vec{j}, \vec{k}$, соответственно. Вторая – сферическая, также с центром в точке S , и с полярной осью, направленной вдоль вектора \vec{v} . Пусть M' – ортогональная проекция точки M на координатную плоскость XZ , β – угол между $\overline{SM'}$ и \vec{i} , γ –

угол между \overrightarrow{SM} и \overrightarrow{SM}' , $a = |MS|$. Тогда связь между координатами в этих системах будет:

$$x = a \cos \beta \cos \gamma, \quad y = a \sin \gamma, \quad z = -a \sin \beta \cos \gamma. \quad (5.21)$$

Положение прямой $D(S, \vec{n})$ на детекторе можно также определять посредством угла β , при этом имеет место

$$p = |\overrightarrow{OS}| \operatorname{ctg} \beta, \quad d\beta / dp = -\sin^2 \beta / |\overrightarrow{OS}|. \quad (5.22)$$

Тогда

$$\frac{\partial SYg(S, p)}{\partial p} = \frac{\partial SYg(S, p(\beta))}{\partial \beta} \frac{d\beta}{dp} = -\frac{\sin^2 \beta}{|\overrightarrow{OS}|} \frac{\partial SYg(S, p(\beta))}{\partial \beta}. \quad (5.23)$$

Из рис. 5.1 видно $\sin^2 \beta = |\overrightarrow{OS} \times \vec{n}|^2 / |\overrightarrow{OS}|^2$, поэтому из (5.23) следует

$$\frac{|\overrightarrow{OS}|^2}{|\overrightarrow{OS} \times \vec{n}|^2} \frac{\partial SYg(S, p)}{\partial p} = -\frac{1}{|\overrightarrow{OS}|} \frac{\partial SYg(S, p(\beta))}{\partial \beta}. \quad (5.24)$$

Выразим функцию SYg в сферической системе координат. Перейдём в (5.19) от интегрирования по параметру q к интегрированию по углу γ . Поскольку $q = |\overrightarrow{SC}| \operatorname{tg} \gamma$, где C – точка пересечения прямой $D(S, \vec{n})$ с плоскостью, проходящей через точки S, O, M' , то имеет место $dq = (|\overrightarrow{SC}| / \cos^2 \gamma) d\gamma$. Тогда

$$SYg(S, p) = \int_{-\pi/2}^{\pi/2} Yg(S, \beta, \gamma) \frac{|\overrightarrow{SC}|}{\cos^2 \gamma} d\gamma. \quad (5.25)$$

Преобразуем выражение (5.17):

$$Yg(S, A) = \frac{|\overline{SO}|}{|\overline{SA}|} f(S, A) = \frac{|\overline{SO}|}{|\overline{SC}|} \frac{|\overline{SC}|}{|\overline{SA}|} f(S, A) = f(S, A) \frac{|\overline{SO}|}{|\overline{SC}|} \cos \gamma. \quad (5.26)$$

Учитывая (5.16) и (5.26), перепишем (5.25) в виде

$$SYg(S, p) = \int_{-\pi/2}^{\pi/2} f(S, \beta, \gamma) \frac{|\overline{SO}|}{\cos \gamma} d\gamma = \int_{-\pi/2}^{\pi/2} \int_0^{\infty} g(a, \beta, \gamma) \frac{|\overline{SO}|}{\cos \gamma} da d\gamma. \quad (5.27)$$

Дифференцируя (5.27) по β получаем утверждение теоремы (5.20). Действительно, учитывая связь между координатами (5.21), для произвольной точки M , принадлежащей плоскости $P(\rho, \vec{n})$ имеем

$$\begin{aligned} \frac{\partial g(M)}{\partial \beta} &= \frac{\partial g(M)}{\partial x} \frac{\partial x}{\partial \beta} + \frac{\partial g(M)}{\partial y} \frac{\partial y}{\partial \beta} + \frac{\partial g(M)}{\partial z} \frac{\partial z}{\partial \beta} = \\ &= -\frac{\partial g(M)}{\partial x} a \cos \gamma \sin \beta - \frac{\partial g(M)}{\partial z} a \cos \gamma \cos \beta. \end{aligned} \quad (5.28)$$

С другой стороны, дифференцирование трёхмерного преобразования Радона по ρ есть дифференцирование в направлении вектора \vec{n} , который лежит в плоскости XZ и имеет координаты $(\sin \beta, 0, \cos \beta)$. Следовательно

$$\frac{\partial g(M)}{\partial \rho} = \frac{\partial g(M)}{\partial x} \sin \beta + \frac{\partial g(M)}{\partial z} \cos \beta. \quad (5.29)$$

Сравнивая (5.28) и (5.29) получаем

$$\frac{\partial g(M)}{\partial \beta} = -\frac{\partial g(M)}{\partial \rho} a \cos \gamma, \quad (5.30)$$

Учитывая (5.30) после дифференцирования (5.27) по β имеем:

$$\frac{1}{|\overline{SO}|} \frac{\partial SYg(S, p)}{\partial \beta} = - \int_{-\pi/2}^{\pi/2} \int_0^{\infty} \frac{\partial g}{\partial \rho} (a, \beta, \gamma) a da d\gamma. \quad (5.31)$$

В правой части (5.31) интегрирование и дифференцирование можно поменять местами. Тогда под знаком дифференцирования будет стоять интеграл по плоскости от функции $g(\vec{r})$, т.е. её преобразование Радона. Тогда, учитывая (5.24), а также то, что уравнение (5.31) выполняется для точек, принадлежащих плоскости $P(\rho, \vec{n})$, т.е. при $p = 0$ и $\rho = \langle \vec{n}, \overline{OS} \rangle$, приходим к (5.20).

По поводу практического использования формулы (5.20) следует заметить, что непосредственно измеряемыми величинами являются, как правило, интегралы вдоль прямых: $f(S, A(q))$. Поэтому для вычисления производной $\partial SYg(S, p)/\partial p$ нужно использовать разностные методы. Но при смещении прямой $D(S, \vec{n})$ параллельно самой себе на поверхности детектора, поворачивается плоскость $P(\rho, \vec{n})$, соответственно, и вектор \vec{n} . Таким образом, данные, используемые для расчёта $\partial SYg(S, p)/\partial p$, относятся к различным направлениям вектора \vec{n} . Это приводит к тому, что при малых расстояниях до источника левая часть выражения (5.20), полученная численно, заметно отличается от правой.

5.3. Алгоритм Фельдкампта.

Рассмотрим следующую схему сбора проекционных данных. Пусть источник движется по окружности радиуса r_s в плоскости $z = 0$. Напротив источника по другую сторону объекта находится двумерный детектор, который синхронно движется вместе с ним. Вращение системы источник-детектор вокруг объекта обеспечивает просвечивание последнего под различными углами. Плоскость, в которой лежит детектор, параллельна оси Z , см. рис. 5.2.

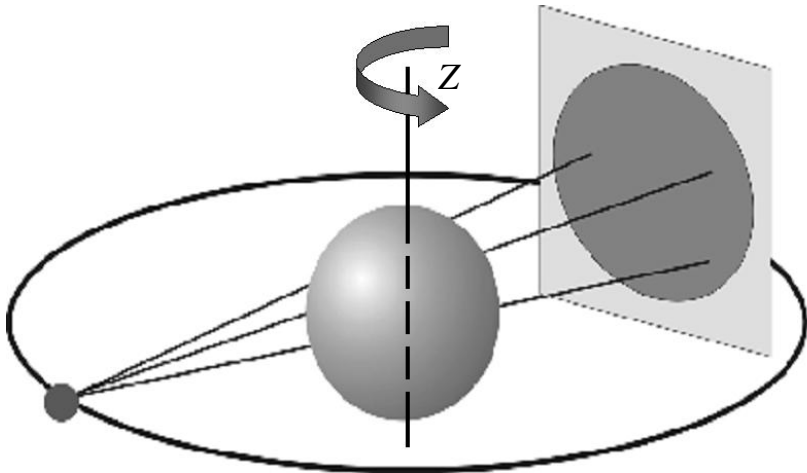


Рис. 5.2: Экваториальная схема сбора данных.

Такая схема измерений (обычно её называют экваториальной) в силу своей простоты часто реализуется в медицинском и промышленном томографическом оборудовании. По этой причине разработка алгоритмов реконструкции по проекционным данным, зарегистрированным при движении источника по окружности, вызывает большой интерес.

Очевидно, что окружность не удовлетворяет третьему из условий Туя, см. раздел 3.4.2. Поэтому использование формулы обращения (5.14) в данном случае невозможно. Одним из подходов к решению является разбиение трёхмерной задачи томографии на ряд двумерных. А именно, производится двумерная реконструкция искомой функции в сечениях, перпендикулярных оси Z . При этом, однако, следует иметь в виду, что ни для одного сечения, за исключением плоскости $z = 0$, не существует набора одномерных проекций. Поэтому для реконструкции слоя $z = z_0$ обычно берутся одномерные проекции, получающиеся в результате пересечения плоскостей, в которых производится регистрация, с плоскостью $z = z_0$. Это приводит к ошибкам, которые увеличиваются при уменьшении радиуса траектории источника.

В работе [33] Фельдкампом (L. A. Feldkamp) с сотрудниками предложен другой подход. Там был разработан алгоритм трёхмерной

реконструкции, получивший широкое распространение. В алгоритме Фельдкамп проекции фильтруются в двух направлениях, после чего производится трёхмерное обратное проецирование. Это обеспечивает достаточно высокую точность восстановления. Отметим однако, что алгоритм Фельдкамп является приближённым.

Дальнейшее изложение в этом разделе ведётся в соответствии с работой [33]. Предположим, что детектор проходит через начало координат. Сначала выведем формулу обращения в плоскости $z = 0$. Пусть φ – угол между направлением на источник и осью X . Рассмотрим прямую, лежащую в плоскости $z = 0$ и проходящую через источник и точку, имеющую на детекторе координаты $(U, 0)$. Пусть l – расстояние от этой прямой до начала координат, а ϕ – угол между перпендикуляром к ней и осью X , см. рис. 5.3,а.

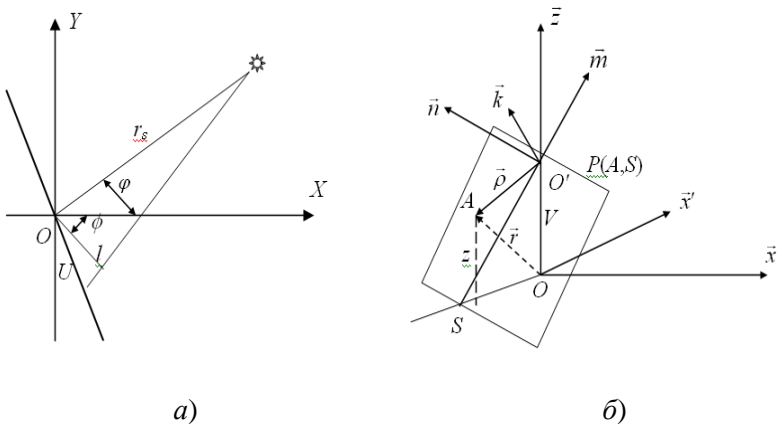


Рис. 5.3: К выводу формулы обращения, полученной Фельдкампом.

