

Г. ДЖЕФФРИС  
Б. СВИРАС

МЕТОДЫ  
МАТЕМАТИЧЕСКОЙ  
ФИЗИКИ



METHODS  
OF  
MATHEMATICAL PHYSICS

by  
Sir Harold Jeffreys  
M. A., D. Sc., F. R. S.

and  
Bertha Swirles (Lady Jeffreys)  
M. A., Ph. D.

**Third Edition**

CAMBRIDGE  
CAMBRIDGE UNIVERSITY PRESS

1966

**Г. ДЖЕФФРИС, Б. СВИРЛС**

**МЕТОДЫ  
МАТЕМАТИЧЕСКОЙ  
ФИЗИКИ**

**ВЫПУСК 1**

Перевод с английского  
под ред. В. Н. Жаркова

**ИЗДАТЕЛЬСТВО «МИР»  
Москва 1969**

Фундаментальное руководство по прикладной математике, написанное известным геофизиком Г. Джеффрисом и его супругой Бертой Свирлс, представляет собой выдающееся явление в мировой литературе, с которым можно сравнить лишь такие труды, как «Методы математической физики» Куранта и Гильберта или «Методы теоретической физики» Морса и Фешбаха, выпущенные издательством «Мир» в русском переводе.

Для удобства советских читателей книга будет разбита на три выпуска: вып. 1 выйдет в 1969 г., 2 и 3 — в 1970 г. В вып. 1 будут рассмотрены функции действительного переменного, скаляры и векторы; тензоры, матрицы, кратные интегралы и теория потенциала и операционные методы.

Книга Г. Джеффриса и Б. Свирлс привлечет внимание физиков, геофизиков и астрономов, имеющих дело с той областью прикладной математики, где наряду с чисто рецептурной вычислительной техникой необходимо строгое понимание методов математической физики. Книга окажет также большую помощь аспирантам и студентам старших курсов.

*Редакция космических исследований, астрономии и геофизики*



## ПРЕДИСЛОВИЕ

Книга Г. Джеффриса и его супруги Берты Свирлс (леди Джеффрис) является одним из основных руководств по «Методам математической физики» повышенного типа на английском языке. Она выдержала три издания, причем последнее издание дважды переиздавалось. Таким образом, мы имеем дело с устоявшимся учебным руководством по прикладной математике.

Сам Джеффрис, являясь крупнейшим современным геофизиком, обобщил в ней свой полувековой опыт плодотворной работы в области теоретической геофизики, смежных областях астрономии и физики сплошной среды и прикладной математики.

В связи с этим в книге чувствуется отбор материала исследователем, который сам на протяжении многих лет использовал математику для прогресса естествознания. В ней описано много рациональных математических методик и приведено решение многих важнейших типовых задач. Короче говоря, она обладает всеми достоинствами английских книг по математической физике.

Предлагаемая вниманию читателя книга достаточно сложна и не может быть рекомендована для первоначального ознакомления с предметом. Она написана весьма своеобразным языком, и подход к изложению многих вопросов не тривиален. Эти обстоятельства создавали заметные трудности для коллектива, работавшего над переводом. Обширное предисловие авторов избавляет нас от обсуждения многих тем, в частности вопроса о терминологии и о той мере строгости, которая необходима в работах по теоретическому естествознанию. Мы во всяком случае нигде не стремились «улучшить» оригинал, а наоборот, старались сохранить все его особенности. В связи с тем, что

человек, который обратится к предлагаемой книге, можно полагать, уже достаточно знаком с предметом, мы решили не приводить дополнительного списка литературы. Обращаем внимание на то, что в книге принята правосторонняя система декартовых координат. В этой системе, если смотреть со стороны оси  $z$ , то поворот оси  $x$  в сторону оси  $y$  происходит против часовой стрелки. Для краткости такое вращение переведено как «вращение вправо».

Для удобства пользования русское издание книги разделено на три выпуска примерно равного объема. Поскольку это деление чисто условно, сохранена нумерация глав английского оригинала. В 1-й выпуск вошли гл. 1—8, во 2-й — гл. 9—15, в 3-й — 16—25. Замечания, помещенные в оригинале в конце книги, в русском переводе разнесены по соответствующим главам. Работа по переводу вып. 1 распределилась следующим образом: А. Л. Левшин перевел предисловие и гл. 8; М. Л. Гервер — гл. 1, 2; В. А. Калинин — гл. 3; Л. В. Никитин — гл. 4; В. Ф. Писаренко — гл. 5, 7; В. Л. Маркушевич — гл. 6.

В заключение я хотел бы от имени всех принимавших участие в работе выразить признательность леди Джеффрис, которая любезно прокомментировала перевод эпиграфов к отдельным главам. Леди Джеффрис также способствовала тому, что сотрудница издательства Кембриджского Университета миссис Мери Паркер прислала нам английский оригинал книги 1966 г. Мне хотелось бы поблагодарить миссис Паркер за эту любезность.

*В. Н. Жарков*

## ОТ АВТОРОВ

В третьей допечатке 3-го издания добавлены следующие разделы: 9.041а об интерполяции в случае, когда даны первые производные, и 9.181а — о достижениях в машинном счете. Было расширено исследование ортогональных преобразований в гл. 4 и сделано добавление в доказательстве леммы Ватсона в 17.03. Сделан ряд других небольших поправок и добавлений в тексте и примерах.

*Март 1966 г.*

Во второй допечатке 3-го издания добавлены следующие разделы: 5.051а о дифференцировании под знаком интеграла, 10.11а о методе, используемом в теории планет, и 23.07а со ссылкой на работу о кулоновских волновых функциях. Пересмотрены разделы 10.01, 10.013 в главе «Вариационное исчисление». Сделаны небольшие исправления и добавления в тексте и примерах.

*Июль 1961 г.*



## ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ

В настоящем издании мы внесли изменения в гл. 1, в основном в результате замечаний, сделанных проф. Безиковичем: более четко сформулирован ряд теорем, добавлено или сокращено несколько доказательств. Мы весьма признательны ему за элементарное доказательство теоремы об ограниченной сходимости Римановых интегралов, данное в примечании. В гл. 6 улучшено доказательство уравнения Пуассона. В гл. 17 мы более подробно рассмотрели интеграл Эйри для комплексного аргумента и сформулировали условия однородности аппроксимации асимптотических решений типа Грина в комплексной области. В гл. 23 мы добавили некоторые замечания об аналитическом продолжении решений и их применении к функциям параболического цилиндра.

Мы благодарим читателей за их сообщения по поводу замеченных ошибок и опечаток.

*Гарольд Джеффрис  
Берта Джеффрис*

*Апрель 1953 г.*

## ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Текст книги для 2-го издания был существенно пересмотрен. Большинство примечаний в конце книги было вставлено в текст. Кроме того, сделаны следующие важнейшие исправления. В гл. 1 теорема Гейне — Бореля и ее модификация Гурса вынесены вперед и использованы для вывода нескольких теорем, которые прежде доказывались другим методом. Кроме того, более полно изложена теория Римановых интегралов.

В гл. 4 добавлено описание блочных матриц и более подробно рассмотрена теорема о характеристических решениях для коммутирующих матриц. Гл. 5 (кратные интегралы) почти полностью переписана и теперь включает теорию функций нескольких переменных, которая частично была изложена в гл. 11. В гл. 9 анализ релаксационных методов расширен настолько, чтобы служить как введение в специальную литературу на эту тему.

В гл. 11 и 12 сделано много улучшений, в частности внесена важная поправка в доказательстве теоремы Коши, улучшено доказательство теоремы Осгуда — Витали и полностью пересмотрена теория обратных функций.

В гл. 17 несколько ослаблены условия для справедливости леммы Ватсона; теперь они выполняются почти во всех физических приложениях. Более полно изложен метод стационарной фазы. В гл. 24 расширен раздел об излучении мультиполя.

Доказательства по возможности сокращены или обобщены, добавлены некоторые новые примеры.

Мы благодарны многочисленным читателям за указания на ошибки. Две наиболее серьезных поправки сделаны проф. Литтлвудом и д-ром Картрайтом.

Мы особенно признательны за замечания проф. Литтлвуду (главы 1, 5, 11, 12), Холлу (гл. 4), проф. Безиковичу и д-ру Баркиллу (гл. 5).

*Гарольд Джеффрис  
Берта Джеффрис*

15 ноября 1948 г.

## ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Цель этой книги — описать те разделы чистой математики, в которых больше всего нуждается физика. Отбор материала довольно труден: книга, содержащая все методы, применяемые в различных разделах физики, была бы непомерно велика. В связи с этим мы обычно включали метод, если он применяется по крайней мере в двух разделах физики; однако мы не следовали этому правилу во всех случаях. Для иллюстрации дополнительных приложений введены специальные задачи.

Нам представляется, что многие студенты, заинтересованные главным образом приложениями, с трудом следят за абстрактными доказательствами не из-за отсутствия способностей, но потому что интерес у них возникает лишь тогда, когда им ясны непосредственные применения.

Мы предполагаем, что читатель знаком с курсом математического анализа. Поясним намечаемые нами уровень строгости и общности. Мы не разделяем широко распространенного взгляда, что любое доказательство хорошо, если оно будет предназначено для использования исследователями в естественных науках. Мы считаем, что для естественных наук, также как и для чистой математики, необходимо, чтобы были сформулированы основные принципы, и заключения вытекали из них. Но в естественных науках также необходимо, чтобы эти принципы были как можно теснее связаны с наблюдениями; для чистой математики несущественно, что взято за основу.

Мы придерживаемся здесь принципа, что строгий анализ в естественных науках не менее важен, чем в чистой математике. Мы также многократно обнаруживали, что легчайший путь сделать какое-либо утверждение достаточно обоснованным — это дать ему строгое доказательство. Некоторые наибо-



более важные результаты (например, теорема Коши) так удивительны на первый взгляд, что ничто более краткое, чем доказательство, не может заставить в нее поверить. С другой стороны, математик обычно неудовлетворен теоремой, если она высказана не в наиболее общей форме.

Приложения же часто ограничиваются несколькими специальными задачами. Поэтому мы часто даем доказательство при условиях, достаточных в большинстве приложений, но излишне ограниченных с точки зрения чистого математика. Общность — хорошая вещь, но иногда она может обойтись слишком дорого. В ряде случаев, если принятые условия не выполняются в данной задаче, то метод обобщения теоремы будет очевиден; иногда же это очень трудно, и мы не думаем, что стоит ломать копья ради редко встречающихся случаев. Для некоторых обширных проблем, которые важны, но требуют долгого обсуждения и хорошо изложены в каком-либо руководстве, мы сочли достаточным дать ссылки.

Мы считаем особенно важным для исследователей знать точную формулировку условий, при которых справедливы используемые ими теоремы. Часто можно встретиться с утверждением, что результат строго доказан; какой-либо контроль за тем, что постулированные при доказательстве условия удовлетворяются в реальной задаче, отсутствует — а очень часто эти условия на самом деле не выполняются. Такое неправильное использование математики может встретиться во многих разделах науки. С другой стороны, многие результаты часто доказаны при достаточных, но не необходимых условиях, и ученые колеблются, использовать ли их, ошибочно предполагая, что условия необходимы. Поэтому мы будем часто давать доказательства при более общих условиях, чем принятые в более элементарных курсах. Обе трудности вызваны главным образом тем, что теоремы разбросаны по многим книгам и статьям, и исследователи не знают, что и где искать.

Эту книгу можно читать последовательно, но некоторые части вполне независимы от предшествующего изложения; таким образом можно и даже полезно изучать отдельные главы параллельно.

В некоторых случаях мы приводим теоремы в частной форме перед более общей формулировкой, если последняя требует более сложного анализа. Это особенно относится к случаям, когда читатель, возможно, встретится с несколькими приложениями для частного вида теоремы до того, как ему потребуется более общая теорема.

Мы сомневались, стоит ли включать главу о теории функций действительного переменного. Обычно более полные работы требуют существенно большего времени для чтения, чем то, которым располагает физик-теоретик. К сожалению, в этих работах иногда сводят важные теоремы к узкому классу или неэффективному примеру, или опускают их целиком.

В итоге мы решили рассмотреть основные методы этой теории, но не доказывать подробно каждый результат; однако полагаем, что для студентов будет полезно самим заполнить некоторые пробелы в доказательствах. Если студент почувствует трудности в достижении того уровня абстракции, на который рассчитана большая часть этой главы, мы советуем ему читать до предела его возможностей, а затем переходить к следующим главам, возвращаясь назад по мере необходимости. В конце концов он обнаружит, что одолел всю эту главу, прежде чем окончит главу 14, и что он знает ее содержание и понимает смысл.

Мы не смогли полностью избежать ссылок на еще не прочитанный материал, но одно забегание вперед, наиболее серьезное — доказательство в гл. 12 теоремы, что алгебраическое уравнение  $n$ -й степени имеет  $n$  корней, использованное в гл. 4, настолько освящено традицией, что несколько менее серьезных забеганий вперед, нам, по-видимому, можно простить.

Обозначение специальных функций в настоящее время перестало быть единообразным, что во многих отношениях неудобно. Так, в квантовой механике проводится полная замена определений, гарантирующая нормализацию, но мы полагаем, что это всего лишь заменяет старые трудности новыми. Мы изменили обычное определение функции Лежандра, чтобы обеспечить более симметричный вид формул и переход к функциям

Бесселя без дополнительных численных множителей. Мы вернулись к определению функции  $K_n$ , данному Хевисайдом, но обозначили ее  $Kh_n$ . Среди других преимуществ, это упрощает ее связь с функциями Лежандра 2-го рода. Мы также отказались от обозначения  $\Gamma$  для факториальных функций, которое, по-видимому, никем и не было рекомендовано.

Непосредственным стимулом к написанию книги послужило то обстоятельство, что второе издание книги «Операционные методы в математической физике», написанной одним из нас, не могло быть выпущено. Большая часть этой работы была использована здесь с добавлением некоторых новых результатов.

Глава о дисперсии не подходила к той книге, поскольку она в основном не зависела от операционного метода. Однако мы включили ее, поскольку понятие групповой скорости уже обсуждалось ранее в связи с методом наискорейшего спуска. Теперь она более естественным образом вошла в главу об асимптотических разложениях. В этой главе описаны некоторые широко используемые методы, сведения о которых разбросаны по отдельным статьям. Большая часть работы «Декартовы тензоры» также включена в эту книгу. Приложения по термодинамике, содержащиеся в первой работе, к гидродинамике и теории упругости целесообразнее рассмотреть в руководствах по этим предметам.

Мы не пытались рассмотреть в деталях каждый раздел физики; это задача специальных учебников.

Мы очень благодарны многим друзьям за их поддержку при написании этой книги. Прежде всего мы должны поблагодарить д-ра Смитиса, чья огромная эрудиция была целиком в нашем распоряжении; он дал неоценимые советы и великодушно помог при сверке. Наш долг — признать, что в некоторых местах мы шли своим путем, несмотря на его строгие протесты.

Д-р Миллер особенно помог нам с гл. 9 и 23, а Бонди — с гл. 24. Много ценных советов дали нам профессора Ньюман, Оффорд, Розенхэд, Торнбалл, а также Безикович, Картрайт и Долзелл.



Мы благодарим также университеты Кембриджа, Лондона и Манчестера за разрешение использовать в качестве примеров экзаменационные вопросы, и работников издательства Кембриджского университета за их внимание к книге и их готовность идти навстречу пожеланиям довольно педантичных авторов.

*Гарольд Джеффрис  
Берта Джеффрис*

1946 г.

Главные разделы каждой главы нумерованы в десятичной системе с шагом 0,01; подразделы показаны следующими десятичными числами. Когда текст раздела или подраздела продолжает предыдущий, нумерация уравнений также продолжается.

Источники, из которых взяты примеры, показаны следующими сокращениями:

M. T.	Mathematical Tripos, Part II and Schedule A.
M. T., Sched. B.	Mathematical Tripos, Part III and Schedule B.
Prelim.	Preliminary Examination in Mathematics.
M/c, III	Manchester, Final Honours in Mathematics.
I. C.	Imperial College, London.

## ДЕЙСТВИТЕЛЬНОЕ ПЕРЕМЕННОЕ

В прошлые времена очень любили мелкую рыбешку.

Даниэль Чандлер Харрис  
«Сказки дядюшки Римуса»

**1.01 Отношение математики к физике.** Простейшее математическое понятие — это понятие числа. Число — это свойство, общее для всех классов, которые могут быть поставлены друг другу в соответствие путем сопоставления их элементов (по одному из каждого класса), так что все элементы образуют пары и ни одного лишнего не остается. Исходя из этого определения, мы можем придать смысл основным операциям — сложению и умножению. Рассмотрим два класса с числами  $a$  и  $b$ , у которых отсутствуют общие члены. Сумма  $a$  и  $b$  — это число класса, состоящего из всех элементов обоих классов, взятых вместе. Произведение  $a$  и  $b$  равно числу всевозможных пар элементов, взятых по одному из каждого класса. Мы не всегда можем придать смысл вычитанию и делению, потому что, например, мы не можем найти класса с числом  $2-3$  или  $7/5$ . Но оказывается очень удобным расширить понятие числа так, чтобы оно включало отрицательные числа, отношения чисел (независимо от того, положительные они или отрицательные) и даже иррациональные числа. Поступая таким образом, мы сможем дать определение всем четырем основным арифметическим действиям, и результат их выполнения всегда будет выражен числом. Нам не нужно больше беспокоиться о том, *возможно* ли некоторое действие внутри частного множества чисел; мы знаем, что оно станет возможным, как только мы придадим достаточную общность понятию числа. Пока мы имеем дело с основными операциями, мы можем пользоваться алгеброй, т. е. мы можем доказывать формулы, которые будут верны при подстановке в них вместо символов любых чисел, с одним лишь исключением, а именно мы не должны делить на 0.

Кроме того, формулы могут остаться верными, если заменять в них буквы не числами, а чем-нибудь другим. Именно этому математическая физика обязана своими возможностями. Поэтому полезно знать, при каких условиях можно перенести

правила алгебры в другую область, где имеют дело не только с числами. Нам придется тогда придать новый смысл основным операциям и знаку « $=$ » (или воспользоваться тем, что сделано для нас другими), но так, чтобы этот новый смысл позволял по-старому оперировать с символами. Подходящий набор условий такой\*). Мы говорим, что элементы системы образуют поле  $F$ , если:

- 1)  $a + b$  и  $ab$  являются однозначно определенными элементами из  $F$  для любых  $a, b$  из поля  $F$ ;
- 2)  $b + a = a + b$  (коммутативный закон сложения);
- 3)  $(a + b) + c = a + (b + c)$  (ассоциативный закон сложения);
- 4)  $ba = ab$  (коммутативный закон умножения);
- 5)  $a(bc) = (ab)c$  (ассоциативный закон умножения);
- 6)  $a(b + c) = ab + ac$  (дистрибутивный закон);
- 7) имеется два элемента 0 и 1 в поле  $F$ , такие, что  $a + 0 = a$ ,  $a1 = a$ ;
- 8) для всякого элемента  $a$  из  $F$  существует элемент  $x$  из  $F$ , такой, что  $a + x = 0$ ;
- 9) для каждого элемента  $a$  из  $F$ , кроме 0, существует элемент  $y$  из  $F$ , такой, что  $ay = 1$ .

Следует обратить внимание на то, что первые семь правил справедливы, если  $F$  состоит только из положительных целых чисел и 0, а последние два правила для такого  $F$  несправедливы, поскольку в этом  $F$  нет такого целого  $x \geq 0$ , что  $a + x = 0$ , если  $a = 1$ , и нет такого целого  $y$ , которое дает  $ay = 1$ , если  $a = 2$ . Восьмое правило вводит отрицательные числа и, следовательно, вычитание. Девятое правило вводит обратные величины и отсюда деление и рациональные дроби. Эти правила справедливы, если  $F$  состоит из всех рациональных чисел, положительных или отрицательных.

Сформулированные правила не устанавливают какого-либо отношения порядка, т. е. хотя они предполагают понятие равенства и, следовательно, неравенства ( $\neq$ ), они не вводят понятия «больше» или «меньше». Мы могли бы упорядочить числа любым способом, сохранив соотношения между ними, согласно условиям 1, 7—9, и правила остались бы по-прежнему верны. В алгебре и чистой геометрии еще можно обойтись без понятий «больше» или «меньше», но ни в высшей математике, ни в какой бы то ни было отрасли физики это невозможно. Измерение не есть установление точного равенства, а установление

---

\*) Впервые он установлен Дедекиндом для случая, когда знаки « $+$ » и « $\times$ » обладают своими обычными арифметическими значениями; в общем виде эти условия сформулированы Х. Вебером.



равенства в пределах определенной погрешности. Поэтому нужны новые правила, касающиеся неравенств, а именно:

10) для любых  $a, b$  в поле  $F$  справедливо  $a > b$ ,  $a = b$  или  $b > a$  (закон сравнимости);

11) для данных  $a, b$  в поле  $F$  может быть справедливо только одно:  $a > b$ ,  $a = b$ ,  $b > a$  (трихотомия);

12) если  $a > b$ , а  $b > c$ , тогда  $a > c$  (свойство транзитивности);

13) если  $a > b$ , тогда  $a + c > b + c$  для любого  $c$  (аддитивность упорядочения);

14) если  $a > b$ ,  $c > 0$ , то  $ac > bc$  (мультипликативность упорядочения);

15) если  $a > b$ , то  $b < a$  (определение знака  $<$ ).

Использование математики в науке — это использование языка, при помощи которого мы можем устанавливать соотношения, слишком сложные, чтобы их можно было достаточно кратко описать обычным языком. Правила, которым подчиняются символы, — грамматика такого языка. В последние годы эта точка зрения получила существенное развитие, в особенности в трудах Карнапа. Но для того чтобы быть пригодным, язык должен удовлетворять двум условиям. Он должен быть таким, чтобы с его помощью можно было сказать то, что нужно, т. е. он должен обладать достаточной общностью. Он также должен быть непротиворечивым, т. е. исходя из самих правил нельзя вывести что-либо согласно этим правилам являющееся ложным. Научное использование математики было бы невозможно, если бы с ее помощью можно было, например, доказать для каких-то  $a$  и  $b$ , что  $a$  одновременно и больше и меньше  $b$ . Вплоть до конца XIX в. считалось само собой разумеющимся, что математика непротиворечива. Но затем неожиданно возник ряд трудностей, указавших на необходимость глубокого анализа основ. В известной книге «Principia Mathematica» Уайтхед и Рассел показали, что математические утверждения о вещественных числах (и не только об отношениях целых чисел, как положительных, так и отрицательных) можно сформулировать по-иному, как предложения об элементарном понятии соответствия классов (которое устанавливается сопоставлением их элементов), и что их можно вывести из аксиом такого сопоставления и аксиом чистой логики.

В последующих работах некоторые из логических аксиом были модифицированы; наилучший выбор аксиом все еще является предметом дискуссии. Еще позднее Гёдель и Карнап показали, что непротиворечивость данной системы математических аксиом не может быть доказана методами, использующими только правила данной системы. Мы вынуждены возвращаться

к чему-то, что в конечном счете подобно обычному языку, когда мы хотим говорить о математике! В этой области, которая является промежуточной между логикой и тем, что мы обычно называем элементами математики, имеется обширная современная литература, и физикам следует знать о ее существовании, хотя ее подробное изучение — дело специалистов.

**1.02. Физические величины.** Общность требует, чтобы в любой конкретной области язык содержал символы для вещей, о которых нам приходится говорить, и символы для процессов, которые мы выполняем. Пастух оказался бы в крайне затруднительном положении, если бы ему пришлось оперировать языком, не содержащим слов «овца» и «стрижка», по сути дела ему пришлось бы выдумать эти слова; именно это мы и делаем в науке. Пока язык представляет собой стройную систему, он несколько не становится хуже от того, что в нем есть множество слов, которыми мы не пользуемся. Чистый математик, специализирующийся в области теории чисел, может использовать обычную алгебру, несмотря на то, что он, возможно, не нуждается в использовании отрицательных чисел и дробей. Для него правила 8 и 9 являются совершенно необязательным обобщением. В настоящее время в физике фундаментальное понятие измерения близко понятию сложения, а большинство физических законов суть утверждения о пропорциональности, что соответствует понятиям умножения и деления. В этом, в конечном счете, и польза математики. Так, например, если два стержня расположены таким образом, что они образуют один прямой стержень, то длина составного стержня является суммой длин двух первоначальных. Это не теорема и не экспериментальный факт, это определение сложения для длин. Кроме того, безразлично, какой из стержней взят первым, т. е. выполняется закон коммутативности сложения. Далее, если мы объединим три стержня, то общая длина не зависит от порядка, следовательно, выполняется закон ассоциативности сложения. Это экспериментальные факты, устанавливаемые сравнением стержней. Этих правил достаточно, чтобы обосновать использование масштабов при измерении длины; при измерениях любая длина сравнивается со стандартной при помощи масштаба, каждое деление которого сравнивалось со стандартным предметом во время изготовления масштаба. Величины, которые можно измерить посредством физического сложения, Кэмпбел назвал *основными величинами* \*). Наиболее важные из них — числа (классов), длина, время и масса; что же касается про-

---

\*) Лучше было бы назвать их «элементарными».

цессов физического сложения, то их можно также ввести для площади и объема, для электрического заряда, потенциала и тока, а также для многих других величин.

Утверждение «расстояние равно 3,7 см» содержит число и единицу измерения. Часто полагают, что алгебра применима лишь к числам и поэтому в математической трактовке символ для обозначения расстояния относится лишь к 3,7, а не к сантиметрам. Единицы существенны, в противном случае оказалось бы, что 10 мм — это не такая же длина, как 1 см, а 1 см и 1 миля — одно и то же. А это противоречит физике, так как единственным обоснованием измерения является прямое физическое сравнение наложением. Мы обойдем эту трудность, если условимся, что символ для обозначения длины относится к самой длине, а не просто к числу, входящему в ее измерение. Например, «1 дюйм = 2,54 см» — полезное утверждение: как 1 дюйм, так и 2,54 см — символы, обозначающие одну и ту же длину. При конкретных измерениях мы, естественно, выражаем фактическую величину символов с помощью именованных чисел, что включает установление единиц измерения; однако в общей теории единицы измерения безразличны. При таком подходе можно сказать, что символы обозначают не числа, а физические величины.

С другой стороны, можно считать, что символы — числа. Тогда может возникнуть и возникает путаница, когда в одной системе производят измерения в разных единицах, по-разному выражающих одно и то же, и в разных системах — в одинаковых единицах, выражающих разные вещи. Измерять высоту в ярдах, длину в футах и глубину в морских саженях не только непоследовательно, но и нелепо. При должном старании этот способ можно использовать правильно, но он страдает некоторыми недостатками; в частности, придается слишком большое значение единицам и малое основным физическим сравнениям, без которых эти единицы были бы бесполезны. Он также наводит на ряд сравнений, бессмысленных, как мы сейчас увидим, с точки зрения физики.

Если мы используем понятие (физической) величины и применим алгебру, немедленно возникнет вопрос, что мы подразумеваем под  $a = b$  и  $a + b$ , если  $a$  — длина, а  $b$  — время или масса. Можно придать значение сумме  $a + b$ , хотя оно будет очень искусственным, но выражению  $a = b$  нельзя придать никакого физического смысла. Напротив, выражение  $a/b$  будет означать соответственно скорость или длину на единицу массы.

Группа правил 10—14 нуждается поэтому в модификации. Правила 1—9 можно не изменять, хотя они вводят много совершенно не нужных сложений и вычитаний, а также, возможно,

умножений и делений. В добавление к трем возможностям, предусмотренным правилом 10, нужно добавить четвертую:  $a$  и  $b$  могут быть несравнимы и, следовательно, принадлежать разным полям, а их произведение и отношение в свою очередь принадлежать другим полям. Это еще один недостаток использования символов, обозначающих только числа, входящие в измерение. Поскольку все числа сравнимы, этот язык не отражал бы того факта, что бессмысленно говорить, будто время больше плотности. Теперь мы можем также сказать, что если  $a$  и  $b$  несравнимы, то  $a + b$  не является физической величиной, и складывать их не надо.

Таким образом, поле всех физических величин разбивается на классы (делянки). Величины, принадлежащие одному классу, сравнимы между собой, но их произведение принадлежит другому классу, если только один из сомножителей не является числом.

В этом языке, необходимый для физики, не вполне совпадает с обычной алгеброй. Но так как алгебра является непротиворечивой системой, а утверждение несравнимости некоторых величин лишь исключает из нее некоторые предложения, не добавляя новых, то этот язык тоже непротиворечив. Мы увидим, что модификация соответствует понятию размерностей. Величины различных размерностей несравнимы; также несравнимы некоторые величины одинаковой размерности. Например, при использовании некоторых определений электрический заряд и «магнитный заряд» имеют одинаковую размерность, и хотя они являются основными величинами, складывать их бессмысленно. Можно считать, что поле физических величин удовлетворяет законам алгебры, но со следующей оговоркой. Сравнимые величины удовлетворяют условию 10, и их можно складывать (по крайней мере при вычислениях), а несравнимые нельзя. Следует, однако, отметить, что невозможность сложения с помощью физического процесса не сводится к несравнимым величинам. Например, нет процесса, позволяющего из двух веществ с плотностью  $1 \text{ г/см}^3$  получить вещество с плотностью  $2 \text{ г/см}^3$ . Плотность измеряется не непосредственно, а вычисляется через основные величины массы и длины. Она называется *производной величиной*. Некоторые величины могут быть как основными, так и производными. Например, электрический ток, измеряемый по его магнитному действию, является основной величиной, а рассматриваемый как заряд, проходящий через сечение проводника за единицу времени, — производной. Многие производные величины являются отношениями двух величин одинаковой размерности. Например, форму треугольника можно задавать двумя отношениями — каждой из двух сторон к третьей.

Эти отношения — просто числа, и правила алгебры применимы к ним без изменений \*).

**1.03. Действительные числа.** Большая часть этой главы будет известна тем, кто изучил какую-нибудь современную книгу по анализу; эта глава и не претендует на соперничество с обычными работами по чистой математике. Однако мы полагаем, что некоторые обсуждения будут здесь уместны по нескольким причинам. Во-первых, чисто математические работы не уделяют достаточного внимания вопросу, почему выбранные аргументы связаны с физикой, а поэтому физики предпочитают полагать, что эти аргументы не связаны с физикой. Во-вторых, эти книги нередко бывают такие толстые, что трудно винить физика, который решает, что у него нет времени, чтобы проработать их. В-третьих, внимание, которое уделяется очень странным функциям, ведет к тому, что рассматривается как бы патология функций. Дело в том, что всякая функция, кроме константы, имеет особенности, и изучение этих особенностей может дать нам важные конструктивные результаты, которые очень трудно получить как-либо по-другому. Однако можно считать, что мы — практики, и ограничиться рассмотрением особенностей, встречающихся в физике, а редкие исключения предоставим исследовать специалистам, в этом случае математикам-профессионалам.

Существо проблемы проявилось в теореме Евклида о том, что отношение гипотенузы к катету в равнобедренном прямоугольном треугольнике не равно никакой рациональной дроби. Евклид, об этом следует помнить, не использовал того, что мы теперь называем численным измерением физических величин. Когда он говорил, что два отрезка равны, он подразумевал, что при наложении они совпадут; это прямое физическое сравнение, не использующее числового описания длины. Когда он говорил, что квадрат гипотенузы равен удвоенному квадрату катета, то подразумевал, что квадрат со стороной, равной гипотенузе, можно разрезать на части так, чтобы из них получилось два квадрата со стороной, равной катету. Он повсюду работал с самими количествами, а не с числами, соответствующими им при измерении в каких-либо специальных единицах. Теорема Евклида показывает, что язык рациональных чисел непригоден для одновременного описания длин катета и гипотенузы треугольника, что легко можно вывести из правил его геометрии.

Измерение с помощью единиц — слишком полезный способ, чтобы легко от него отказываться. Его можно сохранить

---

\*) Подобная трактовка принадлежит Строуду. Дальнейшее обсуждение и вопросы преподавания см. в [1].

согласованно с теоремой Евклида одним из следующих способов:

1) Поскольку можно найти бесконечное число пар целых чисел  $x, y$ , удовлетворяющих уравнению  $x^2 + y^2 = z^2$ , где  $z$  — целое число, причем так, чтобы  $x/y$  было сколь угодно близко к единице, то мы можем предположить, что стороны прямоугольного треугольника в точности удовлетворяют соотношению  $x^2 + y^2 = z^2$ . Но равенство  $x = y$  выполняется лишь приближенно с точностью до ошибки измерения, а стороны всегда соизмеримы (являются точными кратными некоторой определенной длины).

2) Мы могли бы сказать, что  $x/y$  точно равно единице, а уравнение  $x^2 + y^2 = z^2$  выполняется приближенно.

3) Можно сказать, что язык рациональных чисел недостаточно полон и нужен более полный язык, в котором равенства  $x = y$ ,  $x^2 + y^2 = z^2$  выполняются одновременно и непротиворечивы. Последняя альтернатива и была принята повсеместно в результате введения в арифметику иррациональных чисел. Она не противоречит аксиомам Евклида; первая альтернатива противоречит им, так как Евклид предполагает, что отрезок прямой может иметь любую длину; вторая же альтернатива противоречит одному из наиболее известных следствий этих аксиом.

Экспериментально доказать справедливость выбранной альтернативы нельзя, потому что правила 1 и 2 могли бы быть справедливы в пределах ошибок измерения, даже если бы мы приняли, что  $x, y, z$  — только целые числа. Но все нелепо усложнилось бы, и для того чтобы принять одну из первых альтернатив, потребовалось бы не только принять некий неизвестный и неопределяемый эталон (такой, что всякая истинная длина представляет точное кратное этого эталона), но и отказаться от простоты правил Евклида без экспериментально обоснованных причин. В физике обычно принимают альтернативу 3 и создают достаточно общий язык. Мы вводим *действительные числа* и *предполагаем*, что к ним можно применять сложение, вычитание, умножение и деление, причем выполняются те же основные правила, что и для рациональных чисел, и можно ввести отношение порядка, удовлетворяющее правилам 10—15. Действительные числа отличаются от рациональных чисел тем, что обладают неким свойством *полноты*, которое обеспечивает, например, существование действительного числа  $\sqrt{2}$ , квадрат которого равен 2. Отнюдь не очевидно, что это можно сделать без противоречий (и в течение 2000 лет полагали, что действительные числа бессмысленны\*), но в XIX в. исследования Де-

---

\*) Отсюда и название «иррациональные числа».

декинда, Кантора и других установили применимость и возможность практического использования действительных чисел. Для наших целей вполне достаточно этого обоснования. Однако логическое обоснование включает рассмотрение бесконечных совокупностей. В действительности очевидно, что оценка  $\sqrt{2}$  с помощью извлечения корня или методом последовательных приближений к непрерывной дроби, если произведено конечное число шагов, может дать только рациональное число. Для того чтобы придать  $\sqrt{2}$  точное числовое значение, необходимо бесконечное число шагов. За конечное число шагов построение Евклида дает рациональную дробь, которую можно идентифицировать с  $\sqrt{2}$ , но которая не является его описанием в числовой системе, и доказательство непротиворечивости самих аксиом Евклида было получено пока только на пути численного приближения. В школах понятие  $\sqrt{2}$  вводится в основном потому, что мы верим в возможность непротиворечивого измерения физических объектов, а аксиомы Евклида выглядят правдоподобно; однако мы забываем, что евклидовы треугольники не являются действительными треугольниками, или если мы помним об этом, то полагаем, что действительные треугольники — это лишь несовершенные подобию евклидовых. С физической точки зрения треугольник Евклида является идеализированным приближением истинного, и нельзя считать само собой разумеющимся, что идеализация не вносит новых трудностей внутреннего характера.

**1.031 Последовательности вложенных интервалов. Дедекндовы сечения.** Основным свойством действительных чисел является то, что они могут быть сколь угодно точно приближены рациональными числами.

Когда мы говорим, что

$$\sqrt{2} = 1,414 \dots,$$

мы высказываем следующее множество предложений: 1) 2 заключено между  $1^2$  и  $2^2$ ; 2) 2 заключено между  $1,4^2$  и  $1,5^2$ ; 3) 2 заключено между  $1,41^2$  и  $1,42^2$ ; 4) 2 заключено между  $1,414^2$  и  $1,415^2$  и т. д. до любой требуемой точности. На каждом этапе этот процесс можно рассматривать как разделение десятичных дробей с заданным числом знаков после запятой на два класса: на дроби, квадраты которых соответственно больше или меньше 2. На 3-м шаге, например, квадраты 1,414; 1,413; 1,412 меньше 2, а квадраты 1,415; 1,416; 1,417 больше 2. На этом шаге мы ничего не говорим о дробях 1,4141; 1,4142; ...; 1,4149; но на следующем шаге мы скажем, что 2 лежит между

квадратами чисел 1,4142 и 1,4143. Взяв достаточное число десятичных знаков, мы можем сделать нерассматриваемый интервал сколь угодно малым, так как на каждом шаге он уменьшается в десять раз. Таким образом, всякая десятичная дробь конечной длины будет отнесена к одному из двух классов в соответствии с тем, больше или меньше 2 ее квадрат. Этот процесс определяет единственную десятичную бесконечную дробь, которую мы можем принять за  $\sqrt{2}$ , и  $\sqrt{2}$  может рассматриваться как предел, достигаемый последовательными приближениями с обеих сторон.

Этот процесс, который можно сильно обобщить, является примером определения действительного числа с помощью *вложенной последовательности пар рациональных чисел*. Мы берем две последовательности рациональных чисел  $\{a_n\}$  и  $\{b_n\}$ , удовлетворяющие следующим условиям:

- а)  $a_{n+1} \geq a_n$ ,
- б)  $b_{n+1} \leq b_n$ ,
- в)  $a_n \leq b_n$  для всех  $n$ ,
- г) для любого положительного рационального числа  $\varepsilon$  найдется такое число  $N$ , что

$$b_n - a_n < \varepsilon \text{ для всех } n > N.$$

Такая вложенная последовательность  $\{a_n | b_n\}$  может использоваться как определение действительного числа. Элемент  $a_n, b_n$  вложенной последовательности состоит из всех рациональных чисел, больших или равных  $a_n$  и меньших или равных  $b_n$ . Действительное число, определяемое вложенной последовательностью, лежит между концевыми точками всех ее элементов. Вложенная последовательность может определять рациональное число. Например, если мы рассмотрим десятичные дроби, квадраты которых соответственно строго больше или строго меньше чем 2,25, мы получим вложенную последовательность (1; 2); (1,4; 1,6); (1,49; 1,51); (1,499; 1,501) ... Единственной десятичной дробью, лежащей между концевыми точками всех ее элементов, является дробь 1,5. Ее квадрат в точности равен 2,25. Для всякой рациональной дроби мы можем построить подобную вложенную последовательность. Таким образом, рациональные числа сами являются действительными числами.

Отдельное действительное число может быть определено многими различными вложенными последовательностями. Например, вместо деления интервала на десять частей на каждом шаге мы могли бы делить его пополам. Так мы получили бы двоичную дробь. Нужно было бы сделать приблизительно втрое больше шагов для достижения той же точности, однако было бы определено то же самое действительное число, что и раньше.



Две вложенные последовательности  $\{a_n | b_n\}$  и  $\{\alpha_n | \beta_n\}$  определяют одно и то же действительное число тогда и только тогда, когда  $a_n, b_n$  содержит  $\alpha_m, \beta_m$  при достаточно больших  $m$ , и аналогично  $\alpha_n, \beta_n$  содержит  $a_m, b_m$  при достаточно больших  $m$ . Впрочем, достаточно выполнения только одного из этих условий — второе будет из него следовать.

Теперь мы подходим к наиболее важному свойству системы действительных чисел. Отбросим требование рациональности  $a_n, b_n$  и рассмотрим вложенную последовательность  $\{a_n | b_n\}$ , где теперь  $a_n$  и  $b_n$  — действительные числа. Интервал, являющийся элементом такой вложенной последовательности, состоит из всех действительных чисел, больших или равных  $a_n$  и меньших или равных  $b_n$ . В условии «г»  $\varepsilon$  теперь — произвольное действительное положительное число. Можно доказать, что существует одно и только одно действительное число, лежащее в каждом интервале такой вложенной последовательности. Другими словами, применяя к действительным числам процесс, который мы применили к рациональным числам, мы не получим ничего нового и не выйдем из уже определенной системы. Это и есть свойство полноты, упомянутое в 1.03.

Другим важным способом определения действительных чисел являются *дедекиндовы сечения*. Если рациональные числа разделены на два класса  $L$  и  $R$  так, что каждый элемент из  $L$  меньше каждого элемента из  $R$ , то существует единственное действительное число, большее или равное каждому элементу из  $L$  и меньшее или равное каждому элементу из  $R$ . Если это действительное число рационально, то оно является или наибольшим элементом из  $L$ , или наименьшим элементом из  $R$ . Например,  $L$  может состоять из всех отрицательных рациональных чисел, нуля и положительных рациональных чисел, квадраты которых меньше 2, а  $R$  — из положительных рациональных чисел, квадраты которых больше 2. Это сечение определяет действительное число  $\sqrt{2}$ .

Дедекиндовы сечения наиболее естественно возникают при классификации чисел в соответствии с тем, обладают они некоторым свойством или нет. Например, свойство «квадрат  $x$  не больше 2,25» определяет класс  $L$ , наибольший элемент которого равен 1,5; «квадрат  $x$  меньше 2,25» определяет класс  $L$  без наибольшего элемента и 1,5 — наименьший элемент из класса  $R$ . Свойство « $x$  рационально и его квадрат меньше 2» определяет классы  $L$  и  $R$  рациональных чисел без наибольшего и наименьшего элементов соответственно; « $x$  действительно и его квадрат меньше 2» определяет класс  $L$  без наибольшего элемента и класс  $R$  с наименьшим элементом  $\sqrt{2}$ .

На языке дедекиндовых сечений свойство полноты системы действительных чисел эквивалентно тому, что любое сечение в действительных числах определяет действительное число. Таким образом, многие проблемы, которые нельзя решить в системе рациональных чисел, можно решить в системе действительных чисел. Пока мы рассмотрели только  $\sqrt{2}$ , но мы уже готовы к появлению  $\pi$  и  $e$ , и нам не нужно будет каждый раз отыскивать такую формулировку проблемы, чтобы можно было решить ее в рациональных числах. Использование системы действительных чисел позволяет избежать многих осложнений, не прибегая к физике.

Методы вложенных интервалов и дедекиндовых сечений эквивалентны. Если классы  $L$  и  $R$  существуют, то мы можем построить вложенную последовательность интервалов, взяв  $a_1, a_2 \dots$  из  $L$  и  $b_1, b_2$  из  $R$  таким образом, чтобы все условия, наложенные на рассматриваемые последовательности интервалов, выполнялись. Обратно, если существует вложенная последовательность, то имеются и такие рациональные числа  $r$ , для которых существуют  $a_m$ , превосходящие их, и такие рациональные числа, которые больше всех  $a_m$ . Соответствующие неравенства определяют классы  $L$  и  $R$ , и все условия для сечений будут выполнены.

Если вложенная последовательность  $(a_m, b_m)$  определяет положительное действительное число  $x$ , то  $(1/b_m, 1/a_m)$  определяет число  $1/x$ . Если далее вложенные последовательности  $(a_m, b_m)$ ,  $(a'_m, b'_m)$  определяют  $x, x'$ , то  $(a_m a'_m, b_m b'_m)$  определяет  $xx'$ . Вложенная последовательность  $(-b_m, -a_m)$  определяет  $-x$ . Наконец, если  $(a_m, b_m)$  определяет  $x$ , а  $(a'_n, b'_n)$  определяет  $x'$ , то, независимо от того, положительны или отрицательны  $x$  и  $x'$ , последовательность  $(a_m + a'_m, b_m + b'_m)$  определяет  $x + x'$ .

Таким образом, для действительных чисел определены операции сложения, вычитания, умножения и деления. Можно показать, что они удовлетворяют основным правилам. Полное рассмотрение дано в [2].

Ни один из методов не доказывает существования иррациональных чисел, но оба показывают, что их можно использовать, не приходя к противоречиям, и что любое предложение, доказанное с их помощью, можно интерпретировать как истинное предложение о рациональных числах. (Установить его бывает, разумеется, значительно труднее.)

Цель книги «Principia Mathematica» более значительна: действительное число интерпретируется как класс рациональных чисел (фактически класс  $L$  Дедекенда) и алгебраическим зако-

нам придается смысл в терминах операций над этими классами. Доказывается, что установленные так законы справедливы. В этом смысле действительно доказывается существование иррациональных чисел, удовлетворяющих законам алгебры.

**1.032. Число  $\varepsilon$ ; доказательства от противного.** Особенность многих основных теорем о действительных числах состоит в том, что не удается дать их прямого доказательства. Они доказываются с помощью процесса, известного под названием *доказательство от противного*. Мы считаем истинным утверждение, противоположное доказываемой теореме, и доказываем, что это ведет к противоречию. Из этого заключаем, что теорема не может быть ложной и, следовательно, она справедлива. А так как большинство теорем имеет заключение вида  $x = y$ , то противоположными утверждениями будут неравенства вида  $x < y$  или  $x > y$ . Большинство начинающих находят, что правильно обращаться с неравенствами значительно труднее, чем с равенствами, и из всех трудностей математической физики для многих студентов наибольшей является изучение аппроксимаций. Именно по этой причине наиболее низкие оценки на экзамене получают за малые колебания динамических систем и за потенциалы тел, близких к сферическим. Природа не состоит ни полностью, ни по большей части из задач, составленных экзаменатором, которые можно решить абсолютно точно за конечное число шагов. За что бы мы ни взялись, нам необходимо сперва преодолеть робость перед аппроксимацией. Различие между теорией действительного переменного и динамикой состоит в том, что в первой мы хотим рассматривать сколь угодно точные приближения, требующие произвольного числа шагов, а во второй — лишь приближения, достаточно точные для практических целей. Но опыт в одном придает уверенность в другом.

Наиболее простым рассуждением такого типа является следующее: если  $x \geq 0$  и  $x < \varepsilon$ , где  $\varepsilon$  положительно и может быть выбрано сколь угодно малым, то  $x = 0$ . Это верно потому, что не существует числа  $x$ , которое больше нуля и меньше любого положительного  $\varepsilon$ . Очевидное обобщение можно получить после введения понятия *модуля* или *абсолютной величины*  $x$ , которая обозначается  $|x|$  и читается «модуль  $x$ ». Модуль равен  $x$ , если  $x$  неотрицательно, и  $-x$ , если  $x$  отрицательно. Следовательно, всегда  $|x| \geq 0$ . Тогда если  $|x| < \varepsilon$  для любого положительного  $\varepsilon$ , то  $|x| = 0$  и, следовательно,  $x = 0$ . Отметим, что  $|x| + |y| \geq |x + y|$ ,  $|x - y| \geq |x| - |y|$ .

В проведенном рассуждении необходим символ для малой величины. Если мы скажем, что  $\varepsilon = 0,001$ , и докажем вычислениями, что  $|x| < 0,001$ , то можно возразить: «Вы не доказали.

что  $x = 0$ ; может быть,  $x = 0,0001$ ». Символ  $\varepsilon$ , обозначающий произвольно малую величину, дает нам возможность ответить на такое возражение. Коль скоро мы докажем, что  $|x|$  меньше *любого*  $\varepsilon$ , то сумеем отвергнуть любое отличное от нуля значение  $x$ , которое возражающий может предложить.

Существенно, что в общем случае мы имеем дело с процессами, которые можно завершить только за бесконечное число шагов. Примером служит доказательство того, что две вложенные последовательности интервалов определяют одно и то же число. Мы обходим эту трудность и получаем конечное доказательство: пусть  $a \neq b$ , тогда  $|a - b|$  имеет определенное значение  $M$ , отличное от нуля. Если затем мы сумеем показать, что  $M < \varepsilon$  для любого положительного  $\varepsilon$ , то из этого будет следовать, что  $M = 0$  и, следовательно, вопреки предположению,  $a$  и  $b$  должны быть равны между собой.

**1.033. Множества.** *Предельной точкой* множества чисел называется число  $x$ , такое, что для любого  $\varepsilon > 0$  найдется принадлежащее множеству число  $y$ , отличное от  $x$ , и  $|y - x| < \varepsilon$ . Отсюда следует, что существует бесконечно много значений  $y$ , удовлетворяющих этому условию. Действительно, по определению одно значение найдется, назовем его  $y_1$ . Возьмем новое  $\varepsilon$ , скажем  $\varepsilon_1$ , меньшее, чем  $|y_1 - x|$ . Тогда должно существовать другое  $y$  из множества, скажем  $y_2$ , такое, что  $0 < |y_2 - x| < \varepsilon_1$ . Очевидно, этот процесс можно продолжать неограниченно \*).

Очевидно, что конечное множество не имеет предельных точек. Однако бесконечное множество тоже может не иметь предельных точек. Рассмотрим множество всех целых чисел. Для любого его элемента нет других элементов на расстоянии, меньшем 1, а любое не целое число не может иметь более одного целого на расстоянии, не превышающем  $1/2$ . В множестве рациональных чисел все точки предельные, так как любое рациональное число можно как угодно приблизить другим рациональным числом. Аналогичное утверждение имеет место для действительных чисел. Множество может иметь только одну предельную точку. Рассмотрим, например, числа  $n^{-1}$ , где  $n$  — целые. В любой конечной окрестности 0 содержится бесконечно много элементов этого множества и, следовательно, 0 — предельная точка этого множества. Однако для любого другого числа, рационального или нет, можно указать такую окрест-

---

\*) Отметим, что выражения «процесс можно неограниченно продолжать» и «и так далее» подразумевают применение *математической индукции*. Подобные рассуждения из-за недостатка места мы будем редко приводить полностью. Однако читатель может самостоятельно провести их в некоторых случаях для практики.

ность, которая не содержит элементов множества, отличных от самого числа (если последнее является элементом множества).

Предельная точка множества не обязана сама быть элементом множества. Мы можем, например, рассматривая последовательные приближения  $\sqrt{2}$  десятичными дробями, построить множество рациональных чисел, для которого  $\sqrt{2}$  будет предельной точкой, а сам  $\sqrt{2}$  — не рациональное число.

Если все предельные точки множества принадлежат ему, то оно называется *замкнутым*. Интервал  $a \leq x \leq b$ , определенный в 1.031, является замкнутым множеством и называется замкнутым интервалом, или *отрезком*. Соответствующим открытым интервалом является  $a < x < b$ . Мы вернемся к этому в 1.061.

**1.034.** Если множество содержит бесконечно много элементов внутри конечного отрезка  $a \leq x \leq b$ , то оно имеет по меньшей мере одну предельную точку  $x$ , такую, что  $a \leq x \leq b$ . Действительно, если мы поделим отрезок пополам, то по крайней мере в одной из половин будет бесконечное число точек множества. Разделим эту половину пополам. Вновь одна из половин содержит бесконечное число элементов, и мы видим, что, продолжая этот процесс, можем отыскать сколь угодно короткий интервал, содержащий бесконечно много точек множества. Но этот процесс соответствует заданию действительного числа с помощью вложенной последовательности интервалов и, значит, определяет действительное число, в любой окрестности которого содержится бесконечно много точек множества. Следовательно, это действительное число является искомой предельной точкой множества. Это утверждение известно как теорема Больцано — Вейерштрасса.