Как видно из рис. 5.3,а,

$$l = \frac{Ur_s}{(r_s^2 + U^2)^{\frac{1}{2}}}, \quad \phi = \varphi + \frac{\pi}{2} + \arctan\left(\frac{U}{r_s}\right). \quad (5.32)$$

Рассмотрим формулу обращения двумерного преобразования Радона (3.10). Распишем в ней Фурье-образ проекции

$$g(\psi, r) = \frac{1}{8\pi^2} \int_0^{2\pi} \left(\int_{-\infty}^{\infty} |\omega| \left(\int_{-\infty}^{\infty} f(\phi, l) \exp(-il\omega) dl \right) \exp(i\omega r \cos(\phi - \psi)) d\omega \right) d\phi. \quad (5.33)$$

Преобразуем (5.33), учитывая, что $g(\psi, r)$ и $f(l)$ – вещественные функции, а реальная часть Фурье-образа проекции – чётная:

$$g(\psi, r) = \frac{1}{4\pi^2} \operatorname{Re} \left\{ \int_0^{2\pi} \left(\int_0^{\infty} \omega \left(\int_{-\infty}^{\infty} f(l) \exp(i\omega(r \cos(\phi - \psi) - l)) dl \right) d\omega \right) d\phi \right\}. \quad (5.34)$$

Делаем в (5.34) замену переменных согласно (5.32). Этой замене соответствует следующий якобиан:

$$J \begin{pmatrix} \partial(l, \phi) \\ \partial(U, \varphi) \end{pmatrix} = \begin{pmatrix} \partial l / \partial U & \partial l / \partial \varphi \\ \partial \phi / \partial U & \partial \phi / \partial \varphi \end{pmatrix} = \begin{pmatrix} \frac{r_s(r_s^2 + U^2) - r_s U^2}{(r_s^2 + U^2)^{\frac{3}{2}}} & 0 \\ \left(1 + \frac{U^2}{r_s^2}\right)^{-1} & 1 \end{pmatrix} = \frac{r_s^3}{(r_s^2 + U^2)^{\frac{3}{2}}}. \quad (5.35)$$

Далее делаем замену $\omega' = \omega(r \cos(\psi - \varphi) + r_s) / (r_s^2 + U^2)^{1/2}$. После ряда преобразований, учитывая (5.35), получаем из (5.34) следующее выражение:

$$g(\psi, r) = \frac{1}{4\pi^2} \operatorname{Re} \left\{ \int_0^{2\pi} \frac{r_s^2}{(r \cos(\psi - \varphi) + r_s)^2} \left(\int_0^{\infty} \omega \left(\int_{-\infty}^{\infty} f_{\varphi}(U) \frac{r_s}{(r_s^2 + U^2)^{\frac{1}{2}}} \times \right. \right. \right. \\ \left. \left. \left. \times \exp \left(i\omega \left(\frac{r_s r \sin(\psi - \varphi)}{r \cos(\psi - \varphi) + r_s} - U \right) \right) dU \right) \omega d\omega \right) d\varphi \right\}. \quad (5.36)$$

Формула (5.36) представляет один из вариантов обращения веерного преобразования. Другой вариант приведён в разделе 3.4.1,

формула (3.29). Обратим внимание на то, что в (3.29) и (5.36) проекционные данные параметризуются по-разному: в (3.29) – через угол между центральным лучом и текущим, в (5.36) – через координату на детекторе.

Внешнее интегрирование в (5.36) соответствует обратному проецированию, внутренние интегралы – фильтрации одномерной проекции, которая образуется в пересечении детектора с плоскостью $z = 0$. Представим фильтрацию в виде свёртки:

$$g(\psi, r) = \frac{1}{4\pi^2} \int_0^{2\pi} \frac{r_s^2}{(r \cos(\psi - \varphi) + r_s)^2} \tilde{f}_\varphi(U(r, \psi, \varphi)) d\varphi, \quad (5.37)$$

где

$$U(r, \psi, \varphi) = \frac{rr_s \sin(\psi - \varphi)}{r \cos(\psi - \varphi) + r_s}, \quad (5.38)$$

$$\tilde{f}_\varphi(U) = \int_{-\infty}^{\infty} \frac{r_s}{(r_s^2 + U'^2)^{1/2}} f_\varphi(U') \Phi_U(U - U') dU', \quad (5.39)$$

$$\Phi_U(U) = \operatorname{Re} \left\{ \int_0^{\omega_{0U}} \omega \exp(i\omega U) d\omega \right\}. \quad (5.40)$$

Верхний предел интегрирования в (5.40) ограничен разрешением на детекторе в горизонтальном направлении, т.е. соответствующей частотой Найквиста. Однако он может быть взят и меньше её, для того чтобы подавить высокочастотный шум и препятствовать перекрытию спектров в пределах главного частотного диапазона, см. раздел 4.1.

Перейдём к выводу формулы обращения для трёхмерного пространства. Предположим, что функция $g(\vec{r})$ ищется в точке A , имеющей произвольную координату z . В [33] для реконструкции был предложен следующий метод. Проведём через точку A и

источник, расположенный в точке S , плоскость $P(A, S)$, которая пересекает детектор по прямой, перпендикулярной оси Z , см. рис. 5.3,б. Обозначим через \vec{n} единичный вектор, направленный вдоль этой прямой. Пусть O' – точка пересечения плоскости $P(A, S)$ с осью Z . Через \vec{m} обозначим единичный вектор в направлении от источника к O' . Единичный вектор \vec{k} определим как векторное произведение $\vec{k} = \vec{m} \times \vec{n}$. Через $\vec{\rho}$ обозначим вектор $\overline{O'A}$.

Пусть сначала A принадлежит плоскости $z = 0$. Тогда точка O' совпадает с началом координат O , а вектор $\vec{\rho}$ с вектором \vec{r} . В плоскости $z = 0$ имеет место формула обращения (5.36). Перепишем её, используя введённые обозначения:

$$g(\vec{\rho}) = \frac{1}{4\pi^2} \operatorname{Re} \left\{ \int_0^{2\pi} \frac{r_s^2}{(r_s + \langle \vec{\rho}, \vec{m} \rangle)^2} \left(\int_0^\infty \omega \left(\int_{-\infty}^\infty f_\varphi(U, V=0) \frac{r_s}{(r_s^2 + U^2)^{\frac{1}{2}}} \times \right. \right. \right. \\ \left. \left. \left. \times \exp \left(i\omega \left(\frac{r_s \langle \vec{\rho}, \vec{n} \rangle}{r_s + \langle \vec{\rho}, \vec{m} \rangle} - U \right) \right) dU \right) \omega d\omega \right) d\varphi \right\}. \quad (5.41)$$

В (5.41) через V обозначена z -координата точки O' .

Пусть теперь точка A не лежит в плоскости $z = 0$. Согласно работе [33], в этом случае в формуле (5.41) поворот на малый угол $d\varphi$ вокруг оси Z следует заменить поворотом на угол $d\varphi'$ вокруг оси, проходящей через точку O' параллельно вектору \vec{k} . При этом расстояние r_s должно быть заменено расстоянием от источника до новой оси вращения r'_s . Как это нетрудно видеть из рис. 5.3,б, связь между $d\varphi$, r_s и $d\varphi'$, r'_s даётся посредством соотношений

$$d\varphi' = \frac{r_s}{(r_s^2 + V^2)^{1/2}} d\varphi, \quad r'_s = (r_s^2 + V^2)^{\frac{1}{2}}. \quad (5.42)$$

Заменяем в (5.41) $d\varphi$ и r_s на $d\varphi'$ и r_s' . Подставим выражения для $d\varphi'$ и r_s' из (5.42). Для любой точки $A \in P(A, S)$ имеет место $\vec{r} = \vec{\rho} + V\vec{z}$, где \vec{z} – единичный вектор вдоль оси Z . Отсюда получаем $\vec{\rho} = \vec{r} - V\vec{z}$. Это выражение для $\vec{\rho}$ также подставим в (5.41). Учитывая что $\langle \vec{\rho}, \vec{k} \rangle = 0$, поскольку $\vec{\rho}$ лежит в плоскости $P(A, S)$, перпендикулярной вектору \vec{k} , в конечном итоге получаем

$$g(\vec{r}) = \frac{1}{4\pi^2} \operatorname{Re} \left\{ \int_0^{2\pi} \frac{r_s^2}{(r_s + \langle \vec{r}, \vec{x}' \rangle)^2} \left(\int_0^\infty \omega \left(\int_{-\infty}^\infty f_\varphi(U, V) \frac{r_s}{(r_s^2 + U^2 + V^2)^{\frac{1}{2}}} \times \right. \right. \right. \\ \left. \left. \left. \times \exp \left(i\omega \left(\frac{r_s \langle \vec{r}, \vec{y}' \rangle}{r_s + \langle \vec{r}, \vec{x}' \rangle} - U \right) \right) dU \right) \omega d\omega \right) d\varphi \right\}. \quad (5.43)$$

В (5.43) $V = r_s \langle \vec{r}, \vec{z} \rangle / (r_s + \langle \vec{r}, \vec{x}' \rangle)$, через \vec{x}' и \vec{y}' обозначены единичные векторы вдоль осей системы координат, повернутой вокруг оси Z на угол φ .

Формулу (5.43) можно привести к виду, аналогичному (5.37):

$$g(\vec{r}) = \frac{1}{4\pi^2} \int_0^{2\pi} \frac{r_s^2}{(r_s + \langle \vec{r}, \vec{x}' \rangle)^2} \tilde{f}_\varphi(U(\vec{r}, \varphi), V(\vec{r}, \varphi)) d\varphi, \quad (5.44)$$

где

$$U(\vec{r}, \varphi) = \frac{r_s \langle \vec{r}, \vec{y}' \rangle}{(r_s + \langle \vec{r}, \vec{x}' \rangle)}, \quad V(\vec{r}, \varphi) = \frac{r_s \langle \vec{r}, \vec{z} \rangle}{(r_s + \langle \vec{r}, \vec{x}' \rangle)}, \quad (5.45)$$

$$\tilde{f}_\varphi(U, V) = \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{r_s}{(r_s^2 + U'^2 + V'^2)^{\frac{1}{2}}} f_\varphi(U', V') \Phi_U(U - U') \Phi_V(V - V') dV' dU'. \quad (5.46)$$

Функция $\Phi_U(U)$ в (5.46) по-прежнему выражается формулой (5.40), а для функции $\Phi_V(V)$ имеет место выражение

$$\Phi_V(V) = \sin(\omega_{0V}V)/\pi V. \quad (5.47)$$

Для параметра ω_{0V} справедливы те же замечания, что и для ω_{0U} .

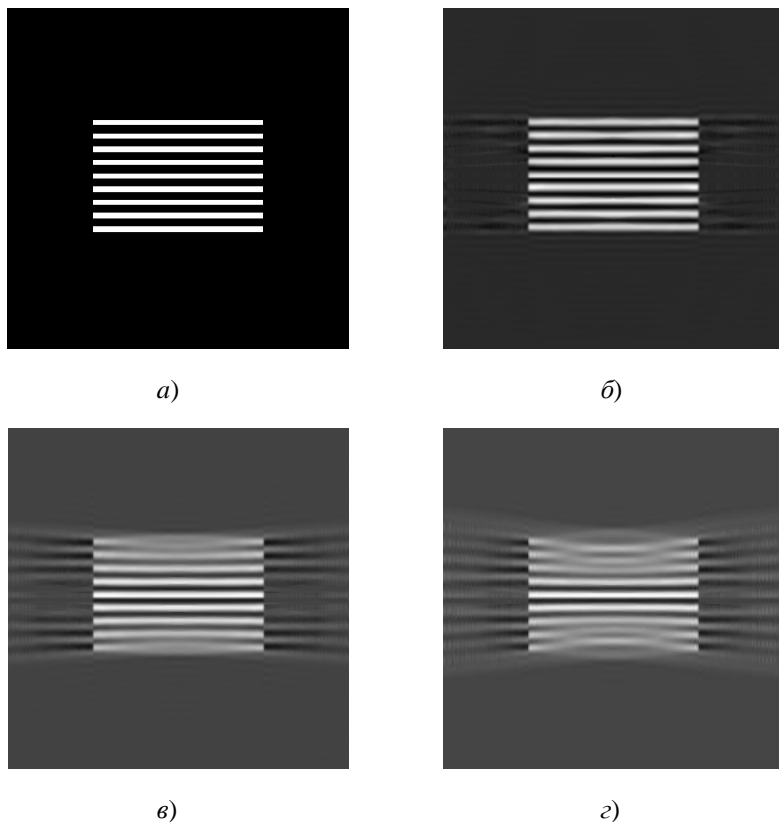


Рис. 5.4: сечения $y = 0$: *a*) – математический фантом; *б*), *в*), *з*) – результаты реконструкции, алгоритм Фельдкамп, разные расстояния до источника; *б*) $r_s=15$; *в*) $r_s=5$; *з*) $r_s=3$.

Алгоритм (5.44) – (5.47) может быть реализован в виде следующей последовательности операций. Двумерные проекции умножаются на весовую функцию, зависящую от координат точки на проекции и расстояния до источника. После этого проекции фильтруются перпендикулярно и параллельно оси Z путём свёртки с функциями $\Phi_U(U)$ и $\Phi_V(V)$. Для реконструкции искомой функции $g(\vec{r})$ в точке \vec{r} при каждом положении источника берётся значение соответствующей фильтрованной проекции в точке $(U(\vec{r}, \varphi), V(\vec{r}, \varphi))$, координаты которой зависят от угла поворота источника φ и точки, в которой производится реконструкция. Эти значения умножаются на коэффициент, также зависящий от \vec{r} и φ . После чего производится интегрирование по углу φ . Таким образом, для реконструкции алгоритмом Фельдкампта искомой функции в одной точке требуется три интегрирования.

Алгоритм Фельдкампта обеспечивает тем лучшую точность реконструкции, чем больше радиус окружности, по которой движется источник. На рис. 5.4 приведён пример результатов вычислительного эксперимента. Сечение математического фантома, состоящего из тонких дисков, перпендикулярных оси Z , представлено рис. 5.4,а. Сечения функций, полученных в результате реконструкции при различных расстояниях до источника, даны на рис. 5.4,б – рис. 5.4,г.

5.4. Спиральное движение источника.

В отличие от окружности, спираль, полностью охватывающая объект, удовлетворяет условиям Туя. Следовательно, по проекционным данным, полученным при прохождении по ней источника, возможно точное решение. С другой стороны, такая схема измерений легко может быть реализована в томографическом оборудовании. Эти два фактора обуславливают высокий интерес к задачам томографии со спиральной траекторией источника. Для этого случая построен ряд алгоритмов реконструкции. Условно их можно разделить на две группы. К первой относятся алгоритмы, в которых из совокупности конусных проекций путём интерполяции получают набор веерных проекций. По ним производится

двумерная реконструкция (например, по формуле (3.29)) параллельных слоёв объекта, которые, как правило, выбираются перпендикулярными оси спирали. В англоязычной литературе этот подход называется *rebinning* метод. Процедура *rebinning* часто производится в Фурье-пространстве. Одна из её реализаций подробно изложена в работе [34].

Во вторую группу входят алгоритмы обратного проецирования с фильтрацией. Первое осуществляется как интегрирование проекционных данных по параметру, изменением которого задаётся спираль. Для построения фильтров используются различные подходы. В частности, это модификации фильтрации, предложенной Смитом [35], или обобщения фильтрации Фельдкамп, см. выше. Методы построения более сложных фильтров основываются на теореме Гранжа. Среди работ, посвящённых этой теме, можно рекомендовать, например, [36].

Ниже приводится один из наиболее простых приближённых алгоритмов обратного проецирования с фильтрацией. Изложение ведётся в соответствии с работой [37]. Пусть траектория источника – круглая спираль радиуса R с шагом H и осью, совпадающей с осью Z , её уравнение: $\vec{C}(\lambda) = (R \cos \lambda, R \sin \lambda, H \lambda)$. На первом этапе алгоритма производится поворот системы координат на детекторе $(O_1, u, v) \rightarrow (O_1, u', v')$ таким образом, чтобы ось U' стала параллельной и сонаправленной касательному вектору к спирали в точке положения источника см. рис. 5.5. Для круглой спирали с равномерным шагом угол поворота γ не зависит места источника, т.е. является постоянным для любой точки траектории. Он равен:

$$\gamma = \arccos\left(R / \sqrt{R^2 + H^2}\right) \quad (5.48)$$

Поворот системы координат производится для каждой проекции. При этом получается набор данных $f(u', v', \lambda)$.

На втором этапе каждая из повернутых проекций умножается на весовую функцию, аналогичную той, которая используется в алгоритме Фельдкамп. В результате получается функция $f_1(u', v', \lambda)$, которая выражается уравнением

$$f_1(u', v', \lambda) = \frac{D}{\sqrt{u'^2 + v'^2 + D^2}} f(u', v', \lambda), \quad (5.49)$$

где D – расстояние между источником и детектором, которое в данной схеме сканирования предполагается постоянным.

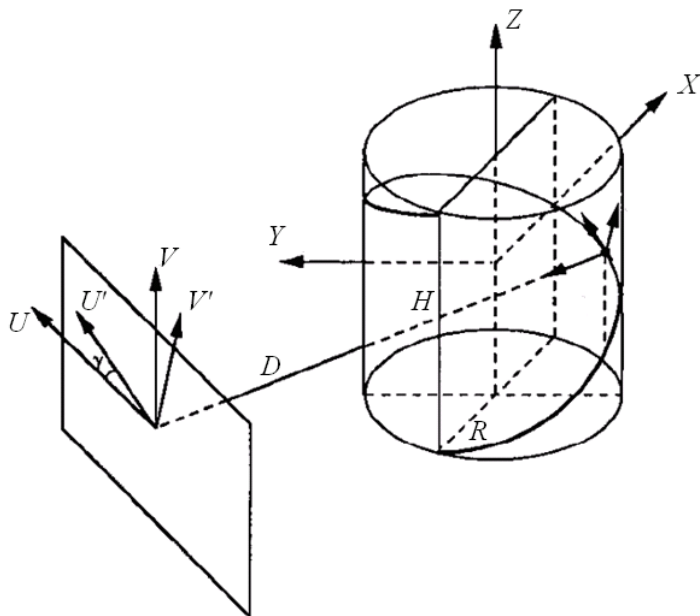


Рис. 5.5: Томография со спиральным движением источника.

Следующим шагом является фильтрация функций $f_1(u', v', \lambda)$. В работе [37] она осуществлялась посредством одномерной свёртки с сингулярным в пределе ядром:

$$\tilde{f}(u', v', \lambda) = \frac{1}{4\pi^2} \lim_{\varepsilon \rightarrow 0} (f_1(u', v', \lambda) * E_\varepsilon(u')), \quad \begin{cases} E_\varepsilon(u') = 1/\varepsilon^2, & u' < \varepsilon \\ E_\varepsilon(u') = 1/u'^2, & u' \geq \varepsilon \end{cases}. \quad (5.50)$$

Ядро $E_\varepsilon(u)$ является одной из реализаций гамп фильтрации. Исходя из изложенного в разделе 4.4, можно заключить что (5.50)

соответствует нерегуляризованной обобщённой функции, получаемой в результате обратного преобразования Фурье от $|\omega|$. Впервые в задачах трёхмерной томографии с источником, движущимся по кривой в трёхмерном пространстве, ядро $E_\varepsilon(u)$ было использовано Смитом, см. [35]. При численной реализации (5.50) величину ε обычно берут равной шагу дискретизации.

На последнем этапе алгоритма производится обратное проецирование фильтрованных функций $\tilde{f}(u', v', \lambda)$, в результате чего получается оценка искомой функции $g(\vec{r})$. Для спиральной геометрии регистрации данных обратное проецирование задаётся формулой

$$g(x, y, z) = \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{D\sqrt{R^2 + H^2}}{(x \cos \lambda + y \sin \lambda - R)^2} \tilde{f}(u', v', \lambda) d\lambda. \quad (5.51)$$

Координаты (u', v') повернутой системы на детекторе связаны с координатами (x, y, z) точки, в которой производится реконструкция следующими соотношениями:

$$\begin{aligned} u' &= \frac{y \cos \lambda - x \sin \lambda}{R - x \cos \lambda - y \sin \lambda} D \cos \gamma + \frac{(z - H \lambda)}{\sqrt{(x - R \cos \lambda)^2 + (y - R \sin \lambda)^2}} D \sin \gamma \\ v' &= \frac{x \sin \lambda - y \cos \lambda}{R - x \cos \lambda - y \sin \lambda} D \sin \gamma + \frac{(z - H \lambda)}{\sqrt{(x - R \cos \lambda)^2 + (y - R \sin \lambda)^2}} D \cos \gamma \end{aligned} \quad (5.52)$$

Уравнения (5.52) получаются путём вычисления координат точки пересечения детектора с прямой, проходящей через источник и точку (x, y, z) .

Алгоритм (5.49) – (5.52) является приближённым. По структуре он сходен с алгоритмом Фельдкампса. Двумерные проекции так же фильтруются гаусс-фильтром в направлении, касательном к траектории источника. Основное отличие состоит в том, что в алгоритме Фельдкампса дополнительно проводится фильтрация в перпендикулярном направлении.

Глава 6

Алгебраические методы реконструкции.

6.1 Дискретизация задачи.

В предыдущих главах были кратко описаны некоторые алгоритмы реконструкции, построенные путём аппроксимации формул обращения интегральных преобразований. Их несомненным преимуществом является то, что они основаны на точных соотношениях и поэтому при определённых условиях могут реконструировать искомую функцию с высокой точностью. Однако у них имеются и недостатки. Во-первых, до настоящего времени эти алгоритмы разработаны только для приближения лучевой томографии. Во-вторых, для удовлетворительного качества реконструкции они требуют сравнительно большого количества проекционных данных.

В связи с этим большое внимание уделяется разработке и исследованию других методов решения томографических задач. Наиболее важными из них являются те, в которых используется предварительная дискретизация, что приводит, в конечном итоге, к поиску решения системы из большого числа уравнений. Основное преимущество этих методов заключается в универсальности. С их помощью может быть произведена реконструкция внутренней структуры объектов, при различных уравнениях томографии, в том числе имеющих сложную форму. Эти методы в дальнейшем будут называться *алгебраическими*.

Алгебраические методы строятся по следующей общей схеме. Производится дискретизация реконструируемой функции g и проекционных данных f , при этом они, становятся векторами из линейных пространств. Самым простым, хотя и не единственным, способом дискретизации функции g является такой, при котором её область определения разбивается на малые (в определенном смысле, диктуемом условиями задачи) подобласти, пересекающиеся только по границам, в которых значение функции считается постоянным. Их называют *пикселями* и *вокселями* соответственно в двумерном и трёхмерном случаях. Обычно пиксели выбирают

квадратной формы, а воксели – кубической. Номер пиксела или воксела j определяет координату в векторном пространстве, а величина функции в нём g_j задает значение координаты. Естественной дискретизацией проекционных данных являются показания отдельных детекторов, или отдельных измерений (например, в одной и той же точке в различные моменты времени). В дальнейшем они будут нумероваться посредством индекса i . Если зарегистрированные данные объединены в проекции, как это часто бывает при использовании линейки или матрицы детекторов, и каждая проекция образуется показаниями q детекторов, то p -ому детектору на m -ой проекции приписывается номер $i = (m - 1)q + p$.