**1.035.** Бесконечное множество называется *счетным*, если его элементы можно так сопоставить с натуральными числами, чтобы каждому элементу множества соответствовало одно и только одно натуральное число и наоборот. Например, квадраты  $1^2, 2^2, \dots, n^2, \dots$  образуют счетное множество, так как каждому  $n$  соответствует  $n^2$  и каждому  $n^2$  соответствует  $n$ . Рациональные дроби между нулем и единицей тоже образуют счетное множество. Их можно упорядочить так:  $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \dots$ , и дробь, стоящую на  $n$ -м месте, сопоставить с  $n$ . Все положительные рациональные числа образуют другое счетное множество. Их можно упорядочить так:  $\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{1}, \frac{2}{4}, \frac{4}{2}, \frac{3}{5}, \frac{5}{3}, \dots$ . Здесь числа

собраны в группы с постоянной суммой числителя и знаменателя, и в каждой группе сумма на единицу больше, чем в предыдущей. А внутри групп числа расположены по возрастанию числителя. В двух последних примерах для установления соответствия с натуральными числами необходимо полное изменение естественного порядка.

Не все бесконечные множества счетны. Наиболее важные примеры несчетных множеств — это множество всех действительных чисел и множество всех действительных чисел из данного конечного интервала. Кантор доказал, что как бы мы ни пытались поставить их элементы во взаимно-однозначное соответствие с натуральными числами, всегда некоторые элементы окажутся пропущенными.

**1.036. Необходимость. Достаточность.** Пусть два предложения, обозначенные I и II, соотносятся между собой так, что если I истинно, то II истинно. Мы говорим тогда, что I — *достаточное* условие для II, а II — *необходимое* условие для I, т. е. I не может быть истинным, если II неистинно. Если II истинно тогда и только тогда, когда I истинно, то I — *необходимое и достаточное* условие для II и наоборот. В этом случае мы можем также сказать, что I и II *эквивалентны*.

Вообще говоря, может существовать несколько необходимых и достаточных условий для истинности данного предложения. Например, необходимым и достаточным условием того, что действительная величина  $x$  равна нулю, является условие  $|x| < \epsilon$  для любого положительного  $\epsilon$ , но, кроме этого, такими условиями будут равенства  $x^2 = 0$ ,  $x^3 = 0$ . Необходимым и достаточным условием  $ax^2 - 2bx + c > 0$  для всех действительных  $x$  будет  $a > 0$ ,  $ac - b^2 > 0$ , а также и  $c > 0$ ,  $ac - b^2 > 0$ .

Необходимое и достаточное условие может содержать излишнюю информацию. Например, если  $ax^2 - 2bx + c > 0$  для всех  $x$ , то  $a > 0$ ,  $c > 0$ ,  $ac - b^2 > 0$  и наоборот. Здесь  $a > 0$ ,  $c > 0$ ,  $ac > b^2$  — необходимое и достаточное условие. Но если  $ac > b^2$ , то из  $a > 0$  следует, что  $c > 0$ , а из  $c > 0$  следует, что  $a > 0$ , так что одно из требований  $a > 0$ ,  $c > 0$  избыточно в том смысле, что оно вытекает из другой данной информации. С другой стороны, требования  $a > 0$ ,  $c > 0$  или  $ac > b^2$  сами по себе недостаточны для того, чтобы  $ax^2 - 2bx + c$  было больше 0 для всех  $x$ ; ни одно из этих условий не является достаточным. Множество необходимых и достаточных условий истинности данного предложения называется *минимальным*, если после отбрасывания любой части этих условий оставшиеся условия не являются достаточными.

**1.04. Последовательности \*).** Рассматривая свойства множества, мы не считали, что его члены расположены в определенном порядке. В доказательстве **1.034**, например, точки, лежащие в любом промежутке, определяются указанием самого множества. (Представим себе, что мы положили несколько мячей в коробку; тогда вопрос, какие там лежат мячи, не имеет отношения к их перемещениям из-за встряхивания коробки.)

Когда мы начинаем изучать свойства, существенно связанные с конкретным порядком, мы имеем дело с *последовательностями*. Числа 1, 2, 3, ..., расположенные в порядке возрастания, образуют последовательность. Если они перегруппированы, но так, что мы знаем, как найти любое из них, то они образуют то же множество, но другую последовательность. Если мы обозначим  $n$ -й член данной последовательности через  $s_n$ , то свойство  $s_{n+1} - s_n = 1$  будет выполняться для первоначально заданного порядка и не будет выполняться ни для какого другого порядка. Вообще, если  $s_n$  вполне определяется заданием  $n$ , то  $s_n$  может рассматриваться как функция натурального аргумента  $n$ , а значения  $s_1, s_2, \dots, s_n, \dots$  для последовательных значений  $n$  образуют последовательность. (Те, кто знает кое-что о рядах, часто полагают на первых порах, что назначение последовательности быть просуммированной, но это не так.) И

$$1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots, \quad (1)$$

и

$$1, 2, 1, 2, 1, 2, \dots \quad (2)$$

— последовательности. Элементы первой — это элементы бесконечного множества, упорядоченные определенным образом. Элементы второй — это элементы конечного множества, повторяемые вновь и вновь.

Последовательность с общим членом  $s_n$  можно обозначить через  $\{s_n\}$ .

**1.041. Ограниченные, неограниченные, сходящиеся и осциллирующие последовательности.** Пусть  $M$  — произвольное положительное число. Может случиться, что какое бы  $M$  ни взять, найдется хоть один член последовательности  $s_n$ , такой, что  $|s_n| > M$ . Такая последовательность называется *неограниченной*. Очевидный пример:  $s_n = n$ ; в самом деле, в качестве  $n$  мы можем взять любое целое число, большее  $M$ . Справедливо

---

\*) Более подробное рассмотрение последовательностей можно найти в [2, 3].

следующее утверждение, аналогичное теореме о предельной точке: неограниченная последовательность должна содержать бесконечно много членов, модули которых  $|s_n|$  больше любого наперед заданного  $M$ .

Если можно выбрать такое  $M$ , что все  $s_n$  по модулю меньше  $M$ , то последовательность называется *ограниченной*. Обе последовательности, приведенные в конце 1.04, ограничены; условие  $|s_n| < M$  выполняется для них обеих при  $M = 3$ .

Пусть существует такое число  $s$ , что для любого положительного числа  $\varepsilon$  можно выбрать  $m$  так, что для всех  $n > m$

$$|s_n - s| < \varepsilon. \quad (1)$$

Тогда говорят, что последовательность *сходится* и имеет предел  $s$ . Записывается это как [4]

$$s_n \rightarrow s \quad (n \rightarrow \infty), \quad (2)$$

или

$$\lim_{n \rightarrow \infty} s_n = s.$$

Стрелка читается „стремится к“. Можно просто писать

$$\lim s_n = s, \quad (3)$$

если при этом не возникает разночтений.

В вышеприведенных примерах последовательность 1.04 (1) сходится и имеет предел 0, нужно только взять  $m > 1/\varepsilon$ ; последовательность 1.04 (2) не является сходящейся, потому что, каковы бы ни были  $s$  и  $m$ , если  $\varepsilon < 1/2$ , то найдутся члены с номерами  $n > m$ , такие, что  $|s_n - s| \geq \frac{1}{2} > \varepsilon$ .

Самое важное свойство сходящейся последовательности таково: если у нас есть правило для вычисления каждого члена, то мы можем вычислить предел с любой желаемой точностью. Некоторые методы приближения (см. гл. 9 и 17) показывают, что предел лежит в заданном промежутке, но этот промежуток не сколь угодно мал. Точность может оказаться достаточной для применений, но ее нельзя беспредельно повышать.

Последовательность, ограниченная, но не сходящаяся, называется *осциллирующей с конечным размахом* или просто *осциллирующей*. В качестве примера можно привести последовательность 1.04 (2); другой пример

$$s_n = (-1)^n + \frac{1}{n}. \quad (4)$$

Здесь, в отличие от 1.04 (2), все  $s_n$  различны.



Последовательность (4) ограничена, потому что  $|s_n| < 2$  для любого  $n$ ; но она не сходится, поскольку для больших  $n$  ее члены попеременно близки то к 1, то к  $-1$ , так что неравенство (1) не может выполняться, если  $\varepsilon < \frac{1}{2}$ .

Если для любого  $M$  существует такое  $m$ , что  $s_n > M$  для всех  $n > m$ , то мы пишем

$$s_n \rightarrow \infty. \quad (5)$$

Примеры:  $s_n = n$  и  $s_n = n^2$ .

Если для любого  $M$  существует такое  $m$ , что  $s_n < -M$  для всех  $n > m$ , то мы пишем

$$s_n \rightarrow -\infty. \quad (6)$$

Примеры:  $s_n = -n$  и  $s_n = -n^2$ .

Ниже представлены другие типы неограниченных последовательностей:

$$s_n = (-1)^n n, \quad s_n = n \cos \frac{1}{2} \pi n, \quad s_n = n (1 - \cos \pi n).$$

Про эти последовательности нельзя сказать, что они стремятся к какому-нибудь пределу, в частности они не стремятся и к бесконечности, и их называют иногда *бесконечно осциллирующими*. Неограниченные последовательности можно называть *расходящимися*; однако разные авторы вкладывают в этот термин различный смысл: некоторые (например, Бромвич и Харди) исключают из числа расходящихся бесконечно осциллирующие последовательности, некоторые (например, Кноп) включают в их число последовательности, осциллирующие с конечным размахом.

Полезно классифицировать последовательности в зависимости от того, обладают или не обладают они следующими свойствами:

1) для любого  $M$  и для любого положительного  $M$  существует такое  $n > m$ , что  $s_n > M$ ;

2) для любого  $M$  и для любого положительного  $M$  существует такое  $n > m$ , что  $s_n < -M$ .

Последовательности, не обладающие ни одним из этих свойств, ограничены. Если последовательности обладают первым свойством, но не обладают вторым, то они *ограничены снизу* и *не ограничены сверху*. Аналогично описываются два оставшихся случая.

Отметим, что бесконечности как таковой никакого определенного значения не приписывается. Мы лишь придаем смысл всем выражениям, которые содержат слово бесконечность или символ  $\infty$ . Выражение  $s_n \rightarrow \infty$  есть краткая запись свойства

$\{s_n\}$ , сформулированного в соответствующем определении. Она не подразумевает существования какой-либо действительной величины, обозначаемой символом  $\infty$ .

Бесконечность исключена из правил алгебры не потому, что есть какое-нибудь противоречие в понятии бесконечных чисел, а потому, что эти числа подчиняются другим правилам. Действительно, понятие бесконечного множества проникает в большую часть нашей теории: в каждом интервале  $x$  принимает бесконечно много значений. Непротиворечивая алгебра для положительных бесконечных чисел была построена Кантором и позднее неоднократно обобщалась другими авторами, но эта алгебра отличается от обычной. Если  $a$  и  $b$  — положительные бесконечные числа, мы можем определить  $a + b$  и  $ab$  однозначно. Однако  $a + b$  не обязательно больше  $a$ , в действительности  $a + b$  обычно равняется  $a$  или  $b$ . Невозможно определить однозначно  $a - b$  и  $a/b$ . Следовательно, алгебра, включающая и конечные, и бесконечные числа, должна все же различать их в своих правилах.

**1.042.** Если бесконечное множество имеет предельную точку  $s$ , то мы можем образовать последовательность из его элементов, имеющую пределом  $s$ . Если у множества больше одной предельной точки, то можно образовать последовательности, сходящиеся к любой из них.

Мы показали (начало 1.033), что на расстоянии от предельной точки, не превышающем заданное, существует бесконечно много элементов множества. Если выписывать элементы множества в таком же порядке, какой указан в 1.033, то получим последовательность, обладающую нужным свойством.

**1.043.** Любая последовательность, образованная из различных элементов ограниченного бесконечного множества, имеющего единственную предельную точку  $s$ , стремится к пределу  $s$ . Ясно, что при построении последовательности из элементов множества мы имеем выбор на каждом этапе.

Следовательно, мы можем образовать бесконечно много различных последовательностей. Нам нужно показать, что все они имеют одинаковый предел. Для любого  $t$  число элементов  $s_n$  последовательности с  $n > t$  бесконечно. Но поскольку множество ограничено и имеет единственную предельную точку, то в любом интервале, не содержащем предельной точки, имеется лишь конечное число элементов множества. Следовательно, для любого  $\varepsilon$  только конечное число элементов последовательности лежит вне отрезка  $s \pm \frac{1}{2}\varepsilon$ . Пусть это  $s_{a_i}$ ,

$s_\beta, \dots, s_\mu$ . Возьмем  $m$  равным наибольшему из  $\alpha, \beta, \dots, \mu$ . Тогда для всех  $n$ , больших  $m$ , справедливо, что  $|s_n - s| \leq \frac{1}{2}\varepsilon < \varepsilon$  и, таким образом, последовательность сходится к  $s$ .

Этот результат нельзя получить, если элементы последовательности не обязаны быть различными и могут повторяться сколь угодно часто. Например, если множество состоит из чисел, обратных целым, его единственной предельной точкой будет 0. Однако, если повторения разрешены, мы можем образовать из его элементов последовательность

$$1, \frac{1}{2}, 1, \frac{1}{3}, 1, \frac{1}{4}, 1, \dots, \frac{1}{n}, 1, \dots,$$

которая является осциллирующей. Если, однако, каждый элемент может повторяться не больше  $k$  раз, где  $k$  — некоторое фиксированное число, то результат опять можно получить простым обобщением доказательства.

**1.044. Верхняя и нижняя грани.** *Множество (или последовательность), ограниченное сверху, имеет верхнюю грань; множество, ограниченное снизу, имеет нижнюю грань.* При этом величина  $M$  называется *верхней гранью* множества, если ни один элемент множества не превосходит  $M$ , и в то же время для любого положительного числа  $\varepsilon$ , как бы мало оно ни было, существует элемент, превосходящий  $M - \varepsilon$ . *Нижней гранью* называется такая величина  $m$ , что нет элементов, меньших  $m$ , но всегда найдется элемент, меньший  $m + \varepsilon$ .

Воспользуемся методом сечений Дедекинда. Класс чисел  $a$ , ограниченных сверху хоть одним элементом множества, не пуст: можно взять в качестве такого числа любое  $a$ , меньшее, чем некоторый известный элемент множества.

Поскольку множество ограничено сверху, то существуют числа  $b$ , такие, что никакой элемент множества не превосходит  $b$ . Каждое  $b$  больше любого  $a$ , и всякое число есть либо  $a$ , либо  $b$ . Следовательно, числа  $a$  образуют класс  $L$ , а числа  $b$  — класс  $R$  и определяют сечение, которое мы обозначим через  $M$ . Сечение  $M$  принадлежит классу  $R$ . Если бы  $M$  было элементом класса  $L$ , то оно было бы меньше некоторого элемента множества, скажем  $K$ ; но тогда между  $M$  и  $K$  не оказалось бы чисел  $b$  и, следовательно,  $M$  не являлось бы числом, соответствующим сечению. Итак, ни один элемент множества не превосходит  $M$ . Так же проверяется, что  $M - \varepsilon$  принадлежит классу  $L$ , а потому существует элемент множества, превосходящий  $M - \varepsilon$ . Соответствующий результат для нижней грани доказывается аналогично.

Доказательство не предполагает, что множество бесконечно; однако для конечного множества верхней гранью является просто его наибольший элемент. Для бесконечного множества все элементы могут оказаться меньше верхней грани; для множества 1.04 (1) верхняя грань равна 1 и совпадает с первым членом, а нижняя грань равна 0, хотя 0 не является элементом множества.

То, что мы называем верхней гранью, часто называют *точной верхней гранью*, а любое число, такое, что нет элементов множества, превосходящих его, называют тогда просто *верхней гранью*.

Заметим, что если  $s_n < t_n$  для всех  $n$  и  $s_n \rightarrow s$ ,  $t_n \rightarrow t$ , то  $s \leq t$  (и вовсе не обязательно  $s < t$ ). Рассмотрим, например, последовательности  $s_n = 1 - 2^{-n}$ ,  $t_n = 1 - 3^{-n}$ ; здесь  $s = t$ . Мы можем рассматривать  $(s_n, t_n)$  как интервал, длина которого стремится к нулю, но эти интервалы не образуют вложенной последовательности, потому что каждый не является частью предыдущего, и на самом деле каждый интервал лежит целиком по одну сторону от предела.

**1.0441.** Если  $s_n \geq s_{n-1}$  для всех  $n$  и последовательность ограничена, то она сходится. Пусть  $s$  — верхняя грань  $s_n$ . Тогда  $s_n \leq s$  для всех  $n$ . Кроме того, для любого  $\varepsilon$  существует такое  $m$ , что  $s_m > s - \varepsilon$ ; но тогда для всех  $n > m$

$$s \geq s_n \geq s_m > s - \varepsilon,$$

а потому последовательность сходится к пределу  $s$ .

**1.045 Основной принцип сходимости.** Для того чтобы последовательность  $\{s_n\}$  была сходящейся, необходимо и достаточно выполнение следующего условия: для любого положительного числа  $\varepsilon$  существует такое  $m$ , что для всех  $n \geq m$

$$|s_n - s_m| < \varepsilon. \quad (1)$$

Покажем сначала, что это условие является необходимым.

Предположим, что  $s_n \rightarrow s$ . Мы должны показать, что существует  $m$ , такое, что (1) справедливо. Для любого положительного  $\omega$  мы можем взять  $m$  так, что  $|s_n - s| < \omega$  для всех  $n \geq m$ . Тогда  $|s_n - s_m| < 2\omega$  для всех  $n \geq m$ . Возьмем  $\omega = \frac{1}{2}\varepsilon$ , тогда (1) выполняется.

Для доказательства достаточности условия (1) заметим прежде всего, что последовательность ограничена, так как для любого заданного положительного  $\omega$  существует такое  $m$ , что  $|s_n - s_m| < \omega$  для всех  $n \geq m$ , а  $s_1, s_2, \dots, s_m$  все конечны.

Определим  $a_n$  и  $b_n$  как нижнюю и верхнюю грани соответственно для  $s_p$  при  $p \geq n$ . Ясно, что

$$a_n \leq a_{n+1} \leq b_{n+1} \leq b_n.$$

Кроме того,  $|s_p - s_q| < 2\omega$  для  $p$  и  $q$ , больших  $m$ , и мы имеем  $b_n - a_n \leq 2\omega$  при  $n \geq m$ , поскольку  $b_n - a_n$  есть верхняя грань  $s_p - s_q$  при  $p, q \geq n$ . Так как  $\omega$  произвольно мало,  $b_n - a_n \rightarrow 0$ . Следовательно, интервалы  $(a_n, b_n)$  образуют вложенную последовательность, определяющую некоторое действительное число. Назовем его  $s$ . Поскольку

$$\left. \begin{array}{l} a_n \leq s_n \leq b_n \\ a_n \leq s \leq b_n \end{array} \right\} \text{ для всех } n,$$

имеем

$$|s - s_n| \leq 2\omega \text{ для } n \geq m,$$

т. е.  $s_n \rightarrow s$  при  $n \rightarrow \infty$ .

В дальнейшем часто будет встречаться использованный нами прием. Если величину, о которой мы хотим сказать, что она меньше  $\epsilon$ , можно представить в виде суммы нескольких членов, то вводятся дополнительные произвольно малые величины, обычно обозначаемые через  $\omega$ ,  $\delta$  или  $\eta$ , которые потом определяются как доли  $\epsilon$ .

**1.05. Ряды.** Пусть  $u_n$  есть  $n$ -й член бесконечной последовательности. Рассмотрим последовательность, образованную суммами

$$\begin{aligned} s_1 &= u_1, & s_2 &= u_1 + u_2, & s_3 &= u_1 + u_2 + u_3, \dots, \\ s_n &= u_1 + u_2 + \dots + u_n, \dots \end{aligned}$$

Если эта последовательность сумм сходится, то мы говорим, что ряд

$$\sum u_n = u_1 + \dots + u_n + \dots$$

сходится ( $n$  теперь сколь угодно велико); предел последовательности  $s_n$  мы называем *суммой ряда* \*).

Если  $\{s_n\}$  является не сходящейся, а осциллирующей с конечным размахом последовательностью, то мы говорим, что ряд осциллирует с конечным размахом.

---

\*) Как и в случае последовательностей, употребляются различные определения *расходимости*; некоторые авторы ограничиваются применением этого термина к случаю, когда  $s_n \rightarrow \infty$  или  $s_n \rightarrow -\infty$ , другие называют расходящимися все ряды, которые не сходятся.

Каждой теореме о последовательностях соответствует теорема о рядах, так как если  $\{s_n\}$  — некоторая последовательность, то, полагая  $u_1 = s_1$ ,  $u_n = s_n - s_{n-1}$  при  $n > 1$ , имеем  $\sum_1^n u_r = s_n$ .

Геометрическая прогрессия — это ряд

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \dots$$

Здесь если  $x \neq 1$ , то

$$s_n = \frac{1 - x^{n+1}}{1 - x},$$

и  $|x^{n+1}|$  становится как угодно большим с возрастанием  $n$  при  $|x| > 1$ . Если  $|x| < 1$ , то  $x^{n+1}$  стремится к нулю \*).

Таким образом, ряд сходится, если  $|x| < 1$ , и не сходится, если  $|x| > 1$ . Если  $x = 1$ , то сумма  $n$  членов равна  $n$  и ряд не сходится. Если  $x = -1$ , то сумма любого четного числа членов равна 1, а сумма любого нечетного числа членов равна 0. Поэтому ряд осциллирует с конечным размахом. Таким образом, чтобы наш ряд сходился, необходимо и достаточно выполнение условия  $|x| < 1$ .

$\zeta$ -ряд Римана — это ряд вида

$$\sum_{n=0}^{\infty} n^{-x} = 1 + \frac{1}{2^x} + \frac{1}{3^x} + \dots$$

Возьмем сначала  $x > 1$ . Мы можем объединить члены ряда в группы

$$s_n = 1 + \left(\frac{1}{2^x} + \frac{1}{3^x}\right) + \left(\frac{1}{4^x} + \frac{1}{5^x} + \frac{1}{6^x} + \frac{1}{7^x}\right) + \dots + \left(\dots + \frac{1}{n^x}\right).$$

Суммы в скобках, идущих вслед за первым членом, соответственно меньше, чем

$$\frac{2}{2^x}, \quad \frac{4}{4^x}, \dots = \frac{1}{2^{x-1}}, \quad \frac{1}{(2^{x-1})^2}, \dots$$

Далее, если  $m = 2^{r-1}$  и  $n \geq m$ , то

$$|s_n - s_m| < \frac{2^{-r(x-1)}}{1 - 2^{-(x-1)}};$$

---

\*) Точные доказательства этих совершенно очевидных утверждений можно найти в [3, стр. 134, 135].

правая часть может быть сделана  $< \varepsilon$ , если взять  $r$  достаточно большим. Следовательно,  $\xi$ -ряд сходится, если  $x > 1$ .

Если  $x = 1$ , напомним

$$s_n = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots$$

Все суммы в скобках превосходят  $1/2$ . Следовательно, ряд не сходится:  $s_n \rightarrow \infty$ . Если  $0 < x < 1$ , то все члены ряда после первого возрастают; следовательно, опять  $s_n \rightarrow \infty$ .

Следующий ряд связан с  $\ln 2$ :

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

Здесь

$$s_n - s_m = \pm \left[ \left( \frac{1}{m+1} - \frac{1}{m+2} \right) + \left( \frac{1}{m+3} - \frac{1}{m+4} \right) + \dots \right],$$

и сумма в скобках положительна независимо от четности  $n - m$ . Но в то же время

$$s_n - s_m = \pm \left[ \frac{1}{m+1} - \left( \frac{1}{m+2} - \frac{1}{m+3} \right) - \left( \frac{1}{m+4} - \frac{1}{m+5} \right) - \dots \right],$$

и каждое выражение в круглых скобках положительно. Следовательно,

$$0 < |s_n - s_m| < \frac{1}{m+1},$$

а это меньше  $\varepsilon$  для всех  $n > m$ , если только  $(m+1) > 1/\varepsilon$ . Отсюда следует, что ряд сходится.

Можно сразу же применить это к доказательству следующего факта: если  $u_n > 0$ ,  $u_n > u_{n+1}$  для всех  $n$  и  $u_n \rightarrow 0$ , то ряд

$$u_1 - u_2 + u_3 - u_4 + \dots$$

сходится.

**1.051. Абсолютная сходимость.** Если ряд  $\sum |u_n|$  сходится, то  $\sum u_n$  сходится, потому что сумма любой части ряда от  $u_m$  до  $u_n$  по модулю не может быть больше, чем сумма соответствующих членов от  $|u_m|$  до  $|u_n|$ . В этом случае говорят, что  $\sum u_n$  *абсолютно сходится*; если  $\sum u_n$  сходится, а  $\sum |u_n|$  нет, то говорят, что ряд  $\sum u_n$  *сходится условно*. (Иногда используют название *полусходимость*; однако этой приставкой злоупотребляют: то же название используют и для асимптотических рядов, которые лучше вообще не рассматривать в качестве бесконечных рядов. Вот почему следует избегать употребления этого названия.)

Мы видели, что если взять все члены ряда для  $\ln 2$  с положительными знаками, то полученный ряд не сходится. Следовательно, ряд для  $\ln 2$  условно сходящийся. Геометрическая прогрессия, если сходится вообще, то сходится абсолютно.

**1.052. Перестановки ряда.** Сумма абсолютно сходящегося ряда не меняется, если брать члены ряда в любом порядке. Пусть  $\sum u_n$  абсолютно сходится и сумма его равна  $s$ ;  $\sum v_n$  — тот же ряд, но члены его расположены в другом порядке. Понятно, что каждый член любого ряда встречается в другом, но, вообще говоря, не на том же месте. Возьмем произвольное положительное число  $\omega$  и выберем  $m$  так, чтобы  $\sum_{n=m+1}^{\infty} |u_n| < \omega$ ; тогда сумма модулей для любого набора членов, взятых из первого ряда после  $m$ -го, меньше  $\omega$ . Возьмем  $m'$  так, чтобы все члены  $u_n$ , предшествующие  $u_m$ , встречались во втором ряду под номерами  $n'$ , меньшими  $m'$ . Напишем

$$s_m = u_1 + \dots + u_m, \quad s'_{m'} = v_1 + \dots + v_{m'}.$$

Тогда  $s'_{m'} - s_m$  есть сумма членов первого ряда, идущих после  $m$ -го, и ее модуль меньше  $\omega$ . Точно так же, если взять  $n' \geq m'$ , то  $s'_{n'} - s'_{m'}$  есть сумма другого множества членов первого ряда, идущих после  $m$ -го, и поэтому ее модуль тоже меньше  $\omega$ . Следовательно, второй ряд сходится. Пусть его сумма равна  $s'$ . Тогда  $|s - s'| = |(s - s_m) - (s' - s'_{m'}) - (s'_{m'} - s_m)| < 3\omega$  и может быть сделан меньше произвольного  $\varepsilon$ , если взять  $\omega = \frac{1}{3}\varepsilon$ . Следовательно, оба ряда имеют одну и ту же сумму.

Теорема не верна для условно сходящихся рядов. Можно показать, что если ряд  $\sum u_n$  сходится условно, то, переставляя его члены, мы можем сделать его сумму, какой захотим. Сумма условно сходящегося ряда имеет точное значение, если порядок его членов задан и только в этом случае. Обычно они сходятся слишком медленно, чтобы их можно было использовать для вычислений, однако они могут найти применение в теоретической работе.

Признаки сходимости, основанные на использовании метода „сравнения рядов“, настолько тесно связаны с признаками равномерной сходимости, что мы отложим их рассмотрение до обсуждения этого последнего свойства (1.115, 1.117).

**1.053. Двойные ряды.** Аналогичные замечания применимы к *двойным рядам*, общим членом которых является  $u_{m,n}$ . На этот раз условие сходимости состоит в том, что мы можем



выбрать  $m, n$  так, что для всех  $p$ , больших  $m$ , и для всех  $q$ , больших  $n$ , суммы  $\sum_{r=1}^p \sum_{s=1}^q u_{r,s}$ ,  $\sum_{r=1}^m \sum_{s=1}^n u_{r,s}$  отличаются друг от друга на величину, по модулю меньшую  $\varepsilon$ . Аналогично можно определить абсолютную и условную сходимость, и опять верно, что абсолютно сходящиеся двойные ряды имеют одну и ту же сумму, как бы ни переставлять их члены. Доказательства отличаются от соответствующих доказательств для простых рядов только тем, что более громоздки.

**1.06. Пределы функций. Непрерывность.** В самом общем смысле, когда мы говорим, что  $f(x)$  является функцией  $x$  в некоторой области значений  $x$ , то имеем в виду, что для каждого значения  $x$  из этой области имеется одно (или более) значение  $f(x)$ . Например, можно говорить о функции от  $x$ , которая равна 1, если  $x$  рационально, а при  $x$  иррациональном равна 0.

Физик справедливо рассматривает такую функцию как патологическую, его интересует гораздо более узкий класс функций: грубо говоря, такие функции, которые можно представить с помощью графика\*). Обычным (хотя и не обязательным) будет также требование *однозначности* функции. Так, для окружности

$$x^2 + y^2 = a^2$$

мы имеем

$$y = \pm \sqrt{a^2 - x^2}$$

и  $y$  является функцией  $x$ , но получим ее значения для всей окружности, только если возьмем оба знака перед корнем. Однозначная функция  $x$  в некоторой области — это функция, которая принимает в точности одно значение для каждого значения  $x$ . Сначала мы будем рассматривать только однозначные функции.

Идея, лежащая в основе понятия предела функции, похожа на идею, связанную со сходимостью последовательности; дело в том, что члены последовательности  $\{s_n\}$  являются значениями функции целого положительного переменного  $n$ , которое может

---

\*) Функция, приведенная в качестве примера, часто используется для предостережения от поспешных выводов. Можно сразу же использовать ее с этой целью. Мы могли бы попробовать определить патологические функции как такие, которые сами не являются непрерывными и не могут быть получены из непрерывных предельным переходом. Однако, что может быть проще функции  $\cos^{2m} t! \pi x$ , которая стремится к нашей функции, если сначала  $n$ , а затем  $m$  устремить к бесконечности?

принимать произвольно большие значения. Особенность здесь состоит в том, что для функции  $f(x)$  переменное  $x$  не обязательно должно быть целым: оно может принимать любые значения из некоторого интервала или даже любые, как угодно большие значения.

Когда значения  $x$  заполняют интервал, можно определить *предел*  $f(\xi)$  при  $\xi \rightarrow x$  следующим образом: пусть существует такое число  $c$ , что для любого заданного положительного  $\varepsilon$  найдется такое положительное  $\delta$ , что как только  $0 < |\xi - x| < \delta$ , так  $|\dot{f}(\xi) - c| < \varepsilon$ ; тогда мы говорим, что  $c$  является пределом  $\dot{f}(\xi)$  при  $\xi \rightarrow x$ . (В дальнейшем мы можем ограничить допустимые значения  $\xi$ , например говорить о пределе  $\dot{f}(\xi)$  при  $\xi - x \rightarrow 0$  по положительным или отрицательным значениям.) Если, кроме того,  $c = f(x)$ , мы говорим, что  $\dot{f}(\xi)$  непрерывна при  $\xi = x$ . Следовательно, определение непрерывности можно сформулировать следующим образом: *если для любого положительного  $\varepsilon$  мы можем выбрать такое положительное  $\delta$ , что как только  $|h| < \delta$ , так*

$$|\dot{f}(a+h) - \dot{f}(a)| < \varepsilon,$$

*тогда  $\dot{f}(a)$  называется непрерывной при  $x = a$ .* Если это условие выполняется и если взять произвольную последовательность  $\{h_n\}$ , стремящуюся к 0, то для любого  $\delta$  найдется такое  $m$ , что  $|h_n| < \delta$  для всех  $n \geq m$  и, следовательно,  $|\dot{f}(a+h_n) - \dot{f}(a)| < \varepsilon$ . Значит, для всех таких последовательностей  $\dot{f}(a+h_n) \rightarrow \dot{f}(a)$ .

Большинство функций, встречающихся практически, непрерывны или самое большее имеют конечное число точек разрыва. Самый типичный вид разрыва таков:  $\dot{f}(x+h)$  для некоторого значения  $x$  имеет определенный предел, когда  $h \rightarrow 0$ , пробегая любое множество положительных значений, и имеет другой предел, если  $h \rightarrow 0$  по любому множеству отрицательных значений. Этот случай называется *обыкновенным*, или *простым*, *разрывом*. Например, если

$$\dot{f}(x) = 0 \quad (x < 0), \quad \dot{f}(x) = 1 \quad (x > 0),$$

то предел  $\dot{f}(h)$  при  $h \rightarrow 0$  по любому множеству положительных значений равен 1, а при  $h \rightarrow 0$  по любому множеству отрицательных значений равен 0. Эта функция очень типична для физических приложений, поскольку, например, она представляет силу, которая начинает действовать на систему в некоторый определенный момент, а затем становится постоянной. Ее обычно называют *единичной функцией Хевисайда*. Почтовые расходы как функция веса письма имеют простые разрывы.

Если иметь в виду экспериментальные приложения, то значение функции при  $x = 0$  обычно не нуждается в уточнении.

Пусть  $x$  — координата, описывающая положение какого-либо объекта. Чтобы объект можно было увидеть, он должен иметь некоторую протяженность. Мы не в состоянии наблюдать величину, связанную с точным значением  $x$ , а можем найти лишь ее среднее значение в некоторой области. Аналогично если  $x$  — время, то мы не можем наблюдать величину в некоторой отдельный момент, а лишь в течение ненулевого интервала. Обычная тенденция чистой математики настаивать, что функцию следует точно определять для всех значений независимого переменного; в физике же обычно достаточно, чтобы был определен ее интеграл. Так как значение функции в единственной точке в предположении, что оно конечно, не влияет на интеграл, то обычно нет смысла говорить о нем в физических приложениях, и если точное значение указывается, то это делается просто ради удобства.

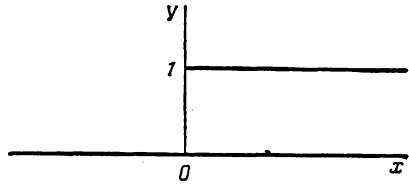


Рис. 1. Функция Хевисайда имеет в нуле простой разрыв.

Часто употребляют такие обозначения: при  $h > 0$

$$\lim_{h \rightarrow 0} f(x + h) = f(x +), \quad \lim_{h \rightarrow 0} f(x - h) = f(x -),$$

следовательно, в случае, который мы рассматривали,

$$f(0 +) \neq f(0 -).$$

Может случиться, что  $f(a +) = f(a -)$ , но не равно  $f(a)$ . Про такую функцию говорят, что она имеет *устраняемый разрыв*, но так как  $f(a)$  не влияет на интеграл, то такие разрывы не имеют большого значения. Это, разумеется, невозможно проиллюстрировать графически.

Предел вообще не существует, если функция неограничена в окрестности значения  $x$ , как, например,  $f(x) = 1/x$  около  $x = 0$ . Для любой последовательности значений  $x$ , стремящейся к нулю,  $f(x)$  неограничена.

Опять-таки, если  $f(x) = \sin(1/x)$  и  $x$  стремится к нулю по значениям  $1/n\pi$ , где  $n$  — целое, то предел равен 0. Если же  $x$  стремится к нулю по значениям  $1/(n + \frac{1}{2})\pi$ , то  $f(x)$  стремится к  $+1$ , если ограничиться четными  $n$ , и к  $-1$ , если ограничиться нечетными  $n$ . При сложном способе задания функции особенности такого рода труднее всего обнаружить; кроме того, этот тип особенностей легко забывается.

Поведение  $f(x)$  при  $x \rightarrow \infty$  даже еще сильнее похоже на поведение последовательности, поскольку, вообще говоря,  $f(\infty)$  не определено непосредственно, и нас всецело интересует именно предел, если он существует. Отметим только определение и основной критерий. Пусть существует такое  $c$ , что для любого  $\varepsilon > 0$  найдется  $X$ , такое, что для всех  $x \geq X$  имеем  $|f(x) - c| < \varepsilon$ , тогда говорят, что  $f(x)$  стремится к  $c$  при  $x \rightarrow \infty$ .

По аналогии с основным принципом сходимости рядов можно показать, что необходимым и достаточным условием стремления  $f(x)$  к некоторому пределу при  $x \rightarrow \infty$  является следующее: для любого положительного  $\varepsilon$  существует такое  $X$ , что для всех  $x \geq X$  выполняется соотношение  $|f(x) - f(X)| < \varepsilon$ .

**1.061. Непрерывность на интервале.** Говорят, что  $f(x)$  непрерывна на интервале, если она непрерывна в каждой точке этого интервала. Функция  $f(x)$  непрерывна на открытом интервале  $a < x < b$ , если она непрерывна для каждого значения  $x$ , такого, что  $a < x < b$ . Она непрерывна на замкнутом интервале  $a \leq x \leq b$ , если выполняется указанное условие, и, кроме того,

$$f(a+) = f(a), \quad f(b-) = f(b).$$

Отметим, что утверждение о непрерывности  $f(x)$  при  $x = c$  включает условие, что  $f(c)$  конечно; в противном случае выражение  $f(c+h) - f(c)$  вообще не имело бы смысла. Аналогично если  $f(x)$  непрерывна на интервале, то она конечна во всех точках этого интервала. Мы докажем, что при выполнении последнего условия функция ограничена на интервале, если он замкнут, но не обязательно ограничена, если он открыт.

Мы обозначаем любой интервал с концами в точках  $a, b$  через  $(a, b)$ . В случае необходимости мы будем точно указывать, какой именно интервал имеется в виду:  $a < x < b$ ,  $a \leq x \leq b$ ,  $a < x \leq b$  или  $a \leq x < b$  \*).

Заметим, что каждая точка  $x$  открытого интервала является внутренней точкой, т. е. существуют такие точки интервала  $y, z$ , что  $a < y < x < z < b$ . Для замкнутого интервала это неверно, поскольку  $x$  может быть равным  $a$  или  $b$ . Но если  $a$  и  $b$  конечны, то они являются предельными точками множества  $a < x < b$ .

Другой способ выразить различие между замкнутым и открытым интервалами — это сказать, что все предельные точки множеств, расположенных в замкнутом интервале, принадлежат этому интервалу, в то время как предельные точки под-

---

\*) Для открытых и замкнутых интервалов используют специальные обозначения; обычно открытый интервал обозначают через  $(a, b)$ , а замкнутый интервал через  $[a, b]$ .

множества открытого интервала могут ему и не принадлежать, поскольку для некоторых подмножеств они совпадают с концевыми точками. Когда мы говорим, что  $x$  лежит *внутри* интервала (или в последних главах внутри области), мы подразумеваем при этом, что  $x$  является внутренней точкой; если же просто сказать, что  $x$  — точка замкнутого интервала или области, то она может быть концевой или граничной точкой. Функции, непрерывные всюду, кроме конечного числа точек с простыми разрывами, называются *кусочно непрерывными*.

Если функция дифференцируема, то она непрерывна; обратное неверно, что можно увидеть на примере функции  $\sqrt{x}$  на интервале  $0 \leq x \leq 1$ . Она непрерывна на интервале, включая концевые точки, но не дифференцируема в точке  $x = 0$  \*).

Фактически построены примеры функций, которые непрерывны всюду на интервале, но нигде не дифференцируемы. Как правило, мы будем интересоваться функциями, которые дифференцируемы всюду, кроме, быть может, отдельных точек, однако в физике кристаллов такие точки весьма многочисленны. Существует теорема Вейерштрасса о том, что непрерывную функцию в любой конечной области можно сколь угодно точно приблизить полиномом или суммой синусов и косинусов с подходящими коэффициентами (см. 14.08). Следовательно, хотя непрерывная функция и не обязательно дифференцируема, ее можно заменить с любой желаемой степенью точности функцией, которая дифференцируема.

**1.062. Теоремы о покрытии.** Мы видим, что требование непрерывности эквивалентно следующему утверждению: каждая точка  $x$  интервала  $(a, b)$  принадлежит некоторому интервалу  $(x - \delta, x + \eta)$  (где  $\delta$  и  $\eta$ , возможно, зависят от  $x$ ), так что 1)  $x$  является внутренней точкой этого интервала (кроме случая,  $x = a$  или  $b$ , когда она может оказаться концевой точкой); 2) интервал имеет ненулевую длину; 3) каждая точка  $\xi$  этого интервала обладает некоторым свойством, в нашем случае  $|f(\xi) - f(x)| < \epsilon$ .

Можно показать, что в подобной ситуации имеется возможность выбрать *конечное* число интервалов, так что каждый

---

\*) В подобных случаях разные авторы употребляют различные термины. Какие бы положительные и стремящиеся к нулю  $x_n$  мы ни выбрали,  $(\sqrt{x_n} - 0)/x_n$  обязательно превзойдет любое заданное положительное значение. Если  $f(x) = x \sin(1/x)$ , то  $f(x_n)/x_n$  при подходящем выборе  $x_n$  может стремиться к любому пределу между  $-1$  и  $1$ . В последнем случае говорят, что  $f'(x)$  не существует при  $x = 0$ . Если  $f(x) = \sqrt{x}$ , то многие авторы говорят, что  $f'(x)$  бесконечно при  $x = 0$ . Является вопросом определения, считать ли  $f(x)$  существующей, когда  $[f(x+h) - f(x)]/h \rightarrow \infty$ . Мы обычно будем считать, что она не существует.

интервал удовлетворяет указанным условиям и каждая точка  $(a, b)$  принадлежит по крайней мере одному из них.

Для разных целей бывает нужно сделать этот выбор несколькими различными способами, и чтобы показать возможность этого, нам понадобятся две теоремы.

**1.0621. Теорема Гейне — Бореля.** *Если каждая точка замкнутого интервала  $(a, b)$  лежит внутри некоторого интервала  $I$  из семейства  $F$ , то существует такое конечное подмножество  $F$ , что каждая точка  $(a, b)$  лежит внутри по крайней мере одного интервала из этого подмножества.* Мы говорим, что  $I$  покрывает  $(c, d)$ , если каждая точка  $(c, d)$  является внутренней точкой  $I$  (т. е. не является концевой точкой).

Может случиться, что некоторый интервал  $I$ , принадлежащий  $F$ , целиком покрывает  $(a, b)$ . Если это так, то нечего и доказывать. Если это не так, то разделим  $(a, b)$  пополам. Возможно, что найдется пара интервалов  $I_1, I_2$ , так что каждая точка  $(a, \frac{1}{2}a + \frac{1}{2}b)$  окажется внутренней для  $I_1$ , а каждая точка  $(\frac{1}{2}a + \frac{1}{2}b, b)$  — внутренней для  $I_2$ . Если хоть одна из половин не принадлежит никакому интервалу  $I$ , разделим пополам эту половину. Мы утверждаем, что после конечного числа шагов каждая из полученных частей  $(a, b)$  будет лежать по крайней мере в одном из интервалов  $I$ . Пусть это не так. Тогда последовательные деления пополам дадут последовательность вложенных интервалов, длина каждого из них вдвое меньше длины предыдущего и ни один из них не содержится ни в каком  $I$ . Последовательность вложенных интервалов определяет число  $x_0$ , общее всем ее членам. Но по предположению  $x_0$  покрыто некоторым интервалом, скажем  $I_0$ , и, следовательно, существует такое положительное  $\delta$ , что все точки интервала  $(x_0 - \delta, x_0 + \delta)$  принадлежат  $I_0$ . Значит, все интервалы вложенной последовательности с длиной, меньшей  $\delta$ , покрываются  $I_0$ , и мы приходим к противоречию. Следовательно, процесс деления пополам за конечное число шагов разбивает  $(a, b)$  на такие части, каждая из которых принадлежит некоторому  $I$ . Взяв для каждой части содержащий ее интервал  $I$ , получим доказательство теоремы.

Небольшое изменение происходит в случае, когда концевая точка, скажем  $a$ , является концевой точкой интервала  $I_a$  семейства. Пусть длина интервала  $I_a$  равна  $\delta_a > 0$ . Тогда  $a$  лежит внутри интервала  $J_a(a - \delta_a, a + \delta_a)$  и доказательство применимо к семейству интервалов  $J$ , отличающемуся от семейства интервалов  $I$  только тем, что  $I_a$  заменен на  $J_a$ . Каждая

точка  $(a, b)$  является внутренней точкой по крайней мере для одного из интервалов  $I$ . Таким образом, теорема доказана и в том случае, когда  $a$  может быть концевой точкой интервала  $I_a$  или  $b$  — концевой точкой интервала  $I_b$ , если только  $a$  принадлежит  $I_a$  и  $b$  принадлежит  $I_b$ .

Из этой теоремы как частный случай выводится теорема Больцано — Вейерштрасса (1.034). Допустим, что  $(a, b)$  не содержит предельных точек множества. Тогда каждая точка  $(a, b)$  содержится в интервале  $I$ , содержащем не более одного элемента множества. Поскольку  $(a, b)$  можно покрыть конечным множеством таких интервалов, то в нем, вопреки предположению, содержится только конечное число элементов множества.

В проведенном нами доказательстве на каждом этапе пополам делились только те интервалы, которые еще не были покрыты  $I$ . Можно было бы, однако, делить все интервалы. Действительно, если  $I$  покрывает  $(c, d)$ , то он покрывает и обе его половины. Следовательно,  $(a, b)$  при указанных предположениях можно разделить на конечное число *равных* интервалов, каждый из которых покрыт некоторым  $I$ .

**1.0622. Модифицированная теорема Гейне — Бореля.** В теореме Гейне — Бореля интервалы  $I$  можно определить с помощью любого правила, лишь бы каждый из них имел ненулевую длину и любая точка  $(a, b)$  была внутренней точкой по крайней мере одного из них (исключение: точки  $a$  и  $b$  могут быть концевыми). Иногда, однако, накладывается следующее добавочное ограничение: каждая точка  $x$  из  $(a, b)$  определяет свой интервал  $I_x$ , для которого она является внутренней точкой. Тогда справедлива следующая теорема. *Предположим, что любая точка  $x$  из отрезка  $a \leq x \leq b$  лежит внутри некоторого интервала  $I_x(x - \delta_x, x + \eta_x)$ , где  $\delta_x > 0$ ,  $\eta_x > 0$ , за тем исключением, что  $I_a$  и  $I_b$  могут иметь вид  $a \leq x < a + \eta_a$  и  $b - \delta_b < x \leq b$ ; тогда  $(a, b)$  можно разбить на конечное число интервалов, каждый из которых является частью  $I_x$ , определенной одной из точек этого интервала.* Доказательство, как и раньше, получается с помощью последовательного деления пополам. Допустив, что теорема неверна, мы показываем, что существует вложенная последовательность интервалов, сходящихся к некоторому  $x_0$ , и таких, что ни один из них не является частью  $I_x$  ни для какого  $x$ , лежащего внутри этого интервала. Однако все эти интервалы, начиная с некоторого номера, являются частями  $I_{x_0}$ . Так как все они, кроме того, содержат  $x_0$ , то мы приходим к противоречию. В рассматриваемом случае отсюда не следует, однако, что  $(a, b)$  можно

разбить на *равные* интервалы, обладающие требуемым свойством.

Если  $(c, d)$  покрыт интервалом  $I_x$ , где  $x$  принадлежит  $(c, d)$ , то  $x$  является внутренней точкой лишь для одной половины  $(c, d)$ ; вторая половина не обязательно покрыта интервалом  $I_y$ , ибо на  $y$  наложено теперь новое ограничение; точка  $y$  должна принадлежать этой второй половине.

Имеется важное приложение к дифференцируемым функциям. Пусть  $f(x)$  дифференцируема во всех точках  $a \leq x \leq b$ , т. е. для любого  $\omega$  и для любого  $x$  из  $(a, b)$  существует такое положительное  $\delta(\omega, x)$ , что при  $|h| < \delta$

$$|f(x+h) - f(x) - hf'(x)| < \omega|h|. \quad (1)$$

Тогда  $(a, b)$  можно разбить на конечное число интервалов  $(x_r, x_{r+1})$ , так что для всех  $x$  из  $(x_r, x_{r+1})$

$$|f(x) - f(\xi_r) - (x - \xi_r)f'(\xi_r)| < \omega|x - \xi_r|, \quad (2)$$

где  $\xi_r$  сама является точкой  $(x_r, x_{r+1})$ .

Как бы ни зафиксировать  $\eta$ , формула (1) останется верной, если на все  $\delta(\omega, x)$  наложить ограничение  $\delta(\omega, x) \leq \eta$ . При этом все  $x_{r+1} - x_r$  окажутся  $\leq 2\eta$ .

Гейне [5] доказал, что непрерывная функция равномерно непрерывна (1.071). Его метод по существу являлся методом сечений Дедекнда и годился для доказательства общей теоремы Гейне — Бореля; именно так он был использован в доказательстве Лебега. Борелю принадлежит идея использования перекрывающихся интервалов \*). Приведенная здесь форма теоремы Гейне — Бореля принадлежит В. Юнгу [6]. Метод деления пополам применил Больцано. Гурса (см. 11.043) с помощью этого метода существенно упростил условия теоремы Коши, обнаружив важность ограничения, состоящего в том, что каждый отрезочек можно покрывать только таким интервалом  $I_x$ , который соответствует точке  $x$ , принадлежащей этому отрезочку. Он не доказал, однако, модифицированную теорему в ее общей форме и не отметил, что ее можно доказать тем же методом. Впервые это сделал Бейкер [7] в заметке, содержащей только формулировку.

**1.063.** *Функция, непрерывная на замкнутом интервале  $a \leq x \leq b$ , ограничена на  $a \leq x \leq b$ . Возьмем произвольное положительное  $\varepsilon$ . Для любого  $x$  из  $(a, b)$  найдется интервал*

---

\*) Интервалы  $I$  обязательно перекрываются; в самом деле, если  $I_x$  не содержит внутренних точек никакого другого интервала, то его концевые точки вообще не лежат внутри никакого  $I$ .



$I_x = (x - \delta_x, x + \eta_x)$ , такой, что для каждого  $\xi$  из этого интервала

$$|f(\xi) - f(x)| < \varepsilon. \quad (1)$$

Здесь  $\delta_a = 0$ ,  $\eta_b = 0$ , а в остальных случаях  $\delta_x, \eta_x > 0$ . Поэтому для любых  $\xi_1, \xi_2$  из этого интервала

$$|f(\xi_1) - f(\xi_2)| < 2\varepsilon. \quad (2)$$

Далее можно разделить  $(a, b)$  на конечное число, скажем на  $n$ , интервалов  $(x_r, x_{r+1})$ , так что для всех  $\xi_{r1}, \xi_{r2}$  из каждого интервала, включая концевые точки,

$$|f(\xi_{r1}) - f(\xi_{r2})| < 2\varepsilon. \quad (3)$$

Следовательно, для любого  $x$  из  $(a, b)$

$$|f(x) - f(a)| < 2n\varepsilon, \quad (4)$$

и поэтому  $f(x)$  ограничена (и сверху, и снизу) на  $a \leq x \leq b$ .