Затем строится *проецирующий оператор*, переводящий дискретизованную искомую функцию в набор проекционных данных. В результате рассматриваемая задача сводится к системе уравнений, которая, вообще говоря, может быть нелинейной и состоять не из алгебраических уравнений. Однако многие практически важные постановки описываются системой линейных алгебраических уравнений, в частности, медицинская рентгеновская томография и эмиссионная томография плазмы, основанные на соотношениях (2.11), (2.13), соответственно. В обоих случаях проецирующий оператор аппроксимирует интегралы вдоль прямых конечной суммой, поэтому можно записать:

$$\sum_j a_{ij} g_j = f_i \quad \text{или} \quad \mathbf{A}g = f. \quad (6.1)$$

В качестве элемента a_{ij} матрицы системы (6.1) обычно берётся длина пересечения луча, связанного с i -ым измерением, с j -ым элементом разбиения области определения функции g (пикселом или вокселом, соответственно, для \mathbf{R}^2 и \mathbf{R}^3).

В (6.1) $g \in \mathbf{R}^J$, $f \in \mathbf{R}^I$, где J – число пикселов (или вокселов), на которое разбита область определения искомой функции, I – полное число известных интегралов вдоль прямых. Соответственно, матрица \mathbf{A} имеет порядок $I \times J$. В типичных задачах томографии I и J весьма велики, в двумерном случае они порядка миллионов; а когда речь идет о трехмерных постановках, они могут достигать

нескольких миллиардов Матрица \mathbf{A} является, как правило, неквадратной. Часто бывает $J > I$, т.е. имеется недостаток проекционных данных. Именно в этом случае алгебраические методы имеют преимущество над методами, основанными на аппроксимации точных формул обращения. Однако иногда рассматриваются задачи, в которых $J \approx I$ или даже $J < I$. Далее, матрица \mathbf{A} является сильно разреженной, т.е. большинство ее элементов – нули. Недоопределённость ($J > I$) приводит к тому, что система (1.6) может иметь бесконечное множество решений. Но возможны и такие случаи, когда она несовместна. Обращение подобных систем требует специальных методов и алгоритмов, некоторые из которых будут кратко описаны ниже.

6.2 Аддитивные алгоритмы.

В силу указанных в предыдущем разделе свойств матрицы \mathbf{A} , систему уравнений (6.1) обычно решают итерационными методами. Наиболее простыми из них являются алгоритмы, которые условно можно назвать аддитивными. Эти алгоритмы имеют следующую общую форму:

$$g^{(k+1)} = g^{(k)} + c_k (g^{(k)}, f, \mathbf{A}). \quad (6.2)$$

Значениями функций c_k являются векторы пространства \mathbf{R}^J .

Часто функции c_k строятся на основе невязки. При этом невязка может вычисляться как со всем вектором f , так и с отдельными его компонентами. В первом случае алгоритмы принято называть алгоритмами типа SIRT (Simultaneous Iterative Reconstruction Technique), т.е. алгоритмами с одновременной коррекцией итерационных решений для всех уравнений системы. Простейшим из алгоритмов SIRT является следующий:

$$g^{(k+1)} = g^{(k)} + \lambda^{(k)} (f - \mathbf{A}g^{(k)}). \quad (6.3)$$

Здесь $\lambda^{(k)}$ – положительные параметры, влияющие на сходимость алгоритма. Обычно их называют *параметрами релаксации*.

Алгоритм (6.3) применяется не только для обращения матриц, но и для решения ряда других задач. В теории вычислительных методов он известен как *метод простой итерации*. Применительно к решению систем линейных уравнений его иногда называют *алгоритмом Ричардсона*.

Для использования итерационной схемы (6.3) необходимо, чтобы матрица \mathbf{A} была квадратной. Однако в задачах томографической реконструкции она чаще бывает прямоугольной. Переход к квадратной матрице можно осуществить путём умножения уравнения (6.1) слева на матрицу \mathbf{A}^T , транспонированную к матрице \mathbf{A} . Поскольку элементы матрицы \mathbf{A} вещественные, $\mathbf{A}^T = \mathbf{A}^*$, где \mathbf{A}^* – матрица, сопряжённая к \mathbf{A} . Таким образом, если матрица \mathbf{A} неквадратная, итерационный процесс (6.3) можно записать в виде

$$g^{(k+1)} = (\mathbf{E} - \lambda^{(k)} \mathbf{A}^* \mathbf{A} g^{(k)}) g^{(k)} + \lambda^{(k)} \mathbf{A}^* f, \quad (6.4)$$

где \mathbf{E} – единичная матрица размерности $J \times J$. В разделе 3.5 говорилось о том, что алгоритмы поиска регуляризованного решения могут быть получены, исходя из формулы (3.33). Положим в (6.4), что параметры $\lambda^{(k)}$ не зависят от номера итерации, т.е. $\lambda^{(k)} = \lambda$ для любого k . В этом случае прямой расчет показывает, что итерационный процесс (6.4) получается из (3.33) при

$$\mathcal{A}(\mathbf{A}^* \mathbf{A}, \alpha) = \left(\mathbf{E} - (\mathbf{E} - \lambda \mathbf{A}^* \mathbf{A})^{\frac{1}{\alpha}} \right) (\mathbf{A}^* \mathbf{A})^{-1}, \quad \alpha = 1/k. \quad (6.5)$$

В данном случае регуляризация осуществляется путём варьирования числа итераций. Для того, чтобы итерационный процесс (6.4) сходился, параметр λ должен удовлетворять следующему условию: $0 < \lambda < 2 / \|\mathbf{A}^* \mathbf{A}\|$. Подробности получения итерационных процессов из (3.33) можно найти в монографии [28].

В алгоритме (6.3) коррекция итерационного решения производится сразу с учётом всех уравнений системы (6.1). Другим подходом является последовательная коррекция в соответствии с

каждым из уравнений. Часто она осуществляется путём следующей процедуры:

$$g^{(k+1)} = \begin{cases} g^{(k)} + \lambda^{(k)} \frac{f_{i(k)} - \langle a^{i(k)}, g^{(k)} \rangle}{\|a^{i(k)}\|^2} (a^{i(k)})^T, & \|a^{i(k)}\| \neq 0 \\ g^{(k)}, & \|a^{i(k)}\| = 0 \end{cases}. \quad (6.6)$$

В формуле (6.6) через $\langle \circ, \circ \rangle$ обозначено скалярное произведение в \mathbf{R}^J ; через $a^{i(k)}$ – $i(k)$ -ая строка матрицы \mathbf{A} ; через $(a^{i(k)})^T$ – вектор-столбец, полученный в результате транспонирования строки $a^{i(k)}$. Производится циклический перебор строк: $i(k) = [k(\text{mod } I) + 1]$ (квадратными скобками обозначена целая часть числа). Среди алгебраических методов реконструкции итерационный процесс (6.6) является одним из наиболее распространенных. Он называется алгоритмом ART (Algebraic Reconstruction Technique). Как указывается в [4], ART в виде (6.6) был впервые предложен Качмаржем (S. Kaczmarz) в 1937 году для решения совместных систем линейных алгебраических уравнений.

Отметим, что в (6.3) при каждом значении k используются все проекционные данные, а в (6.6) только одно их значение. В связи с этим итерацией алгоритма ART принято называть цикл из I шагов (6.6), в котором перебираются все проекционные данные. То же самое относится и к другим алгоритмам, в которых корректировка решения производится в соответствии с отдельными уравнениями, в частности к алгоритму MART, описанному в разделе 6.4.

Пусть L_i множество всех решений i -го уравнения системы (6.1):

$$L_i = \left\{ g \mid \langle a^i, g \rangle = f_i \right\}. \quad (6.7)$$

Через L обозначим пересечение всех L_i .

Согласно (6.7) множества L_i представляют гиперплоскости в пространстве \mathbf{R}^J . Тогда, исходя из (6.6) заключаем, что вектор

$g^{(k+1)}$ получается путём проецирования вектора $g^{(k)}$ на гиперплоскость $L_{i(k)}$. Можно показать, что при $0 < \lambda^{(k)} < 2$ и любом начальном приближении $g^{(0)} \in \mathbf{R}^J$ алгоритм ART даёт последовательность оценок, которая сходится к вектору из множества L , при условии, что L непустое.

В общем случае множество L может содержать больше одного элемента. Обозначим через C_A линейную оболочку векторов $(a^i)^T$:

$$C_A = \left\{ g \mid g = \sum_{i=1}^I c_i (a^i)^T \right\}. \quad (6.8)$$

В (6.8) c_i – любые вещественные числа. Имеет место следующая *теорема о минимальной норме*. Если L – непустое множество, то в $L \cap C_A$ существует один и только один элемент g^* , причём для всех элементов множества L , отличных от g^* , справедливо неравенство $\|g^*\| < \|g\|$.

Если $g^{(0)} \in C_A$, то нетрудно видеть, что и все $g^{(k)}$ принадлежат C_A . Тогда и g^* , являющийся пределом последовательности $g^{(k)}$, принадлежит C_A . В силу сходимости итерационного процесса (6.6), g^* также принадлежит множеству L . Следовательно, g^* является решением системы (6.1), имеющим минимальную норму. В случае, если множество L не содержит ни одного элемента, алгоритм ART сходится к вектору, минимизирующему норму невязки $\|Ag - f\|$. Отметим, что нулевой вектор принадлежит множеству C_A . Обычно именно его берут в качестве начального приближения.

6.3 Использование квадратичного целевого функционала.

Типичным подходом к решению системы уравнений (6.1) является следующий. Ставится задача о нахождении экстремума (минимума или максимума) некоторого *целевого функционала*, на

определенном множестве векторов, выделяемом дополнительными условиями. В [4] приведен общий вид квадратичного целевого функционала:

$$F(g) = c_1 (f - \mathbf{A}g)^T \mathbf{B}_1 (f - \mathbf{A}g) + (g - g_0)^T (c_2 \mathbf{B}_2 - c_3 \mathbf{B}_3^{-1}) (g - g_0). \quad (6.9)$$

В (6.9) \mathbf{B}_1 – симметричная $I \times I$ матрица, \mathbf{B}_2 и \mathbf{B}_3 – $J \times J$ матрицы, c_1, c_2, c_3 – положительные постоянные, $g_0 \in \mathbf{R}^J$, значок "Т" на месте верхнего индекса означает транспонирование. Очевидно, что следует искать минимум функционала (6.9). В случае если существует несколько векторов, на которых он имеет минимум, то из них выбирается тот, который минимизирует дополнительный функционал $\|\mathbf{B}_4^{-1} g\|$, где \mathbf{B}_4 положительно определенная $J \times J$ матрица. Матрицы \mathbf{B}_i и константы c_i определяются, исходя из имеющейся априорной информации. В частности, если функционалу (6.9) придавать статистический смысл, то первый член можно рассматривать как автокорреляционную функцию невязки, а второй – как автоковариацию вектора решения g , при этом g_0 будет его математическим ожиданием.

Пусть \mathbf{B}_3 – положительно определённая матрица, \mathbf{B}_2 – неотрицательно определённая матрица. Напомним, что матрица \mathbf{B} называется *положительно определенной*, если она симметрична и для любого ненулевого вектора-столбца g справедливо $g^T \mathbf{B} g > 0$. Если же имеет место $g^T \mathbf{B} g \geq 0$, то \mathbf{B} – *неотрицательно определённая* матрица. Как любая положительно определённая матрица, \mathbf{B}_3 имеет квадратный корень, т.е. существует матрица $\mathbf{B}_3^{1/2}$, такая что $\mathbf{B}_3 = \mathbf{B}_3^{1/2} \mathbf{B}_3^{1/2}$. При этом $\mathbf{B}_3^{1/2}$ также положительно определена, поэтому существует матрица $\mathbf{B}_3^{-1/2}$ обратная ей, причём $\mathbf{B}_3^{-1} = \mathbf{B}_3^{-1/2} \mathbf{B}_3^{-1/2}$.

Введём новый вектор

$$u = \mathbf{B}_3^{-1/2} (g - g_0). \quad (6.10)$$

Тогда

$$g = \mathbf{B}_3^{1/2} u + g_0. \quad (6.11)$$

Подставляем (6.11) в выражение для квадратичного целевого функционала (6.9):

$$\begin{aligned} F(g) &= c_1 (f - \mathbf{A} \mathbf{B}_3^{1/2} u - \mathbf{A} g_0)^T \mathbf{B}_1 (f - \mathbf{A} \mathbf{B}_3^{1/2} u - \mathbf{A} g_0) + (u \mathbf{B}_3^{1/2})^T (c_2 \mathbf{B}_2 + c_3 \mathbf{B}_3^{-1}) (\mathbf{B}_3^{1/2} u) = \\ &= c_1 u^T \mathbf{B}_3^{1/2} \mathbf{A}^T \mathbf{B}_1 \mathbf{A} \mathbf{B}_3^{1/2} u - 2c_1 u^T \mathbf{B}_3^{1/2} \mathbf{A}^T \mathbf{B}_1 (f - \mathbf{A} g_0) + \\ &+ c_1 (f - \mathbf{A} g_0)^T \mathbf{B}_1 (f - \mathbf{A} g_0) + c_2 u^T \mathbf{B}_3^{1/2} \mathbf{B}_2 \mathbf{B}_3^{1/2} u + c_3 u^T \mathbf{B}_3^{1/2} \mathbf{B}_3^{-1} \mathbf{B}_3^{1/2} u \end{aligned} \quad (6.12)$$

В (6.12) учитывается, что $\mathbf{B}_3^{1/2}$ положительно определена и, следовательно, симметрична, т.е. $\mathbf{B}_3^{1/2} = (\mathbf{B}_3^{1/2})^T$. Замечаем, что

$$\mathbf{B}_3^{1/2} \mathbf{B}_3^{-1} \mathbf{B}_3^{1/2} = \mathbf{B}_3^{1/2} (\mathbf{B}_3^{-1/2} \mathbf{B}_3^{-1/2}) \mathbf{B}_3^{1/2} = (\mathbf{B}_3^{1/2} \mathbf{B}_3^{-1/2}) (\mathbf{B}_3^{-1/2} \mathbf{B}_3^{1/2}) = \mathbf{E}, \quad (6.13)$$

где \mathbf{E} – единичная матрица. Введём обозначения:

$$\begin{aligned} \mathbf{P} &= c_1 \mathbf{B}_3^{1/2} \mathbf{A}^T \mathbf{B}_1 \mathbf{A} \mathbf{B}_3^{1/2} + c_2 \mathbf{B}_3^{1/2} \mathbf{B}_2 \mathbf{B}_3^{1/2} + c_3 \mathbf{E} \\ z &= c_1 \mathbf{B}_3^{1/2} \mathbf{A}^T \mathbf{B}_1 (f - \mathbf{A} g_0). \end{aligned} \quad (6.14)$$

Отметим, что \mathbf{P} – квадратная матрица размера $J \times J$. Используя обозначения (6.14), с учётом (6.13), перепишем (6.12) в виде:

$$F(g) = u^T \mathbf{P} u - 2u^T z + c_1 (f - \mathbf{A} g_0)^T \mathbf{B}_1 (f - \mathbf{A} g_0) \quad (6.15)$$

Введём новый функционал:

$$H(u) = \frac{1}{2} u^T \mathbf{P} u - u^T z \quad (6.16)$$

Комбинируя (6.16) и (6.17), получаем

$$H(u) = \frac{1}{2} F(g) - \frac{1}{2} c_1 (f - \mathbf{A}g_0)^T \mathbf{B}_1 (f - \mathbf{A}g_0) \quad (6.17)$$

Из (6.17) следует, что g^* минимизирует функционал $F(g)$ тогда и только тогда, когда $g^* = \mathbf{B}_3^{1/2} u^* + g_0$, где u^* минимизирует функционал $H(u)$.

Смысл перехода от задачи минимизации функционала (6.9) к минимизации функционала (6.16) состоит в следующем. Имеет место утверждение: для любой квадратной неотрицательно определённой матрицы \mathbf{P} и произвольного вектора z вектор u^* минимизирует (6.16) тогда и только тогда, когда он является решением системы уравнений

$$\mathbf{P}u = z. \quad (6.18)$$

Можно показать, что матрица \mathbf{P} , определённая согласно (6.14), является положительно определённой, следовательно, и неотрицательно определённой. Таким образом от задачи минимизации функционала (6.9) мы перешли к решению системы уравнений с квадратной матрицей. Заметим, что, поскольку \mathbf{P} в данном случае положительно определена, то существует обратная к ней матрица \mathbf{P}^{-1} . Поэтому, если система (6.18) совместна, то она имеет единственное решение. Соответственно, существует единственный вектор g^* , который минимизирует функционал (6.9).

6.4 Метод максимума энтропии.

Ряд авторов указывает на то, что целесообразно применять метод максимизации целевого функционала энтропии:

$$F(g) = - \sum_j g_j \ln g_j \quad (6.19)$$

при дополнительных условиях. Часто такими условиями являются $\mathbf{A}g = f$, $g \geq 0$ (неравенство понимается покомпонентно). Для решения этой задачи был предложен алгоритм, получивший название MART (Multiplicative Algebraic Reconstruction Technique). Согласно [38], $k+1$ -ый шаг алгоритма MART имеет вид

$$g_j^{(k+1)} = \left(\frac{f_{i(k)}}{\langle a^{i(k)}, g^{(k)} \rangle} \right)^{\lambda^{(k)} a_j^{i(k)}} g_j^{(k)}, \quad j = 1, 2, \dots, J. \quad (6.20)$$

Здесь обозначения те же, что и в (6.6). Начальное приближение выбирается в виде $g^{(0)} = e^{-1} \vec{1}$ где e – основание натурального логарифма, $\vec{1}$ – вектор, все компоненты которого равны единице. В [38] рекомендуется $0 < \lambda^{(k)} \leq 1$. Там же отмечается, что итерационный процесс (6.20) сходится к решению рассматриваемой задачи в случае, если система (6.1) имеет неотрицательное решение.

В алгоритме (6.20) вместо невязки используется частное $d_{i(k)} = f_{i(k)} / \langle a^{i(k)}, g^{(k)} \rangle$. В случае, если значение очередного приближения оказывается завышенным, $d_{i(k)}$ становится меньше единицы, что приводит к уменьшению величины последующего приближения. Аналогично, если $g^{(k)}$ занижено, то будет иметь место $g_j^{(k)} \leq g_j^{(k+1)}$, $j = 1, 2, \dots, J$. При этом степень изменения решения в j -ом пикселе (вокселе) зависит от величины элемента $a_{i(k)j}$ проецирующей матрицы.

Алгоритм MENT, предложенный в работах Минербо (G. N. Minerbo) так же максимизирует функционал энтропии. Ниже приведён вариант алгоритма MENT для реконструкции функции двух переменных по набору из M одномерных параллельных проекций. Имеются также модификации алгоритма MENT для вверной проекционной геометрии и трёхмерного случая.

Обозначим через U_m матрицу поворота на угол φ_m , который соответствует m -ой проекции. Выражение для проекции $f_m(l)$ может быть записано в виде

$$f_m(l) = \int g \left(U_m^{-1} \begin{bmatrix} l \\ p \end{bmatrix} \right) dp, \quad U_m = \begin{pmatrix} \cos \varphi_m & \sin \varphi_m \\ -\sin \varphi_m & \cos \varphi_m \end{pmatrix}. \quad (6.21)$$

Обратим внимание на то, что интегралы всех проекций равны. Действительно, нетрудно видеть, что интеграл любой проекции равен интегралу функции, от которой была взята эта проекция. Не уменьшая общности можно предполагать, что имеет место $\int_{-\infty}^{\infty} f_m(l) dl = 1$, тогда и $\int_{-\infty}^{\infty} g(x, y) dx dy = 1$. Такая нормировка обеспечивает сходимость итерационного процесса при выбранном начальном приближении, см. ниже формулу (6.27).

Перепишем функционал энтропии в виде:

$$F(g) = - \iint g \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) \ln g \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) dx dy \quad (6.22)$$

Используя метод неопределённых множителей Лагранжа можно показать, что функция, минимизирующая (6.22), имеет вид:

$$g \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \frac{1}{V} \prod_{m=1}^M h_j \left([U_{m,11}, U_{m,12}] \cdot \begin{bmatrix} x \\ y \end{bmatrix} \right). \quad (6.23)$$

Для рассматриваемой нормировки параметр V в (6.23) равен площади носителя функции $g(x, y)$. Подставим (6.23) в выражение для проекции (6.21):

$$\begin{aligned} f_m(l) &= \frac{1}{V} \int \prod_{m'=1}^M h_{m'} \left([U_{m',11}, U_{m',12}] U_m^{-1} \begin{bmatrix} l \\ p \end{bmatrix} \right) dp = \\ &= h_m(l) \frac{1}{V} \int \prod_{m'=1, m' \neq m}^M h_{m'} \left([U_{m',11}, U_{m',12}] U_m^{-1} \begin{bmatrix} l \\ p \end{bmatrix} \right) dp. \end{aligned} \quad (6.24)$$

Последнее равенство в (6.24) имеет место потому, что для $m' = m$

$$\begin{aligned}
h_{m'} \left(\left[U_{m',11}, U_{m',12} \right] U_m^{-1} \begin{bmatrix} l \\ p \end{bmatrix} \right) &= h_m \left(\left[\cos \varphi_m, \sin \varphi_m \right] \begin{bmatrix} \cos \varphi_m & \sin \varphi_m \\ -\sin \varphi_m & \cos \varphi_m \end{bmatrix} \begin{bmatrix} l \\ p \end{bmatrix} \right) = \\
&= h_m \left(\left[\cos^2 \varphi_m + \sin^2 \varphi_m, \cos \varphi_m \sin \varphi_m - \sin \varphi_m \sin \varphi_m \right] \begin{bmatrix} l \\ p \end{bmatrix} \right) = \\
&= h_m \left(\left[1, 0 \right] \begin{bmatrix} l \\ p \end{bmatrix} \right) = h_m(l). \tag{6.25}
\end{aligned}$$

Функции h_m ищутся блочно-циклическим методом Гаусса-Зейделя:

$$h_m^{(k+1)}(l) = \begin{cases} \frac{f_m(l)V}{\int \prod_{m' \neq m} h_{m'}^{(k)} \left(\left[U_{m',11}, U_{m',12} \right] U_m^{-1} \begin{bmatrix} l \\ p \end{bmatrix} \right) dp}, & m = k \bmod M + 1 \\ h_m^{(k+1)}(l) = h_m^{(k)}(l), & m \neq k \bmod M + 1 \end{cases}. \tag{6.26}$$

Для итерационного процесса (6.26) берётся следующее начальное приближение:

$$h_m^{(0)}(l) = \begin{cases} 1, & f_m(l) \neq 0 \\ 0, & f_m(l) = 0 \end{cases}. \tag{6.27}$$

Вычислительные эксперименты показывают, что при реконструкции гладких распределений по малому числу проекций, $M = 3 \div 7$, МЕНТ обеспечивает меньшую ошибку по сравнению с другими алгебраическими алгоритмами. Однако когда реконструируются разрывные функции, решения, получаемые посредством МЕНТ, в некоторых случаях оказываются неустойчивыми.