Отсюда следует, что  $f(x)$  имеет верхнюю и нижнюю грани на  $(a \leq x \leq b)$ . Если  $f(x)$  имеет верхнюю грань  $M$  и нижнюю грань  $m$  на некотором интервале (независимо от того, является ли  $f(x)$  непрерывной функцией), то мы называем  $M - m$  скачком  $f(x)$  на этом интервале \*).

**1.064.** *Функция, непрерывная на замкнутом интервале  $(a, b)$ , достигает на этом интервале своей верхней и нижней грани и принимает все промежуточные между ними значения.*

Пусть  $m$ ,  $M$  — нижняя и верхняя грани  $f(x)$  на  $(a, b)$ . Пусть  $c$  — произвольное значение, которое  $f(x)$  не принимает на  $(a, b)$ . Тогда для любого  $x$  из  $(a, b)$  найдется интервал  $I_x = (x - \delta_x, x + \delta_x)$ , такой, что для всех  $\xi$ , общих для  $(a, b)$  и  $I_x$ ,

$$|f(\xi) - f(x)| < \frac{1}{2} |f(x) - c|,$$

поскольку функция  $f(x)$  непрерывна, а выражение  $|f(x) - c|$  положительно. Следовательно, в точках, общих для  $(a, b)$  и  $I_x$ ,

$$|f(\xi) - c| > \frac{1}{2} |f(x) - c| > 0$$

и  $f(\xi) - c$  имеет тот же знак, что  $f(x) - c$ . Следовательно, по теореме Гейне — Бореля  $(a, b)$  можно покрыть конечным числом

---

\*) Употребляется название *колебание*. Оно кажется нам крайне неудачным, поскольку применяется к функциям, которые не колеблются. Когда мы говорим *колеблющаяся* последовательность, то тут есть некоторое сходство с тем, что физики называют колебаниями. Хобсон использует термин *размах*, а Ньюмен — *скачок*.

перекрывающихся интервалов  $I_x$ , назовем их  $I_{x_1}, \dots, I_{x_m}$ . Следовательно, 1) нижняя грань  $|f(\xi) - c|$  на  $(a, b)$  больше или равна  $\frac{1}{2}|f(x_r) - c|$  для некоторого  $x_r$ ; ни при одном  $r$  эта разность не нуль, а потому  $c$  не является верхней или нижней гранью  $f(x)$  на  $(a, b)$ ; 2)  $f(\xi) - c$  сохраняет на всем  $(a, b)$  один и тот же знак, а потому  $c$  не лежит между  $m$  и  $M$ .

Необходимость ограничиться непрерывными функциями станет яснее, если рассмотреть функцию  $f(x) = x$  ( $0 \leq x < 1/2$ ),  $f(x) = 0$  ( $1/2 \leq x \leq 1$ ). Она имеет верхнюю грань  $1/2$ , но  $f(x)$  нигде не равна  $1/2$ .

В 1.063 и 1.064 интервал должен быть замкнутым. Возьмем  $f(x) = 1/x$  на  $0 < x < 1$ ; эта функция непрерывна в каждой точке интервала, но неограничена. Если  $f(x) = x$  при  $0 \leq x < 1$ , то верхняя грань равна 1, поскольку для любого  $\eta < 1$  существует такое  $x < 1$ , что  $f(x) > \eta$ , однако  $f(x)$  не равна 1 ни при каком  $x < 1$ .

**1.065. Возрастающие и убывающие функции.** Функция называется *возрастающей* на интервале  $a < x < b$ , если для любых  $x_1, x_2$ , таких, что  $a < x_1 < x_2 < b$ , справедливо  $f(x_1) < f(x_2)$ . Она называется *убывающей*, если  $f(x_1) > f(x_2)$ , когда  $a < x_1 < x_2 < b$ . *Неубывающая* функция — это такая функция, что  $f(x_1) \leq f(x_2)$ ; *невозрастающая* функция определяется аналогично. Такие функции могут быть константами на некоторых частях интервала. Возрастающие и убывающие функции нигде не постоянны; они вместе называются *монотонными* \*).

**1.066. Обратные функции.** Если  $y = f(x)$  непрерывна и монотонна в замкнутом интервале  $(a, b)$ , то она принимает один и только один раз каждое значение, заключенное между ее верхней и нижней гранями. Следовательно, существует однозначная обратная функция  $x = g(y)$ , которая тоже монотонна. [Условие монотонности  $f(x)$  является необходимым, потому что если бы  $f(x)$  была постоянна на некотором интервале или если бы она убывала на части интервала и возрастала на другой части, то она могла бы одно и то же значение принимать более чем один раз и  $g(y)$  была бы неоднозначной.]

*Обратная функция  $g(y)$  непрерывна.* Это означает, что для любого  $y$  и любого заданного  $\varepsilon$  найдется такое  $\delta$ , что если

\*) Во многих работах то, что мы называли возрастающей функцией, называется *строго* возрастающей функцией, а то, что мы называли неубывающей функцией, называется возрастающей. Аналогично, то что мы называли монотонной функцией, называется *строго* монотонной функцией.

$|\eta - y| < \delta$ , то  $|g(\eta) - g(y)| < \varepsilon$ ; т. е. для любого  $x$  и любого заданного  $\varepsilon$  найдется такое  $\delta$ , что если  $|f(\xi) - f(x)| < \delta$ , то  $|\xi - x| < \varepsilon$ . Чтобы доказать это, возьмем для определенности возрастающую  $f(x)$ . Как только  $\xi_1 \leq x - \varepsilon$ , а  $\xi_2 \geq x + \varepsilon$ , так мы имеем

$$f(\xi_1) \leq f(x - \varepsilon) < f(x) < f(x + \varepsilon) \leq f(\xi_2).$$

Тогда если  $|f(\xi) - f(x)|$  меньше меньшего из чисел  $|f(x) - f(x - \varepsilon)|$  и  $|f(x) - f(x + \varepsilon)|$ , то из этого следует

$$x - \varepsilon < \xi < x + \varepsilon.$$

Многозначные функции часто можно рассматривать как набор однозначных. Так, для любого  $x > 0$  существуют два допустимых значения  $\sqrt{x}$ . Но если мы условимся брать всегда положительный или всегда отрицательный знак при корне, то в любом из этих случаев мы получим однозначную непрерывную функцию от  $x$ . Теоремы о непрерывных функциях будут тогда применимы к любой из них по отдельности, только уж, приняв решение, какую функцию взять, мы не должны менять его.

**1.07. Равномерная непрерывность.** Если мы выберем  $\delta$  так, что

$$|f(x + h) - f(x)| < \varepsilon \quad \text{при } |h| < \delta \quad (1)$$

для некоторого частного значения  $x$ , то, вообще говоря, окажется, что для других значений  $x$  и для того же самого  $\varepsilon$  неравенство не выполняется для того же значения  $\delta$ . Например, пусть

$$f(x) = x^2 \quad (0 \leq x \leq 1). \quad (2)$$

Если  $x = 0$ , то неравенство (1) будет справедливо при  $\delta = \sqrt{\varepsilon}$ . Однако если  $x = 1$ , то

$$|(1 - h)^2 - 1| = |2h - h^2|,$$

и последнее выражение вовсе не меньше  $\varepsilon$ , если  $h$  равно, скажем,  $\frac{1}{2} \sqrt{\varepsilon}$ , а  $\varepsilon$  достаточно мало. Но если взять  $\delta = \frac{1}{2} \varepsilon$ , то (1) будет верно для всех  $x$  из рассматриваемой области.

Это приводит нас к идее *равномерности*, с которой в дальнейшем мы будем встречаться снова и снова. Неравенство (1) выполняется для некоторого  $\delta$  и для любых  $\varepsilon$  и  $x$ , но  $\delta$  при заданном  $\varepsilon$ , возможно, зависит от  $x$ , так что можно написать  $\delta(\varepsilon, x)$ . Говорят, что некоторое предложение (здесь  $|f(x + h) - f(x)| < \varepsilon$ ) выполняется *равномерно* относительно выбора переменного (здесь  $x$ ), если условие для его справедливости можно

высказать так, чтобы оно не зависело от  $x$ ; так, в данном случае  $|h| < \delta(\epsilon)$ , где  $\delta(\epsilon)$  может зависеть от  $\epsilon$ , но не от  $x$ .

**1.071. Непрерывная функция равномерно непрерывна в любом замкнутом интервале.** В доказательстве **1.063**, которое применимо при любом значении  $\epsilon$ , заменим  $\epsilon$  на  $\omega$ , и пусть  $\frac{1}{2}\delta$  означает длину кратчайшего из интервалов  $(x_r, x_{r+1})$ . Тогда любые две точки  $\xi_1, \xi_2$  из  $(a, b)$ , такие, что  $|\xi_2 - \xi_1| < \delta$ , должны принадлежать либо одному и тому же, либо соседним интервалам  $(x_r, x_{r+1})$ , и поэтому

$$|f(\xi_2) - f(\xi_1)| < 4\omega.$$

Возьмем  $\omega = \frac{1}{4}\epsilon$ ; тогда найдено такое  $\delta$ , что как только  $\xi_1, \xi_2$  суть точки  $(a, b)$ , удовлетворяющие соотношению  $|\xi_2 - \xi_1| < \delta$ , так

$$|f(\xi_2) - f(\xi_1)| < \epsilon.$$

**1.08. Порядки величин.** Если при  $x$ , стремящемся к пределу,  $\varphi(x)$  стремится к 0 или к  $\infty$ , а  $f(x)/\varphi(x)$  ограничено, то мы говорим, что  $f(x) = O[\varphi(x)]$  или что  $f(x)$  — величина того же порядка, что  $\varphi(x)$ . Если  $f(x)/\varphi(x) \rightarrow 0$  при  $\varphi(x) \rightarrow 0$ , то мы пишем  $f(x) = o[\varphi(x)]$ . Если  $f(x)$  ограничена, то можно написать  $f(x) = O(1)$ . Эти обозначения нужно отличать от общеупотребительных в физике, когда мы можем сказать, что массы Юпитера и Сатурна — величины одного порядка, подразумевая под этим, грубо говоря, что они отличаются друг от друга не более чем в 10 раз, причем нет речи ни о каком предельном переходе. Величины, сравнимые с точки зрения физики, должны иметь одну и ту же размерность. С точки зрения математики это не является необходимым. Например,  $x$  может быть интервалом времени, а  $f(x)$  — расстоянием, которое прошла звуковая волна. Тогда  $f(x) = O(x)$ , потому что  $f(x)/x$  — скорость звука и предполагается, что она конечна.

Отметим, что

$$O(x^m) O(x^n) = O(x^{m+n}), \quad o(x^m) O(x^n) = o(x^{m+n}).$$

**1.09. Функции ограниченной вариации.** Если функция  $f(x)$  определена на замкнутом интервале  $(a, b)$  и существует такое число  $M$ , что

$$v = |f(x_1) - f(x_0)| + |f(x_2) - f(x_1)| + \dots + |f(x_n) - f(x_{n-1})| \leq M$$

для любого разбиения  $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ , то говорят, что  $f(x)$  — функция *ограниченной вариации* на интервале  $(a, b)$ ; верхняя грань сумм  $v$  по всевозможным разбиениям

называется *полной вариацией*  $f(x)$  на интервале \*). Полная вариация интересна тем, что имеет отношение к условию существования интеграла Стильбеса (1.102) и к условию существования длины кривой, она используется также в теории рядов Фурье и интегралов Фурье.

Мы неоднократно используем то, что сумма и произведение двух, а следовательно, и любого конечного числа непрерывных функций непрерывны и что при сложении и перемножении двух функций ограниченной вариации снова получается функция ограниченной вариации. Доказательства несложны. В качестве примера докажем утверждение о произведении двух функций ограниченной вариации. Заметим, что

$$\begin{aligned} f(x_{r+1})g(x_{r+1}) - f(x_r)g(x_r) &= \\ &= \frac{1}{2} [f(x_r) + f(x_{r+1})] [g(x_{r+1}) - g(x_r)] + \\ &\quad + \frac{1}{2} [g(x_r) + g(x_{r+1})] [f(x_{r+1}) - f(x_r)]. \end{aligned}$$

Следовательно, если  $M, N$  — верхние грани  $|f(x)|, |g(x)|$ , а  $U, V$  — полные вариации  $f(x), g(x)$  в интервале, то полная вариация  $f(x)g(x)$  не превышает  $MV + NU$ .

Заметим, что определение полной вариации как предела соответствующих сумм не удовлетворительно, поскольку предела может не быть для некоторых способов разбиения или пределы могут быть различными для разных способов. Возьмем, например,

$$f(x) = 0 \left( 0 \leq x < \frac{1}{2} \right); \quad f\left(\frac{1}{2}\right) = 1; \quad f(x) = 0 \left( \frac{1}{2} < x \leq 1 \right).$$

Однако в случае непрерывной или монотонной функции предел, если он существует, дает полную вариацию.

**1.091.** *Функция ограниченной вариации не обязательно непрерывна*, и наоборот. В самом деле, если  $f(x) = 0$  при  $x \leq 0$  и  $f(x) = 1$  при  $x > 0$ , то вариация не превосходит 1 ни на каком интервале, однако  $f(x)$  разрывна. Обратно, если  $f(x) = x \cos(1/x)$  при  $x \geq 0$  и  $f(0) = 0$ , то

$$f\left(\frac{1}{n\pi}\right) = \frac{1}{n\pi} (-1)^n.$$

Поэтому вариация между  $x = \frac{1}{n\pi}$  и  $x = \frac{1}{(n+1)\pi}$  равна по крайней мере  $\frac{1}{n\pi} + \frac{1}{(n+1)\pi}$ , а вариация между  $x = \frac{1}{\pi}$  и  $x = \frac{1}{n\pi}$  больше или равна

$$\frac{1}{\pi} \left( 1 + \frac{2}{2} + \frac{2}{3} + \dots + \frac{2}{n-1} + \frac{1}{n} \right),$$

---

\*) Некоторые авторы употребляют вместо термина *вариация* термин *изменение*.

что стремится к бесконечности вместе с  $n$ . Следовательно,  $f(x)$  не является функцией ограниченной вариации. Легко видеть, однако, что  $f(x)$  непрерывна даже при  $x=0$ , поскольку  $|f(x)| \leq x$ ,  $f(0)=0$ .

**1.092.** Любая функция ограниченной вариации на  $(a, b)$  представима в виде разности двух ограниченных неубывающих функций.

Для любого замкнутого интервала  $(a, x)$  рассмотрим сумму

$$p = \sum [f(x_r) - f(x_{r-1})],$$

где

$$x_0 = a \leq x_1 \leq x_2 \leq \dots \leq x_{k-1} \leq x_k = x,$$

взятую по всем членам  $f(x_r) - f(x_{r-1})$ , которые положительны, и сумму

$$-n = \sum [f(x_r) - f(x_{r-1})],$$

взятую по всем отрицательным членам. Верхнюю грань сумм  $p$  по всевозможным разбиениям назовем положительной вариацией  $P(a, x)$  на  $(a, x)$ , а верхнюю грань всех сумм  $n$  — отрицательной вариацией  $N(a, x)$ . Пусть  $v = p + n$ ; тогда значения  $v$  образуют ограниченное множество, поскольку все они  $\leq V(a, b)$ . Их верхняя грань  $V(a, x)$  есть полная вариация  $f(x)$  на  $(a, x)$ . Взяв верхние грани, имеем

$$V(a, x) = P(a, x) + N(a, x).$$

Очевидно, что  $P(a, x)$ ,  $N(a, x)$ ,  $V(a, x)$  все являются неубывающими функциями от  $x$  и ограничены на  $(a, b)$ .

Для любого разбиения и для любого фиксированного  $x$

$$p - n = f(x) - f(a), \quad p + n = v.$$

Следовательно,

$$p = \frac{1}{2} v + \frac{1}{2} [f(x) - f(a)],$$

$$n = \frac{1}{2} v - \frac{1}{2} [f(x) - f(a)].$$

Возьмем верхнюю грань по всевозможным разбиениям; тогда

$$P(a, x) = \frac{1}{2} V(a, x) + \frac{1}{2} [f(x) - f(a)],$$

$$N(a, x) = \frac{1}{2} V(a, x) - \frac{1}{2} [f(x) - f(a)].$$

Следовательно,

$$\dot{f}(x) = [f(a) + P(a, x)] - N(a, x),$$

так что  $\dot{f}(x)$  выражена в виде разности двух ограниченных неубывающих функций.

**1.093.** Все разрывы функции ограниченной вариации или простые, или устранимые. Характерная особенность простого разрыва при  $x=a$  состоит в том, что  $f(a-)$  и  $f(a+)$  существуют и различны. А особенность устранимого разрыва в том, что они существуют и равны между собой, но не равны  $f(a)$ . Допустим, что возможен случай, когда хоть один из этих пределов не существует, т. е. имеются два числа  $M, m$  ( $M > m$ ), такие, что в любом как угодно коротком интервале по одну сторону от  $a$  найдутся точки, где  $f(x) > M$ , и точки, где  $f(x) < m$ . Пусть  $\xi_1$  — точка, где  $f(x) > M$ . Далее, найдется точка между  $a$  и  $\xi_1$ , скажем  $\xi_2$ , где  $f(x) < m$ ; затем найдется  $\xi_3$  между  $a$  и  $\xi_2$ , где  $f(x) > M$ , и т. д. Отсюда следует, что полная вариация на интервале  $(a, \xi_1)$  неограничена, и поэтому  $f(x)$  не является функцией ограниченной вариации.

Можно рассуждать иначе. Пусть  $f(x)$  — функция ограниченной вариации на  $(a, b)$ . Рассмотрим положительную вариацию  $P(x)$  на  $(a, x)$ . Это ограниченная неубывающая функция  $x$  и поэтому имеет пределы (не обязательно равные между собой), когда  $x \rightarrow c$  [ $c$  принадлежит  $(a, b)$ ] со стороны больших или меньших значений. То же справедливо для отрицательной вариации.

Вычитая одну из другой, находим, что  $\dot{f}(x)$  имеет пределы, когда  $x \rightarrow c$  со стороны больших или меньших значений, а потому  $c$  является либо точкой непрерывности, либо простым или устранимым разрывом.

Если  $\dot{f}(c+)$  существует, то можно говорить о вариации на полуоткрытом интервале  $(c, d)$ , расположенном справа от  $c$ , подразумевая вариацию  $g(x)$  на  $(c, d)$ , где  $g(c) = f(c+)$ , а в остальных точках  $g(x) = f(x)$ . Тогда эта вариация стремится к нулю при  $d \rightarrow c$ . Аналогично можно определить вариацию на интервале слева от  $c$ , обладающую таким же свойством.

**1.094. Скачок в точке разрыва.** Пусть  $f(x)$  разрывна в точке  $a$ , но ограничена в некотором интервале, который содержит  $a$  в качестве внутренней точки. Тогда для некоторого положительного  $\delta$  функция  $f(x)$  имеет верхнюю и нижнюю грани  $M$  и  $m$  на  $(a - \delta, a + \delta)$ . Если  $\delta' < \delta$ , то верхняя грань на  $(a - \delta', a + \delta')$  не может быть больше, а нижняя грань не может быть

меньше, чем на интервале  $(a - \delta, a + \delta)$ . Следовательно, скачок на  $(a - \delta, a + \delta)$  имеет неотрицательный предел, когда  $\delta \rightarrow 0$ . Если этот предел равен нулю, то функция непрерывна в точке  $a$ ; если он положителен, то мы называем этот предел скачком функции в точке  $a$ .

Если  $f(x) = 0$  ( $x < 0$ ),  $f(x) = 1$  ( $x \geq 0$ ), то скачок в 0 равен 1. Если  $f(x) = 0$  ( $x \neq 0$ ),  $f(x) = 1$  ( $x = 0$ ), то скачок в 0 снова равен 1.

Если  $f(x) = \sin(1/x)$ , то скачок при  $x = 0$  равен 2, поскольку значения, произвольно близкие к 1 и к  $-1$ , встречаются в любом интервале около 0.

**1.10. Интегрирование по Риману и Стильтесу.** В этой книге будут использованы два различных определения интеграла.

Пусть  $x_1, x_2, \dots, x_n$  — множество возрастающих значений  $x$ , расположенных между  $a$  и  $b$ ; предположим, что для всех  $r$  разность  $x_{r+1} - x_r < \delta$  (для удобства мы полагаем  $x_0 = a$ ,  $x_{n+1} = b$ ). Возьмем в каждом интервале некоторое  $\xi_r$ , так что  $x_r \leq \xi_r \leq x_{r+1}$ , и образуем сумму

$$S_n = f(\xi_0)(x_1 - a) + f(\xi_1)(x_2 - x_1) + \dots + f(\xi_n)(b - x_n). \quad (1)$$

Эта сумма зависит и от значений  $x_r$ , и от значений  $\xi_r$ , если только  $f(x)$  не константа; однако если мы возьмем стремящуюся к нулю последовательность значений  $\delta$ , выбирая на каждом шагу  $x_r$  и  $\xi_r$  согласно приведенным неравенствам и образуя каждый раз сумму  $S_n$ , то эти суммы, возможно, будут стремиться к пределу, а этот предел, возможно, не будет зависеть от выбора  $x_r$  и  $\xi_r$  на каждом шагу. Если это так, то этот предел называется *интегралом Римана* и обозначается

$$\int_a^b f(x) dx. \quad (2)$$

Можно также интегрировать по функции. Если  $f(x)$  и  $g(x)$  — ограниченные функции  $x$ , то мы образуем сумму

$$S_n = f(\xi_0)[g(x_1) - g(a)] + f(\xi_1)[g(x_2) - g(x_1)] + \dots + f(\xi_n)[g(b) - g(x_n)], \quad (3)$$

выбирая  $\xi_n$  так же, как в (1). Если эта сумма стремится к единственному пределу, когда наибольший из интервалов  $(x_r, x_{r+1})$  стремится к нулю, то предел этот называется *интегралом Стильтеса* [8–10] и обозначается

$$\int_{x=a}^b f(x) dg(x). \quad (4)$$



Следует обратить внимание на запись пределов интегрирования, ибо функция  $g(x)$  может не быть монотонной. Она может вернуться к своему первоначальному значению, но мы не должны при этом писать, что область интегрирования распространяется от  $g(a)$  до  $g(a)$ , интеграл тогда, очевидно, обратился бы в нуль. Постоянное возрастание в области интегрирования требуется не от  $g(x)$ , а от  $x$ .

**1.101.** Интеграл Римана  $\int_a^b f(x) dx$  существует тогда и только тогда, когда  $f(x)$  ограничена на  $(a, b)$ , и для любых положительных  $\omega$  и  $\eta$  интервал  $(a, b)$  можно так разбить на конечное число интервалов, что те из них, где скачок  $f(x) \geq \omega$ , имеют общую длину меньше  $\eta$ .

Прежде всего ограниченность  $f(x)$ , очевидно, необходима для существования интеграла. Ибо если  $f(x)$  неограничена на  $(a, b)$ , то всегда найдется хоть один интервал  $(x_r, x_{r+1})$ , на котором она неограничена; а следовательно, неограничено множество значений  $S_n$ , которые образуются при различном выборе  $\xi_r$  на этом интервале. Поэтому, как бы ни выбирать интервалы, не может случиться, чтобы суммы  $S_n$  стремились к единственному пределу независимо от выбора  $\xi_r$  на каждом шагу.

Предположим теперь, что верхняя и нижняя грани  $f(x)$  на  $(a, b)$  равны  $M$  и  $m$ . Предположим также, что верхняя и нижняя грани на  $(x_r, x_{r+1})$  равны  $M_r$  и  $m_r$ , так что при любом выборе  $\xi_r$  имеем  $m_r \leq f(\xi_r) \leq M_r$ . Образует суммы

$$h_n = \sum m_r (x_{r+1} - x_r), \quad H_n = \sum M_r (x_{r+1} - x_r). \quad (1)$$

Эти суммы называются нижней и верхней суммами для разбиения, определяемого точками  $x_r$ , и являются нижней и верхней гранями  $S_n$  при этом разбиении.

Далее, в любом интервале  $(x_r, x_{r+1})$  найдется такое значение  $x$ , где  $f(x) \geq \frac{3}{4} M_r + \frac{1}{4} m_r$ , и такое значение  $x$ , где  $f(x) \leq \frac{1}{4} M_r + \frac{3}{4} m_r$ . Следовательно, если  $M_r - m_r \geq \omega$ , то можно такими двумя способами выбрать  $\xi_r$ , что соответствующие значения  $f(\xi_r) (x_{r+1} - x_r)$  будут отличаться друг от друга по крайней мере на  $\frac{1}{2} \omega (x_{r+1} - x_r)$ .

Поскольку выбор  $\xi_r$  во всех интервалах производится независимо, то, если интервалы, где  $M_r - m_r \geq \omega$ , имеют общую длину  $\geq \eta$ , где  $\eta$  — некоторое положительное число, мы сможем тогда такими двумя способами выбрать  $\xi_r$  в каждом

интервале, что соответствующие значения  $S_n$  будут отличаться по крайней мере на  $\frac{1}{2}\omega\eta$ . Если найдутся  $\omega > 0$  и  $\eta > 0$ , такие, что для любого разбиения  $(a, b)$  общая длина интервалов, где скачок  $f(x) \geq \omega$ , всегда по крайней мере равна  $\eta$ , то тогда суммы  $S_n$  не могут иметь единственный предел. Следовательно, указанное условие является необходимым.

Поскольку  $m_r \leq M$ ,  $M_r \geq m$ , то всегда

$$h_n \leq M(b-a), \quad H_n \geq m(b-a). \quad (2)$$

Следовательно, значения  $h_n$ , которые получаются при всевозможных разбиениях, имеют верхнюю грань, назовем ее  $h$ ; а значения  $H_n$  имеют нижнюю грань, назовем ее  $H$ . Покажем прежде всего, что  $h \leq H$ .

Если в произвольном интервале  $(x_r, x_{r+1})$  взять еще одну точку разбиения, скажем  $x_{r1}$ , и снова образовать нижнюю и верхнюю суммы, то в каждом из частичных интервалов  $(x_r, x_{r1})$  и  $(x_{r1}, x_{r+1})$  верхняя грань  $f(x)$  может оказаться меньше, чем  $M_r$ , но не больше. Тем самым добавление новых точек разбиения может уменьшить верхнюю сумму, но не может ее увеличить; и аналогично оно может увеличить нижнюю сумму, но не может ее уменьшить.

Рассмотрим теперь два различных разбиения, определяемые точками  $x_r$  и  $x'_s$ . Пусть соответствующие суммы равны  $H_n, h_n, H'_r, h'_r$ . Рассмотрим разбиение, которое образуется, если объединить вместе все точки обоих разбиений. Пусть для него суммы равны  $H''_q, h''_q$ . Можно считать, что новое разбиение получено добавлением точек как к множеству  $x_r$ , так и к множеству  $x'_s$ .

Следовательно, согласно предыдущему абзацу,

$$H_n \geq H''_q \geq h''_q \geq h'_r. \quad (3)$$

Таким образом, никакая нижняя сумма не может оказаться больше какой-нибудь верхней суммы, а следовательно,  $H_n \geq h$  для всех  $n$  и, значит,  $H \geq h$ .

Далее, если мы сможем найти такой способ разбиения, при котором  $H_n - h_n < \epsilon$ , то из этого будет следовать, что  $H - h < \epsilon$ ; в самом деле,  $H_n - h_n = (H - h) + (H_n - H) + (h - h_n)$  и  $H_n - H \geq 0$ ,  $h - h_n \geq 0$ . Предположим теперь, что все интервалы разбиты на два класса: класс  $A$ , в котором  $M_r - m_r < \omega$ , и класс  $B$ , где  $M_r - m_r \geq \omega$ . Для интервалов из класса  $B$  имеется все же оценка сверху:  $M_r - m_r \leq M - m$ . Если  $\alpha$  — общая длина всех  $B$ -интервалов, то

$$H_n - h_n < (b - a - \alpha)\omega + (M - m)\alpha. \quad (4)$$

Допустим теперь, что для любого  $\omega$  общая длина всех  $B$ -интервалов может быть сделана произвольно малой. Тогда для любого положительного  $\varepsilon$  можно подобрать  $\omega$  и  $\alpha$  так, что

$$(b-a)\omega < \frac{1}{2}\varepsilon, \quad (M-m)\alpha < \frac{1}{2}\varepsilon. \quad (5)$$

Отсюда следует, что  $0 \leq H-h < \varepsilon$ , а стало быть, поскольку  $H$  и  $h$  не зависят от  $\varepsilon$ ,

$$H = h. \quad (6)$$

Нам осталось показать еще, что если мы возьмем такие разбиения, для которых длина наибольшего интервала равна  $\delta$  и  $\delta \rightarrow 0$ , то  $H_n \rightarrow H$ ,  $h_n \rightarrow h = H$ . Пусть  $x_r$  — множество точек разбиения, удовлетворяющего условию (5), так что  $H_n - h_n < \varepsilon$ . Пусть самый короткий интервал этого разбиения равен  $\delta$ ; рассмотрим другое разбиение с помощью точек  $x'_s$ , такое, что длина его наибольшего интервала меньше  $\delta$ . Пусть для этого разбиения верхняя и нижняя суммы равны  $H'_p$ ,  $h'_p$ . Тогда любые последовательные точки  $x'_s, x'_{s+1}$  либо принадлежат одному и тому же интервалу разбиения  $x_r$ , либо двум соседним интервалам.

Если эти последние оба являются  $A$ -интервалами, то скачок  $f(x)$  на  $(x'_s, x'_{s+1})$  меньше  $2\omega$ ; если оба они  $B$ -интервалы или один  $A$ -, а другой  $B$ -интервал, то скачок не превосходит  $M-m$ . Но если  $B$ -интервал имеет длину  $\mu \geq \delta$ , то длина  $x'_s$ -интервалов, имеющих с ним общие точки, не превосходит  $\mu + 2\delta \leq 3\mu$ . Следовательно, скачки  $f(x)$  на  $x'_s$ -интервалах  $< 2\omega$  для всех интервалов, кроме, быть может, некоторого подмножества с общей длиной  $\leq 3 \sum \mu = 3\alpha$ . Таким образом,

$$H'_p - h'_p < (b-a-3\alpha)2\omega + 3(M-m)\alpha < \frac{5}{2}\varepsilon. \quad (7)$$

Кроме того,  $H'_p \geq H$ ,  $h'_p \leq H$ ; следовательно,

$$H'_p - H < \frac{5}{2}\varepsilon, \quad H - h'_p < \frac{5}{2}\varepsilon. \quad (8)$$

Поскольку это справедливо для любого разбиения с длиной наибольшего интервала, меньшей  $\delta$ , то это и есть искомый результат.

Наконец, так как  $h_n \leq S_n \leq H_n$ , то  $S_n$  тоже стремится к  $H$ .

**1.1011.** Вышеуказанное условие принадлежит Дюбуа-Раймону. Его можно высказать в другой форме, иногда более удобной. *Необходимое и достаточное условие существования*

интеграла  $\int_a^b f(x) dx$  состоит в следующем:  $f(x)$  ограничена, и для любых положительных  $\omega$  и  $\eta$  точки разрыва, в которых скачок  $\geq \omega$ , можно покрыть конечным числом интервалов с общей длиной  $< \eta$ .

Ясно, что данное условие вытекает из условия Дюбуа-Раймона. Наоборот, если только что сформулированное условие выполнено, то на оставшихся интервалах нет точек разрыва со скачком  $\geq \omega$ . Тогда около любой точки, принадлежащей оставшимся интервалам, можно найти интервал, на котором скачок  $< \omega$ . Следовательно, по модифицированной теореме Гейне — Бореля оставшиеся интервалы можно разбить на конечное число частей, в каждой из которых скачок  $< \omega$ .

**1.1012.** Немедленным следствием является тот факт, что любая непрерывная функция имеет интеграл Римана; в самом деле, она ограничена и вообще не имеет точек разрыва. Точно так же любая функция с конечным числом разрывов (в которых она имеет конечные скачки) обладает интегралом Римана. То же относится к любой функции ограниченной вариации. Ибо если для некоторого  $\omega$  существует бесконечное множество точек разрыва, в которых скачки функции больше  $\omega$ , то она не является функцией ограниченной вариации.

Отметим, что это условие не требует конечности числа точек разрыва. Положим  $f(x) = 1$  при  $x = 1/n$  для всех целых положительных  $n$  и  $f(x) = 0$  в остальных точках. Функция разрывна при всех  $x = 1/n$ , а также при  $x = 0$ . Однако для любого  $\eta$  интервал  $(0, \frac{1}{2}\eta)$  содержит бесконечно много точек разрыва, а оставшихся, таких, что  $\frac{1}{2}\eta \leq x \leq 1$ , — лишь конечное число, и их можно покрыть конечным числом интервалов с общей длиной  $\frac{1}{2}\eta$ . Таким образом, бесконечное множество точек разрыва иногда может быть покрыто конечным числом интервалов с произвольно малой общей длиной.

Если  $f(x) = 0$  для  $x$  иррациональных и  $f(x) = 1/n$  для  $x = m/n$ , где  $m/n$  — правильная несократимая дробь, то  $f(x)$  разрывна при всех рациональных значениях  $x$  из  $(0, 1)$ , но непрерывна при всех иррациональных значениях. В самом деле, любое иррациональное  $x_0$  можно покрыть интервалом длины  $1/n!$ , не содержащим рациональных дробей со знаменателем, меньшим  $n$ , и, следовательно, значения  $f(x)$  в достаточно малом интервале, содержащем  $x_0$ , произвольно малы.

В рассматриваемом случае функция  $f(x)$  имеет точки разрыва в каждом сколь угодно коротком интервале. Тем не менее она обладает интегралом Римана, ибо число точек разрыва, в которых скачок  $f(x)$  превышает  $\varepsilon$ , не больше, чем сумма целых чисел, меньших  $1/\varepsilon$ , и, значит, конечно. Интеграл этой функции равен нулю.

Если  $f(x)=1$  для  $x$  рациональных и  $f(x)=0$  для иррациональных, то для каждого  $x_0$ , неважно, рационального или нет, найдутся значения  $x$ , произвольно близкие к  $x_0$ , где  $f(x)=1$ , и другие, где  $f(x)=0$ . Следовательно, любое значение  $x$  является точкой разрыва со скачком, равным 1, и такие точки на  $(0, 1)$  нельзя покрыть никаким множеством интервалов с длиной  $< 1$ . В этом случае  $H_n=1$ ,  $h_n=0$ , как бы мы ни разбивали интервал.

Прямое практическое значение особенностей такого типа невелико. Однако, коль скоро мы хотим заниматься обобщениями, эти особенности очень важны как сигналы опасности. Существуют другие определения интеграла (наиболее важное принадлежит Лебегу), которые придают значения некоторым интегралам, не существующим по Риману (включая только что упомянутый случай). В этих определениях с самого начала рассматриваются разбиения на бесконечное множество частей. Они упрощают формулировки и заметно увеличивают общность некоторых из последних теорем. Читатель может познакомиться с этими определениями у Баркилла [11] и Титчмарша [12]. Однако оказывается, что случаи, в которых эти методы применимы, а метод Римана нет, так редки в физике, что трудности, связанные с их введением, не оправдываются.

Если  $f(x)$  имеет интеграл Римана, то  $\{f(x)\}^n$  ( $n > 0$ ) и  $|f(x)|$  также обладают интегралом Римана на том же интервале. Поскольку если  $f(x)$  ограничена и ее точки разрыва, в которых скачок превосходит  $\omega$ , могут быть покрыты интервалами с произвольно малой общей длиной, то это же применимо к  $\{f(x)\}^n$  и к  $|f(x)|$ . Обратное неверно. Рассмотрим  $f(x)=1$  при рациональных значениях  $x$ ,  $f(x)=-1$  при иррациональных значениях;  $\{f(x)\}^2$  и  $|f(x)|$  интегрируемы, а  $f(x)$  нет.

**1.1013. „Мера нуль“, „почти всюду“.** Множество точек, которое можно покрыть интервалами с произвольной малой общей длиной, называют множеством *меры нуль*, а предложение, справедливое всюду, кроме такого множества, называют верным *почти всюду*. Любое конечное множество точек имеет меру нуль; тем же свойством обладают и все целые числа, поскольку каждое целое  $n$  можно покрыть интервалом

длины  $2^{-|n|}\alpha$ , где  $\alpha$  произвольно мало, и поскольку ряд  $\sum 2^{-|n|}$  сходится. Мера нуль имеют и все рациональные числа на  $(0, 1)$ .

В самом деле, если  $p$  и  $q$  — целые и такие, что  $p < q$ , то  $p/q$  можно покрыть интервалом длины  $\alpha/q^3$ , где  $\alpha$  положительно. Существует  $q-1$  дробей со знаменателем  $q$ , исключая 0 и 1. Но 0 и 1 можно покрыть интервалами длины  $\alpha$ , а остальные дроби — интервалами с суммарной длиной, меньшей  $\alpha/q^2$ . Суммируя теперь по  $q$ , мы видим, что все рациональные дроби можно покрыть интервалами с общей длиной, меньшей чем  $2\alpha + \alpha \sum_2^{\infty} q^{-2}$ ; ряд сходится, и, следовательно, суммарная длина может быть сделана столь малой, как мы пожелаем, с помощью подходящего выбора  $\alpha$ . То же справедливо для любого счетного множества.

Рассмотрим убывающую последовательность положительных чисел  $\omega_1, \omega_2, \dots$ , стремящуюся к нулю. Если  $f(x)$  обладает интегралом Римана, то точки, в которых скачок  $\geq \omega$  (если таковые вообще имеются), можно покрыть конечным числом интервалов произвольно малой длины; следовательно, точки разрыва, в которых скачок  $< \omega_{n-1}$ , но  $\geq \omega_n$ , можно покрыть конечным множеством интервалов длины  $2^{-n}\eta$ , а все точки разрыва — множеством длины  $\eta$ . Это множество интервалов счетно, ибо каждый интервал может быть достигнут за конечное число шагов; следовательно, точки разрыва интегрируемой функции можно покрыть счетным множеством интервалов с произвольно малой общей длиной. Эти интервалы могут пересекаться.

**1.102. Существование интегралов Стильеса.** Определение интеграла этого типа, данное в 1.10, позволяет, чтобы функция  $g(x)$  была разрывной. Если  $g(x)$  не убывает и ограничена при  $a \leq x \leq b$  и если  $f(x)$  тоже ограничена, то необходимое и

достаточное условие существования  $\int_{x=a}^b f(x) dg(x)$  состоит в сле-

дующем: для любых  $\omega$  и  $\delta$  интервал можно разбить на конечное число подынтервалов, так что в тех подынтервалах, где скачок  $f(x)$  больше  $\omega$ , полная вариация  $g(x)$  меньше  $\delta$ .

Доказательство по существу такое же, как для интеграла Римана. Если  $g(x)$  имеет ограниченную вариацию, то мы получим тот же результат, выражая  $g(x)$  в виде разности двух неубывающих функций  $\varphi(x) - \psi(x)$  и рассматривая отдельно  $\int f(x) d\varphi(x)$  и  $\int f(x) d\psi(x)$ .

В частности, интеграл Стильеса существует на любом конечном интервале, если  $g(x)$  имеет ограниченную вариацию, а  $f(x)$  непрерывна. Он не существует, если  $f(x)$  и  $g(x)$  имеют точку разрыва при одном и том же значении  $x$ , так как в любом интервале, содержащем точку разрыва, ни скачок  $f(x)$ , ни полная вариация  $g(x)$  не являются произвольно малыми. Отсюда следует, что для существования интеграла Стильеса не достаточно, чтобы  $f(x)$  и  $g(x)$  обе были ограниченной вариации.

Мы не даем условий существования интеграла Стильеса для общего случая, когда  $g(x)$  не является функцией ограниченной вариации: достаточно, чтобы  $g(x)$  была непрерывна, а  $f(x)$  имела ограниченную вариацию, но не достаточно, чтобы они обе были непрерывны (это будет доказано ниже).

Если  $a < b < c$  и если ввести обозначение  $I(d, e) = \int_{x=d}^e f(x) dg(x)$ , то из существования  $I(a, c)$  следует:  $I(a, b)$  и  $I(b, c)$  оба существуют, и их сумма равна  $I(a, c)$ . Обратное верно не всегда. Если

$$\begin{aligned} f(x) &= 0 \quad (x \leq 0), & g(x) &= 1 \quad (x < 0), \\ f(x) &= 1 \quad (x > 0), & g(x) &= 0 \quad (x \geq 0), \end{aligned}$$

то  $\int_{x=-1}^0 f dg$  и  $\int_{x=0}^1 f dg$  оба существуют и равны нулю, а  $\int_{x=-1}^1 f dg$  не существует. Обратное делается верным для слегка измененного определения интеграла Стильеса, данного Поллардом.

### 1.103. Дифференцирование.

а) Если  $f(x)$  непрерывна и  $\int_a^x f(u) du = F(x)$ , то

$$\frac{d}{dx} F(x) = f(x)$$

и  $F(x)$  является непрерывной функцией  $x$ .

Это почти очевидно.

б) Если

$$\frac{d}{dx} F(x) = f(x)$$

и  $f(x)$  интегрируема, то

$$\int_a^x f(u) du = F(x) - F(a).$$

Поскольку  $F(x)$  дифференцируема на  $(a, b)$ , то, как мы знаем из **1.0622**, для любых положительных  $\omega$  и  $\delta$  можно разделить  $(a, b)$  на конечное множество интервалов  $(x_r, x_{r+1})$  длины  $\leq \delta$  и в каждом из них взять некоторую точку  $\xi_r$  так, что для любой точки из  $(x_r, x_{r+1})$

$$|F(x) - F(\xi_r) - (x - \xi_r) F'(\xi_r)| < \omega |x - \xi_r| \quad (1)$$

и, следовательно,

$$|F(x_{r+1}) - F(x_r) - (x_{r+1} - x_r) F'(\xi_r)| < \omega (x_{r+1} - x_r). \quad (2)$$

Суммируя, получаем

$$|F(b) - F(a) - \sum (x_{r+1} - x_r) f(\xi_r)| < \omega (b - a). \quad (3)$$

Поскольку  $f(x)$  интегрируема, можно для любого заданного положительного  $\varepsilon$  выбрать такое  $\delta$ , чтобы сумма в формуле (3) отличалась от интеграла меньше, чем на  $\varepsilon$ . Таким образом,

$$|F(b) - F(a) - \int_a^b f(x) dx| < \varepsilon + \omega (b - a),$$

и, следовательно, левая часть равна нулю, ибо  $\varepsilon$  и  $\omega$  произвольно малы.

Заметим, что  $F(x)$  может быть дифференцируемой, а ее производная может не быть интегрируемой; например,

$$F(x) = x^2 \sin \frac{1}{x^2} \quad (0 < x \leq 1), \quad F(0) = 0.$$

Производная существует даже при  $x=0$ , но неограничена в любой окрестности 0.

в) Если  $f(x)$  имеет интеграл Римана  $\int_a^b f(x) dx$ , то  $\int_a^x f(u) du$  существует и является непрерывной функцией для всех  $x$ , таких, что  $a \leq x \leq b$ ; производная этой функции равна  $f(x)$  всюду, кроме, быть может, множества меры нуль, а именно кроме точек разрыва  $f(x)$ .

Пусть  $x$  — точка непрерывности  $f(x)$ . Тогда в интервале  $(x-h, x+h)$  скачок  $f(\xi)$  равен  $\omega$ , где  $\omega \rightarrow 0$  вместе с  $h$ .



Кроме того,

$$f(x) - \omega \leq \frac{1}{h} \int_x^{x+h} f(u) du = \frac{1}{h} \left[ \int_a^{x+h} f(u) du - \int_a^x f(u) du \right] \leq f(x) + \omega.$$

Устремляя  $h$  к нулю, имеем

$$\frac{d}{dx} \int_a^x f(u) du = f(x)$$

во всех точках, где  $f(x)$  непрерывна.

Отсюда следует, что если

$$\int_a^x f(u) du = \int_a^x g(u) du, \quad a \leq x \leq b,$$

то  $f(x) = g(x)$  почти всюду на  $a \leq x \leq b$ ; исключения представляют точки разрыва  $f(x)$  или  $g(x)$  (если таковые имеются).

Исключение, упомянутое в пункте „В“, довольно важное. Если, например,  $f(x) = 0$  при  $x \leq 0$  и  $f(x) = 1$  при  $x > 0$ , то

$$F(x) = \int_{-1}^x f(u) du \begin{cases} = 0 & (x \leq 0), \\ = x & (x > 0), \end{cases}$$

и  $\frac{d}{dx} F(x)$  не существует при  $x = 0$ . Если же  $f(x) = 0$  при  $x \neq 0$

и  $f(x) = 1$  при  $x = 0$ , то  $\int_a^x f(u) du = 0$  на любом интервале и всюду имеет производную, равную нулю; при  $x = 0$  эта производная не равна  $f(x)$ .

**1.1031. Интегрирование по частям для интеграла Стильеса.**  
Положим

$$S_n = \sum_{r=1}^n f(\xi_r) [g(x_r) - g(x_{r-1})], \quad (1)$$

при этом

$$x_0 = a, \quad x_n = b, \quad x_0 \leq \xi_1 \leq x_1 \leq \dots \leq x_{r-1} \leq \xi_r \leq x_r \leq \dots \leq x_n. \quad (2)$$

Тогда

$$S_n = f(x_n) g(x_n) - f(x_0) g(x_0) - \sum_n, \quad (3)$$

где

$$\begin{aligned} \sum_n = & g(x_0) [f(\xi_1) - f(x_0)] + \\ & + \sum_{r=1}^{n-1} g(x_r) [f(\xi_{r+1}) - f(\xi_r)] + g(x_n) [f(x_n) - f(\xi_n)]. \end{aligned} \quad (4)$$

Допустим, что  $I = \int_{x=a}^b f dg$  существует, т. е. для любого  $\varepsilon > 0$  можно подобрать  $\delta$  так, что для всех разбиений, в которых длина наибольшего подынтервала  $< \delta$ ,

$$|S_n - I| < \varepsilon. \quad (5)$$

Тогда для любого множества  $a, \xi_1, \xi_2, \dots, b$ , такого, что  $\xi_1 - a, \dots, \xi_{r+1} - \xi_r, \dots, b - \xi_n$  все меньше  $\frac{1}{2}\delta$ , если для  $x_r$  выполняются неравенства (2), то выполняются также неравенства  $x_r - x_{r-1} < \delta$  при всех  $r$ . Следовательно, для такого множества

$$|\sum_n - f(b)g(b) + f(a)g(a) + I| < \varepsilon, \quad (6)$$

и потому  $\sum_n$  стремится к пределу при  $\delta \rightarrow 0$ , а этот предел по определению равен  $\int g df$ . Следовательно,  $\int g df$  существует и

$$\int_{x=a}^b g df = [fg]_a^b - \int_{x=a}^b f dg. \quad (7)$$

В частности, поскольку  $\int f dg$  существует, когда  $f$  непрерывна, а  $g$  имеет ограниченную вариацию, то он существует также и в том случае, когда  $f$  имеет ограниченную вариацию, а  $g$  непрерывна. Если  $g(x)$  является интегралом Римана  $\left[ g(x) = \int_a^x h(u) du \right]$ , то она и непрерывна, и имеет ограниченную вариацию; следовательно,  $\int f dg$  существует, если  $f$  обладает хоть одним из этих свойств.

Для интегралов Римана результат обычно формулируют в виде

$$\int_a^b f(x) g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx; \quad (8)$$

таким образом, чтобы обе функции и  $f$ , и  $g$  были дифференцируемы для всех  $x$ , требуется  $a \leq x \leq b$ . Если производные существуют и интегрируемы, то этот результат немедленно следует из (7) и 1.1032. Формула (7) верна, однако, в более общих условиях. Простейший путь использования (8) состоит в том, чтобы сначала проинтегрировать  $g'$ , а потом переписать (8) в виде (7).

**1.1032. Замена переменного в интеграле.** Если  $x = h(y) =$   
 $= \int_a^y g(u) du$  и если

$$I = \int_{y=a}^b f(x) dx, \quad J = \int_a^b f(x) g(y) dy$$

оба существуют, то  $I = J$ . Поскольку по предположению оба интеграла существуют, то достаточно доказать, что интегральные суммы стремятся к одному и тому же пределу хотя бы при одном способе их образования. Возьмем  $x_r = h(y_r)$ ,  $\xi_r = h(\eta_r)$ , тогда

$$I_n = \sum f(\xi_r)(x_{r+1} - x_r), \quad J_n = \sum f(\xi_r) g(\eta_r)(y_{r+1} - y_r). \quad (1)$$

Поскольку  $g(y)$  интегрируема, интервалы на оси  $y$  можно выбрать так, чтобы те из них, где скачок  $g(y)$  больше чем  $\omega$ , имели общую длину  $\leq \delta$ , где  $\omega$  и  $\delta$  произвольно малы. Далее,  $|g(y)|$  ограничен, скажем  $< G$ ; следовательно, если длина наибольшего из интервалов  $y_{r+1} - y_r$  меньше  $\lambda$ , то наибольшее из чисел  $|x_{r+1} - x_r|$  меньше  $G\lambda$ . Следовательно,  $I_n \rightarrow I$ ,  $J_n \rightarrow J$ . Пусть теперь  $G_r$  и  $g_r$  — верхняя и нижняя грани  $g(y)$  при  $y_r \leq y \leq y_{r+1}$ , тогда

$$g_r(y_{r+1} - y_r) \leq x_{r+1} - x_r \leq G_r(y_{r+1} - y_r), \quad (2)$$

$$g_r \leq g(\eta_r) \leq G_r, \quad (3)$$

и, следовательно,

$$|x_{r+1} - x_r - g(\eta_r)(y_{r+1} - y_r)| \leq (G_r - g_r)(y_{r+1} - y_r), \quad (4)$$

$$|I_n - J_n| \leq \sum f(\xi_r)(G_r - g_r)(y_{r+1} - y_r). \quad (5)$$

Пусть скачок  $g(y)$  на всем интервале равен  $N$ , и пусть  $G_r - g_r \leq \omega$  всюду, кроме конечного множества интервалов с общей длиной  $\delta$ . Тогда

$$|I_n - J_n| \leq FN\delta + \omega F(b - a - \delta), \quad (6)$$

где  $F$  — верхняя грань  $|f(x)|$ . Правую часть можно сделать произвольно малой; следовательно,

$$I_n - J_n \rightarrow 0, \quad I = J. \quad (7)$$

Общепринятая формулировка

$$\int_{x(a)}^{x(b)} f(x) dx = \int_a^b f(x) \frac{dx}{dy} dy, \quad \frac{dx}{dy} > 0 \quad (8)$$

имеет несколько меньшую общность, поскольку предполагает, что  $dx/dy$  всюду существует. Если это условие выполнено, то

теорема доказывается совсем легко с помощью применения теоремы Ролля (1.13). Более общая формулировка нужна, однако, для преобразования интегралов вдоль кривых, которые могут иметь углы, т. е. точки, где не определена касательная. Формуле (8) можно придать следующий смысл: в каждой точке  $c$ , где  $dx/dy$  не существует, ее можно заменить произвольным значением, заключенным между пределами, к которым при  $\delta \rightarrow 0$  стремятся верхняя и нижняя грани  $dx/dy$  на  $(c - \delta, c + \delta)$ .