6.5 Метод байесовской оценки.

Для построения целевого функционала используется также метод байесовской оценки см., например, [4]. Пусть вектор g и

вектор ошибки в проекционных данных ξ являются выборками случайных переменных, характеризуемых плотностями вероятности, p_g и p_ξ , соответственно. Оценка Байеса максимизирует функционал

$$F(g) = p_\xi(f - \mathbf{A}g)p_g(g). \quad (6.28)$$

Предположим, что распределения статистически независимы и оба описываются гауссовым законом. Плотность вероятности $p_g(g)$ имеет вид:

$$p_g(g) = \frac{1}{(\sqrt{2\pi})^J (\det \mathbf{V}_g)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(g - m_g)^T \mathbf{V}_g^{-1}(g - m_g)\right), \quad (6.29)$$

где m_g – математическое ожидание вектора решения, \mathbf{V}_g – его ковариационная матрица, размерности $J \times J$. Для вектора ошибки плотность вероятности выглядит аналогично (6.29), при этом показатель степени при $\sqrt{2\pi}$ равен $-I$, а ковариационная матрица \mathbf{V}_ξ имеет размерность $I \times I$.

Будем считать, что компоненты каждого из векторов g и ξ независимы между собой, тогда ковариационные матрицы \mathbf{V}_g и \mathbf{V}_ξ пропорциональны единичной матрице: $\mathbf{V}_g = t^2 \mathbf{E}$, $\mathbf{V}_\xi = s^2 \mathbf{E}$. Обозначим $r = t/s$. Также предположим, что $m_\xi = 0$, т.е. математическое ожидание ошибки равно нулю. Учитывая вышесказанное, перепишем функционал (6.28):

$$F(g) = \frac{1}{(\sqrt{2\pi})^I s^{2I}} \exp\left(-\frac{1}{2s^2}(f - \mathbf{A}g)^T (f - \mathbf{A}g)\right) \times \\ \times \frac{1}{(\sqrt{2\pi})^J t^{2J}} \exp\left(-\frac{1}{2t^2}(g - m_g)^T (g - m_g)\right) = \quad (6.30)$$

$$= \frac{1}{(\sqrt{2\pi})^{I+J} s^{2I} t^{2J}} \exp\left(-\frac{1}{2t^2} \left(r^2 \|f - \mathbf{A}g\|^2 + \|g - m_g\|^2\right)\right)$$

Очевидно, что максимум в (6.30) достигается, когда величина функционала

$$H(g) = r^2 \|f - \mathbf{A}g\|^2 + \|g - m_g\|^2 \quad (6.31)$$

минимальна. Таким образом, мы приходим к задаче минимизации квадратичного целевого функционала, которая была исследована в разделе 6.3. Здесь она будет решена другим способом.

В задаче имеется $J + I$ неизвестных (компоненты векторов g и ξ). Введём вектор столбец $\begin{bmatrix} u \\ z \end{bmatrix}$, где u имеет I компонент, а z имеет J компонент. Мы также используем обозначение $[\mathbf{E} \ r\mathbf{A}]$ для матрицы размерностью $I \times (I + J)$, первые I столбцов которой образуют единичную матрицу размерностью $I \times I$, а последние J столбцов – матрицу \mathbf{A} , умноженную на r .

Система уравнений

$$[\mathbf{E} \ r\mathbf{A}] \begin{bmatrix} u \\ z \end{bmatrix} = r(f - \mathbf{A}m_g) \quad (6.32)$$

совместна. Действительно, возьмём произвольный вектор z и положим

$$u = r(f - \mathbf{A}m_g - \mathbf{A}z). \quad (6.33)$$

Подставив эту пару векторов в систему (6.32), убеждаемся, что они ей удовлетворяют.

Система уравнений (6.32) была введена в связи со следующим утверждением. Пусть u^* и z^* – векторы, для которых $\begin{bmatrix} u^* \\ z^* \end{bmatrix}$

является решением системы (6.32) с минимальной нормой. Тогда $g^* = z^* + m_g$ минимизирует функционал (6.31).

Докажем это. Сначала отметим очевидное равенство:

$$\left\| \begin{bmatrix} u \\ z \end{bmatrix} \right\|^2 = \sum_{i=1}^I u_i^2 + \sum_{j=1}^J z_j^2 = \|u\|^2 + \|z\|^2. \quad (6.34)$$

Возьмём произвольный вектор g_1 , пусть $z_1 = g_1 - m_g$. Как было указано выше, если u и z связаны соотношением (6.33), то эта пара является решением системы (6.32). Подставив данное z_1 в (6.33), получим $u_1 = r(f - Ag_1)$. Поскольку $\begin{bmatrix} u_1 \\ z_1 \end{bmatrix}$ решение системы (6.32), учитывая (6.34), имеем

$$\|u^*\|^2 + \|z^*\|^2 \leq \|u_1\|^2 + \|z_1\|^2 = r^2 \|f - Ag_1\|^2 + \|g_1 - m_g\|^2. \quad (6.35)$$

С другой стороны $z^* = g^* - m_g$, возьмём $u_2 = r(f - Ag^*)$, тогда $\begin{bmatrix} u_2 \\ z^* \end{bmatrix}$ – решение системы (6.32). Поэтому

$$[\mathbf{E} \quad r\mathbf{A}] \begin{bmatrix} u^* \\ z^* \end{bmatrix} = [\mathbf{E} \quad r\mathbf{A}] \begin{bmatrix} u_2 \\ z^* \end{bmatrix}. \quad (6.36)$$

Откуда следует что $u^* = u_2$. Выразив в (6.35) u^* и z^* через g^* , получаем

$$r^2 \|f - Ag^*\|^2 + \|g^* - m_g\|^2 \leq r^2 \|f - Ag_1\|^2 + \|g_1 - m_g\|^2. \quad (6.37)$$

Поскольку g_1 – произвольный вектор, (6.37) означает, что g^* минимизирует функционал (6.31). Утверждение доказано.

Таким образом, по решению системы уравнений (6.32), имеющему минимальную норму, однозначно определяется вектор, минимизирующий функционал (6.31). В разделе 6.2 было показано, что алгоритм ART даёт решение совместной системы уравнений, обладающее минимальной нормой. Применяя формулу (6.6) к системе (6.32) получаем:

$$\begin{bmatrix} \mathbf{u}^{(k+1)} \\ \mathbf{z}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{u}^{(k)} \\ \mathbf{z}^{(k)} \end{bmatrix} + c^{(k)} \begin{bmatrix} \mathbf{e}^{i(k)} \\ r\mathbf{a}^{i(k)} \end{bmatrix}, \quad (6.38)$$

где $\mathbf{e}^{i(k)}$ транспонированная $i(k)$ -ая строка единичной матрицы,

$$c^{(k)} = \lambda^{(k)} \frac{r(f_{i(k)} - \langle \mathbf{a}^{i(k)}, \mathbf{m}_g \rangle) - (\mathbf{u}_{i(k)}^{(k)} + r\langle \mathbf{a}^{i(k)}, \mathbf{z}^{(k)} \rangle)}{1 + r^2 \|\mathbf{a}^{i(k)}\|^2}. \quad (6.39)$$

В (6.39) $\mathbf{u}_{i(k)}^{(k)}$ – $i(k)$ -ая компонента вектора $\mathbf{u}^{(k)}$, остальные обозначения те же, что и в (6.6). В качестве начального приближения $\mathbf{u}^{(0)}$, $\mathbf{z}^{(0)}$ выбираются нулевые векторы. Оценка искомого вектора \mathbf{g} определяется на каждой итерации как $\mathbf{g}^{(k)} = \mathbf{z}^{(k)} + \mathbf{m}_g$.

Алгоритм (6.38)-(6.39) можно записать в другом виде, заменив $\mathbf{z}^{(k)}$ на $\mathbf{g}^{(k)}$:

$$\mathbf{u}^{(0)} = \mathbf{0}, \quad \mathbf{g}^{(0)} = \mathbf{m}_g, \quad \mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + c^{(k)} \mathbf{e}_{i(k)}, \quad \mathbf{g}^{(k+1)} = \mathbf{g}^{(k)} + c^{(k)} \mathbf{a}_{i(k)},$$

$$c^{(k)} = \lambda^{(k)} \frac{r(f_{i(k)} - \langle \mathbf{a}_{i(k)}, \mathbf{g}^{(k)} \rangle) - \mathbf{u}_{i(k)}^{(k)}}{1 + r^2 \|\mathbf{a}_{i(k)}\|^2}. \quad (6.40)$$

Алгоритмы, основанные на поиске экстремумов функционалов, подобных (6.28), используются в методах медицинской диагностики СПЕКТ и ПЕТ, а также при недостатке проекционных данных.

Глава 7

Задача эмиссионной томографии в частично поглощающей среде.

7.1 Постоянное поглощение. Формула обращения для двумерного случая.

В разделе 2.2 было выписано соотношение, лежащее в основе метода медицинской диагностики СПЕКТ, см. (2.15). Это уравнение эмиссионной томографии в поглощающей, нерассеивающей среде, которое связывает распределения коэффициентов эмиссии и абсорбции вдоль луча с результатом отдельного измерения. Рассмотрим одно из соответствующих этому уравнению интегральных преобразований. В монографии [22] оно названо *преобразованием Радона с учётом поглощения* и записано следующим образом

$$f_{\eta}(\vec{n}, l) = (\mathbf{R}_{\eta} g)(\vec{n}, l) = \int_{\langle \vec{r}, \vec{n} \rangle = l} g(\vec{r}) \exp(-(\mathbf{D}\eta)(\vec{r}, \vec{n}_{\perp})) d\vec{r} \quad (7.1)$$

где $\vec{n} \in \mathbf{S}^{m-1}$, $\vec{r} \in \mathbf{R}^m$, \vec{n}_{\perp} принадлежит подпространству, перпендикулярному вектору \vec{n} . В уравнении (7.1) оператор \mathbf{D} определяется согласно (3.27).

Выражение (7.1) является более сложным, чем рассмотренные в третьей главе преобразования \mathbf{R} , \mathbf{P} и \mathbf{D} . В первую очередь в него входят распределения двух параметров, характеризующих среду, – коэффициентов эмиссии $g(\vec{r})$ и абсорбции $\eta(\vec{r})$. Наиболее интенсивно исследуется возможности обращения преобразования \mathbf{R}_{η} относительно $g(\vec{r})$ при известном $\eta(\vec{r})$ для случая $m=2$. Отметим, что только в этом случае (7.1) соответствует лучевой эмиссионной томографии в поглощающей среде.

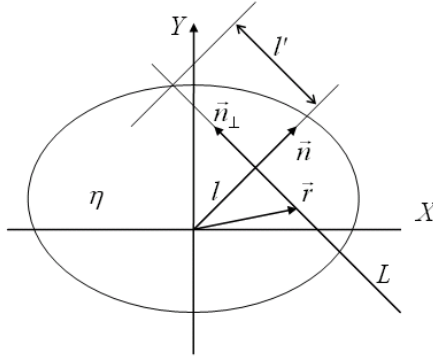


Рис. 7.1: К определению двумерного экспоненциального преобразования Радона.