**1.104. Бесконечные и несобственные интегралы.** Доказательство существования интеграла рушится, если интервал  $(a, b)$  бесконечен или функция, которую следует проинтегрировать, неограничена на интервале. В первом случае  $b - a$  бесконечно, и никаким выбором  $\omega$  нельзя добиться выполнения неравенств  $\omega > 0$ ,  $(b - a)\omega < \varepsilon$ . В последнем случае аппроксимирующая сумма меняется в сколь угодно широких пределах в зависимости от выбора точки, где берется значение  $f(x)$ , для подынтервала, в котором  $f(x)$  неограничена.

В каждом из случаев нужна некоторая специальная схема, чтобы придать значение интегралу. Способ, используемый для интегралов с бесконечным верхним пределом, состоит в первоначальном рассмотрении интеграла с конечным верхним пределом; если этот интеграл стремится к определенному пределу, когда верхний предел интегрирования стремится к бесконечности, то это предельное значение берут в качестве значения для бесконечного интеграла. Использование такой схемы можно проиллюстрировать на примере интеграла

$$I = \int_0^{\infty} \frac{\sin x}{x} dx.$$

Согласно нашему правилу, его можно интерпретировать как

$$\lim_{X \rightarrow \infty} \int_0^X \frac{\sin x}{x} dx.$$

Интеграл до  $X$  существует при всех  $X$ , поскольку интегрируемое выражение всюду непрерывно. Если взять  $Y > X$ , в качестве  $m$  выбрать наименьшее целое число, большее чем  $X/\pi$ , а в качестве  $n$  — наибольшее целое, меньшее  $Y/\pi$ , то

$$\int_X^Y \frac{\sin x}{x} dx = \int_X^{m\pi} \frac{\sin x}{x} dx + \int_{m\pi}^Y \frac{\sin x}{x} dx + \sum_{r=m}^{n-1} \int_{r\pi}^{(r+1)\pi} \frac{\sin x}{x} dx.$$

Поскольку  $|\sin x| \leq 1$ , а  $m\pi - X \leq \pi$ , то первый из интегралов в правой части по модулю  $\leq \pi/X$ . Аналогично второй из интегралов по абсолютной величине  $\leq 1/n$ . Сумма состоит из слагаемых поочередно то положительных, то отрицательных, причем каждое по модулю меньше предыдущего; но существует теорема о том, что если  $u_0 > u_1 > \dots > u_n > 0$ , то

$$u_0 > u_0 - u_1 + u_2 + \dots + (-1)^n u_n > 0.$$

Следовательно, сумма меньше своего первого слагаемого по абсолютной величине, а

$$\left| \int_{m\pi}^{(m+1)\pi} \frac{\sin x}{x} dx \right| < \frac{1}{m}.$$

Итак,

$$\left| \int_X^Y \frac{\sin x}{x} dx \right| < \frac{\pi}{X} + \frac{1}{n} + \frac{1}{m} < \frac{3\pi}{X},$$

что может быть сделано произвольно малым для всех  $Y > X$ , лишь бы  $X$  было достаточно велико.

Следовательно, для любого положительного числа  $\epsilon$ , сколь угодно малого, можно выбрать  $X$  таким образом, что, как бы мы ни увеличивали верхний предел интегрирования по сравнению с  $X$ , мы не можем изменить интеграл больше, чем на  $\epsilon$ . Итак, интеграл до  $X$  имеет некоторое предельное значение, когда  $X$  стремится к бесконечности, и бесконечный интеграл существует в смысле данного выше определения.

**1.1041.** Поскольку интеграл является функцией своего верхнего предела, то можно применить признаки сходимости последовательностей, данные в **1.0441** и **1.045**, как это уже упоминалось в теории непрерывности (см. **1.06**). Доказательства очевидны.

Если  $f(x) \geq 0$  и  $\int_a^x f(x) dx$  ограничен при всех  $X > a$ , то  $\int_a^x f(x) dx$  стремится к пределу при  $X \rightarrow \infty$ .

Необходимое и достаточное условие того, чтобы  $\int_a^x f(x) dx$  стремился к пределу при  $X \rightarrow \infty$ , состоит в том, что для любого положительного  $\epsilon$ , сколь угодно малого, существует такое  $A$ , что  $\left| \int_A^X f(x) dx \right| < \epsilon$  для всех  $X > A$ .

**1.1042.** Связь между бесконечными интегралами и рядами настолько тесна, что их свойства удобно выражать одними и теми же словами:

$\int_a^\infty f(x) dx$  *сходится*, если  $\lim_{X \rightarrow \infty} \int_a^X f(x) dx$  существует.

$\int_a^\infty f(x) dx$  *неограничен*, если  $\left| \int_a^X f(x) dx \right|$  является неограни-

ченным при  $X \rightarrow \infty$ .

$\int_a^\infty f(x) dx = \infty$ , если  $\int_a^X f(x) dx \rightarrow \infty$  при  $X \rightarrow \infty$ .

$\int_a^\infty f(x) dx$  *осциллирует с конечным размахом*, если сущест-

вуют такие положительные числа  $\omega$  и  $M$ , что для любого  $X$  можно выбрать  $Y_1 > X$  так, чтобы  $\left| \int_X^{Y_1} f(x) dx \right| > \omega$ , но нельзя

выбрать  $Y_2$  так, чтобы  $\left| \int_X^{Y_2} f(x) dx \right| > M$ . Примеры сходящихся интегралов:

$$\int_1^\infty \frac{dx}{x^2}, \quad \int_0^\infty e^{-x} dx, \quad \int_1^\infty \frac{\sin x}{x^2} dx, \quad \int_0^\infty \frac{\sin x}{x} dx.$$

Неограниченные интегралы:

$$\int_1^\infty dx, \quad \int_1^\infty \frac{dx}{x}, \quad \int_1^\infty x \sin x dx.$$

Последний из этих интегралов обычно называют „бесконечно осциллирующим“, но у нас нет повода выделять этот случай.

Следующие интегралы осциллируют с конечным размахом:

$$\int_0^\infty \cos x dx, \quad \int_0^\infty \sin x dx.$$

Неограниченные и осциллирующие с конечным размахом интегралы не имеют определенного значения.

Интеграл  $\int_a^\infty f(x) dx$  называется *абсолютно сходящимся*, если  $\int_a^\infty |f(x)| dx$  сходится. Если первый интеграл сходится, а последний нет, то первый называется *условно сходящимся*. В приведенных выше примерах сходящихся интегралов первые три интеграла абсолютно сходятся, а последний условно сходится.

**1.1043.** Если  $f(x)$  положительна и не возрастает при  $x > x_0$ , то интеграл  $I = \int_{x_0}^\infty f(x) dx$  сходится тогда и только тогда, когда сходится ряд  $\sum_{n=n_0}^\infty f(n)$ , где  $n_0$  — наименьшее целое число, большее чем  $x_0$ . Ясно, что ни ряд, ни интеграл не могут сходиться без того, чтобы  $f(x) \rightarrow 0$ ; возьмем целое число  $m > x_0$  и такое, что  $f(m) < \varepsilon$ . Тогда

$$\begin{aligned} f(m) + f(m+1) + \dots + f(n-1) &\geq \int_m^x f(x) dx \geq \\ &\geq f(m+1) + f(m+2) + \dots + f(n-1), \end{aligned}$$

где  $n$  — ближайшее к  $X$  целое число, большее  $X$ . Следовательно,

$$\left| \int_m^X f(x) dx - \sum_{r=m}^{n-1} f(r) \right| \leq f(m) < \varepsilon,$$

где  $\varepsilon$  произвольно мало. Следовательно, если интеграл стремится к определенному пределу, то и сумма тоже, и обратно.

В частности,  $\int_1^\infty x^{-p} dx$  и  $\sum_{n=1}^\infty n^{-p}$  оба сходятся тогда и только тогда, когда  $p > 1$ .

**1.1044.** Если  $f(x)$  стремится к бесконечности при подходе к некоторой точке области определения, то можно, используя аналогичный прием, определить *несобственный интеграл*: изменим сначала область определения, вырезав из нее произвольно короткий интервал, содержащий указанную точку, а затем

заставим длину этого интервала стремиться к нулю. Например,

$$\int_{\epsilon}^1 x^{-1/2} dx = [2x^{1/2}]_{\epsilon}^1 = 2 - 2\epsilon^{1/2}, \quad \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 x^{-1/2} dx = 2.$$

Этот процесс принимают за определение  $\int_0^1 x^{-1/2} dx$ , поскольку определение интеграла как предела сумм в данном случае непосредственно не применимо.

Аналогия между бесконечными и несобственными интегралами в вопросах сходимости настолько тесна, что вся терминология остается без изменения.

**1.1045.** Заменой переменных можно перевести обычный интеграл в бесконечный или несобственный; но значение его при этом не изменится. Например, если для всех  $y$ , как в 1.1032,  $x = h(y)$ , то

$$\int_{h(0)}^{h(y)} f(x) dx = \int_0^y f(x) g(y) dy, \quad (1)$$

когда  $y$  и  $h(y)$  конечны, а  $g(y) \geq 0$ ; обе части равенства имеют один и тот же предел, если либо  $y$ , либо  $h(y)$ , либо и то и другое стремится к бесконечности. Если  $h(y) \rightarrow b$  при  $y \rightarrow \infty$

и если  $\int_{h(0)}^b f(x) dx$  существует, то он является пределом левой части (1); следовательно, предел равен интегралу Римана, когда он существует.

**1.11. Функции двух переменных.** До сих пор мы занимались последовательностями, которые можно рассматривать как функции одного переменного, принимающего только целые значения, и функции одного переменного, изменяющегося непрерывно. В последующем мы займемся функциями, существенно зависящими от двух переменных, которые могут либо принимать целые значения, либо меняться непрерывно. Это приведет к новым трудностям при использовании предельных переходов, ибо не всегда очевидно (и даже не всегда верно), что если менять порядок переходов к пределу, то будет получаться один и тот же результат. Простейшее достаточное условие перестановочности переходов к пределу обеспечивается следующей теоремой об абсолютной сходимости.



1.111. Если  $f(x, y)$  — функция, неубывающая по каждому из переменных  $x$  и  $y$  (причем, быть может, одно из них или оба принимают только целые значения), и если

$$\lim_{x \rightarrow \infty} f(x, y) = g(y), \quad \lim_{y \rightarrow \infty} f(x, y) = h(x), \quad (1)$$

то

$$\lim_{y \rightarrow \infty} g(y) = \lim_{x \rightarrow \infty} h(x) \quad (2)$$

в том смысле, что существование одного из пределов (2) влечет за собой существование другого и равенство обоих пределов между собой.

Заметим сначала, что  $g(y)$  — неубывающая функция  $y$ . Ибо если  $y_2 > y_1$ , то

$$g(y_2) - g(y_1) = \lim_{x \rightarrow \infty} [f(x, y_2) - f(x, y_1)] \geq 0. \quad (3)$$

Аналогично если  $x_2 > x_1$ , то

$$h(x_2) \geq h(x_1). \quad (4)$$

Пусть  $g(y)$  имеет предел  $M$ . Для любого  $\varepsilon$  найдется такое  $Y$ , что для всех  $y \geq Y$

$$M \geq g(y) \geq M - \varepsilon. \quad (5)$$

Для всех  $x, y$  выполняются неравенства  $M \geq g(y) \geq f(x, y)$ . Далее, существует такое  $X$ , что для всех  $x > X$

$$f(x, Y) \geq g(Y) - \varepsilon, \quad (6)$$

поэтому для всех  $y > Y, x > X$

$$M \geq f(x, y) \geq g(Y) - \varepsilon \geq M - 2\varepsilon. \quad (7)$$

Следовательно, если  $y \rightarrow \infty$  и  $x > X$ , то

$$M \geq h(x) \geq M - 2\varepsilon, \quad (8)$$

и поэтому, так как  $\varepsilon$  произвольно,  $h(x)$  тоже имеет предел  $M$  при  $x \rightarrow \infty$ .

Немедленно получаем три следствия. Если

$$f(m, n) = \sum_{r=1}^m \sum_{s=1}^n u_{r,s}, \quad (9)$$

где  $u_{r,s}$  неотрицательны, то  $f(m, n)$  является неубывающей функцией от  $m$  и от  $n$ . Значит, для двойного ряда из неотрицательных членов

$$\sum_{r=1}^{\infty} \sum_{s=1}^{\infty} u_{r,s} = \sum_{s=1}^{\infty} \sum_{r=1}^{\infty} u_{r,s}. \quad (10)$$

Если некоторые из слагаемых  $u_{r,s}$  отрицательны, то можно написать

$$v_{r,s} = |u_{r,s}| + u_{r,s}, \quad w_{r,s} = |u_{r,s}| - u_{r,s}, \quad (11)$$

где все  $v_{r,s}$  и  $w_{r,s}$  уже неотрицательны. Если теперь ряд  $\sum \sum |u_{r,s}|$  сходится при некотором порядке суммирования, то, как легко видеть,  $v_{r,s}$  и  $w_{r,s}$  удовлетворяют условию теоремы о перемене порядка суммирования; вычитая, находим, что  $u_{r,s}$  тоже удовлетворяет этому условию. Следовательно, для любого абсолютно сходящегося двойного ряда можно менять порядок суммирования. Как следствие если  $\sum a_r$  и  $\sum b_s$  абсолютно сходятся, то  $\sum_r a_r \sum_s b_s = \sum_r \sum_s a_r b_s$ , где в правой части члены суммы можно брать в любом порядке.

Если  $u_m(x)$  всюду неотрицательны и если  $\sum_{m=1}^p \int_0^q u_m(x) dx$  существует для всех  $p, q$  и имеет пределы в каждом из двух случаев, когда одно из переменных  $p, q$  стремится к бесконечности, а другое фиксировано, то

$$\sum_{m=1}^{\infty} \int_0^{\infty} u_m(x) dx = \int_0^{\infty} \left\{ \sum_{m=1}^{\infty} u_m(x) \right\} dx. \quad (12)$$

Если  $u_m(x)$  не всегда сохраняют знак, но  $\sum_{m=1}^p \int_0^q |u_m(x)| dx$  существует и удовлетворяет вышеуказанным условиям и если хотя бы одно из

$$\sum_1^{\infty} \int_0^{\infty} |u_m(x)| dx, \quad \int_0^{\infty} \sum_1^{\infty} |u_m(x)| dx$$

существует, то оба предела в формуле (12) существуют и равны между собой.

Если  $\varphi(x, y)$  неотрицательна и существуют соответствующие пределы по каждому из переменных, то

$$\int_0^{\infty} dx \int_0^{\infty} \varphi(x, y) dy = \int_0^{\infty} dy \int_0^{\infty} \varphi(x, y) dx; \quad (13)$$

здесь за  $f(x, y)$  принято общее значение обоих интегралов с верхними пределами  $x$  и  $y$ . Как и раньше, если  $\varphi(x, y)$  не всегда одного знака, то для существования и равенства

двойных пределов достаточно, чтобы существовал один из них для  $|\varphi(x, y)|$ .

**1.112. Равномерная сходимость последовательностей и рядов.** Члены последовательности  $\{f_n(x)\}$  могут быть функциями переменного  $x$ . Тогда если последовательность сходится при всех значениях  $x$  из некоторого интервала, то ее предел является функцией  $x$ , скажем  $f(x)$ . Если выбрать произвольно малое положительное  $\varepsilon$ , то для любого  $x$  можно подобрать  $n(x)$ , так что  $|f_p(x) - f(x)| < \varepsilon$  для всех  $p \geq n(x)$ , поскольку последовательность сходится. Наименьшее значение  $n(x)$ , для которого это верно, вообще говоря, зависит от  $x$ . Но может оказаться возможным выбрать  $n$  независимо от  $x$ , так что  $|f_p(x) - f(x)| < \varepsilon$  для всех  $p > n$  и для всех  $x$  из интервала. Если это возможно при каждом  $\varepsilon$ , то говорят, что  $f_n(x)$  *сходится равномерно* к  $f(x)$  в интервале. Условие равномерной сходимости может нарушаться, если существует такое  $x$ , скажем  $c$ , внутри или на конце интервала, что если взять последовательность значений  $x$ , стремящуюся к  $c$ , то соответствующие данному  $\varepsilon$  значения  $n(x)$  стремятся к бесконечности.

Так как  $f_n(x)$  может быть суммой первых  $n$  членов ряда, то эти же положения справедливы и для рядов  $\sum u_n(x)$ , когда  $x$  пробегает некоторый интервал. Так, ряд  $\sum x^n$  сходится для всех  $x$ , таких, что  $0 \leq x < 1$ , но не является равномерно сходящимся для всех таких  $x$ . В самом деле, если зафиксировать  $\varepsilon$ , а потом выбрать  $n$  так, чтобы

$$x^n + x^{n+1} + \dots + x^{n+p} = \frac{x^n(1 - x^{p+1})}{1 - x} \quad (1)$$

было меньше  $\varepsilon$  при всех  $p \geq 1$ , то тогда должно быть

$$x^n < (1 - x)\varepsilon \quad (2)$$

и, таким образом,

$$n > \frac{\ln[(1 - x)\varepsilon]}{\ln x}, \quad (3)$$

что стремится к бесконечности, когда  $x$  стремится к 1. Поэтому рассматриваемый ряд равномерно сходится на отрезке  $a \leq x \leq b$ , где  $a$  и  $b$  — фиксированные числа между 0 и +1, поскольку можно выбрать  $n$  больше большего из чисел

$$\frac{\ln[(1 - a)\varepsilon]}{\ln a}, \quad \frac{\ln[(1 - b)\varepsilon]}{\ln b},$$

и одно и то же значение  $n$  подойдет тогда для любого промежуточного значения  $x$ . Ряд сходится для любого  $x$  из

$-1 < x < 1$ . Однако ряд не является равномерно сходящимся в интервале  $-1 < x < 1$ , ибо, хотя знаки „ $<$ “ исключают для  $x$  возможность быть равным в точности  $-1$  или  $+1$ , они допускают, чтобы  $x$  принимало произвольное промежуточное значение, сколь угодно близкое к  $1$ , и как бы мы ни выбрали  $n$ , мы всегда сможем найти такое значение  $x$ , чтобы условие (3) нарушилось.

Если  $f_n(x) \rightarrow f(x)$  равномерно на каждом из конечного множества интервалов  $(a_r, b_r)$ ,  $r = 1, \dots, k$ , то сходимость также равномерная на всем этом множестве. В самом деле, для каждого интервала существует такое  $n_r$ , что  $|f_p(x) - f(x)| < \varepsilon$  для  $p \geq n_r$  и  $x$  из  $(a_r, b_r)$ . Возьмем  $m$  равным наибольшему из  $n_r$ ; тогда при  $p > m$  будет  $|f_p(x) - f(x)| < \varepsilon$  для  $x$  в любом из этих интервалов.

Если  $f_n(x) \rightarrow f(x)$  в  $a \leq x \leq b$  и  $f_n(x) \rightarrow f(x)$  равномерно в  $a < x < b$ , то тогда сходимость является равномерной в  $a \leq x \leq b$ . Нужно лишь применить предыдущее доказательство к открытому интервалу  $(a, b)$  и к точкам  $a, b$  и взять  $m$  равным наибольшему из трех соответствующих значений  $n$ . Итак, последовательность, сходящаяся равномерно в открытом интервале, равномерно сходится и в замкнутом интервале, если только она сходится в концевых точках. Пусть, однако,

$$f_n(0) = 2^n, \quad f_n(1) = 2^n, \quad f_n(x) = 2^{-n} \quad (0 < x < 1),$$

тогда  $\{f_n(x)\}$  равномерно сходится в открытом интервале, но не сходится в замкнутом.

**1.113. Непрерывность и интегрируемость равномерно сходящихся рядов.** Сумма ряда из непрерывных функций от  $x$ , равномерно сходящегося в некоторой области, сама является непрерывной функцией  $x$  в этой области.

Интеграл по  $x$  от суммы ряда, равномерно сходящегося, когда  $x$  пробегает некоторую конечную область, равен сумме интегралов от членов ряда, если пределы интегрирования расположены в указанной области.

Чтобы доказать первое утверждение, обозначим сумму ряда через  $S(x)$ . Тогда, поскольку ряд сходится равномерно, если  $\omega$  — положительное число, можно выбрать  $n$  независимо от  $x$  так, что если  $S_n(x)$  — сумма до  $u_n(x)$ , то

$$|S(x) - S_n(x)| < \omega, \quad |S(y) - S_n(y)| < \omega \quad (1)$$

для всех  $x, y$  в рассматриваемой области. Но  $S_n(x)$  — сумма конечного числа непрерывных функций и поэтому непрерывна.

Следовательно, для любого  $x$  можно подобрать положительное, но такое маленькое  $\delta$ , что

$$|S_n(y) - S_n(x)| < \omega \quad (2)$$

для всех  $|y - x| < \delta$ . Поэтому для таких  $y$

$$|S(y) - S(x)| < 3\omega, \quad (3)$$

и, полагая  $\omega = 1/3\varepsilon$ , а затем выбирая  $\delta$  в соответствии с неравенством (2), мы можем добиться, чтобы

$$|S(x+h) - S(x)| < \varepsilon \quad (4)$$

для всех  $h$ , удовлетворяющих неравенствам  $0 \leq h < \delta$ . Следовательно,  $S(x)$  непрерывна (а значит, интегрируема).

Докажем второе утверждение. Если

$$S(x) = S_n(x) + R_n(x) \quad (5)$$

и если  $|R_n(x)| < \omega$  для всех  $x$ , таких, что  $a \leq x \leq b$ , то

$$\int_a^b S(x) dx = \int_a^b S_n(x) dx + \int_a^b R_n(x) dx \quad (6)$$

и

$$\left| \int_a^b R_n(x) dx \right| < \omega(b-a), \quad (7)$$

где правая часть произвольно мала. Следовательно, если взять  $n$  достаточно большим, то выражение

$$\int_a^b S_n(x) dx = \sum_{r=1}^n \int_a^b u_n(x) dx \quad (8)$$

можно сделать сколь угодно близким к  $\int_a^b S(x) dx$ . Эту теорему часто формулируют так: *равномерно сходящийся ряд из непрерывных функций можно интегрировать почленно в любой конечной области.*

Равномерно сходящийся ряд можно интегрировать почленно и в том случае, когда его члены интегрируемы, но не обязательно непрерывны. Если  $S(x)$  интегрируема, то, начиная с формулы (5), проходит прежнее доказательство. Возьмем  $n$  так, чтобы было  $|R_n(x)| < \omega$ . Скачок  $R_n(x)$  ни на каком интервале не превосходит  $2\omega$ . Сумма  $S_n(x)$  интегрируема. Разделим  $(a, b)$  на конечное число интервалов, так чтобы общая длина

тех из них, где скачок  $S_n(x)$  превосходит  $\omega$ , была меньше  $\delta$ . Тогда общая длина тех интервалов, где скачок  $S(x)$  превосходит  $3\omega$ , меньше  $\delta$ ; величины  $\omega$  и  $\delta$  произвольны, следовательно,  $S(x)$  интегрируема.

Если  $f_n(x) \rightarrow f(x)$  равномерно на  $(a, b)$ , то  $\int_a^x f_n(\xi) d\xi$  стремится равномерно к  $\int_a^x f(\xi) d\xi$ , ибо если  $|f_n(\xi) - f(\xi)| < \omega$ , то

$$\left| \int_a^x [f_n(\xi) - f(\xi)] d\xi \right| < \omega(x - a) \leq \omega(b - a).$$

**1.114. Точки разрыва, связанные с неравномерной сходимостью.** Геометрическая прогрессия не сходится ни при  $x = 1$ , ни при  $x = -1$ , и поэтому значения предельной функции в этих точках не определены, а значит, и вопрос о непрерывности не возникает. Может случиться, однако, что ряд сходится при некоторых значениях  $x$ , но не является равномерно сходящимся в прилегающей к ним области. Стокс, первым изучавший это свойство, предложил следующий пример:

$$u_n(x) = \left( \frac{1}{n} - \frac{1}{n+1} \right) + 2 \left( \frac{1}{(n-1)x+1} - \frac{1}{nx+1} \right). \quad (1)$$

Ряд  $\sum u_n(x)$  сходится при всех  $x$ , поскольку

$$\sum_n^{n+p} u_n(x) = \frac{1}{n} - \frac{1}{n+p+1} + 2 \left( \frac{1}{(n-1)x+1} - \frac{1}{(n+p)x+1} \right). \quad (2)$$

Возьмем  $x > 0$ . Тогда

$$\left| \sum_n^{n+p} u_n(x) \right| \leq \frac{1}{n} + \frac{2}{(n-1)x+1}, \quad (3)$$

и можно сделать, чтобы это выражение было меньше  $\varepsilon$ , если взять  $n$  достаточно большим. Если  $x = 0$ , то последняя скобка в выражении (2) равна 0 и сумма меньше  $1/n$ . Ряд сходится поэтому при  $x \geq 0$ . Однако он не является равномерно сходящимся. В самом деле, если величина в правой части (3) больше  $\varepsilon$ , то величину в левой части можно сделать больше  $\varepsilon$ ,

взяв  $p$  достаточно большим; а  $1/n$  всегда положительна. Если теперь

$$\frac{2}{(n-1)x+1} > \varepsilon, \quad (4)$$

т. е.

$$(n-1)x < \frac{2}{\varepsilon} - 1, \quad (5)$$

то левая часть (3) окажется больше  $\varepsilon$  для достаточно больших  $p$ ; а чтобы сделать левую часть (3) меньше  $\varepsilon$  для всех  $p$ , мы должны взять  $n > \frac{2/\varepsilon - 1}{x} + 1$ . Следовательно, если с самого начала зафиксировать  $\varepsilon$ , то соответствующие значения  $n$  безгранично возрастают с уменьшением  $x$ , ряд сходится неравномерно. Стокс характеризовал такие ряды как *сходящиеся бесконечно медленно* вблизи  $x=0$ .

Рассмотрим теперь сумму нашего ряда. Для всех  $x$  мы имеем

$$\sum_1^n u_n(x) = \left(1 - \frac{1}{n+1}\right) + 2 \left(\frac{1}{1} - \frac{1}{nx+1}\right),$$

т. е. сумма ряда, если  $x \neq 0$ , равна 3, поскольку те члены в правой части, которые содержат  $n$ , с возрастанием  $n$  стремятся к 0. Следовательно, предел суммы при  $x$ , стремящемся к 0, равен 3. Если же мы положим  $x$  равным 0, то слагаемые во второй скобке взаимно уничтожаются при всех  $n$  и сумма равна 1.

Этот пример достаточно искусственный, однако функции, использованные в нем, совсем просты; пример служит для иллюстрации того факта, что результаты, которые получаются при выполнении двух предельных переходов, могут быть совершенно различны в зависимости от того, какой переход сделать сначала. Мы заставляем  $x$  стремиться к 0, а  $n$  — к бесконечности. Если мы сначала устремляем  $x$  к 0, а потом  $n$  к бесконечности, то получаем 1; если же мы сначала устремляем  $n$  к бесконечности, а уже потом  $x$  к 0, то получаем 3.

**1.115. Признаки равномерной сходимости.** *Необходимое и достаточное условие равномерной сходимости  $\{u_n(x)\}$  в интервале  $(a, b)$  состоит в следующем: для любого  $\varepsilon$  можно так выбрать  $m$ , что для всех  $n \geq m$   $|u_n(x) - u_m(x)| < \varepsilon$  при всех  $x$  из  $(a, b)$ .* Доказательство, данное для числовых последовательностей, проходит почти без изменений (см. 1.045).

**1.1151. М-признак.** Если для всех  $x$  из рассматриваемого интервала  $|u_n(x)| \leq v_n$ , где  $v_n$  не зависит от выбора  $x$ , и если ряд  $\sum v_n$  сходится, то  $\sum u_n(x)$  сходится равномерно в этом интервале. В самом деле, мы можем так выбрать  $n$ , чтобы для всех  $p \geq 0$  было

$$\sum_n^{n+p} v_n < \varepsilon,$$

поскольку  $\sum v_n$  сходится; а тогда для любого  $x$

$$\left| \sum_n^{n+p} u_n(x) \right| \leq \sum_n^{n+p} |u_n(x)| \leq \sum_n^{n+p} v_n < \varepsilon.$$

Этот признак известен как *М-признак Вейерштрасса*.

На том же принципе основано использование *метода сравнения рядов* при установлении обычной сходимости; соответствующую теорему достаточно только сформулировать. Если для достаточно больших  $n$   $|u_n| \leq v_n$  и  $\sum v_n$  сходится, то  $\sum u_n$  сходится.

М-признак чрезвычайно прост, и мы встретимся с его многочисленными применениями.

**1.1152. Обобщение М-признака.** Некоторую модификацию М-признака иногда можно применять даже к условно сходящимся рядам, т. е. в случае, когда мы не в состоянии указать сходящийся ряд из положительных членов  $v_n$ , больших по модулю, чем  $u_n(x)$ . Предположим, что  $u_n(x)$  равномерно стремятся к 0 при  $n \rightarrow \infty$  (см. ниже), а также что можно объединить члены  $\sum u_n(x)$  в группы по  $m$ , не меняя их порядка, и образовать ряд  $\sum U_v(x)$ , такой, что  $|U_v(x)| < V_v$ , где  $\sum V_v$  сходится. Тогда, согласно М-признаку,  $\sum U_v(x)$  равномерно сходится. Остается показать, что при сформулированных условиях сумма  $\sum u_n(x)$  существует и равна  $\sum U_v(x)$ .

Поскольку  $u_n(x)$  равномерно стремятся к 0, то для любого  $\varepsilon$  можно выбрать такое  $n$ , что для всех  $p \geq n$  и для всех  $x$  из рассматриваемой области  $|u_p(x)| < \varepsilon$ . Далее, если для заданного  $n$  взять  $v$  так, что

$$vm \leq n < (v+1)m,$$

то

$$\left| \sum_1^n u_n(x) - \sum_1^v U_v(x) \right| = |u_{mv+1}(x) + \dots + u_n(x)|.$$



Возьмем  $\nu$  таким, что  $\sum_{\nu+1}^{\infty} U_{\sigma}(x) < 1/2\varepsilon$  и все  $|u_p(x)| < \varepsilon/2m$  при  $p > m\nu$ . Тогда

$$\left| \sum_1^n u_n(x) - \sum_1^{\infty} U_{\nu}(x) \right| < \varepsilon$$

при всех  $x$  и при всех  $n > m\nu$ ; ряд  $\sum u_n(x)$  равномерно сходится.

Указанный признак можно применить к ряду

$$\sum_1^{\infty} (-1)^n \frac{n}{n^2 + x^2}.$$

В самом деле, объединяя члены ряда в пары, мы получим ряд, члены которого  $\leq 1/n(n+1)$  и который удовлетворяет, следовательно,  $M$ -признаку. Кроме того, общий член исходного ряда  $\leq 1/n$  при всех  $x$  и поэтому равномерно стремится к нулю. Следовательно, ряд равномерно сходится.

**1.1153. Лемма Абеля.** Хотя  $M$ -признак фактически является самым распространенным в приложениях, существуют ряды, которые равномерно сходятся, но ему не удовлетворяют. Следующие два более тонких признака основаны на лемме Абеля. Аналогичные признаки имеются для интегралов.

*Если  $\{v_r\}$  — невозрастающая последовательность неотрицательных чисел и если суммы*

$$s_p = a_1 + a_2 + \dots + a_p$$

*удовлетворяют неравенствам  $h \leq s_p \leq H$  при всех  $p$ , то  $h v_1 \leq$*

$$\leq \sum_1^n a_p v_p \leq H v_1 \text{ для всех } n.$$

Мы имеем

$$a_1 = s_1, \quad a_2 = s_2 - s_1, \quad \dots, \quad a_n = s_n - s_{n-1};$$

$$\begin{aligned} S_n = \sum_1^n a_p v_p &= s_1 v_1 + \sum_{p=2}^n (s_p - s_{p-1}) v_p = \\ &= s_1 (v_1 - v_2) + \dots + s_{n-1} (v_{n-1} - v_n) + s_n v_n. \end{aligned}$$

Поскольку все  $v_p - v_{p+1} \geq 0$  и  $v_n \geq 0$ , то последняя сумма не уменьшится, если все  $s_p$  заменить на  $H$ ; поэтому  $S_n \leq H v_1$ . Аналогично сумма не возрастет, если все  $s_p$  заменить на  $h$ ; следовательно,  $S_n \geq h v_1$ .

**1.1154. Признак Абеля.** Если ряд  $\sum a_r$  сходится (не обязательно абсолютно) и если для каждого  $x$  из некоторого интервала  $\{v_r(x)\}$  — последовательность положительных чисел, ограниченная как функция  $x$  и  $r$  и невозрастающая, когда  $x$  задано а  $r$  растет, то  $\sum a_r v_r(x)$  равномерно сходится в указанном интервале. Возьмем в лемме Абеля

$$s_n = \sum_{p=1}^n a_{m+p}, \quad S_n(x) = \sum_{p=1}^n a_{m+p} v_{m+p}(x),$$

выбрав  $m$  так, чтобы  $|s_n| < \omega$  для всех  $n$ . Тогда по лемме Абеля при

$$h = -\omega, \quad H = \omega, \quad v_1 = v_{m+1}(x) \leq M \\ -\omega M \leq S_n(x) \leq \omega M$$

для всех  $x$  из интервала и для всех  $n \geq 1$ . Поскольку  $\omega$  произвольно мало и не зависит от  $x$ , отсюда следует равномерная сходимость \*).

Наиболее важно приложение этой теоремы к степенным рядам  $\sum a_n x^n$ . Если такой ряд сходится при  $x=1$ , то степени  $x^n$  при  $0 \leq x \leq 1$  удовлетворяют условиям, наложенным на  $v_n(x)$ : следовательно, ряд  $\sum a_n x^n$  сходится равномерно вплоть до  $x=1$  и предел его суммы равен  $\sum a_n$ . Это теорема Абеля. Она избавляет нас от массы хлопот: решая задачу, мы часто получаем ответ в виде степенного ряда и хотим узнать, стремимся ли в пределе сумма ряда для  $x < 1$  к сумме при  $x=1$ , когда  $x$  приближают к 1. Теорема дает нам простой ответ: да, стремимся при условии, что для  $x=1$  ряд сходится.

Теорема остается справедливой, если  $a_n = a_n(x)$  и  $\sum a_n(x)$  равномерно сходится. Доказательство проходит без изменений.

**1.1155. Признак Дирихле — Харди. \*\*)** Если в некотором интервале изменения  $x$  суммы  $\sum_1^n a_r(x)$  равномерно ограничены (как функции  $n$  и  $x$ ) и если  $\{v_r\}$  — невозрастающая последовательность положительных чисел, стремящаяся к нулю, то ряд  $\sum a_r(x) v_r$  равномерно сходится в рассматриваемом интервале.

\*) Случай  $v_n(x) = x^n$  ( $0 \leq x \leq 1$ ) доказан и использовался Абелем. Общая форма теоремы принадлежит Харди.

\*\*) Для числовых рядов признак сходимости дан Дирихле; Харди превратил его в признак равномерной сходимости. Харди предложил назвать соответствующие признаки именами Абеля и Дирихле, однако применение к равномерной сходимости в последнем случае полностью принадлежит Харди.

Признак обобщается на случай  $v_r = v_r(x)$  при условии, что  $v_r(x) \rightarrow 0$  равномерно.

Возьмем  $S_n(x) = \sum_{p=1}^n a_{m+p} v_{m+p}$ , где  $m$  выбрано так, что  $v_m(x) < \omega$ . Поскольку для всех  $n$

$$-M \leq \sum_{p=1}^n a_{m+p}(x) \leq M,$$

то по лемме Абеля

$$-M\omega \leq S_n(x) \leq M\omega.$$

Равномерная сходимость доказана.

Замечательная особенность этого признака состоит в том, что он устанавливает равномерную сходимость ряда, не сравнивая его с какими бы то ни было сходящимися рядами. Наиболее важны приложения к рядам вида  $\sum v_n \cos n\theta$ ,  $\sum v_n \sin n\theta$ . Имеем

$$\sum_1^n \cos n\theta = \frac{\sin\left(n + \frac{1}{2}\right)\theta - \sin \frac{1}{2}\theta}{2 \sin \frac{1}{2}\theta},$$

$$\sum_1^n \sin n\theta = \frac{\cos \frac{1}{2}\theta - \cos\left(n + \frac{1}{2}\right)\theta}{2 \sin \frac{1}{2}\theta}.$$

Если  $\sin \frac{1}{2}\delta$  при  $\frac{1}{2}\pi > \delta > 0$  есть наименьшее значение  $|\sin \frac{1}{2}\theta|$  в области изменений  $\theta$ , то ни для какой суммы модуль не превосходит  $\text{cosec } \frac{1}{2}\delta$ , каковы бы ни были  $n$  и  $\theta$ . Если теперь  $v_1 \geq v_2 \geq \dots \geq v_n \rightarrow 0$ , то отсюда следует, что  $\sum v_n \cos n\theta$  и  $\sum v_n \sin n\theta$  равномерно сходятся в любом замкнутом интервале  $a \leq \theta \leq b$ , который не содержит нулей  $\sin \frac{1}{2}\theta$ , т. е. не содержит точек  $\theta = 0, 2\pi, 4\pi, \dots$ .

В частности, ряды

$$1 + \cos \theta + \frac{1}{2} \cos 2\theta + \frac{1}{3} \cos 3\theta + \dots,$$

$$\sin \theta + \frac{1}{2} \sin 2\theta + \frac{1}{3} \sin 3\theta + \dots$$

равномерно сходятся на любом отрезке  $a \leq \theta \leq b$ , который не включает  $0, 2\pi, \dots$ . В действительности первый из рядов расходится при  $\theta = 0$ , второй сходится всюду, но неравномерно в любом интервале, содержащем  $\theta = 0$ . Позже (гл. 14, пример 4) мы увидим, что его сумма совершает скачок от  $-\frac{1}{2}\pi$  до  $\frac{1}{2}\pi$ .

когда  $\theta$ , возрастая, проходит через 0, так что неравномерная сходимость, как и в 1.114, связана с разрывностью.

**1.116. Теорема об ограниченной сходимости.** Равномерная сходимость является достаточным условием непрерывности и интегрируемости суммы при условии, что отдельные члены непрерывны или интегрируемы. Однако это условие отнюдь не является необходимым. На практике обычно бывает проще непосредственно проверить, интегрируема ли предельная функция, чем проверить, равномерна ли сходимость, к тому же во многих случаях переход к пределу под знаком интеграла возможен, хотя сходимость не является равномерной; поэтому возникает необходимость в более общем правиле. Таким правилом является следующее. Оно известно как *теорема об ограниченной сходимости* (см. приложение к гл. 1). Если  $|f_n(x)| \leq M$  для всех  $n$  и  $x$  при  $a \leq x \leq b$ , все  $\int_a^b f_n(x) dx$  интегрируемы и  $\int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx$ , где  $f(x)$  интегрируема, то  $\int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx$ . Доказательство не просто, но результат следует знать.

Поведение  $f_n(x)$  и

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx, \quad \int_0^1 \lim_{n \rightarrow \infty} f_n(x) dx$$

полезно изучить на примерах

$$f_n(x) = xe^{-nx}, \quad f_n(x) = nxe^{-nx}, \quad f_n(x) = n^2xe^{-nx}.$$

**1.117. Ряды, используемые для сравнения.** В дальнейшем самыми важными рядами, используемыми для сравнения, будут  $\sum x^n$  ( $0 \leq x < 1$ ),  $\sum n^{-s}$  ( $s > 1$ ) (их мы уже изучили) и  $\sum n^s a^n$  ( $0 \leq a < 1$ ). Сходимость последнего сразу следует из  $M$ -признака, если  $s < 0$ . Если  $s \geq 0$ , имеем

$$\frac{u_{n+1}}{u_n} = \left( \frac{n+1}{n} \right)^s a = \left( 1 + \frac{1}{n} \right)^s a.$$

С ростом  $n$  это отношение стремится к  $a$ . Следовательно, поскольку  $a < 1$ , можно взять  $m$  настолько большим, чтобы для всех  $n > m$  отношение было меньше  $b$ , где  $a < b < 1$ . Тогда при  $n > m$

$$|u_n| < u_m b^{n-m},$$

а  $\sum b^{n-m}$  — сходящийся ряд из положительных членов. Следовательно,  $\sum n^s a^n$  сходится при  $0 \leq a < 1$ .

Сравнение с рядом  $\sum n^{-s}$  часто можно упростить. Если  $v_n = n^{-s}$  ( $1 < s$ ), то

$$n \left( 1 - \frac{v_n}{v_{n-1}} \right) = n \left[ 1 - \left( \frac{n-1}{n} \right)^s \right] \rightarrow s.$$

Если  $u_n$  положительны при всех  $n$  и

$$n \left( 1 - \frac{u_n}{u_{n-1}} \right) \rightarrow t > 1,$$

то можно взять  $s = \frac{1}{2}(t+1)$ , а потом выбрать  $m$  так, чтобы для всех  $n > m$  было

$$n \left( 1 - \frac{u_n}{u_{n-1}} \right) > n \left[ 1 - \left( \frac{n-1}{n} \right)^s \right];$$

тогда

$$u_n < u_m \left( \frac{m}{n} \right)^s$$

и ряд  $\sum u_n$  сходится. Аналогично если  $t$  существует и меньше 1, то  $\sum u_n$  расходится. При  $t=1$  нужен более тонкий признак, однако мы не встретимся с таким случаем в этой книге.

Окончательный вывод: *если  $u_n > 0$ , то  $\sum u_n$  сходится при выполнении каждого из следующих условий:*

$$\frac{u_n}{u_{n-1}} \rightarrow k < 1$$

или

$$\frac{u_n}{u_{n-1}} \rightarrow 1, \quad n \left( 1 - \frac{u_n}{u_{n-1}} \right) \rightarrow t > 1.$$

### 1.12. Равномерная сходимость бесконечных интегралов.

Если подынтегральное выражение зависит от  $x$  и еще от некоторого параметра  $y$ , то так же, как для рядов, возникает вопрос о *равномерной сходимости*. Мы будем всегда предпо-

лагать, что  $\int_a^X f(x) dx$  существует, как бы велико ни было  $X$ .

Это замечание необходимо по следующей причине. Утверждение „бесконечный интеграл сходится“ означает, что сходится любая последовательность интегралов с конечным верхним пределом, когда этот предел стремится к бесконечности. Таким образом, нельзя говорить о сходимости бесконечного интеграла, не предположив, что все интегралы от  $a$  до  $X$  существуют. Ниже мы дадим условия сходимости бесконечного интеграла в предположении, что конечные интегралы существуют. В частном случае

если  $\int_0^X f(x) dx$  существует и  $\int_X^Y |f(x)| dx < \epsilon$  для всех  $Y > X$ , то

$\int_0^{\infty} f(x) dx$  сходится. Отметим, что существование интеграла  $\int_0^X |f(x)| dx$ , напротив, не гарантирует существования  $\int_0^X f(x) dx$ .

Интеграл  $\int_0^{\infty} f(x, y) dx$  равномерно сходится в области  $b_0 \leq y \leq b_1$ , если для любого  $\varepsilon$  можно так подобрать  $X$ , что для всех  $Y > X$  и для всех  $y$  из этой области

$$\left| \int_X^Y f(x, y) dx \right| < \varepsilon.$$

Это свойство позволяет менять порядок интегрирования в повторном интеграле даже в том случае, когда один из пределов бесконечен. Под повторным интегралом мы подразумеваем интеграл вида

$$\int_{b_0}^{b_1} dy \int_{a_0}^{a_1} f(x, y) dx,$$

где сначала  $f(x, y)$  интегрируется по  $x$  от  $a_0$  до  $a_1$ , а затем результат интегрируется по  $y$  от  $b_0$  до  $b_1$ . Рассмотрим интеграл, в котором  $f(x, y)$  непрерывна по обоим переменным  $x$  и  $y$ :

$$I = \int_{b_0}^{b_1} dy \int_a^{\infty} f(x, y) dx = \int_{b_0}^{b_1} dy \left\{ \int_a^X f(x, y) dx + \int_X^{\infty} f(x, y) dx \right\}. \quad (1)$$

Поскольку все пределы интегрирования конечны, то

$$\int_{b_0}^{b_1} dy \int_a^X f(x, y) dx = \int_a^X dx \int_{b_0}^{b_1} f(x, y) dy. \quad (2)$$

(Это доказывается просто\*). Так как  $X$  находится в нашем распоряжении, выберем его так, чтобы для всех  $Y > X$  было

$$\left| \int_X^Y f(x, y) dx \right| < \omega. \quad (3)$$

---

\*) Для непрерывной функции  $f(x, y)$ ; в гл. 5 это утверждение будет установлено при несколько более широких предположениях.

Тогда абсолютная величина второго слагаемого формулы (1) не больше чем  $(b_1 - b_0)\omega$ , и

$$\left| I - \int_a^x dx \int_{b_0}^{b_1} f(x, y) dy \right| < (b_1 - b_0)\omega. \quad (4)$$

Так как для произвольно малого  $\omega$  можно найти такое значение  $X$ , что выполняется (3), получаем

$$I = \lim_{x \rightarrow \infty} \int_a^x dx \int_{b_0}^{b_1} f(x, y) dy = \int_a^\infty dx \int_{b_0}^{b_1} f(x, y) dy, \quad (5)$$

что и доказывает теорему.

Эта теорема может быть сформулирована следующим образом: *в равномерно сходящемся интеграле можно производить интегрирование под знаком интеграла.* Отсюда следует, что

$\int_0^\infty f(x, y) dx$  можно дифференцировать по  $y$  под знаком интеграла при условии, что существует производная  $\partial f / \partial y$  и интеграл от нее по  $x$  равномерно сходится в окрестности рассматриваемой точки  $y$ . Доказательство немедленно получится, если в последней теореме заменить  $f(x, y)$  на  $\partial f / \partial y$ .

Обобщение на случай равномерно сходящихся интегралов не обязательно от непрерывных функций  $f(x, y)$  может быть сделано на пути, указанном в конце разд. 1.113.

**1.121. М-признак.** Простейший признак равномерной сходимости аналогичен М-признаку для рядов. Если для всех  $y$ , таких, что  $b_0 \leq y \leq b_1$ ,

$$|f(x, y)| < g(x),$$

где  $\int_a^\infty g(x) dx$  сходится, то  $\int_a^\infty f(x, y) dx$  равномерно и абсолютно сходится при  $b_0 \leq y \leq b_1$ .

**1.122. Лемма Абеля для интегралов.** Если  $v(x)$  — неотрицательная, ограниченная при  $a \leq x \leq b$  и невозрастающая функция  $x$  и если  $h, H$  — нижняя и верхняя грани

$$F(\xi) = \int_a^\xi f(x) dx$$

при  $a \leq \xi \leq b$ , то

$$hv(a) \leq \int_a^b f(x)v(x)dx \leq Hv(a).$$

Положим

$$I = \int_a^b f(x)v(x)dx = \int_{x=a}^b v(x)dF(x) = v(b)F(b) - \int_{x=a}^b F(x)dv(x).$$

Равенство справедливо, потому что  $F(x)$  — интеграл, и, следовательно,  $F(x)$  непрерывна, а  $v(x)$  — функция ограниченной вариации. Далее, поскольку  $v(x)$  нигде не возрастает,  $I$  не уменьшится, если  $F(x)$  всюду заменить ее верхней гранью, и не увеличится, если  $F(x)$  заменить ее нижней гранью; итак,

$$h \left[ v(b) - \int_{x=a}^b dv(x) \right] \leq I \leq H \left[ v(b) - \int_{x=a}^b dv(x) \right],$$

т. е.

$$hv(a) \leq I \leq Hv(a).$$

В интегральной лемме Абеля необходимо специально требовать ограниченности  $v(x)$ ; будучи неотрицательной, функция  $v(x)$  должна иметь нижнюю грань, однако она может быть неограниченной вблизи  $x=a$ , если не оговорить это отдельно.

**1.123. Признак Абеля для бесконечных интегралов.** Если  $\int_0^\infty f(x)dx$  сходится (не обязательно абсолютно) и если для каждого значения  $y$  при  $b_0 \leq y \leq b_1$  функция  $v(x, y)$  неотрицательна, ограничена для всех  $x, y$  и никогда не возрастает по  $x$ , то  $\int_a^\infty f(x)v(x, y)dx$  сходится равномерно по  $y$  при  $b_0 \leq y \leq b_1$ .

Мы имеем  $0 \leq v(x, y) \leq M$ ; возьмем  $X$  так, что

$$\left| \int_X^{X'} f(x)dx \right| < \omega \quad \text{для всех } X' > X.$$

Тогда по лемме Абеля при  $b_0 \leq y \leq b_1$

$$\left| \int_X^{X'} f(x)v(x, y)dx \right| \leq \omega M,$$



откуда следует равномерная сходимость, поскольку  $\omega$  можно взять произвольно малым.

Например,  $\int_0^{\infty} \frac{\sin x}{x} dx$  сходится, а  $e^{-xy}$  положительна, ограничена и не возрастает по  $x$  при  $0 \leq y \leq \infty$ . Следовательно,  $\int_0^{\infty} e^{-xy} \frac{\sin x}{x} dx$  сходится равномерно при  $y \geq 0$ .

**1.124. Признак Дирихле—Харди для бесконечных интегралов.** Если  $\int_a^x f(x, y) dx$  ограничен при всех  $X > a$  и при  $b_0 \leq y \leq b_1$  и если  $v(x)$  ограничена, положительна, не возрастает и стремится к 0 при  $x \rightarrow \infty$ , то  $\int_a^{\infty} f(x, y) v(x) dx$  равномерно сходится при  $b_0 \leq y \leq b_1$ . Можно взять  $X$  так, что  $v(X) < \omega$  и для всех  $X' > X$  есть такое  $M$ , что

$$-M < \int_X^{X'} f(x, y) dx < M.$$

Тогда при  $b_0 \leq y \leq b_1$  и при всех  $X' > X$

$$\left| \int_X^{X'} f(x, y) v(x) dx \right| \leq M\omega.$$

Равномерная сходимость следует отсюда, как и раньше.

Отметим, что от  $\int_a^x f(x, y) dx$  не требуется, чтобы он стремился к пределу при  $X \rightarrow \infty$ ; он может осциллировать с конечным размахом.

Например,

$$\int_0^x \sin xy dx = \frac{1 - \cos Xy}{y},$$

и если  $|y| > \delta > 0$ , то это выражение по модулю меньше, чем  $2/\delta$ . Функция  $1/x$  положительна и стремится к 0 с возрастанием  $x$ . Следовательно,  $\int_a^{\infty} \frac{\sin xy}{x} dx$  равномерно сходится, если

$a > 0$ , в любой области, где  $|y| > \delta > 0$ . Ни в какой области, включающей  $y = 0$ , этот интеграл не является равномерно сходящимся. В действительности он равен  $+1/2\pi$  при  $y > 0$ ,  $-1/2\pi$  при  $y < 0$  и 0 при  $y = 0$ , так что, как и для рядов, неравномерная сходимость интеграла может быть связана с тем, что он является разрывной функцией параметра.

Равномерная сходимость интеграла от непрерывной функции — очень полезное условие, достаточное для непрерывности интеграла и для возможности интегрировать под знаком интеграла. Мы уже рассмотрели пример, когда неравномерная сходимость связана с разрывностью интеграла. Следующий пример, данный Курантом [13], показывает, что она может быть связана с невозможностью менять порядок интегрирования. Если

$$f(x, y) = (2 - xy) xye^{-xy},$$

то мы находим

$$\int_0^1 dx \int_0^\infty f(x, y) dy = 0, \quad \int_0^\infty dy \int_0^1 f(x, y) dx = 1.$$

Имеем

$$\int_0^y f(x, u) du = xy^2e^{-xy}.$$

Для любого  $x \neq 0$  это выражение стремится к 0 при  $y \rightarrow \infty$ , а при  $x = 0$  функция  $f(x, y) = 0$  при всех  $y$ . Итак,  $\int_0^\infty f(x, y) dy$  сходится, но неравномерно вблизи  $x = 0$ ; в самом деле, если  $\eta$  — наибольшее значение  $y$ , при котором  $xy^2e^{-xy} = \varepsilon$ , то  $xy^2e^{-xy} < \varepsilon$  при всех  $y > \eta$ , однако  $\eta$  стремится к бесконечности при  $x \rightarrow 0$ . Действительно,  $\int_0^y f(x, u) du$  неограничен по  $y$  при  $x \rightarrow 0$ .

Распространение этих результатов на случай, когда интегралы зависят не от двух, а от трех или более переменных, не дает ничего принципиально нового.

Иногда полезно следующее приложение признака Дирихле. Пусть

$$I = \int_a^\infty \cos[f(x)] dx,$$

где  $f'(x)$  — положительная возрастающая функция при  $x \geq a$  и  $f'(x) \rightarrow \infty$ , когда  $x \rightarrow \infty$ . Положим  $f(x) = y$ ,  $f'(x) = 1/g(y)$ ,  $f(a) = b$ . Тогда  $y$  — возрастающая функция  $x$  и

$$I = \int_b^{\infty} \cos y \cdot g(y) dy.$$

Но  $g(y) > 0$  и является убывающей функцией с пределом, равным 0. Следовательно, для сходимости интегралов  $\int_0^{\infty} \cos [f(x)] dx$ ,  $\int_0^{\infty} \sin [f(x)] dx$  достаточно, чтобы  $f'(x)$  была возрастающей функцией и стремилась к  $\infty$ . Например,

$$\int_0^{\infty} \cos x^2 dx, \quad \int_0^{\infty} \cos (x^3 - mx) dx \quad (m - \text{действительное})$$

сходятся. [Для последнего интеграла, если  $m > 0$ , возьмем  $a > \left(\frac{1}{3} m\right)^{1/2}$ .]

**1.125. Интегралы с верхним пределом, стремящимся к бесконечности.** Если  $f(x, n) \rightarrow g(x)$  и  $\lambda_n \rightarrow \infty$  при  $n \rightarrow \infty$ , то иногда нужно бывает найти условие, при котором

$$\int_a^{\lambda_n} f(x, n) dx \rightarrow \int_0^{\infty} g(x) dx. \quad (1)$$

Ясно, что вопрос этот связан с равномерной сходимостью; действительно, можно определить функцию

$$h(x, n) = f(x, n) \quad (a \leq x \leq \lambda_n), \quad h(x, n) = 0 \quad (\lambda_n < x). \quad (2)$$

Тогда

$$\int_a^{\lambda_n} f(x, n) dx = \int_a^{\infty} h(x, n) dx. \quad (3)$$

Следовательно, для выполнения (1) достаточно, чтобы  $h(x, n)$  удовлетворяла любому из признаков 1.121, 1.123 или 1.124.

Детальные доказательства признаков 1.123 и 1.124 в нужной форме и аналогичных признаков для рядов даны Бромвичем [14, стр. 123, 438, 443], называются они *теоремой Тэннери*.

**1.126. Изменение порядка суммирования для двойных рядов и повторных интегралов.** В случае равномерной сходимости

теорема, аналогичная 1.111, формулируется следующим образом (эта теорема принадлежит Муру):

Если  $f(x, y) \rightarrow h(x)$  при  $y \rightarrow \infty$  и  $f(x, y) \rightarrow g(y)$  равномерно при  $x \rightarrow \infty$ , то пределы

$$\lim_{x \rightarrow \infty} h(x), \quad \lim_{y \rightarrow \infty} g(y) \quad (1)$$

оба существуют и равны между собой. Возьмем такое  $X$ , что  $|f(x, y) - g(y)| < \omega$  при  $x \geq X$  и при всех  $y$ . Затем возьмем такое  $Y$ , что  $|f(X, y) - h(X)| < \omega$  при  $y \geq Y$ . Тогда при  $x > X$  и  $y \geq Y$

$$f(x, y) - h(X) = [f(x, y) - g(y)] - [f(X, y) - g(y)] + [f(X, y) - h(X)], \quad (2)$$

$$|f(x, y) - h(X)| < 3\omega, \quad |g(y) - h(X)| \leq 3\omega,$$

а следовательно, если  $y_1 > Y$ , то

$$|g(y_1) - g(Y)| \leq 6\omega. \quad (3)$$

Поскольку  $\omega$  произвольно,  $g(y)$  имеет предел  $F$ .

Следовательно, можно взять такое  $Y'$ , что  $|g(y) - F| \leq \omega$  при всех  $y \geq Y'$ . Если  $Y''$  — большее из чисел  $Y$  и  $Y'$  и если  $x > X$ ,  $y > Y''$ , то

$$|f(x, y) - g(y)| + |g(y) - F| \leq 2\omega. \quad (4)$$

Пусть  $y \rightarrow \infty$ ; тогда  $|h(x) - F| \leq 2\omega$  и  $h(x) \rightarrow F$ .

Ниже приводятся следствия для сумм двойных рядов, для сумм бесконечных интегралов и для повторных интегралов, аналогичные следствиям из 1.111.

Для рядов, если

$$f(m, n) = \sum_{r=1}^m \sum_{s=1}^n u(r, s), \quad (5)$$

условия таковы: когда  $n \rightarrow \infty$ ,  $f(m, n)$  сходится при любом  $m$ ; когда  $m \rightarrow \infty$ ,  $f(m, n)$  равномерно сходится к  $g(n)$  при всех  $n$ .

Если

$$f(m, x) = \sum_{r=1}^m \int_0^x u_m(\xi) d\xi, \quad (6)$$

условия таковы: интегралы сходятся при  $x \rightarrow \infty$  для любого  $m$ ; сумма для конечного  $x$  сходится равномерно при всех  $x$ , больших некоторого  $x_0$ . Если, наоборот, взять

$$f(m, x) = \int_0^x \left[ \sum_{r=1}^m u_m(\xi) \right] d\xi,$$

то условия таковы: ряд сходится, когда  $\xi$  изменяется на любом конечном интервале; интеграл сходится равномерно для всех  $m$ , больших  $m_0$ .

Если

$$f(x, y) = \int_0^x d\xi \int_0^y \Phi(\xi, \eta) d\eta = \int_0^y d\eta \int_0^x \Phi(\xi, \eta) d\xi, \quad (7)$$

то условия таковы: интеграл по  $\eta$  сходится на любом интервале по  $\xi$ ; второй интеграл сходится равномерно при всех  $y$ , больших некоторого  $Y$ .

**1.13. Теоремы о среднем.** Мы видели, что непрерывная функция на любом отрезке достигает своих верхней и нижней граней. Пусть  $f(x)$  непрерывна при  $a \leq x \leq b$  и имеет производную  $f'(x)$  при  $a < x < b$ , и пусть  $f(a) = f(b) = 0$ , а в некоторой внутренней точке  $f(\xi) > 0$ . Далее, пусть  $x = \eta$  соответствует верхней грани  $f(x)$  в интервале. Так как  $f(\eta) \geq f(\xi) > 0$ , то  $\eta$  не равно ни  $a$ , ни  $b$ . Далее,

$$f'(\eta) = \lim_{h \rightarrow 0} \frac{f(\eta + h) - f(\eta)}{h}.$$

Если  $h > 0$ , то  $f(\eta + h) \leq f(\eta)$ , и, следовательно,  $f'(\eta) \leq 0$ . Если  $h < 0$ , то  $f(\eta + h) \leq f(\eta)$ , и, следовательно,  $f'(\eta) \geq 0$ . Эти условия совместны только в случае, когда  $f'(\eta) = 0$ . При этом не требуется, чтобы  $f'(x)$  была непрерывна.

Если  $f(x)$  имеет на интервале отрицательную нижнюю грань, то мы получаем тот же результат, применяя доказательство к  $-f(x)$ . Следовательно, если  $f(x)$  имеет производную при  $a < x < b$  и непрерывна при  $a \leq x \leq b$ , то производная обращается в нуль между любыми двумя нулями  $f(x)$ . Результат этот известен под названием *теоремы Ролля*.

Пусть  $c, d$  — два произвольных значения  $x$ , такие, что  $f(x)$  непрерывна при  $c \leq x \leq d$  и имеет производную при  $c < x < d$ ; рассмотрим тогда функцию

$$g(x) = f(x) - f(c) - \frac{x-c}{d-c} [f(d) - f(c)].$$

Она обращается в нуль при  $x = c$  и  $x = d$ . Ее производная обращается поэтому в нуль при некотором  $x$  между  $c$  и  $d$ , скажем при  $x = \eta$ ; тогда

$$0 = g'(\eta) = f'(\eta) - \frac{f(d) - f(c)}{d - c}.$$

Таким образом,

$$f(d) - f(c) = (d - c) f'(\eta),$$

где  $c < \eta < d$ . Это *теорема о среднем* для производной. Геометрический смысл ее следующий: возьмем произвольную хорду,

соединяющую две точки гладкой кривой, тогда в некоторой промежуточной точке касательная параллельна этой хорде.

Наиболее важное приложение: *если  $f(x)$  непрерывна при  $a \leq x \leq b$  и  $f'(x) = 0$  при  $a < x < b$ , то  $f(x)$  постоянна на  $(a, b)$ .* В самом деле,  $f(x) - f(a) = (x - a)f'(\xi)$ , где  $a < \xi < x$ ; но  $f'(\xi) = 0$ , следовательно,  $f(x) = f(a)$ . Отметим недостаточность условия  $f'(x) = 0$  почти всюду. Известна функция, непрерывная на  $(0, 1)$ , с производной, почти всюду на  $(0, 1)$  равной нулю, и непостоянная на  $(0, 1)$ . Достаточно, однако, чтобы  $f'(x) = 0$  всюду, кроме конечного числа точек, в которых  $f(x)$  непрерывна.

Если  $f(x)$  непрерывна и  $F(x) = \int_a^x f(u) du$ , то

$$\frac{dF(x)}{dx} = f(x).$$

Если  $G'(x) = f(x)$ , то  $F(x) - G(x)$  является непрерывной функцией с производной, всюду равной нулю, следовательно, эта функция постоянна. Соответствующая теорема, в которой дана лишь интегрируемость  $f(x)$ , приведена в 1.103. Каждую из этих теорем можно использовать для проверки справедливости метода интегрирования, согласно которому сначала ищется функция с производной, равной  $f(x)$ . Другое следствие полезно в некоторых случаях, когда нужно узнать производную в точке  $x = a$ , но вычислить ее почему-либо трудно (например, из-за того, что интеграл или ряд, представляющий ее, не сходится). Если  $f(x)$  непрерывна и  $f'(x)$  существует всюду, кроме, быть может,  $x = a$ , то

$$\frac{f(a+h) - f(a)}{h} = f'(a + \theta h),$$

где  $0 < \theta < 1$ . Пусть теперь  $h$  стремится к нулю, тогда если левая часть стремится к пределу, то этот предел равен  $f'(a)$ . Но если  $f'(x)$  стремится к пределу, когда  $x$  стремится к  $a$ , то к этому пределу стремится правая часть, а левая часть, будучи равна правой, имеет тот же предел. Следовательно, *если  $f(x)$  непрерывна и  $f'(x)$  имеет предел при  $x$ , стремящемся к  $a$ , то  $f'(a)$  существует и равна этому пределу.*

Если  $\frac{1}{h} [f(a+h) - f(a)]$  имеет предел при  $h \rightarrow 0$  по положительным значениям, то этот предел можно назвать *производной  $f(x)$  справа* в точке  $a$  и обозначить через  $f'(a+)$ . Последнее доказательство также годится для того, чтобы показать, что если  $f(x)$  непрерывна справа в точке  $a$  и  $f'(a+h)$  имеет предел при  $h \rightarrow 0$  по положительным значениям, то  $f'(a+)$  су-

существует и равна этому пределу. Аналогичным свойством обладают производные слева. Утверждение о том, что  $f'(a)$  существует, эквивалентно утверждению о существовании и равенстве  $f'(a+)$  и  $f'(a-)$ .

**1.131. Первая теорема о среднем для интегралов.** Если для всех  $x$ , таких, что  $a \leq x \leq b$ ,  $m \leq f(x) \leq M$ , то

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a),$$

и, следовательно,

$$\int_a^b f(x) dx = N(b-a),$$

где  $N$  такое, что  $m \leq N \leq M$ . В частности, если  $f(x)$  непрерывна, то существует такое  $\xi$ , что  $f(\xi) = N$  и

$$\int_a^b f(x) dx = (b-a)f(\xi).$$

**1.132. Обобщение первой теоремы о среднем.** Чтобы получить теорему Тейлора, нам понадобится специальное обобщение первой теоремы о среднем; а именно если  $g(x) \geq 0$  при  $a \leq x \leq b$  и  $m \leq f(x) \leq M$ , то

$$m \int_a^b g(x) dx \leq \int_a^b f(x) g(x) dx \leq M \int_a^b g(x) dx.$$

Это можно переписать в виде

$$\int_a^b f(x) g(x) dx = \begin{cases} N \int_a^b g(x) dx & (m \leq N \leq M), \\ f(\xi) \int_a^b g(x) dx & (a \leq \xi \leq b); \end{cases}$$

последнее верно при некотором  $\xi$ , если  $f(x)$  непрерывна.

**1.133. Теорема Тейлора.** Пусть  $f(x)$  имеет производные до  $n$ -го порядка; обозначим  $n$ -ю производную через  $f^{(n)}(x)$ . Рассмотрим функцию

$$g(x) = f(x) - f(0) - xf'(0) - \frac{x^2}{2!} f''(0) - \dots - \frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0).$$

Дифференцируя последовательно  $(n-1)$  раз, находим, что  $g(x)$  и ее производные до  $(n-1)$ -го порядка обращаются в 0 при  $x=0$ . Кроме того,

$$g^{(n)}(x) = f^{(n)}(x).$$

Рассмотрим теперь интеграл

$$h(x) = \int_0^x \frac{(x-u)^{n-1}}{(n-1)!} f^{(n)}(u) du,$$

где  $f^{(n)}(u)$  предполагается ограниченной и интегрируемой. (Почему выбирается такое выражение, выясняется при обсуждении операционных методов.)

Повторным интегрированием по частям находим

$$\begin{aligned} h(x) &= \int_0^x \frac{(x-u)^{n-1}}{(n-1)!} df^{n-1}(u) = \left[ \frac{(x-u)^{n-1}}{(n-1)!} f^{(n-1)}(u) \right]_0^x + \\ &+ \int_0^x \frac{(x-u)^{n-2}}{(n-2)!} f^{(n-1)}(u) du = -\frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0) + \\ &+ \int_0^x \frac{(x-u)^{n-2}}{(n-2)!} f^{(n-1)}(u) du = -\frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0) - \dots \\ &\dots - \frac{x^2}{2!} f''(0) - xf'(0) + \int_0^x f'(u) du = g(x). \end{aligned}$$

Следовательно,

$$\begin{aligned} f(x) &= f(0) + xf'(0) + \frac{x^2}{2!} f''(0) + \dots \\ &\dots + \frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0) + \int_0^x \frac{(x-u)^{n-1}}{(n-1)!} f^{(n)}(u) du. \end{aligned}$$

Это точная форма теоремы Тейлора. Не требуется, чтобы  $f^{(n)}(u)$  была непрерывна. Если  $m$  и  $M$  — нижняя и верхняя грани  $f^{(n)}(u)$  при  $0 \leq u \leq x$ , то, поскольку

$$\int_0^x \frac{(x-u)^{n-1}}{(n-1)!} du = \frac{x^n}{n!},$$

интеграл в формуле Тейлора лежит между  $mx^n/n!$  и  $Mx^n/n!$ . Следовательно, он равен  $Nx^n/n!$ , где  $m \leq N \leq M$ . Если далее



$f^{(n)}(u)$  непрерывна, то найдется значение  $\theta x$  ( $0 < \theta < 1$ ), при котором она равна  $N$ . Интеграл можно тогда записать в виде

$$\frac{x^n}{n!} f^{(n)}(\theta x).$$

Остаточный член записан здесь в форме Лагранжа. Форма  $Nx^n/n!$  требует лишь интегрируемости  $n$ -й производной.

Другая форма остаточного члена, принадлежащая Коши, получится, если заметить, что в случае, когда  $f^{(n)}(u)$  непрерывна, существует  $\theta$ , такое, что  $0 \leq \theta \leq 1$  и

$$\int_0^x (x-u)^{n-1} f^{(n)}(u) du = x(x-\theta x)^{n-1} f^{(n)}(\theta x).$$

Это равенство вытекает из первой теоремы о среднем для интегралов; отсюда остаточный член можно записать в виде

$$\frac{x^n}{(n-1)!} (1-\theta)^{n-1} f^{(n)}(\theta x).$$

Мы увидим, что приведенные формы теоремы Тейлора не предполагают сходимости бесконечного ряда. Однако их широко используют при доказательстве сходимости (показывая с их помощью, что остаточный член стремится к нулю для больших  $n$ ) и при оценке ошибки, возможной для заданного конечного числа членов.

**1.134. Вторая теорема о среднем для интегралов.** Выведем сначала простое следствие из леммы Абеля для интегралов. Пусть  $f(x)$  интегрируема в интервале  $(a, b)$ , а  $\varphi(x)$  имеет ограниченную вариацию. Пусть  $P(x)$  и  $N(x)$  — положительная и отрицательная вариации  $\varphi(x)$  в  $(a, x)$ , и пусть максимум из  $P(b)$  и  $N(b)$  равен  $\omega$ . Тогда  $\omega - P(x)$  и  $\omega - N(x)$  удовлетворяют условиям, наложенным на  $v(x)$  в лемме Абеля, и если  $h, H$  — нижняя и верхняя грани  $\int_a^x f(t) dt$  при  $a \leq x \leq b$ , то

$$\omega h \leq \int_a^b [\omega - P(x)] f(x) dx \leq \omega H, \quad (1)$$

$$\omega h \leq \int_a^b [\omega - N(x)] f(x) dx \leq \omega H. \quad (2)$$

Вычитая, получаем

$$\left| \int_a^b f(x) [\varphi(x) - \varphi(a)] dx \right| \leq \omega(H - h). \quad (3)$$

Следовательно,  $\int_a^b f(x) \varphi(x) dx$  можно приближенно заменить на  $\varphi(a) \int_a^b f(x) dx$ ; ошибка, которую мы при этом совершим, оценена сверху. Этот результат особенно важен в теории рядов и интегралов Фурье.

Под второй теоремой о среднем обычно подразумевают одну из следующих теорем.

**Форма Боннэ:** если  $\varphi(x)$  положительна и не возрастает при  $a \leq x \leq b$ , то существует  $\eta$  из  $a \leq \eta \leq b$ , такое, что

$$\varphi(a) \int_a^{\eta} f(x) dx = \int_a^b f(x) \varphi(x) dx. \quad (4)$$

**Форма дю Буа-Раймона:** если  $\varphi(x)$  монотонна при  $a \leq x \leq b$ , то существует  $\xi$  из  $a \leq \xi \leq b$ , такое, что

$$\int_a^b f(x) \varphi(x) dx = \varphi(a) \int_a^{\xi} f(x) dx + \varphi(b) \int_{\xi}^b f(x) dx. \quad (5)$$

Обе теоремы легко выводятся из леммы Абеля, однако, как отмечает Бромвич [14, стр. 426 и 427], они не несут информации, которая не содержалась бы в самой лемме, поскольку в них не указывается способ оценки  $\xi$  и  $\eta$  и они менее информативны, чем приведенная выше формула (3) (см. приложение к гл. 1).

#### 1.14. Бесконечные произведения. Если

$$\prod_n = (1 + a_1)(1 + a_2) \dots (1 + a_n), \quad (1)$$

то предел  $\prod_n$  при  $n$ , стремящемся к бесконечности, если он существует и не равен нулю, обозначается через

$$\prod_1^{\infty} (1 + a_n); \quad (2)$$

в этом случае говорят, что бесконечное произведение сходится. Оно равно нулю, если любой из множителей — нуль. (В по-

следнем случае часто говорят, что произведение *сходится к нулю*, чтобы отличить его от такого произведения, как

$$\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{2}\right) \dots,$$

которое стремится к нулю, хотя все сомножители отличны от нуля. Про такие произведения, как последнее, говорят, что они *расходятся к нулю*.)

Теория сходимости бесконечных произведений тесно связана с теорией сходимости для бесконечных рядов. Действительно, если все  $a_n$  положительны или если все они отрицательны, то для сходимости произведения (2) необходимо и достаточно, чтобы сходилась ряд  $\sum a_n$ . Имеем

$$S_n = \ln \prod_n = \sum_1^n \ln(1 + a_r). \quad (3)$$

Ясно, что ни  $\prod (1 + a_n)$ , ни  $\sum a_n$  не могут сходиться без того, чтобы  $a_n \rightarrow 0$ . Поэтому мы должны рассмотреть только случай  $a_n \rightarrow 0$ . В этом случае

$$\frac{\ln(1 + a_n)}{a_n} \rightarrow 1. \quad (4)$$

Следовательно, отношения соответствующих членов двух рядов  $\sum \ln(1 + a_n)$  и  $\sum a_n$  ограничены, и эти ряды сходятся или нет одновременно. Но если  $S_n$  имеет предел  $S$ , то  $\prod_n$  имеет предел  $e^S$ , и наоборот, если  $\prod_n$  стремится к пределу, отличному от  $\infty$  и 0, то  $S_n$  тоже стремится к пределу; утверждение доказано.

Если  $a_n$  не все одного знака и  $\sum |a_n|$  сходится, то легко показать, что  $\prod (1 + a_n)$  сходится. Если  $\sum |a_n|$  не сходится, но  $a_n \rightarrow 0$ , то можно выбрать  $m$  так, что при  $n > m$

$$\ln(1 + a_n) = a_n - \frac{1}{2} \lambda_n a_n^2,$$

где  $c < |\lambda_n| < d$ , а  $c$  и  $d$  фиксированы; поэтому если  $\sum a_n^2$  сходится и  $\sum a_n$  условно сходится, то  $\prod (1 + a_n)$  сходится.

Доказательство не проходит для такого произведения, как  $\prod [1 + (-1)^n / \sqrt{n}]$ . Здесь  $\sum a_n$  сходится, а  $\sum a_n^2$  нет. Легко доказать, что указанное произведение не сходится.

**1.15. Условие Липшица.** Если для данного  $x$  и для всех  $\xi$ , удовлетворяющих неравенству  $|\xi - x| < \delta$ ,

$$|f(\xi) - f(x)| \leq A |\xi - x|^\alpha,$$

где  $A$  и  $\alpha$  не зависят от  $\xi$  и  $\alpha > 0$ , то говорят, что  $f(\xi)$  удовлетворяет при  $\xi = x$  условию Липшица порядка  $\alpha$ . Если  $f(\xi)$  удовлетворяет условию Липшица, то она непрерывна при  $\xi = x$ ; если она удовлетворяет этому условию при всех  $x$  из отрезка  $a \leq x \leq b$ , то она непрерывна при  $a \leq x \leq b$ . Но даже при  $\alpha = 1$  функция не обязана быть дифференцируемой или иметь ограниченную вариацию. Например, положим

$$f(0) = 0, \quad f(x) = x \sin \frac{1}{x}.$$

Эта функция удовлетворяет условию Липшица порядка 1 даже при  $x = 0$ ; однако она не является дифференцируемой при  $x = 0$  и не является функцией ограниченной вариации ни в каком интервале, содержащем  $x = 0$ . Функция  $|x|$  удовлетворяет условию Липшица порядка 1 при  $x = 0$  и в любом интервале имеет ограниченную вариацию. Однако она не дифференцируема при  $x = 0$ . В ряде теорем обнаруживается, что условие Липшица является достаточным, в то время как требование непрерывности недостаточно, а условие дифференцируемости достаточно, но не является необходимым.

Если в точке  $x$  удовлетворяется условие Липшица порядка  $\alpha > 1$ , то, очевидно,  $f'(x) = 0$ . Если это условие выполняется в каждой точке интервала для некоторого  $\alpha > 1$ , то  $f'(x) = 0$  на всем интервале. Следовательно,  $f(x) = \text{константа}$ ; тем самым интерес представляет только случай  $0 < \alpha \leq 1$ .

Важный класс функций составляют функции  $f(x)$ , удовлетворяющие условию Липшица порядка 1 равномерно на  $(a, b)$ ; имеется в виду, что существует такая константа  $A$ , что  $|f(x_2) - f(x_1)| \leq A|x_2 - x_1|$  для всех  $x_1, x_2$  из  $(a, b)$ . Ясно, что в этом случае  $f(x)$  удовлетворяет обоим условиям: непрерывна и имеет ограниченную вариацию. Она не обязательно дифференцируема во всех точках  $(a, b)$ , как видно на примере  $f(x) = |x|$  при  $x$  в интервале  $(-1, 1)$ . Можно доказать, однако, что  $f(x)$  имеет производную почти всюду и является интегралом Лебега от ограниченной функции. Это утверждение является частным случаем принадлежащей В. и Г. Юнгам [15] довольно трудной теоремы о том, что функция ограниченной вариации имеет производную почти всюду\*). Указанный выше специальный случай эквивалентен утверждению о том, что кривая конечной длины почти всюду имеет касательную; элементарное доказательство этого факта дал Безикович [16].

---

\*) В классе всех непрерывных функций это неверно.

**1.16. Неравенство Коши \*).** Если  $a_1, \dots, a_n, b_1, \dots, b_n$  — произвольные числа, то

$$(a_1^2 + \dots + a_n^2)(b_1^2 + \dots + b_n^2) - (a_1 b_1 + \dots + a_n b_n)^2 = (a_1 b_2 - a_2 b_1)^2 + \dots$$

Это равенство является обобщением тождества Лагранжа. Таким образом,

$$\sum_1^n a_r^2 \sum_1^n b_r^2 \geq \left( \sum_1^n a_r b_r \right)^2.$$

Это соотношение называется неравенством Коши. Из него следует, что если  $\varphi(x)$  и  $\psi(x)$  — две функции, то

$$\int_a^b \varphi^2(x) dx \int_a^b \psi^2(x) dx \geq \left[ \int_a^b \varphi(x) \psi(x) dx \right]^2.$$

Это неравенство Шварца. В частности, если  $\psi = 1$ ,  $\varphi(x) = |f(x)|$ , то

$$(b-a) \int_a^b f^2(x) dx \geq \left[ \int_a^b |f(x)| dx \right]^2.$$

Точно так же при  $n = 2$  если  $b_1^2 + b_2^2 = 1$ , то  $(a_1 b_1 + a_2 b_2)^2 \leq a_1^2 + a_2^2$ . Все эти и сходные неравенства находят многочисленные применения.

### ПРИМЕРЫ

1. Какие из аксиом поля, сформулированных в 1.01, не выполняются для 1) множества всех целых положительных чисел, 2) множества, состоящего из 0 и всех корней квадратных из целых положительных чисел?

2. Один студент (разумеется, не из Кембриджа) будто бы отыскал такой маршрут из дома на лекции и обратно, который и туда и назад идет под гору. Какой аксиоме это противоречит?

3. Если  $s_n \geq s_{n-1}$ ,  $t_n \geq t_{n-1}$ , если для каждого  $m$  существует такое  $n$ , что  $t_n > s_m$ , и для каждого  $m$  существует такое  $p$ , что  $s_p > t_m$ , и если  $\{s_n\}$  ограничена, то  $\{t_n\}$  сходится к тому же пределу, что  $\{s_n\}$ .

4. Доказать, что если  $a_0 = 3$ ,  $a_{n+1} = 3 - \frac{2}{a_n}$ , то  $a_n \rightarrow 2$ . Показать графически, используя кривые  $y = 3 - \frac{2}{x}$ ,  $y = x$ , что последовательность  $\{a_n\}$  имеет предел, равный 2, при всех значениях  $a_0$ , кроме  $a_0 = 1$ . (Разрешается, чтобы  $a_n$  принимали бесконечные значения.) (I.C., 1938.)

5. Если  $s_{n+1} = \sqrt{2s_n + a}$ , где  $s_1$  и  $a$  положительны и берется положительное значение квадратного корня, то при  $n$ , стремящемся к бесконечности,  $s_n$  стремится к пределу, равному  $1 + \sqrt{a+1}$ . Доказать. (М. Т., 1940.)

\*) В русской литературе Коши — Буняковского. — Прим. ред.

6. Доказать, что при фиксированном  $s > 1$  и при  $n \rightarrow \infty$

$$\sum_{v=1}^n \left( \frac{s}{v} - \frac{1}{v^s} \right) - s \ln n$$

стремится к пределу; если этот предел равен  $\varphi(s)$ , то

$$0 < \varphi(s) + \frac{1}{s-1} \leq s-1. \quad (\text{М. Т., 1938.})$$

7. Показать, что положительный корень уравнения

$$x^4 + 4x - 1 = 0$$

можно найти, рассматривая сходящуюся последовательность

$$x_{n+1} = \frac{1}{x_n^3 + 4}.$$

Определить четыре десятичных знака этого корня. (И. С., 1937.)

8. Решить следующие разностные уравнения

1)  $y_{n+1} - 11y_n + y_{n-1} = 0$ , где  $y_0 = 0$ ,  $y_1 = 1$ ,

2)  $y_{n+1} + 5y_n + 2z_n = 0$ ,  $z_{n+1} + 2z_n + 2y_n = 0$ . (И. С., 1941.)

9. Выразив  $\sin m\theta$  и  $\operatorname{sh} mu$  через экспоненты, доказать тождество

$$\sum_{m=1}^{\infty} \frac{\sin m\theta}{\operatorname{sh} mu} = \sum_{n=1}^{\infty} \frac{\sin \theta}{\operatorname{ch} (2n-1)u - \cos \theta} \quad (u > 0, \theta - \text{действительное}).$$

(М. Т., 1938.)

10. Если  $f(x)$  равна 0 при иррациональных  $x$  и равна 1 при рациональных  $x$ , то  $xf(x)$  непрерывна при  $x=0$  (и нигде больше не является непрерывной), а  $x^2f(x)$  дифференцируема при  $x=0$  (и нигде больше не является дифференцируемой). Доказать.

11. Доказать, что при  $a^2 \neq 1$

$$\frac{1-a^{2n}}{1-a^2} = \prod_{r=1}^{n-1} \left( 1 - 2a \cos \frac{r\pi}{n} + a^2 \right).$$

Пусть  $a$  действительно; доказать, что

$$\int_0^{\pi} \ln(1 - 2a \cos x + a^2) dx = 0$$

при  $|a| < 1$ , и найти значение этого интеграла при  $|a| > 1$ . (М. Т., 1939.)

12. Доказать, что сумма ряда

$$\frac{|x|}{1+|x|} + \frac{|x|}{(1+|x|)^2} + \frac{|x|}{(1+|x|)^3} + \dots$$

существует при всех действительных  $x$ , но является разрывной функцией. (М. Т., 1938.)

13. При каких значениях  $x$  каждый из следующих рядов равномерно сходится?

$$1) \sum_1^{\infty} \frac{(-1)^n}{n^x}, \quad 2) \sum \frac{1}{n} \left( 1 + \frac{1}{2} + \dots + \frac{1}{n} \right) \sin nx. \quad (\text{М/с, III, 1928.})$$

14. Доказать, что

$$\ln 3 = 1 + \frac{1}{2} - \frac{2}{3} + \frac{1}{4} + \frac{1}{5} - \frac{2}{6} + \frac{1}{7} + \frac{1}{8} - \frac{2}{9} + \dots,$$

и получить аналогичные ряды для  $\ln a$ , где  $a$  — натуральное число. (Указание: использовать признак Абеля для равномерной сходимости.) (М/с, III, 1930.)

15. Показать, что для положительных  $n$

$$\lim_{n \rightarrow \infty} x \left( \frac{1}{n+x} + \frac{1}{n+2x} + \dots + \frac{1}{n+nx} \right) = \ln(1+x). \quad (\text{I. С., 1938.})$$

16. Доказать, что биномиальный ряд

$$1 + \sum_{r=1}^{\infty} \frac{n(n+1) \dots (n+r-1)}{r!} x^r$$

сходится при  $|x| < 1$  и неограничен при  $|x| > 1$ . Доказать далее, что 1) если  $x = 1$ , то ряд сходится при  $n \leq 0$  и неограничен при  $n > 0$ , 2) если  $x = -1$ , то ряд сходится при  $n < 1$ , неограничен при  $n > 1$  и осциллирует при  $n = 1$ .

17. Показать, что  $\sum u_n$  сходится, если  $|u_n|^{1/n} < k < 1$  при  $n > m$ . Применить это правило к рядам

$$u_{2n} = 2^{-2n}, \quad u_{2n+1} = 3^{-2n-1}.$$

Устанавливает ли сходимость этого ряда правило 1.117? Если нет, то как это правило нужно обобщить, чтобы из него следовала сходимость?

18. Показать, что ряд

$$\sum_1^{\infty} \sum_1^{\infty} \frac{l+m}{(l^2+m^2)^s}$$

сходится или не сходится в зависимости от того, какое из неравенств справедливо:  $s > \frac{3}{2}$  или  $s \leq \frac{3}{2}$ . (М/с, III, 1932.)

19. Для двух рядов

$$S = \sum_1^{\infty} n^{-4}, \quad T = \sum_1^{\infty} 2^{-n}$$

найти, как велико должно быть  $n$ , чтобы ошибка, возникающая при замене ряда  $n$ -й частичной суммой, оказалась 1)  $< 0,005$ , 2)  $< 0,0000005$ .

20. Доказать, что ряд

$$\sum \frac{1}{n (\ln n)^p}$$

сходится при  $p > 1$  и стремится к бесконечности при  $p \leq 1$ .

Вывести отсюда, что если

$$u_n > 0, \quad \frac{u_n}{u_{n-1}} \rightarrow 1, \quad n \left( 1 - \frac{u_n}{u_{n-1}} \right) \rightarrow 1$$

и

$$\ln n \left[ n \left( 1 - \frac{u_n}{u_{n-1}} \right) - 1 \right] \rightarrow k > 1,$$

то ряд  $\sum u_n$  сходится.

21. Исследовать сходимость ряда

$$1 + \frac{ax}{b} + \frac{a(a+2)}{b(b+1)} x^2 + \frac{a(a+2)(a+4)}{b(b+1)(b+2)} x^3 + \dots$$

при положительных  $a, b, x$ . (И. С., 1944.)

22. Показать, что

$$1 + \frac{\alpha\beta}{2\gamma} x + \dots + \frac{\alpha(\alpha+1) \dots (\alpha+n-1) \beta(\beta+1) + (\beta+n-1)}{n! \gamma(\gamma+1) \dots (\gamma+n-1)} x^n + \dots$$

сходится, если  $0 \leq x < 1$  и если  $x=1$  и  $\gamma > \alpha + \beta$ .

23. Доказать, что произведение двух функций, интегрируемых по Риману, есть функция, интегрируемая по Риману.

24. Доказать результат 1.104 с помощью интегрирования по частям.

25. Пусть

$$S = \sum_{n=1}^{\infty} \frac{1}{n^2 + 1},$$

доказать, что

$$\frac{1}{4} + \frac{1}{4} \pi < S < \frac{1}{2} + \frac{1}{4} \pi.$$

26. Пусть

$$f(x) = \sum_{n=0}^N a_n \cos \lambda_n x \pi, \quad g(x) = \sum_{n=0}^N a_n \cos \left( \lambda_n x + \frac{1}{x} \right) \pi,$$

где  $a_n$  и  $\lambda_n$  — действительные числа. Доказать, что если одна из функций:  $f(x)$  или  $g(x)$  не обращается в нуль на интервале  $\frac{1}{m+1} \leq x \leq \frac{1}{m}$  (где  $m$  — натуральное число), то другая имеет по крайней мере один нуль внутри этого интервала. (М. Т., 1938.)

27. Доказать, что

$$\int_1^{\infty} dx \int_1^{\infty} \frac{x-y}{(x+y)^3} dy = -\frac{1}{2}, \quad \int_1^{\infty} dy \int_1^{\infty} \frac{x-y}{(x+y)^3} dx = \frac{1}{2}. \quad (\text{М. Т., 1913.})$$

28. Исследовать сходимость бесконечных произведений  $\prod (1 + u_n)$ , если

$$1) u_n = \frac{(-1)^n}{n^{1/3}}, \quad 2) u_n = \frac{(-1)^n}{\ln(n+1)}. \quad (\text{М/с, III, 1928.})$$



29. Рассмотреть  $f(x) = \sqrt{x} \cos(1/x)$ ,  $g(x) = \sqrt{x} \sin(1/x)$  и показать, что непрерывность  $f(x)$  и  $g(x)$  не достаточна для существования интеграла

$$\int_{x=0}^1 f(x) dg(x).$$

30. Пусть  $f(x)$  имеет производные вплоть до  $(n-1)$ -й при  $-a < x < a$  и  $f^{(n)}(0)$  существует. Доказать, что

$$f(x) = f(0) + \sum_{r=1}^n f^{(r)}(0) \frac{x^r}{r!} + o(x^n).$$

31. Пусть при  $a \leq x \leq b$   $f(x)$  имеет производную  $f'(x)$ ; тогда если  $f'(a) < p < f'(b)$ , то имеется такое  $\xi$ , что  $a < \xi < b$ , и  $f'(\xi) = p$  [ $f'(\xi)$  не предполагается непрерывной].

32. Пусть  $f_n(x)$  — убывающие функции  $x$  на  $(a, b)$  или  $(-\infty, \infty)$  и равномерно ограничены по  $n$ , а также  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ . Доказать равномерную сходимость в любом интервале, не содержащем разрывов  $f(x)$ .

## ЛИТЕРАТУРА

1. *Henderson J. B.*, Engineering, **116**, 409—410 (1923).
2. *Knopp K.*, Theorie und Anwendung der unendlichen Reihen, Springer, 1947.
3. *Hardy G. H.*, Pure Mathematics. (Русский перевод: Г. Харди, Курс чистой математики, М., ИЛ, 1949.)
4. *Leathem J. G.*, Volume and Surface Integrals used in Physics, 1905.
5. *Heine E. J.*, Reine angew. Math., **74**, 188 (1871).
6. *Young W. H.*, Proc. Lond. Math. Soc., (1) **35**, 384—388 (1903).
7. *Baker H. F.*, Proc. Lond. Math. Soc., (1), **35**, 459 (1903).
8. *Stieltjes T. J.*, Ann. d. Fac. d. Sci., Toulouse, **8**, 68—75 (1894).
9. *Widder D. V.*, The Laplace Transform, 1941, ch. 1.
10. *Pollard S.*, Quart. J. Math., **49**, 73—138 (1923).
11. *Burkill J. C.*, Cambridge Math. Tracts, № 40, 1951.
12. *Titchmarsh E. C.*, The Theory of Functions, 1932, ch. X, XI, XII.
13. *Courant R.*, Differential and Integral Calculus, **2** (1936), p. 316. (Русский перевод: Р. Курант, Курс дифференциального и интегрального исчисления, М.-Л., ОНТИ, 1934.)
14. *Bromwich T. J.*, Theory of Infinite Series, 1908, p. 123, 443, 426, 427.
15. *Young W. H.*, *Young G. C.*, Proc. Lond. Math. Soc., (2), **9**, 325—335 (1911).
16. *Besicovitch A. S.*, J. Lond. Math. Soc., **19**, 205—207 (1944).

## ПРИЛОЖЕНИЕ К ГЛАВЕ 1

**1.116а. Теорема об ограниченной сходимости.** Сначала нам понадобятся несколько определений. Для любого конечного множества  $I$  неперекрывающихся интервалов обозначим через  $lI$  сумму их длин. Для бесконечного множества  $I$  интервалов, расположенных внутри  $(a, b)$ , обозначим через  $lI$  верхнюю грань сумм длин, взятую по всем подмножествам  $I$ , состоящим

из конечного числа интервалов \*). Если  $E$  — некоторое множество точек из  $(a, b)$ , то дополнительное множество  $CE$  определяется как множество точек  $(a, b)$ , не принадлежащих  $E$ . Дополнение к конечному множеству замкнутых неперекрывающихся интервалов состоит из конечного числа открытых неперекрывающихся интервалов. Если все точки  $E_1$  являются точками  $E_2$ , мы пишем  $E_1 \subset E_2$  (читается:  $E_1$  принадлежит  $E_2$ ).

*Лемма 1. Если  $I$  — множество неперекрывающихся интервалов, принадлежащих  $(a, b)$ , и  $\delta$  — произвольное (как угодно малое) положительное число, то существует конечное множество  $J$  интервалов из  $I$ , такое, что  $II \geqslant IJ > II - \delta$ . Это сразу следует из того факта, что  $II$  является верхней гранью  $IJ$  по всем конечным множествам.*

*Лемма 2. Если выполнены условия предыдущей леммы, то существует конечное множество замкнутых интервалов  $K$ , принадлежащее  $I$  и такое, что  $IK > II - 2\delta$ .*

В самом деле, пусть  $J$  из леммы 1 состоит из  $m$  интервалов. Для любого интервала  $(\alpha_i, \beta_i)$  из  $J$ , такого, что  $\beta_i - \alpha_i > \delta/m$ , определим интервал из  $K$  следующим образом: возьмем замкнутый интервал  $(\alpha_i + \delta/2m, \beta_i - \delta/2m)$ . Такое множество интервалов  $K$  обладает требуемым свойством.

Множество  $I$  является объединением  $K$  и другого множества неперекрывающихся интервалов  $K'$ , такого, что  $II = IK + IK'$ .

*Лемма 3. Пусть  $\{I_n\}$  — последовательность множеств неперекрывающихся интервалов, принадлежащих  $(a, b)$ , такая, что  $I_n \subset I_{n-1}$  и ни одно  $x$  из  $(a, b)$  не принадлежит всем  $I_n$ . Тогда  $II_n \rightarrow 0$ . По заданному  $\delta > 0$  можно найти принадлежащее  $I_n$  множество  $K_n$ , состоящее из конечного числа замкнутых интервалов и такое, что  $IK_n > II_n - 2^{-n}\delta$ . Множество точек  $I_n$ , не принадлежащих  $K_n$ , обозначим через  $K'_n$ . Пусть  $L_n$  — множество точек, общих для  $K_1, K_2, \dots, K_n$ . Тогда каждая точка  $I_n$  является точкой  $I_1, I_2, \dots, I_n$  и, следовательно, либо точкой  $L_n$ , либо точкой по крайней мере одного из множеств  $K'_1, K'_2, \dots, K'_n$ . Значит,  $II_n \leqslant IL_n + IK'_1 + \dots + IK'_n < IL_n + \delta$ . Следовательно,  $L_n$  состоит из конечного числа неперекрывающихся замкнутых интервалов (и, следовательно, замкнуто)*

$$L_n \subset L_{n-1} \text{ и } IL_n > II_n - \delta.$$

Так как никакая точка  $(a, b)$  не принадлежит всем  $I_n$ , то каждая точка принадлежит некоторому  $CI_n$ , а следовательно,

---

\*) Предполагается, конечно, что  $I$  состоит из неперекрывающихся интервалов. — Прим. перев.

и некоторому  $CL_n$ , поскольку  $CI_n \subset CL_n$ . Она не может быть концевой точкой ни для какого интервала из  $CL_n$ , потому что концевые точки  $CL_n$  являются одновременно концевыми точками  $L_n$  и, следовательно, принадлежат  $L_n$ , поскольку  $L_n$  замкнуто. Значит, для каждой точки  $x$  из  $(a, b)$  найдется такое  $n$ , что  $x$  лежит внутри некоторого интервала из  $CL_n$ . Рассмотрим множество всех таких интервалов. По теореме Гейне — Бореля существует конечное число интервалов  $d_1, d_2, \dots, d_k$  (каждый из которых есть часть некоторого  $CL_n$ ), которые покрывают  $(a, b)$ . Пусть  $d_r$  — часть  $CL_{n_r}$ , и пусть  $n_0$  — наибольшее из чисел  $n_1, \dots, n_k$ . Так как  $CL_n \subset CL_{n+1}$ , то  $CL_{n_0}$  включает все  $d_r$ ; следовательно,  $CL_{n_0}$  — это весь интервал  $(a, b)$ , и  $L_n$  пусто для всех  $n \geq n_0$ . Отсюда  $lL_n = 0$ ,  $lL_n < \delta$  при всех  $n \geq n_0$ . Поскольку  $\delta$  произвольно,  $lL_n \rightarrow 0$ .

*Лемма 4.* Пусть  $f_n(x)$  неотрицательны и интегрируемы на  $(a, b)$ ,  $f_n(x) < M$  при всех  $x$  и  $n$  и  $f_n(x) \rightarrow 0$  при всех  $x$ . Тогда  $\int_a^b f_n(x) dx \rightarrow 0$ . Для каждого  $n$  произведем такое разбиение  $(a, b)$ , чтобы соответствующая нижняя интегральная сумма (см. 1.101) отличалась от  $\int_a^b f_n(x)$  меньше, чем на  $1/n$ . В каждой точке разбиения положим  $g_n(x) = 0$ , а в каждом частичном интервале разбиения возьмем  $g_n(x)$  равной нижней грани  $f_n(x)$  по точкам этого интервала. Тогда

$$f_n(x) \geq g_n(x) \geq 0; \quad 0 \leq \int_a^b \{f_n(x) - g_n(x)\} dx < \frac{1}{n};$$

$$g_n(x) \rightarrow 0.$$

При заданном  $\varepsilon > 0$  обозначим через  $I_n$  множество всех  $x$ , где  $g_p(x) > \varepsilon$  хотя бы для одного  $p \geq n$ . Внутри любого интервала, построенного для  $f_p(x)$ , функция  $g_p(x)$  постоянна; следовательно,  $g_p(x) > \varepsilon$  во всем частичном интервале, если она больше  $\varepsilon$  в одной точке. Таким образом,  $I_n$  — множество, состоящее из неперекрывающихся интервалов, и  $I_n \subset I_{n-1}$ . Ни одно  $x$  не принадлежит всем  $I_n$ , иначе оказалось бы, что  $f_p(x) \geq g_p(x) > \varepsilon$  для бесконечной последовательности значений  $p$  и последовательность  $f_n(x)$  не стремилась бы к 0. Следовательно,  $I_n$  удовлетворяет условиям леммы 3 и  $lI_n \rightarrow 0$ . Возьмем  $n_0$ , так чтобы при  $n \geq n_0$  выполнялось неравенство

$II_n < \delta$ . Тогда при  $n \geq n_0$

$$\int_a^b g_n(x) dx \leq \varepsilon(b-a-\delta) + M\delta,$$

$$\int_a^b f_n(x) dx \leq \varepsilon(b-a-\delta) + M\delta + \frac{1}{n}.$$

Но  $\varepsilon$ ,  $\delta$  и  $1/n$  можно взять произвольно малыми. Следовательно, лемма доказана.

*Теорема.* Если  $f_n(x)$  интегрируемы на  $(a, b)$ ,  $|f_n(x)| < M$  и  $f_n(x) \rightarrow f(x)$ , где  $f(x)$  интегрируема, то  $\int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx$ . Возьмем  $h_n(x) = |f_n(x) - f(x)|$ . Тогда  $h_n(x)$  удовлетворяет условиям леммы 4, и, следовательно,  $\int_a^b h_n(x) dx \rightarrow 0$ . Но

$$\left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| \leq \int_a^b |f_n(x) - f(x)| dx,$$

откуда

$$\int_a^b f_n(x) dx \rightarrow \int_a^b f(x) dx.$$

Следует отметить, что  $\int_a^x f_n(u) du \rightarrow \int_a^x f(u) du$  равномерно на  $(a, b)$ .

Приведенное выше доказательство — одно из трех (кое-где пересекающихся) доказательств, предложенных нам Безиковичем. Независимое доказательство дал Смитсис.

В качестве иллюстрации возьмем  $f_n(x) = 0$  для всех иррациональных  $x$  из  $(0, 1)$ . Если  $x$  рационально и равно  $m/k$  (где  $m/k$  — несократимая дробь), положим  $f_n(x) = 1$  при  $n = k$  и  $f_n(x) = 0$  при  $n \neq k$ . Тогда  $f_n(x) \rightarrow 0$  всюду, и

$$\lim_{n \rightarrow \infty} \int_0^1 f_n(x) dx = \int_0^1 \lim_{n \rightarrow \infty} f_n(x) dx.$$

Однако в любом интервале по  $x$  и для любого  $n$  найдутся рациональные дроби, знаменатели которых превосходят  $n$ ;

следовательно,  $f_n(x)$  ни в каком интервале по  $x$  не стремится к нулю равномерно. Причина, по которой в доказательстве были использованы нижние суммы (вместо того чтобы, как обычно, использовать верхние суммы), состоит в том, что верхние суммы не приводят к определению множества интервалов  $I_n$ , которые обладали бы нужными свойствами.

**1.134а.** Наиболее важными являются приложения к интегралам вида

$$\int_a^b [e^{-tx}, \cos tx, \sin tx] v(x) dx \quad (b > a),$$

где  $t$  — большое положительное число. Прежде всего отметим, что если про функцию  $v(x)$  известно только, что она ограничена, скажем  $|v(x)| < A$ , то уже тогда (и без применения леммы Абеля) можно утверждать, что

$$\left| \int_a^b e^{-tx} v(x) dx \right| < \frac{Ae^{-ta}}{t}. \quad (1)$$

Без дополнительных ограничений нельзя утверждать, что

$$\int_a^b [\cos tx, \sin tx] v(x) dx = O\left(\frac{1}{t}\right).$$

Однако если  $v(x)$  удовлетворяет нижеследующим условиям, то можно доказать последнее утверждение, а кроме того, вывести результат, заменяющий (1) в случае несобственного интеграла.