Когда поглощение постоянно, $\eta(\vec{r}) = const$, (7.1) носит специальное название: *экспоненциальное преобразование Радона*. Следуя [22], будем обозначать его через \mathbf{T}_η . Рассмотрим двумерный случай. Обозначим через φ – угол между осью X и вектором \vec{n} . Введём вращающуюся систему координат $(\vec{n}(\varphi), \vec{n}_\perp(\varphi))$, $\vec{n}(\varphi) = (\cos\varphi, \sin\varphi)$, $\vec{n}_\perp(\varphi) = (-\sin\varphi, \cos\varphi)$. Пусть поглощение равно константе η в выпуклой области D , содержащей носитель функции g . Обозначим через l' расстояние между точкой $l\vec{n}$ и точкой пересечения прямой L с границей области D , см. рис. 7.1. Тогда, как это не трудно видеть, для любой точки \vec{r} , принадлежащей прямой L , будет иметь место:

$$(\mathbf{D}\eta)(\vec{r}, \vec{n}_\perp) = (l' - \langle \vec{r}, \vec{n}_\perp \rangle)\eta. \quad (7.2)$$

В (7.2) перед скалярным произведением стоит "минус", поскольку оно отрицательно. Подставляя (7.2) в (7.1) получаем:

$$(\mathbf{R}_\eta g)(\vec{n}, l) = \int_{\langle \vec{r}, \vec{n} \rangle = l} g(\vec{r}) \exp(-(l' - \langle \vec{r}, \vec{n}_\perp \rangle)\eta) d\vec{r} =$$

$$= \exp(-l'\eta) \int_{\langle \vec{r}, \vec{n} \rangle = 1} g(\vec{r}) \exp(\langle \vec{r}, \vec{n}_\perp \rangle \eta) d\vec{r}. \quad (7.3)$$

Интеграл в правой части (7.3) есть по определению двумерное преобразование \mathbf{T}_η . Перепишем его в следующем виде:

$$f_\eta(\varphi, l) \equiv (\mathbf{T}_\eta g)(\varphi, l) = \int_{-\infty}^{\infty} g(\vec{n}(\varphi)l + \vec{n}_\perp(\varphi)p) e^{\eta p} dp. \quad (7.4)$$

Для двумерного преобразования \mathbf{T}_η имеет место теорема, аналогичная теореме о центральном слое, (3.20):

$$(\mathbf{T}_\eta g)^\wedge(\vec{n}, \omega) = \hat{g}(\vec{n}\omega + i\eta\vec{n}_\perp). \quad (7.5)$$

Эта формула не столь удобна в обращении, как (3.20), поскольку позволяет найти Фурье-образ функции g на поверхности в пространстве двух комплексных переменных \mathbb{C}^2 , тогда как обратное преобразование Фурье требует интегрирование по \mathbb{R}^2 .

Двумерное преобразование \mathbf{T}_η имеет формулу обращения:

$$g = \frac{1}{4\pi} \mathbf{T}_{-\eta}^\# \mathbf{I}_\eta^{-1} f_\eta, \quad (7.6)$$

где \mathbf{I}_η^{-1} – обобщённый потенциал Рисса:

$$(\mathbf{I}_\eta^{-1} f)^\wedge(\omega) = \begin{cases} |\omega| \hat{f}(\omega), & |\omega| > \eta \\ 0, & |\omega| \leq \eta \end{cases}. \quad (7.7)$$

Оператор $\mathbf{T}_{-\eta}^\#$ в (7.6) иногда называют экспоненциальным обратным проецированием:

$$\left(\mathbf{T}_{-\eta}^{\#} f\right)(\vec{r}) = \int_0^{2\pi} f\left(\vec{n}(\varphi), \langle \vec{r}, \vec{n}(\varphi) \rangle\right) \exp\left(-\eta \langle \vec{r}, \vec{n}_{\perp}(\varphi) \rangle\right) d\varphi. \quad (7.8)$$

Таким образом, двумерное экспоненциальное преобразование Радона имеет формулу обращения, сходную с формулой обращения для “обычного” преобразования Радона. Проекция фильтруется посредством фильтра (7.7). Затем проводится интегрирование по углу. Но при этом, в отличие от формулы (3.7), каждая проекция умножается на функцию, зависящую как от направления, регистрации, так и от точки, где производится реконструкция.

Фильтр (7.7) зануляет значения Фурье-образа на всех частотах, меньших по модулю, чем η . Следовательно, чем больше значение коэффициента поглощения, тем шире область, в которой фильтр не пропускает. Это вызывает дополнительные сложности при использовании формулы обращения (7.6) на практике. Возникает не только проблема представления фильтра на частотах, близких к частоте Найквиста (см. главу 4), но и проблема его аппроксимации при $|\omega| \leq \eta$. В работе Третьяка и Метца (O. J. Tretiak, C. Metz) [40] было предложено заменить фильтр (7.7) следующей сглаженной функцией, зависящей от двух параметров, ε и σ :

$$\hat{\Phi}_{\eta}(\omega) = |\omega| \exp\left(-\frac{\sigma^2}{2} \left(\omega^2 - \frac{\omega^2}{4\pi^2}\right)\right) \hat{\Phi}_{\varepsilon}(\omega), \quad (7.9)$$

где

$$\hat{\Phi}_{\varepsilon}(\omega) = \begin{cases} 0, & 0 \leq \omega^2 \leq \eta^2 - \varepsilon^2 \\ \cos^4\left(\frac{1}{8\pi\varepsilon} \sqrt{(\eta^2 - \omega^2)}\right), & \eta^2 - \varepsilon^2 < \omega^2 \leq \eta^2 \\ 1, & \omega^2 < \eta^2 \end{cases}. \quad (7.10)$$

В выражениях (7.9), (7.10) предполагается, что $|\varepsilon| < \eta$. Можно показать, что функция, определяемая выражениями (7.9), (7.10), при $\varepsilon \rightarrow 0$, $\sigma \rightarrow 0$ переходит в (7.7).

Формула обращения экспоненциального преобразования Радона (7.6) может быть приведена к виду, аналогичному (3.14):

$$g(\vec{r}) = \frac{1}{4\pi^2} \int_0^{2\pi} \left(\exp(-\eta \vec{r} \vec{n}_\perp) \int_{-\infty}^{\infty} \frac{\text{ch}(-i\eta(\vec{r}\vec{m} - l))}{\vec{r}\vec{m} - l} \frac{\partial}{\partial l} f_\eta(\varphi, l) dl \right) d\varphi. \quad (7.11)$$

Другие способы решения задачи эмиссионной томографии в поглощающей среде приведены, например, в монографии [23].

7.2 Переменное поглощение. Итерационные методы решения задачи.

В настоящее время рассматриваемая задача активно исследуется для случая произвольного распределения абсорбции. Получены формулы обращения для двумерных задач с параллельной и веерной геометрией сбора данных в предположении, что распределении коэффициента абсорбции $\eta(\vec{r})$ известно. Здесь эти формулы не приводятся виду их большой сложности. Вопрос о том, возможно ли одновременное точное восстановление эмиссии и абсорбции из соотношения (7.1) до сих пор остаётся открытым.

Данный раздел посвящён итерационным методам реконструкции функции $g(\vec{r})$ при заданной функции $\eta(\vec{r})$. Для определённости будем рассматривать двумерную задачу, переход к трёхмерному случаю очевиден. Пусть объект исследуется при помощи набора линеек детекторов, сориентированных к нему под различными углами. В приближении лучевой томографии сигнал на каждом детекторе определяется согласно (2.15). Проведём дискретизацию задачи. Область реконструкции разобьём на квадратные пиксели, которым будем приписывать значения эмиссии g_j и абсорбции η_j . Значением i -ой компоненты вектора измерений будем считать показание i -го детектора f_i .

Определим элемент проецирующей матрицы a_{ij} для. В разделе 6.1 указывалось, что для эмиссионной задачи вклад от j -го пикселя в измеряемую величину f_i предполагается пропорциональным

длине его пересечения с i -ым лучом, которую будем обозначать через r_{ij} . Тогда эмиссия из j -го пиксела на i -ый детектор равна $\varepsilon_{ij} = g_j r_{ij}$. По пути к детектору интенсивность излучения ослабляется по экспоненциальному закону, см. (2.15). Представим интеграл в экспоненте в виде конечной суммы. При этом значение коэффициента абсорбции в каждом пикселе будем суммировать с весом также равным длине соответствующего пересечения r_{ik} . Таким образом, элемент a_{ij} будет выражаться уравнением

$$a_{ij} = r_{ij} \exp\left(-\sum_k \eta_k r_{ik}\right). \quad (7.12)$$

Суммирование в (7.12) ведётся по пикселям, которые пересекает отрезок i -ого луча, ограниченный точкой детектирования и точкой пересечения с границей j -ого пиксела. На рис. 7.2, поясняющем уравнение (7.12), эти пиксеты заштрихованы.

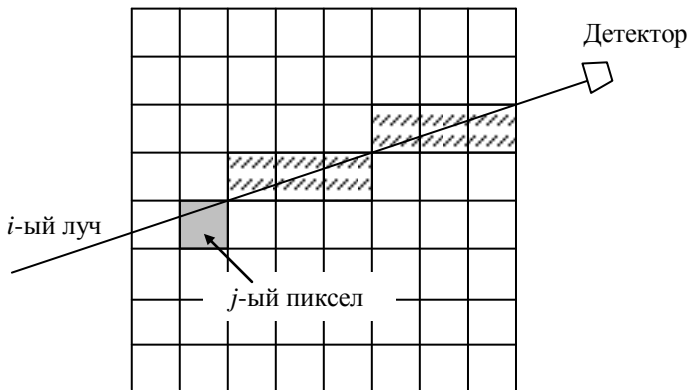


Рис. 7.2: К вычислению элемента проецирующей матрицы для задачи эмиссионной томографии в частично поглощающей среде.

В результате дискретизации мы пришли к системе уравнений (6.1), в которой элементы проецирующей матрицы a_{ij}

определяются согласно (7.12). Поскольку распределение коэффициента абсорбции $\eta(\vec{r})$ предполагается известным, a_{ij} могут быть вычислены. Отметим, что структура матрицы не изменяется по сравнению с эмиссионной задачей: нулевые элементы по-прежнему будут равны нулю, ненулевые – также останутся ненулевыми, но их величина уменьшится. Поэтому следует ожидать, что алгоритмы, описанные в главе 6, будут эффективны и для решения полученной системы.

Рассмотрим другой подход. Пусть \mathbf{T}^{-1} – какой-либо оператор решения двумерной задачи эмиссионной томографии без учёта поглощения. Это может быть как алгебраический алгоритм реконструкции, так и алгоритм, реализующий формулы обращения, см. раздел 3.1. Подействуем оператором \mathbf{T}^{-1} на проекционные данные (7.1). В результате будет реконструирована некоторая функция $g^{(0)}$, которую мы примем за начальное приближение итерационного процесса. Запишем для $g^{(0)}$ невязку с проекционными данными с поглощением:

$$f^{(1)} = f_\eta - \mathbf{R}_\eta g^{(0)} = f_\eta - \mathbf{R}_\eta \mathbf{T}^{-1} f_\eta. \quad (7.13)$$

Оператор \mathbf{R}_η , определённый согласно (7.1), нам известен, так как известна функция $\eta(\vec{r})$.