а) Если  $t$  положительно и  $\left| \int_a^{\xi} f(x) dx \right| < M$  при  $a \leq \xi \leq b$ , то  $e^{-tx}$  удовлетворяет условиям, наложенным на  $v(x)$  в лемме Абеля, и мы имеем

$$\left| \int_a^b e^{-tx} f(x) dx \right| < Me^{-ta}. \quad (2)$$

б) Если  $t$  положительно, а функция  $v(x)$  неотрицательна, ограничена и не возрастает при  $a \leq x \leq b$ , то

$$\left| \int_a^{\xi} \cos tx dx \right| \leq \frac{2}{t}, \quad \left| \int_a^{\xi} \sin tx dx \right| \leq \frac{2}{t}.$$

Отсюда, полагая в лемме Абеля  $f(x) = \cos tx$  или  $\sin tx$ , имеем

$$\begin{aligned} \left| \int_a^b \cos tx v(x) dx \right| &\leq \frac{2}{t} v(a), \\ \left| \int_a^b \sin tx v(x) dx \right| &\leq \frac{2}{t} v(a). \end{aligned} \quad (3)$$

в) Пусть  $t$  положительно, полная вариация  $v(x)$  на интервале  $a \leq x \leq b$  равна  $V$ , а положительная и отрицательная вариации  $v(x)$  на  $(a, x)$  равны  $P(x)$  и  $N(x)$ . Тогда

$$\begin{aligned} v(x) &= v(a) + P(x) - N(x), \\ v(x) &= v(a) + [N(b) - N(x)] - [P(b) - P(x)] + P(b) - N(b), \\ v(x) &= v(b) + [N(b) - N(x)] - [P(b) - P(x)]. \end{aligned} \quad (4)$$

Здесь  $v(b)$  — константа.  $N(b) - N(x)$  и  $P(b) - P(x)$  удовлетворяют условиям, наложенным на  $v(x)$  в лемме Абеля, и превращаются в  $N(b)$  и  $P(b)$  при  $x = a$ . Отсюда

$$\left| \int_a^b \cos tx v(x) dx \right| \leq \frac{2}{t} [|v(b)| + N(b) + P(b)] = \frac{2}{t} [|v(b)| + V], \quad (5)$$

$$\left| \int_a^b \sin tx v(x) dx \right| \leq \frac{2}{t} [|v(b)| + V]. \quad (6)$$

## СКАЛЯРЫ И ВЕКТОРЫ

Мораль здесь такова:  
 позаботься о смысле,  
 а звуки позаботятся  
 о себе сами.

Льюис Кэррол „Алиса в стране чудес“

**2.01. Декартовы координаты. Правило суммирования.** Любое физическое измерение — это определение отдельной величины. Физику можно определить как науку, изучающую такие соотношения между величинами, что из одного множества измерений можно предсказать другие, если только даны условия наблюдения. Наиболее элементарными измерениями, кроме простого счета, являются измерения расстояний. Как мы видели, экспериментально было установлено, что расстояния вдоль заданной прямой аддитивны в некотором смысле и удовлетворяют законам ассоциативности и коммутативности сложения. Но если два расстояния взяты не вдоль одной прямой, то их сумма не может быть определена однозначно без дополнительных условий. Если  $P$ ,  $Q$ ,  $R$  — три точки, то расстояние  $PR$  определяется не только расстояниями  $PQ$  и  $QR$ . Между расстояниями вдоль любых *двух* пересекающихся прямых  $PQQ'$  и  $PRR'$  имеется экспериментально проверяемое соотношение

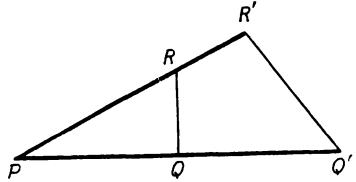


Рис. 2.

$$\frac{PQ^2 + PR^2 - QR^2}{2PQ \cdot PR} = \frac{PQ'^2 + PR'^2 - Q'R'^2}{2PQ' \cdot PR'}, \quad (1)$$

когда  $P$  не лежит между  $Q$  и  $Q'$  и  $R$  и  $R'$ . Это отношение (число) обозначается  $\cos \theta$ , и  $\theta$  называется углом между прямыми. Величина  $\cos \theta$  не может быть меньше  $-1$  или больше  $+1$ . Теперь, когда измерение в основном заменило методы Евклида и проповедуются экспериментальный подход к геометрии, желательно, чтобы одним из первых шагов в преподавании была прямая проверка этого закона и чтобы он был положен в основу всего дальнейшего изложения предмета. Его значительно легче проверять, чем некоторые из обычных аксиом. Он делает угол

производной величиной, и свойство аддитивности углов на плоскости, взятое Евклидом как постулат, может быть выведено из него. Это и к лучшему, так как плоскость — более сложное понятие, чем прямая линия. Из (1) можно развить всю теорию Евклида вплоть до введения прямоугольных координат [1, гл. 7]\*).

В методах Евклида понятие наложения играет выдающуюся роль. Он все время говорит о реальных вещах, которые сравниваются. Современному преподаванию свойственно стремление избежать понятия наложения. Но наложение прямо относится к физическим методам, и язык Евклида не позволяет перепутать, скажем, длину и площадь. На языке физических величин легко можно выразить то, что имеет физический смысл и что было бы трудно или невозможно выразить языком Евклида. Однако попытка свести его систему к чистой математике уничтожает то, что с физической точки зрения представляется наиболее ценным.

Прямоугольные координаты обладают тем свойством, что расстояния между двумя точками выражаются в них симметрично через сумму трех квадратов их разностей. Никакой другой способ задания положения не обладает этим свойством. Общее утверждение, что прямоугольные координаты не имеют особого физического значения, вздорно. В проективной геометрии отвергается понятие расстояния, потому что оно метрическое. В физике мы не можем обойтись без него. Но когда оно введено, мы можем рассмотреть кратчайшее расстояние между точкой и точками данной плоскости (которая определена как геометрическое место точек, равноудаленных от двух заданных точек), и это прямо подводит к понятиям перпендикуляра и прямоугольных координат.

Прямоугольная координата — это расстояние от данной плоскости. Условимся, что точкам по одну сторону плоскости соответствуют расстояния с положительным знаком, а по другую — с отрицательным. Теперь мы можем сказать, что смещение от  $P$  к  $Q$  определяется тремя *компонентами*, а именно разностями прямоугольных координат этих двух точек, и эти разности равны проекциям  $PQ$  на три используемые оси. Так как каждая проекция является расстоянием вдоль заданной прямой в заданном направлении, то они обладают аддитивным свойством и их можно взять в любом порядке. Иначе говоря, начиная с  $P$ , мы можем найти точку  $P'$ , у которой координаты  $y$  и  $z$

---

\*) Другие рассмотрения, касающиеся физических величин, можно найти в гл. 4 и 6. В частности, важно понять, что законы науки устанавливаются путем последовательных приближений.



такие же, как у  $P$ , а координата  $x$  такая же, как у  $Q$ , потом точку  $P''$ , у которой координаты  $x$  и  $z$  такие же, как у  $P'$ , а координата  $y$  такая же, как у  $Q$ , и, наконец, получить  $Q$ , изменив координату  $z$ . Хотя можно упорядочить координаты шестью различными способами, мы в три шага всегда попадаем в  $Q$ . Этот факт верен для любой системы координат, но в прямоугольных координатах эти три смещения обладают тем особым свойством, что каждое из них имеет заданное направление и величину. (Это верно также для косоугольных декартовых координат, но они используются только в специальных случаях, и мы не будем их рассматривать до гл. 4.) В известном смысле мы можем рассматривать параллельные смещения одинаковой длины как эквивалентные. Это является частным случаем *правила параллелограмма*. Последнее в своем общем виде используется только в косоугольных координатах. Поэтому мы не будем касаться его сейчас. Мы можем рассматривать указанную эквивалентность как факт, представляющий физический процесс, полагая, что смещения переносят тело целиком к его новым начальным точкам в однородном пространстве. Однако этот процесс употребляется не часто.

Действительно важным свойством прямоугольных координат является то, что они в равной мере хорошо и в одинаковом виде выражают свойства расстояния, какие бы направления осей мы ни выбрали, лишь бы эти направления были взаимно перпендикулярны. Мы часто изменяем направления осей, и нам нужен способ определения координат в одной системе, если они заданы в другой. Компонентами смещения относительно новых осей будут его проекции на эти оси с соответствующими знаками. Если прямая  $RS$  образует углы  $\alpha, \beta, \gamma$  со старыми осями координат и смещение  $PQ$  имеет компоненты  $u, v, w$  в тех же осях, то проекция  $PQ$  на  $RS$  равна

$$u \cos \alpha + v \cos \beta + w \cos \gamma. \quad (2)$$

Такое обозначение очень громоздко. Его обычно сокращают, обозначая косинусы через  $l, m, n$ . Их называют *направляющими косинусами* прямой  $RS$ . Тогда проекция записывается в виде  $lu + mv + nw$ . Дальнейшие упрощения получаются, если мы обозначим оси через  $x_1, x_2, x_3$ , компоненты — через  $u_1, u_2, u_3$ , а направляющие косинусы — через  $l_1, l_2, l_3$ . Тогда проекция запишется в виде

$$\sum l_i u_i \quad (i = 1, 2, 3). \quad (3)$$

Преимущество индексного обозначения состоит в том, что наиболее общие законы физики имеют одинаковый вид для всех

компонент. Следовательно, если мы имеем, скажем, дифференциальные уравнения для трех координат частицы, то достаточно написать одно уравнение в индексном обозначении; следует только помнить, что индекс должен принимать все три значения по очереди. Дальнейшее сокращение записи достигается посредством *правила суммирования*. Мы видим, что в (3) каждый член содержит один и тот же индекс дважды и все члены суммируются. Условимся, что в любом записанном в индексной форме выражении, в котором встречается повторяющийся индекс, этому индексу следует придать все возможные значения и результаты затем просуммировать. Пользуясь правилом суммирования, вместо (3) мы пишем просто  $l_{ii}$ .

**2.02. Преобразование.** Если нам даны две прямоугольные системы координат  $O123$  и  $O1'2'3'$  с общим началом  $O$ , то мы обозначаем соответствующие координаты через  $x_i$  и  $x'_j$ . Их следует рассматривать как два различных описания положения точки  $P$ . Обозначим через  $l_{ij}$  направляющий косинус оси  $x'_j$  относительно оси  $x_i$ . Тогда  $x'_j$  — проекция  $OP$  на соответствующую ось и, следовательно,  $x'_j = l_{ij}x_i$ . Это верно для  $j = 1, 2, 3$ ; следовательно,

$$x'_j = l_{ij}x_i. \quad (4)$$

Здесь записаны три уравнения преобразования, каждое из которых имеет три члена в правой части.

Мы пока не использовали условие взаимной перпендикулярности осей  $x'_j$ . Условие перпендикулярности  $x'_1$  и  $x'_2$  состоит в том, что

$$l_{i1}l_{i2} = 0. \quad (5)$$

Для других пар осей верны два аналогичных соотношения. Поскольку  $l_{i1}$  — направляющие косинусы  $x'_1$  относительно  $O123$ , то выполняется соотношение

$$l_{i1}l_{i1} = 1 \quad (6)$$

и еще два подобных соотношения. (Мы не пишем здесь  $l_{ij}l_{ij}$  потому, что тогда нужно было бы просуммировать и по  $i$ , и по  $j$ , что не входит в наши намерения.) Итак, хотя имеется девять  $l_{ij}$ , они связаны шестью соотношениями вида (5) и (6), и следует ожидать, что только три из них можно выбирать независимо. Это действительно так, хотя указанные соображения нельзя рассматривать как доказательство.

**2.021.  $\delta_{ik}$ .** Рассмотренные выше шесть соотношений могут быть записаны как одно. Введем набор чисел  $\delta_{ik}$ , где  $i$  и  $k$

принимают значения 1, 2 и 3, и  $\delta_{ik} = 1$ , если  $i = k$ , и  $\delta_{ik} = 0$ , если  $i \neq k$ . Тогда мы имеем

$$l_{ij}l_{il} = \delta_{jl}. \quad (7)$$

Этот набор величин называется *тензором подстановки* \*). Если любое выражение, содержащее неповторяющийся индекс  $k$ , умножить на  $\delta_{ik}$  и просуммировать, то, поскольку единственное ненулевое слагаемое будет при  $k = i$ , получаем замену индекса  $k$  индексом  $i$ . Заметим, что следует различать выражение  $\delta_{ik}$  при  $i = k$ , которое равно 1, и  $\delta_{il}$ , которое подразумевает суммирование и равно 3. Отметим также, что в  $l_{il}$  сам вид выражения показывает, что два  $l$  не могут иметь одинаковый смысл. При употреблении индексов неизбежно появление одной и той же буквы в двух смыслах, так как много букв уже имеют специальное назначение, но индекс просто определяет ось и может быть равным 1, 2 и 3, а та же буква на строке обозначает физическую величину. Если иметь это в виду, то никакой путаницы не возникнет. Для первоначальных осей здесь использовались индексы  $i, k, t, p, \dots$ , а для преобразованных —  $j, l, n, q, \dots$ ; буква  $o$  пропущена, так как похожа на цифру. Часто в каждой системе используются свои буквы: в одной — греческие, а в другой — соответствующие им латинские; но иногда приходится употреблять слишком много букв (и нам даже может не хватить алфавита) или вводить буквы со штрихами, что усложняет запись.

**2.022. Обратное преобразование.** Поскольку оси  $x'_j$  ортогональны между собой и косинус угла между осями  $x_i$  и  $x'_j$  равен  $l_{ij}$ , имеем

$$x_i = l_{ij}x'_j, \quad (8)$$

а поскольку и оси  $x_i$  ортогональны между собой, то

$$l_{ij}l_{kj} = \delta_{ik}. \quad (9)$$

Эти соотношения алгебраически выведены из (7) в 2.073. Теперь мы имеем 12 соотношений между девятью величинами. Отсюда следует, что нельзя полностью доверять методу счета констант.

**2.023. Скорость, ускорение, сила. Определение вектора. Единичные, или направляющие векторы.** В инерциальной системе

---

\*) См. 3.03. В данный момент нам не нужно общее определение тензора.  $\delta_{ik}$  — частный случай  $\delta$ -символа Кронекера.

координат уравнение движения материальной точки имеет вид \*)

$$m\ddot{x}_i = X_i \quad (10)$$

(термин *инерциальная* лучше, чем обычный термин *фиксированная*). Если мы рассмотрим другую инерциальную систему координат, то, поскольку  $l_{ij}$  не зависят от времени,

$$\dot{x}'_j = l_{ij}\dot{x}_i, \quad \ddot{x}'_j = l_{ij}\ddot{x}_i. \quad (11)$$

Таким образом, компоненты скорости и ускорения при преобразовании системы координат меняются подобно координатам.

Мы полагаем, что соотношение (10) верно в *любой* инерциальной системе координат. В различных системах координат компоненты силы  $X_i, X'_j$  должны иметь разные значения, но все же должно выполняться соотношение

$$m\ddot{x}'_j = X'_j, \quad (12)$$

как бы ни преобразовывались оси и совершенно независимо от фактических значений  $X_i$ . Но это возможно, только если

$$X'_j = l_{ij}X_i. \quad (13)$$

Таким образом, мы получили, что не только смещение, но и скорость, ускорение и сила преобразуются по правилу (4). Для силы это было выведено из предположения о том, что уравнение движения выполняется для всех инерциальных систем. Так, например, для частицы, двигающейся под действием силы тяжести, в системе координат с осью  $OZ$ , направленной вертикально вверх, мы имеем  $(\ddot{x}_1, \ddot{x}_2, \ddot{x}_3) = (0, 0, -g)$ . Но это неверно для других систем координат, и общая формула имеет вид  $\ddot{x}_i = -gl_i$ , где  $l_i$  — направляющие косинусы оси, направленной вертикально вверх.

Любые три величины  $A_i$ , которые при повороте осей преобразуются по правилу

$$A'_j = l_{ij}A_i, \quad (14)$$

называются *компонентами вектора* по отношению к этим осям. Теперь если мы исходим из набора трех уравнений, которые верны в любой системе координат, и выводим из них какое-то следствие, используя конкретные координаты  $O1, O2, O3$ , то с тем же успехом мы могли бы вывести из этих уравнений,

---

\*) В наши цели не входит детальное обсуждение того, как и в какой мере динамика Ньютона основывается на опыте. Такое обсуждение может быть найдено в [1, гл. 8].

записанных в другой системе координат  $O1', O2', O3'$ , следствие, формально отличающееся только добавлением штрихов ко всем буквам. Если следствие представляет собой систему трех уравнений и левые части суть компоненты вектора, то мы знаем, что они преобразуются при переходе к новой системе координат согласно (14). Но так как обе системы уравнений верны, то правые части также должны преобразовываться по (14) и являются компонентами вектора. Обратно, если три уравнения задают равенство компонент двух векторов в некоторой системе координат, то, поскольку обе части уравнений изменяются по одному и тому же закону, эти уравнения сразу же можно применить к любой другой системе координат просто добавлением штрихов.

Если мы возьмем две прямые, направляющие косинусы которых относительно осей  $x_i$  суть  $m_i$  и  $n_i$ , то косинус угла между ними будет  $m_i n_i$ . Если, в частности, первая прямая — ось  $x'_i$ , то косинус угла между ней и прямой с направляющими косинусами  $n_i$  есть

$$n'_i = l_{ij} n_j. \quad (15)$$

Таким образом, направляющие косинусы данной прямой преобразуются по правилу (14) и являются компонентами вектора. Этот вектор часто называют *единичным вектором*. Мы предпочитаем термин *направляющий вектор*, так как его единственное назначение — определять направление.

Мы можем говорить о компоненте вектора в любом направлении. Если прямая имеет направляющие косинусы  $n_i$  в системе координат  $x_i$ , то компонента вектора  $A_i$  в этом направлении есть  $n_i A_i$ . Теперь предположим, что мы хотели бы отыскать компоненту этого же вектора в том же направлении, используя систему координат  $x'_j$ . Мы получили бы

$$n'_j A'_j = l_{ij} n_i l_{kj} A_k = (l_{ij} l_{kj}) n_i A_k = \delta_{ik} n_i A_k = n_i A_i, \quad (16)$$

и следовательно, компонента вектора в заданном направлении не зависит от выбора осей координат.

Этот результат можно было получить также, используя третью систему координат, у которой одна из осей имеет направление  $n_i$ , и переходя от первоначальной системы к этой, а потом к  $x'_j$ .

Вектор часто определяют как объект, задаваемый совокупностью трех компонент с аддитивным свойством, которое выражается правилом параллелограмма. Последняя часть этого определения предполагает, однако, что рассматриваемые векторы могут быть представлены смещениями в произвольном

скалярном масштабе, как правило имеющем размерность. Пока это не сделано, мы не знаем, что подразумевается под правилом параллелограмма для этих векторов. Введение параллелограмма фактически лишь усложняет дело. Будет лучше перейти прямо к аналитической формулировке требуемого свойства. Итак, это правило требует, чтобы мы знали, что подразумевается под сложением для рассматриваемых векторов. У нас есть естественные интерпретации для смещения, скорости, ускорения и силы. Однако нам будут встречаться более сложные векторы, для которых трудно указать отличное от аналитического определение сложения. К тому же в этом нет необходимости, так как все, что требуется для наших целей, — это правильность наших уравнений. Если мы можем правильно провести выкладку, то для физики не обязательно отыскивать особую физическую интерпретацию для каждого входящего в выкладку члена. Таким образом, мы определяем вектор как совокупность компонент  $A_i$ , определенным образом изменяющихся при преобразовании координат. Это включает утверждение о том, что компонентой вектора с направляющими косинусами  $n_i$  является  $n_i A_i$ . *Скаляр* — это однокомпонентная величина, не меняющаяся при замене осей.

**2.03. Обозначение вектора одной буквой.** Вектор с компонентами  $A_i$  можно еще короче обозначить через **A**. Очень трудно и фактически не обязательно объяснять, что мы называем вектором в отрыве от его компонент. Производя сравнения с наблюдениями, мы интересуемся именно компонентами, однако часто удобно работать с одним символом. Мы определяем сумму двух векторов **A** и **B** как вектор с компонентами  $A_i + B_i$  и обозначаем его через **A + B**. Обозначение  $-\mathbf{B}$  — вектор с компонентами  $-B_i$ , а  $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$ . Мы можем определить умножение вектора **A** на скаляр  $m$  как вектор с компонентами  $m A_i$ . Очевидно, что при таком определении выполняются законы ассоциативности и коммутативности сложения;  $\mathbf{B} + \mathbf{A}$  — вектор с компонентами  $B_i + A_i = A_i + B_i$ , а это компоненты вектора  $\mathbf{A} + \mathbf{B}$ . Аналогично закон ассоциативности

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$$

немедленно следует из определения. Не имеет смысла прибавлять вектор к скаляру, поскольку первый меняется при замене осей, а второй нет. Однако ясно, что  $m\mathbf{A}$  и  $\mathbf{A}m$  представляют один и тот же вектор, и, следовательно, выполняется закон коммутативности умножения. Аналогично если  $m$  и  $n$  — скаляры, то

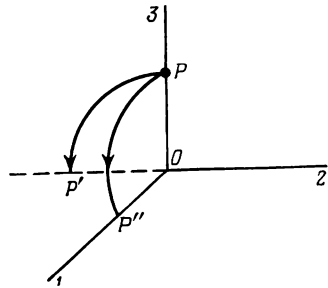
$$(m + n)\mathbf{A} = m\mathbf{A} + n\mathbf{A} = (n + m)\mathbf{A}$$

при условии, что  $m$  и  $n$  имеют одинаковую размерность; в противном случае сложение бессмысленно. Также

$$m(n\mathbf{A}) = (mn)\mathbf{A}.$$

Смещение от  $P$  к  $Q$ , рассматриваемое как вектор, будет обозначаться  $\overline{PQ}$ .

Мы видели, что вектор вполне задается тремя компонентами. Однако легко видеть, что это условие не влечет за собой выполнения закона коммутативности сложения и, следовательно, само по себе не является достаточным для выведения всех свойств векторов \*). Рассмотрим вращение твердого тела на конечный угол вокруг произвольной оси, проходящей через фиксированную точку  $O$  этого тела. Естественной „суммой“ двух последовательных вращений является одно вращение, переводящее тело в то же конечное положение. Одно вращение вполне определяется заданием направления оси вращения, для чего нужны два числа, и углом поворота. Возьмем неподвижную прямоугольную систему координат  $O123$  и рассмотрим два последовательных вращения. Первое на угол  $\pi/2$  вокруг  $O1$ , второе на угол  $\pi/2$  вокруг  $O2$ , оба в правую сторону. Если мы возьмем их в этом порядке, точка  $P$  с координатами  $(0, 0, 1)$  сперва перейдет в  $P'(0, -1, 0)$  и останется там при втором вращении. Но если мы сделаем сперва вращение относительно  $O2$ , точка перейдет в  $P''(1, 0, 0)$  и останется там при следующем вращении. Порядок вращений влияет на результат. Если бы сумма двух вращений получалась по векторным законам, этого не могло бы быть, поскольку для векторов сложение коммутативно. Представление конечных вращений будет рассмотрено более полно в следующей главе.



Р и с. 3.

**2.031.** Из нашего определения можно вывести, что вектор допускает *геометрическое представление*. Если  $\mathbf{A}$  — произвольный вектор, мы можем умножить его компоненты  $A_i$  на константу  $c$ , выбранную так, чтобы они приобрели размерность длины. Тогда если  $x_i = cA_i$ , то проекция  $\mathbf{x}$  на направление  $l_i$  есть

$$l_i x_i = c l_i A_i,$$

и, разделив на  $c$ , мы получаем компоненту  $\mathbf{A}$  в направлении  $l_i$ . Отсюда видно, что сложение векторов одинаковой размерности полностью описывается, если представлять их как смещения в одинаковом масштабе, и, таким образом, правило параллелограмма установлено для векторов общего вида. Дальнейшие свойства векторов могут быть выведены из свойств смещений с помощью этого представления. В частности, каждый вектор

\*) Другими словами, не всякий объект, задаваемый тремя компонентами, является вектором. — *Прим. перев.*

имеет модуль и направление. Действительно, если  $r$  — длина смещения, представляющего его в данном масштабе, то

$$r^2 = x_1^2 + x_2^2 + x_3^2 = c^2(A_1^2 + A_2^2 + A_3^2) = c^2 A^2, \quad (1)$$

где

$$A^2 = A_i A_i \quad (2)$$

и не зависит от выбора осей. Тогда  $A$  (взятое положительным) может быть названо *модулем*  $\mathbf{A}$  и соответствует величине смещения. Также если  $A \neq 0$  и мы напишем

$$\frac{x_i}{r} = \frac{A_i}{A} = m_i,$$

то  $m_i$  будут направляющими косинусами определенной прямой, которая может быть названа направлением  $\mathbf{A}$ . Далее, проекция рассматриваемого смещения на прямую в направлении  $l_i$  есть  $s A_i m_i = s A \cos \theta$ , где  $\theta$  — угол между направлениями  $l_i$  и  $m_i$ . Компонентой  $\mathbf{A}$  в направлении  $l_i$  является  $A \cos \theta$ . Однако компонента в заданном направлении не зависит от взятых осей, и, следовательно,  $A$  и  $\theta$  для всех  $l_i$  не зависят от осей. Значит  $A$  — одна и та же величина, а  $m_i$  — одно и то же направление в любых осях.

В одном случае удобно отказаться от правила брать  $A$  положительным. Прямая линия может быть представлена уравнениями

$$x_i = a_i + s l_i.$$

Если  $l_i$  фиксированы, мы можем получить все точки прямой, изменяя  $s$  от  $-\infty$  до  $+\infty$ . Это соответствует перемещению вдоль прямой в определенном направлении. Если  $l_i$  указаны, то прямую, проходящую через начало координат, удобно записать так:

$$x_i = x l_i,$$

где  $x$  может принимать и отрицательные значения. Расстояние от начала координат, взятое положительным, всегда будет обозначаться через  $r$ . Когда мы берем

$$x_i = r l_i$$

с положительным  $r$ , мы рассматриваем две прямые, выходящие из начала координат в противоположных направлениях, как разные прямые с  $l_i$ , равными по величине и противоположными по знаку. Это иногда удобно, но не всегда.

**2.032. Сравнение систем обозначений.** Важность и полезность векторного обозначения являются предметом обсуждений для занимающихся математической физикой. То, что можно сказать в терминах  $\mathbf{A}$ , можно сказать и в терминах  $A_i$ , выписывая полностью все компоненты. Однако если постоянно иметь в виду геометрическое представление вектора в виде напра-



вленного отрезка, то, по мнению некоторых, содержание многих физических законов наиболее понятно в векторном обозначении. Некоторые усилия нужны, чтобы научиться думать в терминах векторов. Они нужны также, чтобы приобрести уверенность в том, что компактность, получаемая благодаря правилу суммирования, не ведет к ошибкам. Кроме того, следует помнить, что если получен некоторый физический результат о векторе, то для его проверки всегда понадобится произвести три измерения и, независимо от того, как он выражен — с помощью векторного или индексного обозначения, — придется расписать все три компоненты. Расписывание компонент при индексном обозначении всегда проще, чем при векторном. Некоторые общие теоремы записываются короче в одной системе, некоторые — в другой. В конкретных задачах необходимы некоторые размышления для того, чтобы выбрать наилучший момент для расписывания по компонентам, и многие студенты слишком долго откладывают его, хотя условия задачи выделяют одно или два специальных направления. В теории упругости и динамике вязких жидкостей естественнее применять индексное обозначение. В теории относительности векторное обозначение полностью неприменимо, так как не выполняется правило параллелограмма для скоростей. Вследствие этого некоторые специалисты считают, что векторное обозначение — чистая потеря времени, и откладывают знакомство с этим в основном полезным методом. В этой главе мы покажем, насколько это возможно, оба метода параллельно. Умение переводить с одного языка на другой и переходить к расписанной форме абсолютно необходимо для понимания современной физической литературы. Часто полезно представлять себе вектор как вектор смещения. Хотя в определениях мы приводим четкое различие между общими векторами и векторами смещения, мы часто будем применять к общим векторам геометрические термины: например, угол между двумя векторами **A** и **B**, строго говоря, — значит „угол между векторами смещений, представляющими **A** и **B** в соответствующем определенном масштабе“; „два перпендикулярных вектора **A** и **B**“ — значит „два вектора **A** и **B**, такие, что представляющие их смещения перпендикулярны“. Использование этой аналогии не нужно при индексных обозначениях, здесь достаточно аналитических определений.

**2.033. Нулевой вектор.** Нулевой вектор — это вектор, модуль которого равен нулю.

**2.034. Направляющие векторы.** Вектор с модулем 1 (число) в направлении вектора **A** называется единичным или

направляющим вектором в этом направлении. Его компоненты, очевидно, равны  $l_i$  — направляющим косинусам  $\mathbf{A}$  по отношению к координатным осям. В частности, мы будем обозначать направляющие векторы в направлении осей через  $\mathbf{e}_{(1)}$ ,  $\mathbf{e}_{(2)}$ ,  $\mathbf{e}_{(3)}$  соответственно. Таким образом,

$$\mathbf{e}_{(1)} = (1, 0, 0), \quad \mathbf{e}_{(2)} = (0, 1, 0), \quad \mathbf{e}_{(3)} = (0, 0, 1).$$

Индекс взят в скобки, чтобы подчеркнуть, что это не компонента, а некоторый вектор. Любой вектор  $\mathbf{A}$  может быть записан так:

$$A_1\mathbf{e}_{(1)} + A_2\mathbf{e}_{(2)} + A_3\mathbf{e}_{(3)}.$$

В некоторых книгах направляющие векторы, параллельные осям, обозначают через  $\mathbf{i}$ ,  $\mathbf{j}$ ,  $\mathbf{k}$  и пишут

$$\mathbf{A} = A_x\mathbf{i} + A_y\mathbf{j} + A_z\mathbf{k}.$$

**2.04. Линейно зависимые или компланарные векторы.** Если три вектора  $\mathbf{A}$ ,  $\mathbf{B}$  и  $\mathbf{C}$  удовлетворяют соотношению

$$\alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C} = 0, \quad (1)$$

где  $\alpha$ ,  $\beta$ ,  $\gamma$  — действительные числа (не все равные нулю), то эти векторы называются линейно зависимыми. Геометрически это значит, что  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  могут быть представлены смещениями, лежащими в одной плоскости; так, если  $\gamma \neq 0$ , то

$$\mathbf{C} = -\frac{1}{\gamma}(\alpha\mathbf{A} + \beta\mathbf{B}),$$

и, следовательно,  $\mathbf{C}$  представляется смещением, лежащим в плоскости смещений, представляющих  $\mathbf{A}$  и  $\mathbf{B}$ , в предположении, что все они проведены из одной точки. Сами векторы называются тогда компланарными. Мы можем здесь напомнить, что параллельные векторы равной величины эквивалентны в нашей системе. Некоторые авторы различают „свободные векторы“ и „связанные векторы“. Связанный вектор включает, например, задание и силы, и точки ее приложения. В нашей системе связанный вектор — это не один, а два вектора: один — определяющий силу, а другой — точку ее приложения.

Если между тремя векторами нет соотношения вида (1), то они называются линейно независимыми или некомпланарными. Соответствующие результаты в индексных обозначениях могут быть написаны без труда.

**2.041. Выражение любого вектора через три некомпланарных вектора.** Если  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  — три некомпланарных вектора и

$\mathbf{D}$  — произвольный вектор, то  $\mathbf{D}$  можно выразить как  $\alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C}$ , где  $\alpha, \beta, \gamma$  — действительные числа. Пусть  $\overline{PQ}$  (см. рис. 4) представляет  $\mathbf{D}$ . Тогда прямые, проведенные через  $P$  и представляющие  $\mathbf{A}$  и  $\mathbf{B}$ , определяют плоскость. Пусть  $RQ$  — прямая, проведенная через  $Q$  в направлении  $\mathbf{C}$ , пересекает эту плоскость в точке  $R$ . Тогда

$$\overline{PQ} = \overline{PR} + \overline{RQ}.$$

Но  $\overline{RQ}$  представляет  $\gamma\mathbf{C}$ , где  $\gamma$  — скаляр, а  $\overline{PR}$  — вектор, компланарный  $\mathbf{A}$  и  $\mathbf{B}$ . Значит, его можно выразить как  $\alpha\mathbf{A} + \beta\mathbf{B}$ . Следовательно,

$$\mathbf{D} = \alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C}.$$

Поскольку между четырьмя векторами всегда есть соотношение такого типа, то, следовательно, четыре вектора не могут быть линейно независимы. Из данного построения ясно, что для заданных  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  (некомпланарных) и любого  $\mathbf{D}$  величины  $\alpha, \beta, \gamma$  определяются однозначно.

Если  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  — попарно перпендикулярные направляющие векторы, то получаем частный случай, приведенный в 2.034. Если  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  не являются попарно перпендикулярными, то  $\alpha\mathbf{A}, \beta\mathbf{B}, \gamma\mathbf{C}$  называются *косугольными компонентами*  $\mathbf{D}$  в направлениях  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  соответственно.

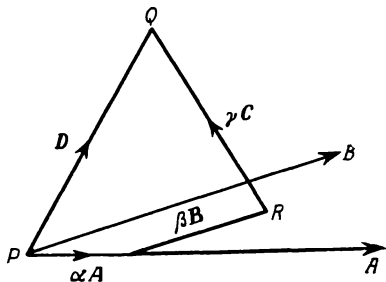


Рис. 4.

## 2.05. Умножение векторов.

Мы рассмотрели умножение вектора на скаляр, но какой смысл можно придать умножению двух векторов и вообще можно ли это сделать, не так уж ясно. Мы можем расположить девять произведений компонент в виде квадратной таблицы следующим образом:

$A_1B_1$	$A_1B_2$	$A_1B_3$
$A_2B_1$	$A_2B_2$	$A_2B_3$
$A_3B_1$	$A_3B_2$	$A_3B_3$

Мы увидим, что все эти произведения появятся в гл. 3. Два произведения, называемые *скалярным* и *векторным* произведениями, к определению которых мы переходим, суть некоторые комбинации этих девяти произведений. Их выбор определяется пользой, которую они приносят в физических приложениях.

**2.06. Скалярное произведение.** Эта функция непосредственно связана с фундаментальным и экспериментально проверяемым соотношением **2.01(1)** между расстояниями, измеряемыми не вдоль одной прямой:

$$PQ \cdot PR \cos \theta = \frac{1}{2} (PQ^2 + PR^2 - QR^2). \quad (1)$$

Это выражение вполне определяется заданием трех длин  $PQ$ ,  $PR$ ,  $QR$ . Поскольку расстояние является основным понятием для всей темы и одинаково для всех систем координат, это выражение определяет величину, не зависящую от системы координат. Мы назвали такие величины *скалярами*. Перейдем теперь к декартовым координатам. Пусть  $x_i$  обозначает  $PQ$ , а  $y_i - \overline{PR}$ , тогда  $y_i - x_i$  обозначает  $\overline{QR}$ , и

$$\frac{1}{2} (PQ^2 + PR^2 - QR^2) = \frac{1}{2} \{ (x_1^2 + x_2^2 + x_3^2) + (y_1^2 + y_2^2 + y_3^2) - [(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2] \}, \quad (2)$$

$$\frac{1}{2} (PQ^2 + PR^2 - QR^2) = x_1 y_1 + x_2 y_2 + x_3 y_3. \quad (3)$$

Это выражение называется *скалярным произведением* векторов  $\mathbf{x}$  и  $\mathbf{y}$ . В общем случае скалярное произведение  $\mathbf{A}$  и  $\mathbf{B}$  определяется так:

$$\mathbf{A} \cdot \mathbf{B} = A_1 B_1 + A_2 B_2 + A_3 B_3 = \sum_{i=1,2,3} A_i B_i = A_i B_i, \quad (4)$$

если использовать правило суммирования. Произведение  $\mathbf{A} \cdot \mathbf{B}$  равно  $AB \cos \theta$ , где  $\theta$  — угол между направлениями этих векторов. Читается это так: „ $A$  на  $B$  скалярно“. Два выражения в координатах (2), (3) могут быть записаны по соглашению о суммировании так:

$$\frac{1}{2} [x_i x_i + y_i y_i - (y_i - x_i)(y_i - x_i)] = x_i y_i, \quad (5)$$

а левую часть можно записать короче:

$$\frac{1}{2} [x_i^2 + y_i^2 - (y_i - x_i)^2]. \quad (6)$$

Напомним, что  $x^2$  было в свое время введено для записи  $xx$ , что мы и подразумеваем под  $x^2$ ; поэтому, когда мы видим выражение, подобное  $x_i^2$ , то интерпретируем его как  $x_i x_i$  и применяем правило суммирования. По этому правилу модуль  $A$  вектора  $\mathbf{A}$  задается равенством

$$A^2 = A_i^2.$$

Это выражение есть скалярное произведение  $\mathbf{A}$  с самим собой. Благодаря правилу суммирования так заметно сокращается запись, что те редкие случаи, когда мы им не пользуемся, оговариваются специально. Без правила суммирования индексное обозначение имело бы мало преимуществ перед выписыванием формулы полностью в декартовых координатах; с его помощью выражение, которое в развернутом виде содержит 9 или 81 член, можно записать как одночлен, и обращаться с ним также легко. Это правило остается полезным в теории относительности и общей динамике, поскольку является просто искусственным приемом и не зависит от правила параллелограмма.

Доказательство того, что  $A_i B_i$  не зависит от системы координат, без представления смещениями проводится так же, как при выводе 2.023 (16):

$$A'_j B'_j = l_{ij} l_{kj} A_i B_k = \delta_{ik} A_i B_k = A_i B_i. \quad (7)$$

*Коммутативность.* Из определения ясно, что порядок сомножителей  $\mathbf{A}$  и  $\mathbf{B}$  в скалярном произведении безразличен:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}.$$

*Ассоциативность.* Поскольку  $\mathbf{A} \cdot \mathbf{B}$  — не вектор, мы не можем продолжить умножение и образовать произведение  $\mathbf{A} \cdot \mathbf{B} \cdot \mathbf{C}$ , так что говорить об ассоциативности бессмысленно. Но  $(\mathbf{A} \cdot \mathbf{B}) \mathbf{C}$  — вектор, имеющий направление такое же, как  $\mathbf{C}$ , а величину, большую в  $\mathbf{A} \cdot \mathbf{B}$  раз.

*Дистрибутивность.* Мы можем, однако, доказать, что

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C},$$

так как это немедленно следует из определения. Из этого сейчас можно вывести, что скалярное произведение двух сумм векторов можно представить суммой скалярных произведений. В частности, так как

$$\mathbf{e}_{(2)} \cdot \mathbf{e}_{(3)} = \mathbf{e}_{(3)} \cdot \mathbf{e}_{(1)} = \mathbf{e}_{(1)} \cdot \mathbf{e}_{(2)} = 0,$$

$$\mathbf{e}_{(1)} \cdot \mathbf{e}_{(1)} = \mathbf{e}_{(2)} \cdot \mathbf{e}_{(2)} = \mathbf{e}_{(3)} \cdot \mathbf{e}_{(3)} = 1,$$

то

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= (A_1 \mathbf{e}_{(1)} + A_2 \mathbf{e}_{(2)} + A_3 \mathbf{e}_{(3)}) \cdot (B_1 \mathbf{e}_{(1)} + B_2 \mathbf{e}_{(2)} + B_3 \mathbf{e}_{(3)}) = \\ &= A_1 B_1 + A_2 B_2 + A_3 B_3. \end{aligned}$$

Таким образом, мы получили определение (4).

Заметим, что, когда определены направляющие векторы вдоль осей, прямоугольные координаты считаются скалярами.

Следует отметить, что если скалярное произведение двух смещений равно нулю, то  $PQ^2 + PR^2 = QR^2$  и, следовательно, или одно из смещений — нуль, или они перпендикулярны. Обращение в нуль  $\mathbf{A} \cdot \mathbf{B}$  не означает, что  $\mathbf{A}$  или  $\mathbf{B}$  — нулевой вектор, а означает, что либо они перпендикулярны, либо один из них нуль. Однако если  $\mathbf{A} \cdot \mathbf{B}_1$ ,  $\mathbf{A} \cdot \mathbf{B}_2$ ,  $\mathbf{A} \cdot \mathbf{B}_3$  все равны нулю, а  $\mathbf{B}_1$ ,  $\mathbf{B}_2$ ,  $\mathbf{B}_3$  — некопланарны, то  $\mathbf{A}$  не может быть перпендикулярным к ним всем и должен быть нулевым.

Косинус угла между двумя прямыми равен скалярному произведению их направляющих векторов, т. е. если  $l_i$ ,  $m_i$  — направляющие косинусы прямых, то

$$\cos \theta = l_i m_i.$$

Компонента вектора  $\mathbf{A}$  вдоль прямой с направляющими косинусами  $l_i$  равна  $l_i A_i$ , что является скалярным произведением  $\mathbf{A}$  и направляющего вектора этой прямой.

**2.07. Векторное произведение.** Векторное произведение двух векторов записывается  $\mathbf{A} \times \mathbf{B}$  (читается  $A$  на  $B$  векторно). Определим его на геометрической модели. Если  $\overline{OP}$  и  $\overline{OQ}$ , представляющие  $\mathbf{A}$  и  $\mathbf{B}$  соответственно, не параллельны, то они определяют плоскость. Пусть  $\overline{OR}$  представляет направляющий вектор  $\mathbf{n}$ , перпендикулярный к этой плоскости. Тогда если  $\theta$  — угол, на который нужно повернуть  $\overline{OP}$  относительно  $\overline{OR}$  вправо, чтобы получить прямую  $\overline{OQ}$ , то

$$\mathbf{A} \times \mathbf{B} = AB \sin \theta \mathbf{n}.$$

Сейчас мы увидим, что определение не зависит от выбора направления  $\mathbf{n}$ . Если мы направим  $\mathbf{n}$  в обратную сторону, то угол поворота  $\theta$  заменится на  $(2\pi - \theta)$ , а так как  $\sin(2\pi - \theta) = -\sin \theta$ , то величина и направление векторного произведения не изменятся.

В векторном произведении порядок сомножителей является существенным. Действительно,

$$\mathbf{B} \times \mathbf{A} = BA \sin(-\theta) \mathbf{n} = -\mathbf{A} \times \mathbf{B}. \quad (1)$$

Векторное умножение не коммутативно. Скоро мы увидим, что оно и не ассоциативно. Однако мы можем доказать его дистрибутивность относительно сложения, т. е.

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}. \quad (2)$$

Рассмотрим сперва два частных случая:

1)  $\mathbf{A}$  перпендикулярен  $\mathbf{B}$  и  $\mathbf{C}$ . Заметим, что если  $\mathbf{A}$  и  $\mathbf{B}$  перпендикулярны, то  $\mathbf{A} \times \mathbf{B}$  получается из  $\mathbf{B}$  умножением на  $\mathbf{A}$

и поворотом относительно  $\mathbf{A}$  вправо на прямой угол. Следовательно, векторы  $\mathbf{A} \times (\mathbf{B} + \mathbf{C})$ ,  $\mathbf{A} \times \mathbf{B}$ ,  $\mathbf{A} \times \mathbf{C}$  представляются отрезками в  $A$  раз большей длины, чем стороны треугольника, представляющего  $\mathbf{B} + \mathbf{C}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  соответственно, и каждый из них повернут на прямой угол. Следовательно, векторы  $\mathbf{A} \times (\mathbf{B} + \mathbf{C})$ ,  $\mathbf{A} \times \mathbf{B}$ ,  $\mathbf{A} \times \mathbf{C}$  представляются сторонами треугольника, подобного треугольнику, представляющему  $\mathbf{B} + \mathbf{C}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ . Следовательно, в этом случае (2) выполняется.

2)  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  — компланарны. Тогда  $\mathbf{A} \times (\mathbf{B} + \mathbf{C})$ ,  $\mathbf{A} \times \mathbf{B}$ ,  $\mathbf{A} \times \mathbf{C}$  перпендикулярны к плоскости, определяемой векторами  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , и нужный результат следует из формулы сложения для синусов.

В общем случае мы полагаем, что  $\mathbf{A}$  и  $\mathbf{B}$  не параллельны и не перпендикулярны. Запишем  $\mathbf{B}$  как  $\mathbf{B}_p + \mathbf{B}_n$ , где  $\mathbf{B}_p$  параллелен  $\mathbf{A}$ , а  $\mathbf{B}_n$  перпендикулярен к  $\mathbf{A}$  в плоскости векторов  $\mathbf{A}$  и  $\mathbf{B}$ . Тогда, поскольку  $B_n = B \sin \theta$ , получаем

$$\mathbf{A} \times \mathbf{B} = \mathbf{A} \times \mathbf{B}_n.$$

Если аналогично  $\mathbf{C} = \mathbf{C}_p + \mathbf{C}_n$ , то  $\mathbf{A} \times \mathbf{C} = \mathbf{A} \times \mathbf{C}_n$ . Поскольку

$$\mathbf{B} + \mathbf{C} = (\mathbf{B}_p + \mathbf{C}_p) + (\mathbf{B}_n + \mathbf{C}_n)$$

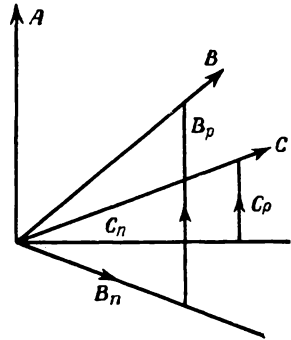


Рис. 5.

и  $\mathbf{B}_p$ ,  $\mathbf{C}_p$  параллельны  $\mathbf{A}$ , а  $\mathbf{B}_n$ ,  $\mathbf{C}_n$  перпендикулярны, то  $\mathbf{B}_n + \mathbf{C}_n$  — компонента вектора  $\mathbf{B} + \mathbf{C}$ , перпендикулярная к  $\mathbf{A}$ , и, следовательно,

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times (\mathbf{B}_n + \mathbf{C}_n). \quad (3)$$

Но, согласно 1-му случаю, это равно

$$\mathbf{A} \times \mathbf{B}_n + \mathbf{A} \times \mathbf{C}_n = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}. \quad (4)$$

Векторное произведение вектора с самим собой или с любым параллельным вектором есть нулевой вектор. Если векторное произведение двух векторов равно нулю, то или один из них нулевой, или они параллельны. Из того, что векторное произведение — нулевой вектор, еще не следует, что один из сомножителей — нулевой вектор. Но если  $\mathbf{A} \times \mathbf{B}_1$  и  $\mathbf{A} \times \mathbf{B}_2$  — нули и  $\mathbf{B}_1$ ,  $\mathbf{B}_2$  не параллельны, то  $\mathbf{A}$  — нулевой вектор.

В частности, для векторов  $\mathbf{e}_{(1)}$ ,  $\mathbf{e}_{(2)}$ ,  $\mathbf{e}_{(3)}$  имеем

$$\mathbf{e}_{(1)} \times \mathbf{e}_{(1)} = \mathbf{e}_{(2)} \times \mathbf{e}_{(2)} = \mathbf{e}_{(3)} \times \mathbf{e}_{(3)} = 0, \quad (5)$$

$$\mathbf{e}_{(2)} \times \mathbf{e}_{(3)} = \mathbf{e}_{(1)} = -\mathbf{e}_{(3)} \times \mathbf{e}_{(2)}. \quad (6)$$

Таким образом,

$$\begin{aligned} \mathbf{A} \times \mathbf{B} &= (A_1 \mathbf{e}_{(1)} + A_2 \mathbf{e}_{(2)} + A_3 \mathbf{e}_{(3)}) \times (B_1 \mathbf{e}_{(1)} + B_2 \mathbf{e}_{(2)} + B_3 \mathbf{e}_{(3)}) = \\ &= (A_2 B_3 - A_3 B_2) \mathbf{e}_{(1)} + (A_3 B_1 - A_1 B_3) \mathbf{e}_{(2)} + (A_1 B_2 - A_2 B_1) \mathbf{e}_{(3)} = \\ &= \begin{vmatrix} \mathbf{e}_{(1)} & \mathbf{e}_{(2)} & \mathbf{e}_{(3)} \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}. \end{aligned} \quad (7)$$

Можно дать геометрическую интерпретацию векторного произведения двух смещений. Если  $P$  — точка, определяемая вектором  $x_i$ , а  $Q$  — вектором  $y_i$ , то их проекции на плоскость  $x_1 = 0$  равны  $P_1(0, x_2, x_3)$ ,  $Q_1(0, y_2, y_3)$ , а площадь треугольника, образованного этими точками и началом координат, равна  $\frac{1}{2}(x_2 y_3 - x_3 y_2)$ . Будем считать, что площадь положительна, если  $OP_1$  переходит в  $OQ_1$  при повороте относительно  $Ox_1$  на угол, меньший  $\pi$ , в положительном направлении. Если проделать эту операцию со всеми тремя проекциями и результаты удвоить, то мы получим

$$x_2 y_3 - x_3 y_2, \quad x_3 y_1 - x_1 y_3, \quad x_1 y_2 - x_2 y_1. \quad (8)$$

Но по теореме из геометрии эти величины представляют собой удвоенную площадь треугольника  $OPQ$ , умноженную на три направляющих косинуса нормали к его плоскости, выбранной так, что вращение относительно нее на угол, меньший  $\pi$ , в положительную сторону переводит  $OP$  в  $OQ$ . Следовательно, удвоенные проекции являются компонентами векторного произведения  $\mathbf{x} \times \mathbf{y}$ .

*Вектор площади.* В определении векторного произведения  $\mathbf{x} \times \mathbf{y}$  направление, взятое для  $\mathbf{n}$ , несущественно. Поворот вправо относительно  $\mathbf{n}$ , переводящий  $\mathbf{x}$  в  $\mathbf{y}$ , может иметь любое значение, меньшее  $2\pi$ . Утверждения последнего раздела остаются верными, если изменить направление всех осей вращения на противоположное, а компоненту брать положительной, когда вращение относительно отрицательного направления оси будет между  $\pi$  и  $2\pi$ . Однако знаки всех компонент изменятся, если  $\mathbf{x}$  и  $\mathbf{y}$  поменять местами. Можно определить направленную площадь треугольника  $OPQ$  так, чтобы она не зависела от обозначения сторон. Это достигается определением *вектора площади*

$$\frac{1}{2} |xy \sin \theta| \mathbf{n} = \frac{1}{2} |\mathbf{x} \times \mathbf{y}| \mathbf{n},$$

где направление  $\mathbf{n}$  выбрано некоторым фиксированным способом. Этот вектор равен векторному произведению, если  $\sin \theta$



положителен, т. е. если угол поворота от  $OP$  к  $OQ$  вокруг  $\mathbf{n}$  в положительную сторону меньше  $\pi$ . Его направление изменится, если изменить направление  $\mathbf{n}$ .

Для однозначного определения вектора площади нужно точно задать направление  $\mathbf{n}$ . Совсем просто это сделать в случае поверхности, составленной из треугольников. С помощью сложения мы можем определить вектор площади для всей такой поверхности; направление  $\mathbf{n}$  определяется так, чтобы оно не пересекало поверхности при переходе с одной грани на смежную с ней грань. В частности, для замкнутого многогранника можно выбрать  $\mathbf{n}$  всегда смотрящим наружу. В этом случае грани, нормаль к которым имеет положительную компоненту  $n_1$ , будут иметь вектор площади, первая компонента которого равна площади проекции многогранника на плоскость  $O23$  с положительным знаком, а грани с отрицательной  $n_1$  будут иметь первую компоненту, равную той же площади, но с обратным знаком. Следовательно, вектор площади многогранника равен нулю.

**2.071. Аналитическое определение векторного произведения:**  
 $\epsilon_{ikm}$ . При аналитическом подходе мы *определяем* векторное произведение как вектор с компонентами

$$(A_2B_3 - A_3B_2, A_3B_1 - A_1B_3, A_1B_2 - A_2B_1).$$

Необходимо доказать, что этот набор величин ведет себя как вектор при преобразованиях координат и что компоненты этого вектора равны  $AB \sin \theta_{n_i}$ . Причина, по которой векторное произведение вообще вводится, заключается в том, что этот набор величин естественно появляется при рассмотрении уравнений динамики, в особенности движения твердого тела и движения заряда под действием магнитной силы, а также в электромагнитной теории.

Рассмотрим набор из 27 чисел  $\epsilon_{ikm}$ , определяемых следующими правилами: 1) если любые два из  $i, k, m$  равны между собой, то  $\epsilon_{ikm} = 0$ ; 2) если все они различны и встречаются в последовательности 12312 ..., т. е. имеют порядок, который мы называем *четным*, то  $\epsilon_{ikm} = 1$ ; 3) если все они различны и встречаются в последовательности 21321 ..., т. е. имеют порядок, который мы называем *нечетным* \*), то  $\epsilon_{ikm} = -1$ . Таким

---

\*) Термины *четный* и *нечетный* обозначают четность числа перестановок индексов, нужных для получения порядка 123. Например, 231 можно превратить одной перестановкой в 132 и затем второй перестановкой в 123. А 213 можно превратить в 123 одной перестановкой.

образом,

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1, \quad (1)$$

$$\varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1, \quad (2)$$

а любые из чисел  $\varepsilon_{111}$ ,  $\varepsilon_{112}$ ,  $\varepsilon_{232}$  и им подобные равны нулю. Рассмотрим сумму

$$\varepsilon_{ikm} A_k B_m. \quad (3)$$

Здесь  $k$  и  $m$  — повторяющиеся индексы. Следовательно, для каждого  $i$  это сумма девяти членов. Но если  $k=i$ ,  $m=i$ , или  $k=m$ , то  $\varepsilon_{ikm}=0$ . Таким образом, единственными членами, которые могут отличаться от нуля, будут два, у которых  $k$  и  $m$  отличаются от  $i$  и друг от друга. Если, например,  $i=1$ , то или  $k=2$ ,  $m=3$  и  $\varepsilon_{ikm}=1$ , или  $k=3$ ,  $m=2$  и  $\varepsilon_{ikm}=-1$ . Следовательно,

$$\varepsilon_{ikm} A_k B_m = A_2 B_3 - A_3 B_2, \quad (4)$$

и аналогично для  $i=2$  и  $3$  мы получим две другие компоненты векторного произведения. Таким образом, (3) дает короткую запись компонент векторного произведения. Чтобы облегчить сравнение с результатами, полученными в векторной форме, мы будем обозначать их  $(\mathbf{A} \times \mathbf{B})_i$ .

Два других важных свойства  $\varepsilon_{ikm}$  заключаются в следующем. Ясно, что для любых  $A_i$

$$\varepsilon_{ikm} A_k A_m = 0, \quad (5)$$

поскольку все слагаемые взаимно уничтожаются. Это аналитическая формулировка того, что *векторное произведение вектора на себя или произвольный параллельный ему вектор равно нулю*. Если  $A_i$ ,  $B_i$ ,  $C_i$  — произвольные наборы из трех величин, то

$$\varepsilon_{ikm} A_i B_k C_m = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} \quad (6)$$

есть определитель, образованный девятью компонентами. Если определитель вырожден, то найдутся такие  $\alpha$ ,  $\beta$ ,  $\gamma$ , что

$$\alpha A_i + \beta B_i + \gamma C_i = 0 \quad (7)$$

для всех  $i$ . Следовательно, вырождение (6) является условием компланарности векторов  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ . Если определитель не вырожден, то уравнения

$$\alpha A_i + \beta B_i + \gamma C_i = D_i \quad (8)$$

имеют единственное решение для любого  $D$ . Мы вновь получили, что *любой вектор может быть линейно выражен через любые три некопланарных вектора*.

Перейдем теперь к *аналитическому* доказательству того, что  $\epsilon_{ikm}A_kB_m$  — компоненты вектора. Это доказательство совершенно независимо от 2.07.

**2.072. Изменение векторного произведения при преобразованиях.** Возьмем две прямые с направляющими косинусами  $l_i, m_i$ . Условие перпендикулярности прямой с направляющими косинусами  $n_i$  к этим двум прямым, записанное полностью, будет

$$\begin{aligned} l_1n_1 + l_2n_2 + l_3n_3 &= 0, \\ m_1n_1 + m_2n_2 + m_3n_3 &= 0. \end{aligned} \quad (1)$$

откуда

$$\begin{aligned} \frac{n_1}{l_2m_3 - l_3m_2} &= \frac{n_2}{l_3m_1 - l_1m_3} = \frac{n_3}{l_1m_2 - l_2m_1} = \\ &= \frac{(n_1^2 + n_2^2 + n_3^2)^{1/2}}{[(l_2m_3 - l_3m_2)^2 + (l_3m_1 - l_1m_3)^2 + (l_1m_2 - l_2m_1)^2]^{1/2}}. \end{aligned} \quad (2)$$

Но сумма квадратов в знаменателе по тождеству Лагранжа равна

$$\begin{aligned} (l_1^2 + l_2^2 + l_3^2)(m_1^2 + m_2^2 + m_3^2) - (l_1m_1 + l_2m_2 + l_3m_3)^2 &= \\ &= 1 - \cos^2 \theta = \sin^2 \theta, \end{aligned} \quad (3)$$

где  $\theta$  — угол между прямыми  $l_i, m_i$ . Поскольку  $n_i$  — направляющие косинусы, числитель в (2) равен  $\pm 1$ . Следовательно, каждое из этих уравнений равно  $\pm \operatorname{cosec} \theta$  и

$$n_i = \pm \operatorname{cosec} \theta \epsilon_{ikm} l_k m_m. \quad (4)$$

Знак определяется выбором направления на прямой  $n_i$ . Если мы выберем это направление так, что  $l_i$  переходит в  $m_i$  при повороте на угол  $\theta$  вправо, то, рассмотрев случай

$$l_i = (1, 0, 0), \quad m_i = (\cos \theta, \sin \theta, 0), \quad n_i = (0, 0, 1),$$

мы видим, что нужно взять положительный знак. Следовательно,

$$n_i = \operatorname{cosec} \theta \epsilon_{ikm} l_k m_m. \quad (5)$$

Теперь, если  $l_i$  и  $m_i$  — данные направления, то перпендикуляр к ним не зависит от системы координат, следовательно,  $n_i$  изменяется по правилу преобразования векторов.

Для двух общих векторов  $\mathbf{A}$  и  $\mathbf{B}$  мы можем определить два направления  $l_i, m_i$  по формулам  $A_i = Al_i, B_i = Bm_i$ . Тогда  $AB \sin \theta$  — скаляр, поскольку  $A, B$  и  $\theta$  не зависят от выбора осей координат; следовательно,  $AB \sin \theta n_i$  — вектор. Но

$$AB \sin \theta n_i = AB \epsilon_{ikm} l_k m_m = \epsilon_{ikm} A_k B_m. \quad (6)$$

Это доказывает, что для двух векторов компоненты векторного произведения преобразуются как компоненты вектора и равняются компонентам, полученным при геометрическом определении.

В индексных обозначениях соотношения 2.07(1) и (2) очевидны, а произведение 2.07(7) не встречается, поскольку нам никогда не приходится рассматривать направляющих векторов координатных осей.

**2.073. Соотношения между  $l_{ij}$ .** Мы можем перейти теперь к доказательству того, что соотношение 2.022(9) следует из 2.021(7). По сути это теорема о непротиворечивости. Если бы она не была верной, то между девятью направляющими косинусами, определяющими преобразование координат, было бы больше шести независимых соотношений. Следовательно, не больше двух параметров преобразования можно было бы выбрать независимо. Однако очевидно, что мы можем повернуть оси на произвольный угол вокруг любой прямой. Направление этой прямой задается двумя параметрами, следовательно, преобразование — тремя. Ясно, что при таком преобразовании система координат остается прямоугольной. Этой информации достаточно для того, чтобы заключить, что 2.022(9) должно следовать из 2.021(7). Однако в нашем рассуждении метрические соотношения 2.021(7) считались независимыми и для проверки желательно прямое доказательство. Обозначим  $l_i$  через  $l_{i1}$ , а  $m_i$  — через  $l_{i2}$ . Поскольку  $OZ'$  перпендикулярно  $Ol_1'$  и  $O2'$  и угол поворота от  $Ol_1'$  к  $O2'$  вправо относительно  $OZ'$  равен прямому\*), то  $\sin \theta = 1$  и

$$l_{13} = \epsilon_{ikm} l_{k1} l_{m2}; \quad (1)$$

аналогично

$$l_{i1} = \epsilon_{ikm} l_{k2} l_{m3}; \quad l_{i2} = \epsilon_{ikm} l_{k3} l_{m1}. \quad (2)$$

Иначе говоря, в определителе

$$L = \begin{vmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{vmatrix} \quad (3)$$

---

\*) См. замечание об ортогональных преобразованиях в конце гл. 4, стр. 276.

каждый элемент равен своему алгебраическому дополнению. Действительно, для каждого элемента любого из столбцов мы получаем это утверждение из соотношений (1) и (2). Если мы разложим  $L$  по элементам первого столбца ( $j=1$ ), то получим  $l_{11}^2$ , что равно 1. Следовательно,  $L=1$ . Но если разложить по элементам первой строки ( $i=1$ ), то получится  $l_{11}^2$ , что, следовательно, тоже равно 1. Аналогично

$$l_{21}^2 = l_{31}^2 = 1. \quad (4)$$

С другой стороны,  $l_{ij}l_{kj}$ , где  $i$  и  $k$  различны, равно определителю, у которого две одинаковые строки, т. е. нулю. Следовательно, для всех  $i, k$

$$l_{ij}l_{kj} = \delta_{ik}. \quad (5)$$

Соотношения (1) и (2) имеют форму, нужную для доказательства теоремы, но они записаны как три уравнения. Их сходство наводит на мысль, что их можно записать как одно, т. е.

$$\epsilon_{jln}l_{ij} = \epsilon_{ikm}l_{kl}l_{mn}. \quad (6)$$

Докажем это. Неповторяющимися индексами будут  $i, l, n$ .

а) Если  $n$  следует за  $l$  в порядке 1231, то единственное значение  $j$ , при котором  $\epsilon_{jln}$  отлично от нуля — это значение, предшествующее  $l$ , и тогда  $\epsilon_{jln}=1$ . Следовательно, в этом случае левая часть сводится к  $l_{ij}$ , где  $j \neq l, n$ , и совпадает с правой частью согласно соотношениям (1), (2).

б) Если  $n$  предшествует  $l$  в порядке 1231, то левая часть равна

$$-l_{ij} \ (j \neq l, n) = -\epsilon_{ikm}l_{kp}l_{ms},$$

где  $jps$  следуют в порядке 12312. Следовательно,  $p=n, s=l$  и правая часть равна

$$-\epsilon_{ikm}l_{kn}l_{ml} = -\epsilon_{imk}l_{mn}l_{kl} = \epsilon_{ikm}l_{kl}l_{mn}.$$

в) Если  $l=n$ , то обе части (6) не изменятся, если  $l$  и  $n$  поменять местами. Но так как знаки обеих частей при этом должны измениться, то они равны нулю. Итак, (6) верно при всех значениях  $l$  и  $n$ . Умножим (6) на  $l_{pl}$ . Тогда

$$\begin{aligned} \epsilon_{jln}l_{ij}l_{pl} &= \epsilon_{ikm}l_{kl}l_{mn}l_{pl} = \\ &= \epsilon_{ikm}l_{mn}\delta_{kp} = \\ &= \epsilon_{ipm}l_{mn}, \end{aligned}$$

и, заменив  $p$  на  $k$ , получаем

$$\epsilon_{ikm}l_{mn} = \epsilon_{jln}l_{ij}l_{kl}. \quad (7)$$

Если определитель записан как

$$\Delta = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix}, \quad (8)$$

где первый индекс обозначает строку, а второй — столбец, то

$$\begin{aligned} \varepsilon_{jln} A_{1j} A_{2l} A_{3n} &= \varepsilon_{jln} A_{2j} A_{3l} A_{1n} = \varepsilon_{jln} A_{3j} A_{1l} A_{2n} = \\ &= -\varepsilon_{jln} A_{2j} A_{1l} A_{3n} = \text{и т. д.}, \end{aligned}$$

и, следовательно,

$$\varepsilon_{ikm} \Delta = \varepsilon_{jln} A_{ij} A_{kl} A_{mn}. \quad (9)$$

Таким образом, определитель, элементы которого занумерованы индексами строки и столбца, можно расписать с помощью символов  $\varepsilon$ . Это можно обобщить для определителей любого порядка (см. [2, гл. 16]).

**2.074. Числа  $\varepsilon_{ikm}\varepsilon_{psm}$ .** Одним из наиболее важных свойств чисел  $\varepsilon_{ikm}$  является то, что 81, число  $\varepsilon_{ikm}\varepsilon_{psm}$ , удовлетворяет некоторому тождеству. Здесь, разумеется, нужно просуммировать по  $m$ , но каждому набору индексов  $i, k, p, s$  соответствует свое выражение. Поскольку каждый из этих четырех индексов может принимать только три значения, то в каждом выражении хотя бы два из них имеют одинаковое значение. Очевидно, что если  $i=k$ , или  $p=s$ , то слагаемые равны нулю. Если  $i \neq k$ , то  $\varepsilon_{ikm}$  не обращается в нуль только при одном значении  $m$ . Тогда, чтобы  $\varepsilon_{psm}$  отличалось от нуля,  $p$  и  $s$  должны принимать те же значения, что  $i$  и  $k$ , в любом порядке. Если порядок одинаковый, то  $\varepsilon_{ikm}$  и  $\varepsilon_{psm}$  равны оба  $+1$  или  $-1$  и произведение равно 1. Если же порядок различен, то произведение равно  $-1$ . Итак,

$$\varepsilon_{ikm}\varepsilon_{psm} = \begin{cases} 0 & (i=k), \\ 0 & (p=s), \\ 1 & (i=p, k=s), \\ -1 & (i=s, k=p), \\ 0 & (i \neq p, s \text{ или } k \neq p, s). \end{cases}$$

Рассмотрим теперь набор чисел

$$\delta_{ip}\delta_{ks} - \delta_{is}\delta_{kp}.$$

Если  $i=k$  или  $p=s$ , то слагаемые взаимно уничтожаются. Если  $i \neq p$  и  $s$  или  $k \neq p$  и  $s$ , то один из сомножителей каждого слагаемого равен 0. Следовательно, ненулевыми будут

выражения, у которых  $i, k$  равны  $p, s$  в любом порядке и  $i \neq k$ ,  $p \neq s$ . Если  $i = p$  и  $k = s$ , то первое слагаемое равно 1, а второе — 0, а если  $i = s$  и  $k = p$ , то первое слагаемое равно 0, а второе 1. Итак, при всевозможных значениях  $i, k, p, s$

$$\varepsilon_{ikm}\varepsilon_{psm} = \delta_{ip}\delta_{ks} - \delta_{is}\delta_{kp}. \quad (1)$$

Мы будем часто встречать это тождество в различных приложениях.

Поскольку

$$\varepsilon_{psm} = \varepsilon_{mps} = \varepsilon_{smp} \quad (2)$$

при всех значениях индексов, то выражения в левой части (1) не изменятся по величине, если мы заменим  $ikm$  на  $kmi$  или  $mik$ . Можно, следовательно, высказать общее правило знаков: *положим, что  $i$  и  $p$  — индексы, которые следуют за повторяющимся индексом в соответствующем сомножителе  $\varepsilon$  (если повторяющийся индекс — последний, полагаем соответственно  $i$  или  $p$  первым); тогда  $\delta_{ip}$  появляется с положительным знаком и формулу можно дописать по симметрии.*

**2.08. Деление векторов.** Деление векторов не может быть определено однозначно и поэтому его избегают. Легко видеть, что если задан ненулевой вектор  $\mathbf{A}$  и скаляр  $M$ , то можно указать такой вектор  $\mathbf{B}$ , что скалярное произведение  $\mathbf{A} \cdot \mathbf{B} = M$ . Однако вектор  $\mathbf{B}$  определен не однозначно, так как скалярное произведение не изменится, если добавить к  $\mathbf{B}$  любой вектор, перпендикулярный  $\mathbf{A}$ . Вообще говоря, нельзя указать такого вектора  $\mathbf{B}$ , что векторное произведение  $\mathbf{A} \times \mathbf{B} = \mathbf{C}$ , а  $\mathbf{A}$  и  $\mathbf{C}$  — данные векторы. Действительно, вектор  $\mathbf{A} \times \mathbf{B}$  перпендикулярен  $\mathbf{A}$ , и если  $\mathbf{C}$  и  $\mathbf{A}$  не перпендикулярны, то вектор  $\mathbf{B}$  вообще нельзя указать. Если, однако,  $\mathbf{C}$  перпендикулярен  $\mathbf{A}$ , то можно добавить к  $\mathbf{B}$  любой вектор, параллельный  $\mathbf{A}$ . Это не изменит векторного произведения, следовательно, частное определено не однозначно.

В теле *кватернионов*, которое является расширением векторной алгебры, деление определено однозначно. До последнего времени лишь немногие физики использовали кватернионы, однако теперь их начинают применять в квантовой теории. (Ср. гл. 4, пример 12.)

**2.09. Тройные произведения.** Скалярное произведение векторов  $\mathbf{B} \times \mathbf{C}$  и  $\mathbf{A}$  называется смешанным произведением векторов  $\mathbf{A}, \mathbf{B}$  и  $\mathbf{C}$ . Мы покажем, что при циклической перестановке векторов смешанное произведение  $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$  не меняется, оно также не меняется, если поменять крест и точку местами.

Имеется, следовательно, шесть возможных способов записать это произведение.

Мы можем также образовать векторное произведение  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$  вектора  $\mathbf{A}$  с векторным произведением векторов  $\mathbf{B}$  и  $\mathbf{C}$ .

Рассмотрим сперва несколько частных случаев.

I.  $\mathbf{B} \cdot (\mathbf{B} \times \mathbf{C})$ . Очевидно, что

$$\mathbf{B} \cdot (\mathbf{B} \times \mathbf{C}) = 0, \quad (1)$$

так как  $\mathbf{B} \times \mathbf{C}$  перпендикулярно  $\mathbf{B}$ .

II.  $\mathbf{B} \times (\mathbf{B} \times \mathbf{C})$ . Вектор  $\mathbf{B} \times (\mathbf{B} \times \mathbf{C})$  перпендикулярен как  $\mathbf{B}$ , так и  $\mathbf{B} \times \mathbf{C}$ , и, следовательно, он лежит в плоскости векторов  $\mathbf{B}$  и  $\mathbf{C}$ , перпендикулярен  $\mathbf{B}$  и его направление получается поворотом  $\mathbf{B} \times \mathbf{C}$  на прямой угол вправо относительно  $\mathbf{B}$ . Его величина в  $B$  раз больше, чем  $\mathbf{B} \times \mathbf{C}$ , т. е. равна  $B^2 C \sin \theta$ .

Из рис. 6 следует, что

$$\begin{aligned} \mathbf{B} \times (\mathbf{B} \times \mathbf{C}) &= B^2 C \sin \theta \operatorname{ctg} \theta \frac{\mathbf{B}}{B} - \\ &- B^2 C \sin \theta \operatorname{cosec} \theta \frac{\mathbf{C}}{C}, \quad (2) \end{aligned}$$

$$\mathbf{B} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{B} \cdot \mathbf{C}) \mathbf{B} - B^2 \mathbf{C}. \quad (3)$$

Аналогично

$$\begin{aligned} \mathbf{C} \times (\mathbf{B} \times \mathbf{C}) &= -\mathbf{C} \times (\mathbf{C} \times \mathbf{B}) = \\ &= C^2 \mathbf{B} - (\mathbf{B} \cdot \mathbf{C}) \mathbf{C}. \quad (4) \end{aligned}$$

$$\begin{aligned} \text{III. } (\mathbf{B} \times \mathbf{C}) \cdot (\mathbf{B} \times \mathbf{C}) &= \\ &= B^2 C^2 - (\mathbf{B} \cdot \mathbf{C})^2. \quad (5) \end{aligned}$$

Это немедленно следует из определений.

2.091. Смешанное произведение  $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ . Из коммутативности скалярного умножения следует, что

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{A}. \quad (6)$$

Далее, поскольку  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{B} \times \mathbf{C}$  некопланарны, если ни один из них не равен нулю, мы можем записать любой вектор  $\mathbf{A}$  так:

$$\mathbf{A} = \alpha \mathbf{B} + \beta \mathbf{C} + \gamma \mathbf{B} \times \mathbf{C}. \quad (7)$$

Тогда

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \gamma (\mathbf{B} \times \mathbf{C})^2 = \gamma [B^2 C^2 - (\mathbf{B} \cdot \mathbf{C})^2]. \quad (8)$$

И также

$$\mathbf{A} \times \mathbf{B} = \beta \mathbf{C} \times \mathbf{B} - \gamma \mathbf{B} \times (\mathbf{B} \times \mathbf{C}) = \beta \mathbf{C} \times \mathbf{B} + \gamma B^2 \mathbf{C} - \gamma (\mathbf{B} \cdot \mathbf{C}) \mathbf{B}, \quad (9)$$

так что

$$\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \gamma [B^2 C^2 - (\mathbf{B} \cdot \mathbf{C})^2] = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}). \quad (10)$$

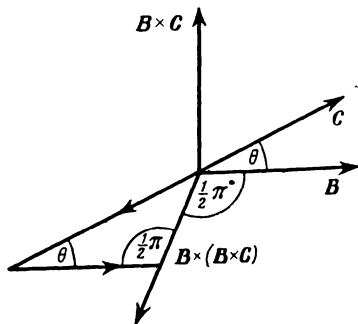


Рис. 6.



Аналогично доказывается, что

$$\mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}). \quad (11)$$

Итак, имеем шесть равных произведений

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} \times \mathbf{C} &= \mathbf{B} \cdot \mathbf{C} \times \mathbf{A} = \mathbf{C} \cdot \mathbf{A} \times \mathbf{B} = \mathbf{A} \times \mathbf{B} \cdot \mathbf{C} = \\ &= \mathbf{B} \times \mathbf{C} \cdot \mathbf{A} = \mathbf{C} \times \mathbf{A} \cdot \mathbf{B}. \end{aligned} \quad (12)$$

Скобки не нужны, поскольку порядок операций определен однозначно. Если взять циклический порядок  $\mathbf{A}$ ,  $\mathbf{C}$ ,  $\mathbf{B}$  вместо  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , то знак изменится.

*Геометрический смысл.* Поскольку модуль  $\mathbf{B} \times \mathbf{C}$  равен площади параллелограмма, натянутого на отрезки, представляющие  $\mathbf{B}$  и  $\mathbf{C}$ , то, если угол между  $\mathbf{A}$  и  $\mathbf{B} \times \mathbf{C}$  острый,  $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$  равно объему параллелепипеда, натянутого на отрезки, представляющие  $\mathbf{A}$ ,  $\mathbf{B}$  и  $\mathbf{C}$ . Если угол тупой, то  $-\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$  равно этому объему.

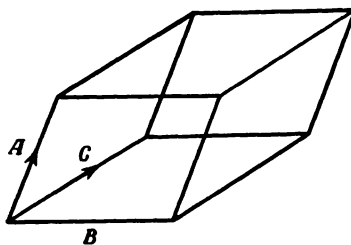


Рис. 7.

Смешанное произведение  $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$  иногда записывают  $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ .

**2.092. Тройное векторное произведение  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ .** Как и выше, запишем

$$\mathbf{A} = \alpha \mathbf{B} + \beta \mathbf{C} + \gamma \mathbf{B} \times \mathbf{C}. \quad (13)$$

Тогда

$$\begin{aligned} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= \alpha \mathbf{B} \times (\mathbf{B} \times \mathbf{C}) + \beta \mathbf{C} \times (\mathbf{B} \times \mathbf{C}) = \\ &= \alpha [(\mathbf{B} \cdot \mathbf{C}) \mathbf{B} - B^2 \mathbf{C}] + \beta [C^2 \mathbf{B} - (\mathbf{B} \cdot \mathbf{C}) \mathbf{C}] = \\ &= [(\alpha \mathbf{B} + \beta \mathbf{C}) \cdot \mathbf{C}] \mathbf{B} - [(\alpha \mathbf{B} + \beta \mathbf{C}) \cdot \mathbf{B}] \mathbf{C} = \\ &= (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{A} \cdot \mathbf{B}) \mathbf{C}. \end{aligned} \quad (14)$$

Аналогично можно показать, что

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{B} \cdot \mathbf{C}) \mathbf{A}. \quad (15)$$

Следует отметить: 1) закон ассоциативности не выполняется, 2) член в правой части, в котором средний вектор левой части появляется в скалярном произведении, имеет отрицательный знак (правило для запоминания знаков).

**2.093.** В индексной системе обозначений эквивалентность разных форм смешанного произведения очевидна. Действи-

тельно,

$$A_i \varepsilon_{ikm} B_k C_m = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix},$$

и выражения  $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ ,  $\mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})$ ,  $\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})$  являются разложениями этого определителя по элементам различных строк. Остальные три выражения следуют из соотношения  $A_i B_i = B_i A_i$ .

**2.094.** Выражение тройного векторного произведения в индексной системе обозначений связано с тождеством **2.074** (1) для  $\varepsilon_{ikm}$ . Заметим вначале, что использование правила суммирования требует, чтобы всякий повторяющийся индекс появлялся *только* дважды. Иначе не однозначен порядок суммирования. Учитывая это, мы запишем  $i$ -ю компоненту тройного векторного произведения  $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$  так:

$$\begin{aligned} \varepsilon_{ikm} A_k \varepsilon_{mps} B_p C_s &= \varepsilon_{ikm} \varepsilon_{mps} A_k B_p C_s = \\ &= (\delta_{ip} \delta_{ks} - \delta_{is} \delta_{kp}) A_k B_p C_s = \\ &= \delta_{ks} A_k B_i C_s - \delta_{kp} A_k B_p C_i = \\ &= B_i A_k C_k - C_i A_k B_k, \end{aligned}$$

следовательно,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{A} \cdot \mathbf{B}) \mathbf{C}.$$

Аналогично

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = -\mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = -(\mathbf{C} \cdot \mathbf{B}) \mathbf{A} + (\mathbf{C} \cdot \mathbf{A}) \mathbf{B}.$$

Запомнить эти формулы непросто, однако их быстро можно вывести из тождества **2.074** (1), которое намного легче запомнить. Кроме того, оно имеет другие применения.

**2.10. Векторные функции скалярного аргумента. Дифференцирование.** Обозначим скалярную переменную через  $t$ . Пусть  $\mathbf{A}'(t)$  — векторная функция. Определим производную  $\mathbf{A}(t)$  по  $t$  следующим образом. Рассмотрим отношение

$$\frac{\delta \mathbf{A}}{\delta t} = \frac{\mathbf{A}(t + \delta t) - \mathbf{A}(t)}{\delta t}.$$

Если  $\delta t$  отлично от нуля, то ясно, что  $\delta \mathbf{A} / \delta t$  — вектор, и если при  $\delta t \rightarrow 0$  он имеет предел, назовем этот вектор производной вектор-функции  $\mathbf{A}(t)$  по  $t$  и будем обозначать его  $d\mathbf{A}/dt$ . Формальное доказательство того, что в пределе получается вектор, легко следует из теоремы о том, что сумма пределов двух функций равна пределу суммы этих функций.

Вектор  $d\mathbf{A}/dt$  имеет компоненты  $(dA_1/dt, dA_2/dt, dA_3/dt)$ . Важно отметить, что, вообще говоря, не только модуль  $d\mathbf{A}/dt$  отличается от модуля  $\mathbf{A}$ , но и *направления* их тоже различны. В частности, производная вектора с постоянным модулем и переменным направлением отлична от нуля.

**Дифференцирование произведения.** Правило дифференцирования произведения двух скалярных функций можно легко обобщить для произведения скалярной и векторной функций, а также для скалярного и векторного произведения. Здесь мы просто сформулируем результаты. Доказательства в каждом случае просты и предоставляются читателю.

1) Если  $\alpha$  — скалярная функция  $t$ , то

$$\frac{d}{dt}(\alpha\mathbf{A}) = \frac{d\alpha}{dt}\mathbf{A} + \alpha\frac{d\mathbf{A}}{dt}, \quad \frac{d}{dt}(\alpha A_i) = \frac{d\alpha}{dt}A_i + \alpha\frac{dA_i}{dt}.$$

2) Если  $\mathbf{A}$  и  $\mathbf{B}$  — две векторные функции  $t$ , то

$$\frac{d}{dt}(\mathbf{A} \cdot \mathbf{B}) = \frac{d\mathbf{A}}{dt} \cdot \mathbf{B} + \mathbf{A} \cdot \frac{d\mathbf{B}}{dt}, \quad \frac{d}{dt}(A_i B_i) = B_i \frac{dA_i}{dt} + A_i \frac{dB_i}{dt}.$$

Порядок сомножителей в произведениях не существен.

$$3) \quad \frac{d}{dt}(\mathbf{A} \times \mathbf{B}) = \frac{d\mathbf{A}}{dt} \times \mathbf{B} + \mathbf{A} \times \frac{d\mathbf{B}}{dt},$$

$$\frac{d}{dt} \varepsilon_{ikm} A_k B_m = \varepsilon_{ikm} \frac{dA_k}{dt} B_m + \varepsilon_{ikm} A_k \frac{dB_m}{dt}.$$

В векторной записи сомножители переставлять нельзя. В индексном обозначении перестановка сомножителей допустима.

Важны некоторые вытекающие из этого частные результаты:

1) Если  $\mathbf{A}$  — вектор с *постоянным модулем*, то

$$\frac{d}{dt}(A^2) \equiv \frac{d}{dt}(\mathbf{A} \cdot \mathbf{A}) = 0,$$

т. е.

$$\mathbf{A} \cdot \frac{d\mathbf{A}}{dt} = 0,$$

из чего следует, что  $d\mathbf{A}/dt$  перпендикулярен  $\mathbf{A}$ .

Ортогональность  $d\mathbf{A}/dt$  и  $\mathbf{A}$  можно усмотреть геометрически. На рис. 8 отрезки, представляющие  $\mathbf{A}$  и  $\mathbf{A} + \delta\mathbf{A}$ , имеют одинаковую длину. При стремлении  $Q$  к  $P$  угол между  $OP$  и  $PQ$  стремится к прямому.

2) Если векторная функция  $\mathbf{A}$  записана как произведение модуля  $A$  и направляющего вектора  $\mathbf{n}$ , то

$$\frac{d\mathbf{A}}{dt} = \frac{dA}{dt} \mathbf{n} + A \frac{d\mathbf{n}}{dt}.$$

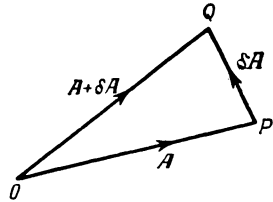


Рис. 8.

Значит, если  $A_i = Al_i$ , то

$$\frac{dA_i}{dt} = l_i \frac{dA}{dt} + A \frac{dl_i}{dt}.$$

Следует отметить, что  $|dA/dt|$ , вообще говоря, не равен  $|dA/dt|$ .

**2.11. Движение частицы в поле силы тяжести с сопротивлением, пропорциональным скорости.** Пусть при  $t=0$  частица находится в начале координат. Сопротивление действует по касательной к траектории движения в направлении, противоположном движению. Следовательно, сопротивление выражается вектором  $-m\kappa\mathbf{v}$ , где  $m$  — масса частицы, а  $\kappa$  — константа. Пусть  $\mathbf{k}$  — направляющий вектор, идущий вертикально вверх. Тогда, приравнявая произведение массы на ускорение величине действующей силы, получаем

$$m\ddot{\mathbf{x}} = -mg\mathbf{k} - m\kappa\dot{\mathbf{x}} \quad (1)$$

или

$$\ddot{\mathbf{x}} + \kappa\dot{\mathbf{x}} = -g\mathbf{k}. \quad (2)$$

Проинтегрируем это уравнение. Его можно записать так:

$$\frac{d}{dt}(\dot{\mathbf{x}}e^{\kappa t}) = -ge^{\kappa t}\mathbf{k}, \quad (3)$$

откуда

$$\dot{\mathbf{x}}e^{\kappa t} = -\frac{g}{\kappa}e^{\kappa t}\mathbf{k} + \mathbf{V} + \frac{g}{\kappa}\mathbf{k}, \quad (4)$$

где  $\mathbf{V}$  — скорость при  $t=0$ . Отсюда

$$\dot{\mathbf{x}} = e^{-\kappa t}\mathbf{V} - \frac{g}{\kappa}(1 - e^{-\kappa t})\mathbf{k} \quad (5)$$

и

$$\mathbf{x} = \frac{\mathbf{V}}{\kappa}(1 - e^{-\kappa t}) - \frac{gt}{\kappa}\mathbf{k} + \frac{g}{\kappa^2}(1 - e^{-\kappa t})\mathbf{k}, \quad (6)$$

так как при  $t=0$   $\mathbf{x}=0$ .

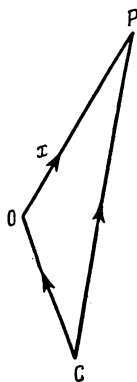
Рис. 9.

Из этого уравнения сразу видно, что в момент времени  $t$  все частицы, выпущенные из точки  $O$  со скоростью  $V$ , окажутся на окружности, центр которой лежит на расстоянии  $\frac{gt}{\kappa} - \frac{g}{\kappa^2}(1 - e^{-\kappa t})$  ниже нуля. Действительно,

$$\left| \mathbf{x} + \frac{gt}{\kappa}\mathbf{k} - \frac{g}{\kappa^2}(1 - e^{-\kappa t})\mathbf{k} \right| = \frac{V}{\kappa}(1 - e^{-\kappa t}), \quad (7)$$

т. е.

$$|\overline{CO} + \mathbf{x}| = CP = \frac{V}{\kappa}(1 - e^{-\kappa t}). \quad (8)$$



Следовательно,  $CP$  равно  $\frac{V}{\kappa}(1 - e^{-\kappa t})$  и не зависит от направления вектора скорости.

Дифференцируя (2) по времени, получаем

$$\dot{\mathbf{a}} + \kappa \mathbf{a} = 0, \quad (9)$$

т. е. если ускорение при  $t = 0$  равно  $\mathbf{a}_0$ , то

$$\mathbf{a} = \mathbf{a}_0 e^{-\kappa t}. \quad (10)$$

Таким образом, направление ускорения не меняется во время движения. Также если  $u$  — горизонтальная составляющая скорости и  $u_0$  — ее начальное значение, то из (5) получаем, что

$$u = u_0 e^{-\kappa t}. \quad (11)$$

Если  $d$  — расстояние, пройденное в горизонтальном направлении за время  $t$ , то, поскольку  $u = \dot{d}$ ,

$$d = \frac{u_0}{\kappa}(1 - e^{-\kappa t}). \quad (12)$$

Следовательно,

$$e^{-\kappa t} = 1 - \frac{\kappa d}{u_0} \quad (13)$$

и (10) можно переписать как

$$\mathbf{a} = \mathbf{a}_0 \left(1 - \frac{\kappa d}{u_0}\right). \quad (14)$$

**2.12. Движение заряженной частицы в перпендикулярных электрическом и магнитном полях.** Пусть  $m$  и  $-e$  масса и электрический заряд частицы;  $c$  — скорость света;  $\mathbf{E}$ ,  $\mathbf{H}$  — электрическое и магнитное поля в гауссовой системе единиц. Тогда уравнение движения имеет вид

$$m\ddot{\mathbf{x}} = -e\mathbf{E} - \frac{e}{c}\dot{\mathbf{x}} \times \mathbf{H}, \quad (1)$$

т. е.

$$\ddot{x}_i = -\frac{e}{m} E_i - \frac{e}{mc} \varepsilon_{ikm} \dot{x}_k H_m. \quad (2)$$

Возьмем

$$\mathbf{E} = (E, 0, 0), \quad \mathbf{H} = (0, 0, H). \quad (3)$$

Тогда

$$\ddot{x}_1 = -\frac{e}{m} E - \frac{e}{mc} \dot{x}_2 H, \quad (4)$$

$$\ddot{x}_2 = \frac{e}{mc} \dot{x}_1 H, \quad (5)$$

$$\ddot{x}_3 = 0. \quad (6)$$

Пусть в начальный момент времени частица находится в начале координат и имеет нулевую скорость. Тогда  $x_3 \equiv 0$ , и движение будет происходить в плоскости, перпендикулярной магнитному полю. Умножим (5) на  $i$  и сложим с (4). Положим  $x_1 + ix_2 = z^*$ . Тогда

$$\ddot{z} = -\frac{eE}{m} + \frac{ieH}{mc} \dot{z}.$$

Пусть  $eH/mc = \omega$ ; тогда

$$\ddot{z} - i\omega \dot{z} = -\frac{eE}{m}, \quad \dot{z} = -\frac{ieE}{m\omega} (1 - e^{i\omega t}),$$

так как  $\dot{z} = 0$  при  $t = 0$ , и

$$z = -\frac{ieE}{m\omega} \left[ t - \frac{e^{i\omega t} - 1}{i\omega} \right] = -\frac{eE}{m\omega^2} (i\omega t - e^{i\omega t} + 1),$$

$$x_1 = -\frac{eE}{m\omega^2} (1 - \cos \omega t),$$

$$x_2 = -\frac{eE}{m\omega^2} (\omega t - \sin \omega t).$$

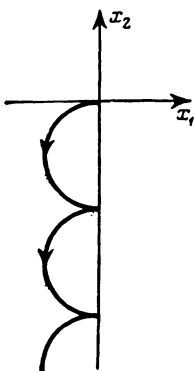


Рис. 10.

Траектория будет циклоидой с точками возврата, лежащими на отрицательной части оси  $x_2$ .

**2.13. Элемент углового смещения; угловая скорость.** Пусть частица первоначально находится в точке  $P(x)$ . Повернем ее на малый угол  $\delta\theta$  (вправо) относительно прямой  $ON$  с направляющими косинусами  $n_i$ . Пусть угол между осью вращения и  $OP$  равен  $\alpha$ . Тогда смещение точки  $P$  с точностью до первого порядка малости по  $\delta\theta$  перпендикулярно к плоскости векторов  $\mathbf{n}$  и  $\mathbf{x}$  и равно по модулю  $r \sin \alpha \delta\theta$ . Итак, смещение  $\delta\mathbf{x}$  задается формулой

$$\delta\mathbf{x} = \delta\theta \cdot \mathbf{n} \times \mathbf{x} + O[(\delta\theta)^2], \quad (1)$$

поскольку  $|\mathbf{n} \times \mathbf{x}| = r \sin \alpha$ ; или если мы положим

$$\delta\theta = \mathbf{n} \delta\theta, \quad (2)$$

то

$$\delta\mathbf{x} = \delta\theta \times \mathbf{x} + O[(\delta\theta)^2]. \quad (3)$$

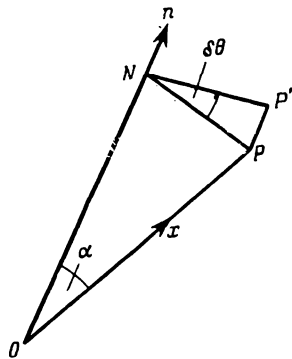


Рис. 11.

Величина  $\delta\theta$  является вектором, потому что  $\delta\theta$  — скаляр, а  $n_i$  — направляющие косинусы данной прямой.

\*) Если читатель еще не знаком с элементами теории функций комплексного переменного, то здесь ему следует прочитать начало гл. 11.

Таким образом, если  $\mathbf{v}$  — скорость точки  $P$  и  $\delta\theta/\delta t$  имеет предел  $\omega$  при  $\delta t \rightarrow 0$ , то

$$\mathbf{v} = \lim \frac{\delta \mathbf{x}}{\delta t} = \omega \times \mathbf{x}, \quad (4)$$

где

$$\omega = \omega \mathbf{n}. \quad (5)$$

Наоборот, если скорость  $\dot{\mathbf{x}}$  задается выражением вида (4) для всех  $t$ , а  $\omega$  постоянно по величине и направлению, то легко показать, что это равномерное движение по окружности. Действительно,  $\dot{\mathbf{x}} \cdot \omega = 0$ , и, следовательно, движение частицы происходит в плоскости, перпендикулярной  $\omega$ . Кроме того,  $\dot{\mathbf{x}} \cdot \mathbf{x} = 0$ , поэтому  $r^2$  постоянно, и частица движется по сфере. Таким образом, движение происходит по окружности радиуса  $r \sin \alpha$ . Наконец,

$$\dot{\mathbf{x}} \cdot \dot{\mathbf{x}} = (\omega \times \mathbf{x}) \cdot (\omega \times \mathbf{x}) = (\omega r \sin \alpha)^2, \quad (6)$$

и, следовательно, скорость постоянна и угловая скорость равна  $\omega$ . Знак может быть проверен отдельно.

Уравнение (4) можно рассматривать как систему трех дифференциальных уравнений, именно

$$\dot{x}_i = \varepsilon_{ikm} \omega_k x_m. \quad (7)$$

Читателю следует самостоятельно продифференцировать (6), чтобы попрактиковаться в использовании  $\varepsilon_{ikm}$ . Уравнения (7) можно использовать для того, чтобы проиллюстрировать метод, который часто бывает полезен, если одна из осей выбрана специально. Возьмем ось  $x_3$  в направлении  $\omega$ . Тогда  $\omega = (0, 0, \omega)$  и

$$\dot{x}_1 = -\omega x_2, \quad \dot{x}_2 = \omega x_1, \quad \dot{x}_3 = 0. \quad (8)$$

Отсюда  $x_3$  — константа. Умножим второе уравнение на  $i$  и прибавим к первому \*). Положим  $\zeta = x_1 + ix_2$ . Тогда

$$\dot{\zeta} = i\omega \zeta, \quad \zeta = C e^{i\beta} e^{i\omega t}, \quad (9)$$

где  $C$  и  $\beta$  — действительные числа. Действительная и мнимая части дают

$$x_1 = C \cos(\omega t + \beta), \quad x_2 = C \sin(\omega t + \beta). \quad (10)$$

Эти уравнения описывают равномерное движение по окружности радиуса  $C$ , и решение, как и должно быть, содержит три параметра, именно  $x_3$ ,  $C$  и  $\beta$ .

\*) См. примечание на стр. 142.

Теперь мы можем решить уравнение 2.12(1) по-другому. Интегрируя один раз, получаем

$$m\dot{\mathbf{x}} = -e\mathbf{E}t - \frac{e}{c} \mathbf{x} \times \mathbf{H}. \quad (11)$$

Положим

$$\mathbf{E} = E\mathbf{e}_{(1)} = E\mathbf{e}_{(2)} \times \mathbf{e}_{(3)}, \quad \mathbf{H} = H\mathbf{e}_{(3)}. \quad (12)$$

Тогда (11) можно переписать в виде

$$\dot{\mathbf{x}} = \frac{e\mathbf{H}}{mc} \times \left( \mathbf{x} + \frac{cEt}{H} \mathbf{e}_{(2)} \right),$$

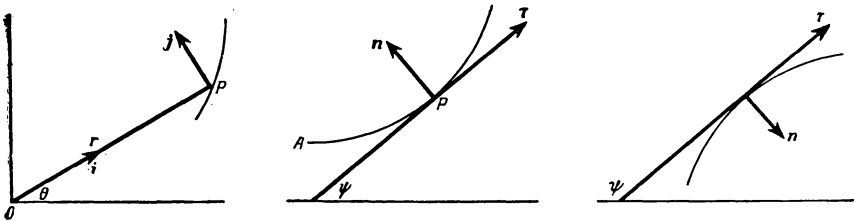
что мы немедленно интерпретируем так: частица движется с угловой скоростью  $eH/mc$  вокруг оси, параллельной  $\mathbf{e}_{(3)}$ , которая в свою очередь движется с постоянной скоростью  $-\frac{cE}{H} \mathbf{e}_{(2)}$ .

**2.131.** В частности, если  $\mathbf{i}$  и  $\mathbf{j}$  — два взаимно перпендикулярных направляющих вектора в плоскости  $O1$ ,  $O2$  и в момент времени  $t$  вектор  $\mathbf{i}$  образует угол  $\theta$ , а  $\mathbf{j}$  — угол  $1/2\pi + \theta$  с  $O1$ , то

$$\frac{d\mathbf{i}}{dt} = \theta\mathbf{j}, \quad (1)$$

$$\frac{d\mathbf{j}}{dt} = -\theta\mathbf{i}. \quad (2)$$

Если в момент времени  $t$  радиус-вектор движущейся частицы равен  $\mathbf{x}$ , то проекции ее скорости  $d\mathbf{x}/dt$  и ускорения  $d^2\mathbf{x}/dt^2$



Р и с. 12.

на оси  $O1$  и  $O2$  будут соответственно  $(\dot{x}_1, \dot{x}_2)$  и  $(\ddot{x}_1, \ddot{x}_2)$ . Если  $\mathbf{i}$  и  $\mathbf{j}$  — направляющие векторы  $\mathbf{x}$  и перпендикулярного к нему направления, то

$$\mathbf{x} = r\mathbf{i}, \quad \frac{d\mathbf{x}}{dt} = \dot{r}\mathbf{i} + r \frac{d\mathbf{i}}{dt} = \dot{r}\mathbf{i} + r\theta\mathbf{j} \quad (3)$$

и

$$\frac{d^2\mathbf{x}}{dt^2} = \ddot{r}\mathbf{i} + \dot{r} \frac{d\mathbf{i}}{dt} + \frac{d}{dt}(r\theta)\mathbf{j} + r\theta \frac{d\mathbf{j}}{dt} = (\ddot{r} - r\theta^2)\mathbf{i} + \frac{1}{r} \frac{d}{dt}(r^2\theta)\mathbf{j}. \quad (4)$$



Выражения (3) и (4) дают разложение скорости и ускорения по направлениям радиуса-вектора и перпендикулярно к нему.

Аналогично если  $\tau$  и  $\mathbf{n}$  — направляющие векторы касательной и внутренней нормали к траектории движения частицы, а  $\psi$  — угол между  $\tau$  и фиксированным направлением, которое выбрано так, чтобы  $\psi$  возрастал при движении по траектории в направлении  $\tau$ , то

$$\frac{d\tau}{dt} = \frac{d\psi}{dt} \mathbf{n} = \frac{d\psi}{ds} \dot{s} \mathbf{n} = \frac{\dot{s}}{\rho} \mathbf{n},$$

где  $s$  — расстояние от фиксированной точки, взятое вдоль траектории, а  $\rho$  — радиус кривизны (здесь он берется положительным). Теперь вектор скорости  $\mathbf{v}$  можно записать

$$\mathbf{v} = v\tau = \dot{s}\tau;$$

следовательно, ускорение  $d\mathbf{v}/dt$  выражается формулой

$$\frac{d\mathbf{v}}{dt} = \dot{v}\tau + \frac{v^2}{\rho} \mathbf{n}.$$

**2.14. Угловая скорость абсолютно твердого тела.** Любое движение абсолютно твердого тела сводится к параллельному переносу, т. е. движению, при котором все частицы смещаются одинаково, и последующему повороту около некоторой оси. По определению абсолютно твердого тела при любых его возможных перемещениях попарные расстояния между точками остаются неизменными. Пусть рассматриваемое перемещение переводит частицу из точки  $O$  в  $O'$ . Сначала предположим, что каждая частица получает смещение  $OO'$ ; при этом попарные расстояния между частицами не меняются. Пусть при этом перемещении частицы  $P, Q$  переходят в  $P', Q'$ , а при действительном перемещении — в  $P'', Q''$ . Согласно метрическому определению, плоскость — геометрическое место точек, равноудаленных от двух заданных точек. Следовательно, точки, равноудаленные от  $P'$  и  $P''$ , образуют плоскость, и  $O'$  лежит в этой плоскости, поскольку  $O'P'$  и  $O'P''$  равны  $OP$ . Аналогично точки, равноудаленные от  $Q'$  и  $Q''$ , образуют плоскость, проходящую через  $O'$ . Эти две плоскости пересекаются по прямой, каждая точка которой  $R'$  удовлетворяет соотношениям  $R'P' = R'P''$ ,  $R'Q' = R'Q''$ . Поэтому, так как расстояние между точками этой прямой и точкой  $O'$  также не меняется, она состоит из одних

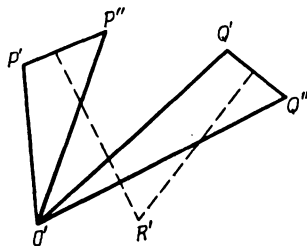


Рис. 13.

и тех же частиц как при смещении на  $OO'$ , так и при действительном перемещении.

Рассмотрим теперь произвольную пару частиц  $S, T$  в этих трех положениях. Углы между плоскостями сохраняются, следовательно, угол между плоскостями  $O'R'S'$  и  $O'R'S''$  равен углу между плоскостями  $O'R'T'$  и  $O'R'T''$ . Таким образом, положение  $P'Q'$  может быть переведено в положение  $P''Q''$  поворотом на определенный угол вокруг  $O'R'$ .