Далее, произведём реконструкцию оператором \mathbf{T}^{-1} по данным $f^{(1)}$. При этом будет получена функция g_1 . В качестве итерационного приближения возьмём сумму $g^{(0)} + \lambda g_1$, где λ – положительный параметр, роль которого будет понятна ниже. Учитывая (7.13), запишем:

$$g^{(1)} = g^{(0)} + \lambda g_1 = \mathbf{T}^{-1} f_\eta + \mathbf{T}^{-1} f^{(1)} = \mathbf{T}^{-1} f_\eta + \lambda \mathbf{T}^{-1} (f_\eta - \mathbf{R}_\eta \mathbf{T}^{-1} f_\eta). \quad (7.14)$$

Определим проекционные данные $f^{(2)}$ как невязку для функции g_1 с данными $f^{(1)}$:

$$f^{(2)} = f^{(1)} - \mathbf{R}_\eta g_1 = f^{(1)} - \mathbf{R}_\eta \mathbf{T}^{-1} f^{(1)}. \quad (7.15)$$

Введём обозначение $g_2 = \mathbf{T}^{-1} f^{(2)}$ и определим приближение $g^{(2)}$:

$$\begin{aligned} g^{(2)} &= g^{(1)} + \lambda^2 g_2 = g^{(0)} + \lambda g_1 + \lambda^2 g_2 = \\ &= \mathbf{T}^{-1} f_\eta + \lambda \mathbf{T}^{-1} (\mathbf{E} - \mathbf{R}_\eta \mathbf{T}^{-1}) f_\eta + \lambda^2 \mathbf{T}^{-1} (\mathbf{E} - \mathbf{R}_\eta \mathbf{T}^{-1})^2 f_\eta. \end{aligned} \quad (7.16)$$

Здесь \mathbf{E} – единичный оператор в функциональном пространстве, которому принадлежат проекционные данные f_η .

Нетрудно видеть, что при повторении описанной выше процедуры, для n -ой итерации получается следующее выражение:

$$g^{(n)} = g^{(n-1)} + \lambda^n g_n = \mathbf{T}^{-1} \sum_{k=0}^n \lambda^k (\mathbf{E} - \mathbf{R}_\eta \mathbf{T}^{-1})^k f_\eta. \quad (7.17)$$

Предположим, что оператор $\mathbf{R}_\eta \mathbf{T}^{-1}$ определён на гильбертовом пространстве \mathbf{H} , и $f_\eta \in \mathbf{H}$. При $n = \infty$ в правой части (7.17) стоит сумма ряда Неймана оператора $\mathbf{R}_\eta \mathbf{T}^{-1}$, которая, как известно, равна в области сходимости обратному оператору. Следовательно,

$$g^{(\infty)} = \mathbf{T}^{-1} \sum_{k=0}^{\infty} \lambda^k (\mathbf{E} - \mathbf{R}_\eta \mathbf{T}^{-1})^k f_\eta = \mathbf{T}^{-1} (\mathbf{R}_\eta \mathbf{T}^{-1})^{-1} f_\eta, \quad (7.18)$$

В случае существования обратного оператора \mathbf{R}_η^{-1} уравнение (7.18) переходит в $g^{(\infty)} = \mathbf{R}_\eta^{-1} f_\eta$. Таким образом, оно даёт точное решение задачи, когда сходится ряд Неймана оператора $\mathbf{R}_\eta \mathbf{T}^{-1}$. Если оператор ограничен по норме, то это всегда может быть достигнуто путём выбора достаточно малого по величине параметра

λ . А именно, сходимость внутри единичного шара в гильбертовом пространстве \mathbb{H} имеет место при условии $|\lambda| \|\mathbf{E} - \mathbf{R}_\eta \mathbf{T}^{-1}\| < 1$.

Отметим, что, как это видно из формулы (7.18), вместо \mathbf{T}^{-1} может быть использован любой оператор, действующий из гильбертова пространства \mathbb{H} в пространство Шварца \mathfrak{S} , которому принадлежит функция g . В частности, им может быть оператор обратного проецирования $\mathbf{R}^\#$, см. (3.8). Однако выбор в качестве \mathbf{T}^{-1} алгоритма реконструкции позволяет ускорить сходимость итерационного процесса.

Рассмотренный выше подход к решению задачи эмиссионной томографии в поглощающей среде, приводящий к разложению в ряд Неймана, носит универсальный характер. Действительно, в формулах (7.13) – (7.18) оператор \mathbf{R}_η может быть заменён проецирующим оператором, построенным исходя из других, возможно более сложных, моделей сбора проекционных данных.

Список литературы.

- [1] **Oldendorf W. H.** Isolated flying-spot detection of radiodensity discontinuities; displaying the internal structural pattern of a complex object. // *IRE Trans. Bio-Med. Electron.* 1961. V.BME-8. P.68-72.
- [2] **Cormack A. M.** Representation of a function by its line integrals, with some radiological applications. // *J. Appl. Phys.* 1963. V.34. P.1722-1727.
- [3] **Hounsfield G. N.** Computerized transverse axial scanning tomography: Part 1, description of the system. // *Br. J. Radiol.* 1973. V.46. P.1016-1022.
- [4] **Хермен Г. Т.** Восстановление изображений по проекциям. Основы реконструктивной томографии. М.: Мир, 1983 (перевод с английского).
- [5] **Del Sole A., Falini A., Ravasi L., Ottobrini L., De Marchis D., Bombardieri E., Lucignani G.** Anatomical and biochemical investigation of primary brain tumours. // *Eur J Nucl Med.* 2001. V.28. P.1851-1872.
- [6] **John C., Elsner E., Müller A., Knauth M., von Kummel R.** Computed tomography in acute cerebral ischemia. // *Radiologe.* 1997. V.37. P.853-858.
- [7] **Hansell D. M.** Computed tomography of diffuse lung disease: functional correlates. // *Eur. Radiol.* 2001. No.9. P.1666-1680.
- [8] **Hughes S. W., D'Arcy T. J., Maxwell D. J., Saunders J. E.** In vitro estimation of foetal liver volume using ultrasound, x-ray computed tomography and magnetic resonance imaging. // *Physiol. Measur.* 1997. V.18, No.4. P.401-410.
- [9] **Ollinger J. M., Fessler J. A.** Positron emission tomography. // *IEEE Signal Processing Magazine.* 1997. V.14, No.1. P.43-55.
- [10] **Majumder D. D., Bhattacharya M. Frank H.** Cybernetic approach to medical technology: application to cancer screening and other diagnostics. // *Kybernetes.* 2000. V.9, No.7-8. P.871-895.
- [11] **Paradiso S., Johnson D. L., Andreasen N. C.** Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. // *American Journal of Psychiatry.* 1999. V.156, No.10. P.1618-1629.

- [12] **Пикалов В. В., Мельникова Т. С.** Томография плазмы. Новосибирск: Наука. Сиб. издат. фирма, 1995.
- [13] **Gustavsson B.** A study of feasible tomographic inversion techniques for ALIS. // IRF Report No.039, April 1992, ISSN 0284-1738.
- [14] **Koslover R., McWilliams R.** Measurement of multidimensional ion velocity distributions by optical tomography. // *Rev. Sci. Instrum.* 1986. V.57, No.10. P.2441-2448.
- [15] **Левин Г. Г., Вишняков Г. Н.** Оптическая томография. М: Радио и связь, 1989.
- [16] **Вест Ч.** Голографическая интерферометрия. М: Мир, 1982 (перевод с английского).
- [17] **Vukićević D., Jäger H., Philipp H., Woisetschläger J.** Tomography reconstruction of the temperature distribution in a convective heat flow using multidirectional holographic interferometry. // *Appl. Optics.* 1989. V.28, No.8. P.1508-1516.
- [18] **Vishnyakov G., Levin G., Zakarin C.** Interferometric computed-microtomography of 3D phase objects. // *Proc. SPIE.* 1997. V.2984. P.67-71.
- [19] **Vishnyakov G. N., Levin G. G., Minaev V. L., Pickalov V. V., Likhachov A. V.** Tomographic interference microscopy of living cells. // *Microscopy and Analysis (UK).* 2004. V.18, No.1. P.15-17.
- [20] **Апресян Л. А., Кравцов Ю. А.** Теория переноса излучения. М.: Наука, 1983.
- [21] **Radon J.** Über die Bestimmung von Functionen durch ihre Integralwerte längs Gewisser Mannigfaltigkeiten. // *Ber. Ver. Sächs. Akad. Wis. Leipsig, Math.-Nat. Kl.* 1917. V.69. P.262-277. (**Радон И.** Об определении функций по их интегралам вдоль многообразий. В книге **Хелгасон С.** Преобразование Радона. М.: Мир, 1983.)
- [22] **Наттерер Ф.** Математические аспекты компьютерной томографии. М.: Мир, 1990. (перевод с английского).
- [23] **Пикалов В. В., Преображенский Н. Г.** Реконструктивная томография в газодинамике и физике плазмы. Новосибирск: Наука, 1987.
- [24] **Вишняков Г. Н.** Восстановление томограмм трехмерных объектов по двумерным проекциям. // *Оптика и спектроскопия.* 1988. Т.65, No.3. С.677-683.
- [25] **Туу Н. К.** An inversion formula for cone-beam reconstruction. // *SIAM J. Applied Mathematics.* 1983. V.43, No.3. P.546-552.

- [26] **Тихонов А. Н., Арсенин В. Я.** Методы решения некорректных задач. М.: Наука, 1986.
- [27] **Тихонов А. Н., Арсенин В. Я., Тимонов А. А.** Математические задачи компьютерной томографии. М: Наука, 1987.
- [28] **Бакушинский А. Б., Гончарский А. В.** Итеративные методы решения некорректных задач. М.: Наука, 1989.
- [29] **Ramachandran G., Lakshminarayanan A.** Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms. // *Proc. Nat. Acad. Sci. U.S.* 1971. V.68. P.2236-2240.
- [30] **Shepp L. A., Logan B. F.** The Fourier reconstruction of a head section. // *IEEE Trans. Nucl. Sci.* 1974. V.21, No.3. P.21-43.
- [31] **Лаврентьев М. М., Зеркаль С. М., Трофимов О. Е.** Численное моделирование в томографии и условно-корректные задачи. Новосибирск: Изд-во ИДМИ НГУ, 1999.
- [32] **Grangeat P.** Mathematical framework of cone-beam 3D reconstruction via the first derivative of the Radon transform. // *Proc. conf. Mathematical methods in tomography.* Oberwolfach, Germany. 1990. P.66-97.
- [33] **Feldkamp L. A., Davis L. C., Kress J. W.** Practical cone-beam algorithm. // *J. Opt. Soc. Amer. A.* 1984. V.1, No.6. P.612-619.
- [34] **Defrise M., Kinahan P., Townsend D., Michel C., Sibomana M., Newport D.** Exact and approximate rebinning algorithms for 3-D PET data. // *IEEE Trans. Med. Imag.* 1997. V.16, No.2. P.145-158.
- [35] **Smith B. D.** Cone-Beam tomography: recent advances and tutorial review. // *Optical Engineering.* 1990. V.29, No.5. P.524-534.
- [36] **Tam K., Samarasekera S., Sauer F.** Exact cone beam CT with spiral scan. // *Phys. Med. Biol.* 1998. V.43, No.4. P.1015-1024.
- [37] **Yan X., Leahy R.** Cone beam tomography with circular, elliptical and spiral orbits. // *Phys. Med. Biol.* 1992. V.37, No.3. P.493-506.
- [38] **Ценсор Я.** Методы реконструкции изображений, основанные на разложении в конечные ряды. // *ТИИЭР.* 1983. Т.71, No.3. С.148-160.
- [39] **Minerbo G., Sanderson J, van Hulsteyn D., Lee P.** Three-dimensional reconstruction of the X-ray emission in laser imploded targets. // *Applied Optics.* 1980. V.19, No.10. P.1723-1728.
- [40] **Tretiak O., Metz C.** The exponential Radon transform. // *SIAM J. Appl. Math.* 1980. V.39, No.2. P.341-354.