Пусть положение точки  $O$  задается вектором  $\mathbf{a}$ , а точки  $O'$  — вектором  $\mathbf{a} + \delta\mathbf{a}$ . Пусть  $n_i$  — направляющие косинусы прямой  $O'R'$ , и пусть вокруг нее производится поворот на малый угол  $\delta\theta$ . Тогда смещение от  $P(\mathbf{x})$  к  $P'$  есть  $\delta\mathbf{a}$  и

$$\overline{P'P''} = \mathbf{n} \delta\theta \times \overline{O'P'} + O(\delta\theta)^2. \quad (1)$$

Положим, как и раньше,  $\mathbf{n} \delta\theta = \delta\boldsymbol{\theta}$ ; тогда

$$\overline{O'P'} = \overline{OP} = \mathbf{x} - \mathbf{a}, \quad (2)$$

$$\overline{PP''} = \overline{PP'} + \overline{P'P''} = \delta\mathbf{a} + \delta\boldsymbol{\theta} \times (\mathbf{x} - \mathbf{a}) + O(\delta\theta)^2, \quad (3)$$

что дает смещение произвольной частицы тела. В этом случае скорость точки  $P$  будет

$$\mathbf{v} = \lim_{\delta t \rightarrow 0} \frac{\overline{PP''}}{\delta t} = \frac{d\mathbf{a}}{dt} + \lim_{\delta t \rightarrow 0} \frac{d\boldsymbol{\theta}}{dt} \times (\mathbf{x} - \mathbf{a}) = \dot{\mathbf{a}} + \boldsymbol{\omega} \times (\mathbf{x} - \mathbf{a}), \quad (4)$$

где

$$\boldsymbol{\omega} = \lim_{\delta t \rightarrow 0} \frac{\delta\boldsymbol{\theta}}{\delta t}, \quad (5)$$

и  $\boldsymbol{\omega}$  называется угловой скоростью тела. В других обозначениях мы можем записать (4) как

$$v_i = \dot{a}_i + \varepsilon_{ikm} \omega_k (x_m - a_m). \quad (6)$$

Чтобы проверить, представляет ли это поле скоростей движение абсолютно твердого тела, вычислим вариацию расстояния между двумя частицами с координатами  $x_i, y_i$ . Имеем

$$\begin{aligned} \frac{d}{dt} (y_i - x_i)^2 &= 2 (y_i - x_i) (\dot{y}_i - \dot{x}_i) = \\ &= 2 (y_i - x_i) \varepsilon_{ikm} \omega_k (y_m - x_m) = \\ &= 0. \end{aligned} \quad (7)$$

Применяя такие же рассуждения к малым вращениям, находим, что с точностью до членов первого порядка малости расстояния при перемещениях, задаваемых формулой (3), не изменяются. Пренебрегая членами второго порядка малости, мы можем сказать, что малые угловые перемещения подчиняются

правилу параллелограмма в том смысле, что сумма смещений, отвечающих малым вращениям вокруг различных прямых, является смещением, соответствующим результирующему вращению вокруг оси, полученной по этому правилу. В частности,

$$(\delta\theta \times \mathbf{x})_i = \varepsilon_{ikm} \delta\theta_k x_m, \quad (8)$$

и компоненты  $(x_3 \delta\theta_2 - x_2 \delta\theta_3, x_1 \delta\theta_3 - x_3 \delta\theta_1, x_2 \delta\theta_1 - x_1 \delta\theta_2)$  являются суммами компонент, соответствующих вращениям  $(\delta\theta_1, \delta\theta_2, \delta\theta_3)$  около осей.

**2.15. Силы, действующие на абсолютно твердое тело.** Движение произвольной частицы описывается обычно следующими уравнениями:

$$m\ddot{\mathbf{x}} = \mathbf{X}, \quad m\ddot{x}_i = X_i. \quad (1)$$

Для каждой частицы абсолютно твердого тела эти уравнения справедливы. Если мы просто сложим их и обозначим через  $S$  суммирование по всем частицам, то получим

$$S(m\ddot{\mathbf{x}}) = S\mathbf{X}. \quad (2)$$

Если мы теперь напомним  $S(m) = M$ , то  $M$  будет полной массой всего тела; и если далее мы напомним

$$S(m\mathbf{x}) = M\bar{\mathbf{x}}, \quad (3)$$

то  $\bar{x}_i$  будут координатами точки, которую мы будем называть центром масс. Получаем

$$M\ddot{\bar{\mathbf{x}}} = S\mathbf{X}, \quad (4)$$

т. е. результирующей всех сил, действующих на тело.

Далее образуем векторное произведение (1) с  $\mathbf{x}$ , тогда

$$m\mathbf{x} \times \ddot{\mathbf{x}} = \mathbf{x} \times \mathbf{X}, \quad m\varepsilon_{ikm} x_k \ddot{x}_m = \varepsilon_{ikm} x_k X_m. \quad (5)$$

При суммировании

$$\begin{aligned} S(m\mathbf{x} \times \ddot{\mathbf{x}}) &= S(\mathbf{x} \times \mathbf{X}), \\ S(m\varepsilon_{ikm} x_k \ddot{x}_m) &= S(\varepsilon_{ikm} x_k X_m). \end{aligned} \quad (6)$$

Центр масс абсолютно твердого тела неподвижен относительно тела. Это обычно считается само собой разумеющимся, но не очевидно. Рассмотрим расстояние  $r_1$  до центра масс от любой данной частицы, скажем первоначально находившейся

в  $\mathbf{x}_1$ , и пусть частица с массой  $m_l$  находится в  $\mathbf{x}_l$ . Тогда

$$(Sm_l)(\mathbf{x}_1 - \bar{\mathbf{x}}) = (Sm_l)\mathbf{x}_1 - S(m_l\mathbf{x}_l) = Sm_l(\mathbf{x}_1 - \mathbf{x}_l), \quad (7)$$

$$\begin{aligned} (Sm_l)^2 r_1^2 &= [Sm_l(\mathbf{x}_1 - \mathbf{x}_l)][Sm_{l'}(\mathbf{x}_1 - \mathbf{x}_{l'})] = \\ &= SS'm_lm_{l'}(\mathbf{x}_1 - \mathbf{x}_l)(\mathbf{x}_1 - \mathbf{x}_{l'}), \end{aligned} \quad (8)$$

где  $S$  обозначает суммирование по  $l$ , а  $S'$  — по  $l'$ . Но

$$(\mathbf{x}_1 - \mathbf{x}_l)(\mathbf{x}_1 - \mathbf{x}_{l'}) = \frac{1}{2} (r_{1l}^2 + r_{1l'}^2 - r_{ll'}^2) \quad (9)$$

в обычных обозначениях. Следовательно,  $r_1^2$  полностью выражено через массы частиц и расстояния между ними и поэтому не меняется при любых перемещениях абсолютно твердого тела.

Но рассмотреть частицу, находившуюся первоначально в  $\bar{\mathbf{x}}$  недостаточно, поскольку, например, у поллой сферы такой частицы нет.

Силы, действующие на частицы, можно разделить на внешние силы и внутренние реакции. В соответствии с принципом, принадлежащим Даламберу, внутренние реакции образуют равновесную систему сил и их вклад в правые части (4) и (6) равен нулю. Этот принцип легче всего понять, если рассматривать абсолютно твердое тело как предельный случай упругого; он очевиден, если рассмотреть тело как такую систему частиц, в которой сила между любой парой частиц направлена вдоль соединяющей их прямой (т. е. компоненты ускорения относятся, как направляющие косинусы этой прямой). Если  $\mathbf{X}'$  — сила, действующая на  $m_l$  со стороны  $m_{l'}$ , то по третьему закону Ньютона реакция равна  $-\mathbf{X}'$ , и

$$-\mathbf{x}_l \times \mathbf{X}' + \mathbf{x}_{l'} \times \mathbf{X}' = (\mathbf{x}_{l'} - \mathbf{x}_l) \times \mathbf{X}' = 0,$$

если  $\mathbf{X}'$  направлена вдоль прямой от  $\mathbf{x}_l$  к  $\mathbf{x}_{l'}$ . Но эти соображения не являются общими, потому что частицы могли бы быть электрическими или магнитными диполями и в этом случае силы не были бы направлены вдоль соединяющей их прямой. Однако можно показать, что сам результат остается верным в значительно более широком классе случаев, а именно когда расстояния между частицами не меняются. Принцип Даламбера является, таким образом, неким приближением, справедливым для реального твердого тела при условии, что деформациями можно пренебречь; если же деформацией пренебречь нельзя, то даже не вполне верно, что центр масс неподвижен относительно тела. Причиной применения принципа Даламбера является в конечном счете то, что на практике он ведет к правильным результатам.

В правых частях равенств (4) и (6) нам нужно, следовательно, рассматривать только внешние силы, действующие на тело. Далее, достаточно шести величин, чтобы задать положение твердого тела: трех координат произвольной заданной частицы и трех углов Эйлера, которые задают его ориентацию. Они рассматриваются в гл. 3. Но (4) и (6) составляют шесть дифференциальных уравнений, и шесть есть число параметров, которое нужно, чтобы определить движение тела.

$\mathbf{S}\mathbf{X}$  является результирующей силой;  $\mathbf{L} = \mathbf{S}(\mathbf{x} \times \mathbf{X})$

называется моментом сил относительно начала координат. Отдельно от уравнений движения момент сил не представляет физического интереса. Следует обратить внимание, что момент силы равен  $\mathbf{x} \times \mathbf{X}$ , в то время как скорость при вращении равна  $\boldsymbol{\omega} \times \mathbf{x}$  и знаки этих выражений очевидны, если пользоваться чертежом, однако довольно распространены ошибки, если используются векторные обозначения. Если  $x_1$  и  $X_2$  положительны, то ясно, что сила стремится повернуть тело от  $O1$  к  $O2$ ; если  $\omega_3$  и  $x_1$  положительны, то  $\dot{x}_2$  положительно. Этого достаточно, чтобы определить знак одного члена в каждой компоненте, а из этого определяются остальные знаки.

Если система эквивалентна одной силе  $\mathbf{X}$ , приложенной в точке с радиусом-вектором  $\mathbf{x}$ , то момент силы равен  $\mathbf{x} \times \mathbf{X} = \mathbf{G}$  и

$$\mathbf{G} \cdot \mathbf{X} = \varepsilon_{ikm} x_k X_m X_i = 0.$$

Следовательно, только в специальных случаях силы, действующие на абсолютно твердое тело, можно заменить одной силой \*).

### ПРИМЕРЫ

1. Доказать, что  $\delta_{ik}\varepsilon_{ikm} = 0$ ,  $\varepsilon_{iks}\varepsilon_{mks} = 2\delta_{im}$ ,  $\varepsilon_{iks}\varepsilon_{iks} = 6$ .
2. Если  $\Delta$  обозначает определитель  $\|u_{ij}\|$ , то доказать, что

$$\varepsilon_{ikm}\Delta = \varepsilon_{jln}u_{ij}u_{kl}u_{mn},$$

$$\varepsilon_{jln}\Delta = \varepsilon_{ikm}u_{ij}u_{kl}u_{mn},$$

$$6\Delta = \varepsilon_{ikm}\varepsilon_{jln}u_{ij}u_{kl}u_{mn}.$$

3. Доказать, что, если  $l_{ij}$  — направляющие косинусы, определяющие преобразование осей, то

$$l_{ij} = \frac{1}{2} \varepsilon_{ikm}\varepsilon_{jln}l_{kl}l_{mn}.$$

---

\*) О приведении общей системы к двум силам или к силе и паре сил см. [3, гл. 5] и [4, гл. 8]. Эта задача, впрочем, появляется только в экзаменационных билетах.

4.  $\mathbf{z}$  — постоянный вектор единичной длины,  $\mathbf{r}$  (радиус-вектор движущейся частицы) — переменный вектор, перпендикулярный  $\mathbf{z}$ . Если скорость в любой момент задается формулой

$$\frac{d}{dt}(\mathbf{r}e^{kt}) = \omega \mathbf{z} \times \mathbf{r}e^{kt}, \quad (1)$$

где  $\omega$  и  $k$  — константы, то показать, что траектория точки — логарифмическая спираль.

Частица движется в плоскости под действием силы, пропорциональной радиусу-вектору, и сопротивления трения, пропорционального скорости частицы. Определить уравнения движения частицы и, отыскивая решения в виде (1) или каком-либо другом, показать, что скорость в любой момент времени есть векторная сумма скоростей двух частиц, описывающих логарифмическую спираль в противоположных направлениях с одинаковыми угловыми скоростями. (M/c, II, 1931.)

5.  $A, B, C$  — любые три точки на сфере единичного радиуса с центром  $O$ . Радиусы-векторы точек  $A, B, C$  относительно точки  $O$  соответственно равны  $\mathbf{u}, \mathbf{v}, \mathbf{w}$ . Покажите, что диаметр, перпендикулярный плоскости  $ABC$ , пересекает сферу в точках, радиусы-векторы которых равны  $\pm \mathbf{d}$ , где

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] \mathbf{d} \sec \theta = \mathbf{v} \times \mathbf{w} + \mathbf{w} \times \mathbf{u} + \mathbf{u} \times \mathbf{v}$$

и  $\theta$  — угол между  $\mathbf{d}$  и  $\mathbf{u}$ .

Рассматривая векторное произведение  $\mathbf{u} \times \mathbf{d}$  или каким-либо другим способом, доказать, что

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] \operatorname{tg} \theta = \pm 4 \sin \frac{a}{2} \sin \frac{b}{2} \sin \frac{c}{2},$$

где  $a, b, c$  — стороны сферического треугольника  $ABC$ . (Prelim., 1941.)

6. Доказать, что если  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$  — четыре произвольных вектора, то

$$\begin{aligned} (\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}), \\ (\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= [\mathbf{C}, \mathbf{D}, \mathbf{A}] \mathbf{B} - [\mathbf{B}, \mathbf{C}, \mathbf{D}] \mathbf{A} = \\ &= [\mathbf{D}, \mathbf{A}, \mathbf{B}] \mathbf{C} - [\mathbf{A}, \mathbf{B}, \mathbf{C}] \mathbf{D}, \end{aligned}$$

где  $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$  обозначает смешанное произведение  $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ .

Выведите правило синуса и косинуса сферической тригонометрии.

(Prelim., 1940.)

7. Две частицы выпущены одновременно из начала координат со скоростями  $\mathbf{v}_1$  и  $\mathbf{v}_2$  соответственно и движутся с постоянным ускорением  $\mathbf{a}$ . Доказать, что если  $\mathbf{v}_1 \cdot \mathbf{v}_2 < 0$ , то существует один и только один момент в последующем движении, когда радиусы-векторы этих частиц образуют прямой угол с вершиной в начале координат.

Показать, что в этот момент радиусы-векторы  $\mathbf{r}_1, \mathbf{r}_2$  частиц удовлетворяют уравнению

$$\mathbf{a}^2 (\mathbf{r}_1 \cdot \mathbf{v}_2 - \mathbf{r}_2 \cdot \mathbf{v}_1) + (\mathbf{a} \cdot \mathbf{r}_2 - \mathbf{a} \cdot \mathbf{r}_1) (\mathbf{a} \cdot \mathbf{v}_1 + \mathbf{a} \cdot \mathbf{v}_2) + 2\mathbf{v}_1 \cdot \mathbf{v}_2 (\mathbf{a} \cdot \mathbf{v}_2 - \mathbf{a} \cdot \mathbf{v}_1) = 0. \quad (\text{Prelim., 1940.})$$

8. Найти выражение для радиуса-вектора  $\mathbf{r}$  в момент времени  $t$  частицы с единичной массой, движущейся под действием постоянной силы  $(n^2 + k^2)\mathbf{b}$  и силы притяжения, равной  $(n^2 + k^2)\mathbf{r}$  и действующей в направлении начала координат ( $n \neq 0$ ). Сила сопротивления равна  $2k\mathbf{v}$ . В момент  $t=0$  частица имела скорость  $\mathbf{v}$  и была в точке  $\mathbf{r}=\mathbf{a}$ .

Выведите, что смешанные произведения векторов  $\mathbf{r}, \mathbf{v}, \mathbf{a}-\mathbf{b}$  и  $\mathbf{a}, \mathbf{b}, \mathbf{v}$  равны между собой. (Prelim., 1941.)

9. Частица, заряд которой равен  $e$ , а масса  $m$ , движется под действием однородного электрического поля, напряженность которого  $(0, E, 0)$ , и однородного магнитного поля, напряженность которого  $(0, 0, H)$  в гауссовых единицах.

Доказать, что это движение можно рассматривать как суперпозицию равномерного движения со скоростью  $(Ec/H, 0, 0)$  и равномерного движения по винтовой линии с угловой скоростью  $-eH/mc$  относительно оси. Изменением массы со скоростью можно пренебречь.

Доказать, что если начало движения частицы совпадает с началом координат, то, какая бы ни была ее начальная скорость, она пересекает каждую прямую

$$x = 2\pi n mc^2 E / e H^2, \quad y = 0, \quad \text{где } n = 1, 2, 3 \dots \quad (\text{М. Т., 1943.})$$

10. На частицу массы  $m$ , радиус-вектор которой  $\mathbf{r}$ , действуют центральная сила  $\mu \mathbf{r}$  и сила  $e(\mathbf{H} \times \mathbf{r})/c$ , где  $\mathbf{H}$  — однородное магнитное поле. Показать, что если  $\mathbf{r}$  и  $\dot{\mathbf{r}}$  первоначально перпендикулярны  $\mathbf{H}$ , то частица будет описывать плоскую кривую.

Показать, что частица может описывать окружность около начала координат под действием этих сил с любой из двух постоянных угловых скоростей. (Л. С., 1942.)

11. Определить вектор  $\mathbf{OC}$ , перпендикулярный  $\mathbf{OA} = a(2, 3, 0)$ ,  $\mathbf{OB} = b(-2, 0, 1)$ , такой, что вращение от  $\mathbf{OA}$  к  $\mathbf{OB}$  относительно  $\mathbf{OC}$  является положительным. Вычислить объем тетраэдра  $OABC$ .

Найти стороны и углы сферического треугольника  $ABC$ , определенного векторами

$$\mathbf{OA} = (1, 0, 0), \quad \mathbf{OB} = \left( \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right), \quad \mathbf{OC} = \left( 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right).$$

## ЛИТЕРАТУРА

1. *Jeffreys H.*, Scientific Inference, Cambridge Univ. Press, 1937.
2. *Durell C. V., Robson A.*, Advanced Algebra, 1937.
3. *Jeffreys H.*, Cartesian Tensors, 1931, ch. 5.
4. *Weatherburn G. E.*, Elementary Vector Analysis.

## ТЕНЗОРЫ

Мы знаем, что интеллектуальная пища иногда усваивается легче, если принимать ее не в слишком концентрированном виде. Можно спросить: „До какой степени можно извлекать пользу, специализируя обозначения, с тем чтобы они усваивались наиболее просто?“ Только практика может дать ответ на этот вопрос.

*F. Cajori „History of Mathematical Notations“, p. 77.*

**3.01.** В этой главе в простом и сжатом виде излагается теория тензоров. Во многих разделах физики тензорные обозначения дают компактное математическое выражение, а успешное освоение тензорного аппарата облегчает использование непрямоугольных осей, криволинейных координат и многомерного пространства. Тензоры служат также введением к идеям матричной алгебры. Общая теория тензоров является необходимым математическим аппаратом теории относительности, квантовой механики, и в классической физике многое в этих обозначениях принимает наиболее четкое выражение. В приложениях, рассмотренных в этой главе, физические идеи просты, и практическое использование в них тензорных обозначений чрезвычайно полезно. Это удобнее, чем применять их в полной мере к теориям, в которых физические идеи сами по себе являются более трудными для понимания.

**3.02. Преобразование координат. Свертывание.** Определим вектор **A** посредством преобразования

$$A'_j = l_{ij} A_i, \quad (1)$$

которое эквивалентно преобразованию

$$A_i = l_{ij} A'_j. \quad (2)$$

Любой вектор можно также назвать *тензором первого порядка*. Любой скаляр является тензором нулевого порядка.



Если теперь рассмотрим набор из девяти произведений  $A_i B_k$ , то заметим, что скалярные и векторные произведения являются частными линейными комбинациями этих произведений. Образует аналогичный набор для компонент, отнесенных к новым осям:

$$A'_i B'_l = l_{il} l_{kl} A_i B_k, \quad (3)$$

$$A_i B_k = l_{il} l_{kl} A'_i B'_l. \quad (4)$$

По аналогии с определением вектора воспользуемся этими соотношениями для определения тензора второго порядка. Набор величин, зависящих от *двух* направлений и определяемых через девять компонент  $K_{ik}$ , относящихся к системе координат  $O 123$ , и  $K'_{il}$ , относящихся к системе координат  $O 1'2'3'$ , образует тензор второго порядка, если для произвольных изменений осей справедливо, что

$$K'_{il} = l_{il} l_{kl} K_{ik} \quad (5)$$

или эквивалентное соотношение

$$K_{ik} = l_{il} l_{kl} K'_{il}. \quad (6)$$

Каждый из двух индексов, обозначающих компоненту  $K$ , относится к одной из координат *той же самой* системы. Направляющие косинусы  $l_{ij}$  тензора *не образуют*, так как два индекса относятся к осям различных систем.

В квадратном расположении

$$\begin{pmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{pmatrix} \quad (7)$$

компоненты  $K_{11}$ ,  $K_{22}$ ,  $K_{33}$  называются диагональными компонентами. Их сумма  $K_{ii}$  называется *следом* и является скаляром. В самом деле,

$$K'_{ii} = l_{il} l_{kl} K_{ik} = \delta_{ik} K_{ik} = K_{ii}. \quad (8)$$

Операция, в которой два индекса в тензоре полагаются равными и затем по ним производится суммирование, называется *свертыванием*. При этом порядок тензора понижается на два.

Сумма двух тензоров  $K$  и  $L$  определяется соотношением

$$(K + L)_{ik} = K_{ik} + L_{ik}$$

и является, очевидно, также тензором.

Тензоры более высокого порядка определяются аналогично, т. е. тензор порядка  $n$  преобразуется как произведение  $A_l B_k C_m \dots$  из  $n$  сомножителей. Мы будем иметь дело в основном с тензорами второго порядка и лишь иногда использовать тензоры третьего и четвертого порядков.

**3.03. Изотропные тензоры.** Можно показать, что набор величин  $\delta_{ik}$  образует тензор. В самом деле, если воспользуемся преобразованием (5), то получим набор величин  $U'_{jl}$ , заданных соотношением

$$U'_{jl} = l_{ij} l_{kl} \delta_{ik} = l_{ij} l_{il} = \begin{cases} 1 & \text{при } j = l, \\ 0 & \text{при } j \neq l. \end{cases} \quad (9)$$

Следовательно, набор  $\delta_{ik}$  преобразуется в  $\delta_{jl}$  при любом вращении осей.

Аналогично можно показать, что тензор третьего порядка с компонентами  $\epsilon_{ikm}$  в системе координат  $O 123$  имеет те же самые компоненты в системе координат  $O 1'2'3'$ . Действительно, при преобразовании для  $jln$ -й компоненты получим  $l_{ij} l_{kl} l_{mn} \epsilon_{ikm}$ . Расписывая более подробно и учитывая, что в неравных нулю членах все  $i, k, m$  различны, получим

$$l_{1j} l_{2l} l_{3n} + l_{2j} l_{3l} l_{1n} + l_{3j} l_{1l} l_{2n} - l_{2j} l_{1l} l_{3n} - l_{3j} l_{2l} l_{1n} - l_{1j} l_{3l} l_{2n}.$$

Если  $j = l$ , то все компоненты взаимно уничтожаются. То же получаем при  $j = n$  или  $l = n$ . Если все  $j, l, n$  различны, то это выражение принимает вид

$$\begin{vmatrix} l_{1j} & l_{1l} & l_{1n} \\ l_{2j} & l_{2l} & l_{2n} \\ l_{3j} & l_{3l} & l_{3n} \end{vmatrix}.$$

Оно равно 1, если  $j, l, n$  отличаются от порядка 1, 2, 3 четным числом перестановок, и  $-1$ , если число перестановок нечетное. Следовательно,  $\epsilon_{ikm}$  преобразуется в  $\epsilon_{jln}$  по правилу преобразования тензоров третьего порядка.

Тензоры, компоненты которых не изменяются при повороте осей, называются *изотропными*. Можно показать [1\*], что изотропного тензора первого порядка не существует, а единственными изотропными тензорами второго и третьего порядков являются  $\delta_{ik}$  и  $\epsilon_{ikm}$ , умноженные на скаляр. Имеются три

---

\*) См. также приложение к гл. 3.

независимых изотропных тензора четвертого порядка, которые равны

$$\begin{aligned} & \delta_{ij}\delta_{mp}, \\ & \delta_{im}\delta_{kp} + \delta_{ip}\delta_{km}, \\ & \delta_{im}\delta_{kp} - \delta_{ip}\delta_{km}. \end{aligned} \quad (10)$$

Последнее выражение используется вместо тензора  $\epsilon_{ijkl}\epsilon_{mpqs}$ . Остальные два встречаются при выводе уравнений движения вязкой жидкости и упругой твердой среды.

**3.031. Изотропные тензоры четвертого порядка.** Если  $u_{ikmp}$  является изотропным тензором четвертого порядка, то для любых поворотов имеем

$$u'_{jlnq} = l_{ij}l_{kl}l_{mn}l_{pq}u_{ikmp} = u_{jlnq}. \quad (1)$$

Учитывая, что у любой компоненты по крайней мере два индекса должны быть равными, замечаем, что эти компоненты относятся к одному из следующих четырех типов:  $u_{1111}$ ,  $u_{1112}$ ,  $u_{1122}$ ,  $u_{1123}$ .

Вначале повернем оси около прямой с направляющими косинусами  $1/\sqrt{3}$ ,  $1/\sqrt{3}$ ,  $1/\sqrt{3}$  так, чтобы ось 1 совпала с первоначальным положением оси 2 и т. д. Так как это приводит к циклической замене индексов, то из свойства изотропности следует, что

$$\begin{aligned} u_{1111} &= u_{2222} = u_{3333}, & u_{1122} &= u_{2233} = u_{3311}, \\ u_{2211} &= u_{3322} = u_{1133}, & u_{1221} &= u_{2332} = u_{3113} \text{ и т. д.} \end{aligned} \quad (2)$$

Затем произведем поворот на  $90^\circ$  вокруг оси  $O3$ . Тогда

$$l_{12} = 1, \quad l_{21} = -1, \quad l_{33} = 1, \quad (3)$$

а остальные  $l_{ik}$  равны нулю. Положим  $i=3$ ,  $l=n=q=1$ . Тогда отличными от нуля будут члены для  $i=3$ ,  $k=m=p=2$  и

$$u_{3111} = -u_{3222}. \quad (4)$$

Точно так же для  $j=3$ ,  $l=n=q=2$  необходимо положить  $i=3$ ,  $k=m=p=1$ , что дает

$$u_{3222} = u_{3111}. \quad (5)$$

Аналогично получим, что все компоненты с тремя одинаковыми индексами и не равным им четвертым индексом равны нулю.

Подобным же способом найдем, что

$$\begin{aligned} j=l=1, \quad n=q=2; \quad i=k=2, \quad m=p=1: & \quad u_{1122} = u_{2211}, \\ j=n=2, \quad l=q=3; \quad i=m=1, \quad k=p=3: & \quad u_{2323} = u_{1313}, \\ j=q=1, \quad l=n=2; \quad i=p=2, \quad k=m=1: & \quad u_{1221} = u_{2112}, \\ j=3, \quad l=n=2, \quad q=1; \quad i=3, \quad k=m=1, \quad p=2: & \quad u_{3221} = -u_{3112}, \\ j=3, \quad l=n=1, \quad q=2; \quad i=3, \quad k=m=2, \quad p=1: & \quad u_{3112} = u_{3221}. \end{aligned} \quad (6)$$

Следовательно, не равны нулю только те компоненты, у которых все индексы одинаковы или попарно одинаковы. При помощи циклической замены индек-

сов получим

$$u_{1111} = u_{2222} = u_{3333} = \kappa, \quad (8)$$

$$u_{1122} = u_{2211} = u_{2233} = u_{3322} = u_{3311} = u_{1133} = \lambda, \quad (9)$$

$$u_{2323} = u_{1313} = u_{3131} = u_{2121} = u_{1212} = u_{3232} = \mu, \quad (10)$$

$$u_{1221} = u_{2112} = u_{2332} = u_{3223} = u_{3113} = u_{1331} = \nu. \quad (11)$$

Все эти соотношения выполняются для кубической симметрии. Теперь можно записать

$$u_{ikmp} = \lambda \delta_{ik} \delta_{mp} + \mu \delta_{im} \delta_{kp} + \nu \delta_{ip} \delta_{km} + (\kappa - \lambda - \mu - \nu) v_{ikmp}, \quad (12)$$

где  $v_{ikmp} = 1$ , если все четыре индекса одинаковы, а в остальных случаях равны нулю. Итак, если  $u_{ikmp}$  является тензором четвертого порядка, то  $u_{ikmp} x_i y_k z_m w_p$  — скаляр и наоборот (см. 3.05). Это выражение приводится к виду

$$\lambda x_i y_i z_m w_m + \mu x_i z_i y_k w_k + \nu x_i w_i y_k z_k + (\kappa - \lambda - \mu - \nu) (x_1 y_1 z_1 w_1 + \dots). \quad (13)$$

Первые три члена являются произведениями скаляров. Последний, если взять все векторы одинаковыми, равен  $x_1^4 + x_2^4 + x_3^4$ , а это не скаляр. Он обладает кубической, но не сферической симметрией. Например, если  $x_1 = x_2 = x_3 = 1/\sqrt{3}$ , то  $x_1^4 + x_2^4 + x_3^4 = 1/3$ , но если  $x'_1 = 1$ ,  $x'_2 = x'_3 = 0$ , то  $x'^4_1 + x'^4_2 + x'^4_3 = 1$ , а это изменение компонент можно получить вращением осей. Следовательно, если тензор является изотропным, то

$$\kappa - \lambda - \mu - \nu = 0 \quad (14)$$

и наиболее общий вид изотропного тензора четвертого порядка дают первые три члена (12). Это можно переписать в виде суммы трех тензоров 3.03 (10).

Упругие константы любого твердого тела образуют тензор четвертого порядка, который может быть изотропным, если изотропно твердое тело. Член  $v_{ikmp}$  является более общим и позволяет выразить упругие свойства кубического кристалла (модули Юнга могут иметь различные значения для деформаций вдоль диагонали и параллельно ребру куба).

**3.04. Аффинорное обозначение.** Иногда бывает удобно обозначать тензор второго порядка одной буквой, как это делается для вектора. Если все компоненты тензора  $K_{ik}$  умножить на компоненты вектора  $A_m$ , то получим тензор третьего порядка. Из этого произведения можно образовать два различных вектора, положив  $m$  равным  $i$  или  $k$  и произведя суммирование по одинаковым индексам, а именно  $K_{ik} A_i$  и  $K_{ik} A_k$ . В аффинорном обозначении они записываются соответственно в виде  $\mathbf{A} \cdot \mathbf{K}$  и  $\mathbf{K} \cdot \mathbf{A}$ . Запомним правило, которое состоит в том, что сомножители располагаются в таком порядке, что *суммирование производится всегда по смежным индексам*, таким образом

$$(\mathbf{A} \cdot \mathbf{K})_k = A_i K_{ik}, \quad (\mathbf{K} \cdot \mathbf{A})_i = K_{ik} A_k. \quad (1)$$

Доказательство того, что в результате свертывания тензора порядка  $n$  получается тензор  $n - 2$ -го порядка, подобно доказательству 3.02 (8) и нет необходимости приводить его целиком. Использование жирного шрифта указывает на то, что один или более индексов немые.

Аналогично можно образовать произведение со свертыванием двух тензоров второго порядка  $\mathbf{K}$  и  $\mathbf{L}$ , а именно

$$(\mathbf{K} \cdot \mathbf{L})_{ik} = K_{im} L_{mk}, \quad (\mathbf{L} \cdot \mathbf{K})_{ik} = L_{im} K_{mk}. \quad (2)$$

Кроме того, результат обычно зависит от порядка расположения сомножителей. Этот тип умножения не коммутативен. В данных обозначениях тензор  $A_i B_k$  записывается в виде  $\mathbf{AB}$ . Отсутствие точки отличает тензор как от скалярного, так и от векторного произведений. Аффинорное обозначение очень похоже на матричное обозначение, о котором говорится в следующей главе. Упрощение записи, получаемое за счет опускания индексов, компенсируется той сверхосторожностью, которую надо проявлять для сохранения порядка сомножителей, а также из-за того, что иногда не требуется производить свертывание. В частности, константы упругости кристалла образуют тензор четвертого порядка.

**3.05. Правило частного.** Если имеется система уравнений

$$K_{ik} A_k = B_i, \quad (1)$$

где  $A_k$  и  $B_i$  — известные тензоры первого порядка, или

$$K_{ik} T_{km} = S_{im}, \quad (2)$$

где  $T_{km}$  и  $S_{im}$  — тензоры второго порядка, то можно ли утверждать обратное, т. е. что  $K_{ik}$  является тензором второго порядка? Ответ будет утвердительным при условии, что все компоненты  $A_k$  или  $T_{km}$  могут изменяться независимо. Рассмотрим начнем с простейшего случая (1). Преобразуем оси, но при этом не известно, как преобразуются  $K_{ik}$ . Должен получиться набор  $K'_{il}$  с девятью компонентами, определяемыми соотношениями  $\partial B'_i / \partial A'_l$ . Тогда получим

$$K'_{il} A'_l = B'_i = l_{ij} B_j = l_{ij} K_{jk} A_k = l_{ij} l_{kl} K_{ik} A'_l. \quad (3)$$

Следовательно,

$$(K'_{il} - l_{ij} l_{kl} K_{ik}) A'_l = 0. \quad (4)$$

Но если это верно для любого  $j$ , когда  $A'_l$  изменяется независимо, то

$$K'_{il} = l_{ij} l_{kl} K_{ik} \quad (5)$$

и  $K_{ik}$  — тензор второго порядка.

Важный частный случай этой общей теоремы состоит в том, что если  $K_{ikm} \dots T_{ikm} \dots$  — скаляр и  $\mathbf{T}$  — произвольный тензор порядка  $n$ , то и  $\mathbf{K}$  также является тензором порядка  $n$ . В частности, коэффициенты  $a_{ik}$  в квадратичной форме  $a_{ik} x_i x_k$ , где  $a_{ik} = a_{ki}$ , в этих координатах образуют тензор второго порядка.

**3.06. Дифференцирование по координатам скалярных и векторных функций.** Скаляр, компоненты вектора или тензора в различных точках пространства могут иметь различные значения, даже если они рассматриваются в данный момент времени и при этом не производится преобразования координатных осей. Так, различные точки тела отличаются расстоянием от начала координат, но расстояние данной точки не зависит от направления осей. Такая функция называется *скалярной функцией, зависящей от пространственного положения точки*. С другой стороны, скорость частицы жидкости является вектором, который обычно изменяется с изменением пространственного положения частицы, и его можно назвать *векторной функцией пространственного положения*. Существование таких функций ставит вопрос об их дифференцировании.

Если  $\phi$  является скалярной функцией, то набор из трех производных  $\partial\phi/\partial x_i$  определяет вектор, который обозначается  $\text{grad}\phi$  или  $\nabla\phi$ . Для доказательства того, что это вектор, повернем координатные оси, тогда

$$\frac{\partial\phi}{\partial x'_j} = \frac{\partial x_i}{\partial x'_j} \frac{\partial\phi}{\partial x_i} = l_{ij} \frac{\partial\phi}{\partial x_i}, \quad (1)$$

что и доказывает высказанное утверждение. При этом предполагается, что  $\phi$  дифференцируема по трем пространственным координатам в смысле Штольца и Юнга, т. е. достаточным условием дифференцируемости является непрерывность частных производных от  $\phi$  по всем координатам (сравни 5.04).

Если  $u_i$  — векторная функция, то

$$\frac{\partial u'_l}{\partial x'_j} = \frac{\partial x_i}{\partial x'_j} \frac{\partial u'_l}{\partial x_i} = l_{ij} \frac{\partial}{\partial x_i} (l_{kl} u_k) = l_{ij} l_{kl} \frac{\partial u_k}{\partial x_i}, \quad (2)$$

а следовательно,  $\partial u_k/\partial x_i$  — тензор второго порядка.

Из этого следует, что  $\partial u_i/\partial x_i$  — скалярная функция. Ее обычно обозначают  $\text{div } \mathbf{u}$  или  $\nabla \cdot \mathbf{u}$ . Далее, если функция  $\phi$  такова, что  $u_i = \partial\phi/\partial x_i$ , то

$$\frac{\partial u_i}{\partial x_i} = \frac{\partial^2\phi}{\partial x_i^2} = \frac{\partial^2\phi}{\partial x_1^2} + \frac{\partial^2\phi}{\partial x_2^2} + \frac{\partial^2\phi}{\partial x_3^2}. \quad (3)$$

Эта комбинация вторых производных имеет важное значение в математической физике, уступая в этом лишь производной по времени. Ее обозначают  $\nabla^2\phi$ .

Величины

$$\begin{aligned}\epsilon_{ikm} \frac{\partial u_m}{\partial x_k} &= \left( \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}, \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}, \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right) = \\ &= \left( \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)\end{aligned}\quad (4)$$

в декартовых координатах определяют вектор. Его обозначают  $\text{rot } \mathbf{u}$  (используется также обозначение  $\nabla \times \mathbf{u}$ ). Отметим, что если  $\mathbf{u}$  является градиентом скаляра, то  $\text{rot } \mathbf{u} = 0$ ; действительно,

$$(\text{rot grad } \varphi)_i = \epsilon_{ikm} \frac{\partial}{\partial x_k} \frac{\partial \varphi}{\partial x_m} = 0. \quad (5)$$

Кроме того, если  $\mathbf{u}$  является произвольной векторной функцией, то

$$\text{div rot } \mathbf{u} = \frac{\partial}{\partial x_i} \epsilon_{ikm} \frac{\partial u_m}{\partial x_k} = \epsilon_{ikm} \frac{\partial^2 u_m}{\partial x_i \partial x_k} = 0, \quad (6)$$

потому что все члены взаимно уничтожаются. Применяя дважды операцию  $\text{rot}$ , получим полезное соотношение, а именно

$$\begin{aligned}(\text{rot rot } \mathbf{u})_i &= \epsilon_{ikm} \frac{\partial}{\partial x_k} \epsilon_{mps} \frac{\partial u_s}{\partial x_p} = (\delta_{ip} \delta_{ks} - \delta_{is} \delta_{kp}) \frac{\partial^2 u_s}{\partial x_k \partial x_p} = \\ &= \frac{\partial^2 u_s}{\partial x_s \partial x_i} - \frac{\partial^2 u_i}{\partial x_k^2} = (\text{grad div } \mathbf{u} - \nabla^2 \mathbf{u})_i.\end{aligned}\quad (7)$$

**3.07. Свойства симметрии.** Покажем, что если  $K_{ik}$  — тензор, то и  $K_{ki}$ , полученный заменой строк на столбцы, также является тензором. Это означает, что если  $K_{ik}$  преобразуется в  $K'_{jl}$ , согласно правилу

$$K'_{jl} = l_{ij} l_{kl} K_{ik}, \quad (1)$$

и если записать  $L_{ik} = K_{ki}$ , то  $L_{ik}$  должен преобразовываться по тому же правилу. Тогда

$$l_{ij} l_{kl} L_{ik} = l_{ij} l_{kl} K_{ki}. \quad (2)$$

Здесь  $i$  и  $k$  — дважды встречающиеся индексы, поэтому не важно, какой из них назвать  $i$ , а какой  $k$ . Поменяв эти индексы, получим

$$l_{kj} l_{li} K_{ik} = l_{li} l_{kj} K_{ik} = K'_{lj}. \quad (3)$$

Таким образом, преобразованная система отличается от  $K'_{jl}$  тем, что строки заменены столбцами.

Если  $K_{ik} = K_{ki}$ , то говорят, что тензор *симметричный*. Если же  $K_{ik} = -K_{ki}$ , то тензор называется *антисимметричным*. Кроме того, если  $K_{ik}$  — тензор, то  $K_{ik} + K_{ki}$  и  $K_{ik} - K_{ki}$  также являются

тензорами. Первый из них не изменяется при замене  $i$  на  $k$  и является *симметричным* тензором. Второй при этом изменяет знаки у всех компонент и является *антисимметричным* тензором. Так как любой тензор  $K_{ik}$  можно записать в виде

$$K_{ik} = \frac{1}{2} (K_{ik} + K_{ki}) + \frac{1}{2} (K_{ik} - K_{ki}), \quad (4)$$

то это означает, что любой тензор можно представить в виде суммы симметричного и антисимметричного тензоров.

Поскольку  $K_{ik}A_k$  и  $A_iK_{ik}$  — векторы, то, умножая их скалярно на другой вектор  $\mathbf{B}$ , получим  $B_iK_{ik}A_k$  и  $A_iK_{ik}B_k$ . Эти произведения, вообще говоря, не равны друг другу. Однако если  $K_{ik}$  симметричен, то они равны. В самом деле,

$$\mathbf{B} \cdot \mathbf{K} \cdot \mathbf{A} = B_iK_{ik}A_k = B_iK_{ki}A_k = A_kK_{ki}B_i = \mathbf{A} \cdot \mathbf{K} \cdot \mathbf{B}. \quad (5)$$

С другой стороны, если  $K_{ik}$  антисимметричен, то в третьем равенстве этого выражения знак изменится; в результате получим

$$\mathbf{B} \cdot \mathbf{K} \cdot \mathbf{A} = -\mathbf{A} \cdot \mathbf{K} \cdot \mathbf{B}. \quad (6)$$

Важно отметить, что если  $K_{ik}$  симметричен, а  $u_i$  — компоненты вектора, то каждый член в выражении  $K_{ik}u_iu_k$  с  $i \neq k$  встречается *дважды*. Так, для  $i=1, k=2$  получим член  $K_{12}u_1u_2$ , но имеется другой член с  $i=2, k=1$ , который равен  $K_{21}u_2u_1$ . Если  $K_{12}=K_{21}$ , то эти члены равны. Таким образом, выражение

$$T = K_{ik}u_iu_k$$

равно

$$K_{11}u_1^2 + 2K_{12}u_1u_2 + K_{22}u_2^2 + 2K_{13}u_1u_3 + \dots,$$

а

$$\frac{\partial T}{\partial u_i} = 2(K_{11}u_1 + K_{12}u_2 + \dots) = 2K_{1k}u_k,$$

или в общем виде

$$\frac{\partial T}{\partial u_i} = 2K_{ik}u_k.$$

**3.071. Произведение  $\varepsilon_{ikm}K_{km}$ , вес  $\mathbf{K}$ .** Рассмотрим тройку чисел  $\varepsilon_{ikm}K_{km}$ . Это произведение тензоров третьего и второго порядков свернуто по двум индексам и, следовательно, является вектором; или, по-другому, изменяя оси координат, получим

$$\varepsilon_{jln}K'_{ln} = \varepsilon_{jln}l_{kl}l_{mn}K_{km} = \varepsilon_{ikm}l_{ij}K_{km}. \quad (7)$$



Это же можно записать в виде  $2 \text{vec K}$ , где

$$(\text{vec K})_i = \frac{1}{2} \varepsilon_{ikm} K_{km},$$

или в компонентах

$$\left[ \frac{1}{2} (K_{23} - K_{32}), \quad \frac{1}{2} (K_{31} - K_{13}), \quad \frac{1}{2} (K_{12} - K_{21}) \right]. \quad (8)$$

**3.072. Связь между антисимметричным тензором и вектором.** Компоненты антисимметричного тензора  $W_{ik}$  с  $i = k$  должны равняться нулю, и так как для остальных компонент  $W_{ik} = -W_{ki}$ , то для определения антисимметричного тензора необходимо задать лишь три независимые величины. В таком случае этот тензор имеет вид

$$\begin{pmatrix} 0 & W_{12} & -W_{31} \\ -W_{12} & 0 & W_{23} \\ W_{31} & -W_{23} & 0 \end{pmatrix}. \quad (9)$$

Но  $W_{23}$ ,  $W_{31}$ ,  $W_{12}$  являются компонентами  $\text{vec W}$ . Обозначим этот вектор через  $\mathbf{w}$ , так что

$$w_i = \frac{1}{2} \varepsilon_{ikm} W_{km}. \quad (10)$$

Тот факт, что число компонент вектора равно числу независимых компонент антисимметричного тензора, справедлив лишь для трехмерного пространства. В случае  $n$  измерений антисимметричный тензор имеет  $\frac{1}{2} n(n-1)$  независимых компонент, тогда как вектор имеет  $n$  компонент.

Из этого следует, что совокупность величин в (9) та же, что и совокупность

$$\begin{pmatrix} 0 & w_3 & -w_2 \\ -w_3 & 0 & w_1 \\ w_2 & -w_1 & 0 \end{pmatrix}, \quad (11)$$

т. е.  $W_{12} = w_3$ ,  $W_{21} = -w_3$ , или в более общем виде

$$W_{ik} = \begin{cases} 0 & (\text{при } i = k), \\ w_m & (\text{при четном числе перестановок в } ikm), \\ -w_m & (\text{при нечетном числе перестановок в } ikm), \end{cases}$$

поэтому

$$W_{ik} = \varepsilon_{ikm} w_m. \quad (12)$$

Не всегда удобно пользоваться вектором, иногда удобнее представление в виде антисимметричного тензора. Связь между ними определяется уравнениями (10) и (12).

Рассмотрим в качестве примера векторное произведение  $\mathbf{w} \times \mathbf{A}$ , что дает

$$\begin{aligned} (\mathbf{w} \times \mathbf{A})_i &= \varepsilon_{ikm} w_k A_m = \frac{1}{2} \varepsilon_{ikm} \varepsilon_{kps} W_{ps} A_m = \\ &= \frac{1}{2} \varepsilon_{mik} \varepsilon_{psk} W_{ps} A_m = \frac{1}{2} (\delta_{mp} \delta_{is} - \delta_{ms} \delta_{ip}) W_{ps} A_m = \\ &= \frac{1}{2} (W_{mi} - W_{im}) A_m = -W_{im} A_m. \end{aligned} \quad (13)$$

Следовательно, векторное умножение на вектор  $\mathbf{w}$  можно представить как умножение на антисимметричный тензор —  $W_{ik}$ . С другой стороны, этот результат можно получить, непосредственно выписывая все члены векторного произведения для определенной компоненты, скажем  $i = 1$ .

В физических приложениях используются и тот и другой способы, но первый, по-видимому, более естественный. Например, вывод уравнений для момента импульса начнем с уравнения  $m\ddot{x}_i = X_i$ . Умножим  $k$ -е уравнение на  $x_m$ , а  $m$ -е уравнение на  $x_k$  и вычтем одно из другого. В результате получим девять уравнений

$$m(\ddot{x}_k x_m - \ddot{x}_m x_k) = X_k x_m - X_m x_k,$$

в которых обе стороны являются антисимметричными тензорами. Причина преобразования этих уравнений к векторному виду состоит в том, что при этом исключаются три уравнения, имеющие вид  $0 = 0$ , а три других получаются переменной знака.

**3.08. Симметричный тензор. Главные оси.** Мы видели, что антисимметричный тензор тесно связан с вектором. Симметричный тензор можно связать с квадратичной формой. Если  $K_{ik}$  является симметричным тензором с действительными компонентами, то уравнение

$$K_{ik} x_i x_k = \text{const} \quad (1)$$

определяет нормальную квадратичную форму с центром в начале координат. В то время как

$$x_i = x l_i \quad (2)$$

определяет прямую, проходящую через начало координат, соотношение

$$K_{ik} l_k x_i = \text{const} \quad (3)$$

определяет полярную плоскость. Данная плоскость перпендикулярна прямой (2), если

$$K_{ik}l_k = \lambda l_i, \quad (4)$$

где  $\lambda$  одинаково для  $i = 1, 2, 3$ . Условие совместности этих уравнений приводит к кубическому уравнению относительно  $\lambda$ , любой корень которого, вообще говоря, должен давать допустимые отношения  $l_i$ . Они будут действительными, если  $\lambda$  действительно, и в таком случае прямая будет перпендикулярна полярной плоскости. В частности, эта прямая будет перпендикулярна к касательной плоскости в точке, в которой прямая пересекает поверхность, определяемую квадратичной формой. Такая прямая называется *главной осью*.

Прежде всего покажем, что если имеются два решения, соответствующие различным значениям  $\lambda$ , например  $\lambda_1$  и  $\lambda_2$ , то они дают значения  $l_i$ , скажем  $l_{i1}$  и  $l_{i2}$ , удовлетворяющие соотношению  $l_{i1} l_{i2} = 0$ . Если  $\lambda_1$  и  $\lambda_2$  являются действительными, то эти прямые перпендикулярны друг другу. В таком случае имеем

$$K_{ik}l_{k1} = \lambda_1 l_{i1}, \quad (5)$$

$$K_{ik}l_{k2} = \lambda_2 l_{i2}. \quad (6)$$

Умножая их соответственно на  $l_{i2}$  и  $l_{i1}$  и свергывая, получим

$$K_{ik}l_{k1}l_{i2} = \lambda_1 l_{i1}l_{i2}, \quad (7)$$

$$K_{ik}l_{k2}l_{i1} = \lambda_2 l_{i2}l_{i1}. \quad (8)$$

Но так как  $K_{ik}$  симметричен, то левые части этих соотношений равны. Следовательно,

$$(\lambda_1 - \lambda_2) l_{i1}l_{i2} = 0, \quad (9)$$

и если  $\lambda_1 \neq \lambda_2$ , то

$$l_{i1}l_{i2} = 0. \quad (10)$$

Условие совместности уравнений (4) при  $l_i \neq 0$  состоит в том, что детерминант  $\|K_{ik} - \lambda \delta_{ik}\| = 0$ , т. е.

$$\begin{vmatrix} K_{11} - \lambda & K_{12} & K_{13} \\ K_{21} & K_{22} - \lambda & K_{23} \\ K_{31} & K_{32} & K_{33} - \lambda \end{vmatrix} = 0. \quad (11)$$

Это уравнение должно иметь по крайней мере один действительный корень; назовем его  $\lambda_1$ . Примем получающиеся при этом  $l_{i1}$  за направляющие косинусы новой оси  $x'_1$  и выберем две другие оси  $x'_2$  и  $x'_3$ , перпендикулярные первой. Тогда,

отмечая штрихом направляющие косинусы относительно новых осей, имеем  $l'_{11} = 1$ ,  $l'_{21} = l'_{31} = 0$  и, согласно (4), получим

$$K'_{11} = \lambda_1, \quad K'_{12} = 0, \quad K'_{13} = 0. \quad (12)$$

Уравнение (11) теперь примет вид

$$\begin{vmatrix} \lambda_1 - \lambda & 0 & 0 \\ 0 & K'_{22} - \lambda & K'_{23} \\ 0 & K'_{23} & K'_{33} - \lambda \end{vmatrix} = 0. \quad (13)$$

Следовательно,

$$\lambda = \lambda_1 \quad (14)$$

или

$$\lambda^2 - (K'_{22} + K'_{33})\lambda + K'_{22}K'_{33} - (K'_{23})^2 = 0. \quad (15)$$

Уравнение (15) имеет действительные корни, так как

$$(K'_{22} + K'_{33})^2 - 4[K'_{22}K'_{33} - (K'_{23})^2] = (K'_{22} - K'_{33})^2 + 4(K'_{23})^2 \geq 0. \quad (16)$$

Если это выражение равно нулю, то корни будут равными, и наоборот. Если корни различные, то имеются три взаимно перпендикулярных направления, удовлетворяющих (4); они называются *главными осями*, а величины  $\lambda$  называются *главными значениями* тензора. Когда тензор отнесен к главным осям  $x'_i$ , то он имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad (17)$$

и говорят, что он приведен к *диагональному виду*. В таком случае относительно исходных осей имеем

$$K_{ik} = l_{ij}l_{kl}K'_{jl} = \lambda_1 l_{i1}l_{k1} + \lambda_2 l_{i2}l_{k2} + \lambda_3 l_{i3}l_{k3}. \quad (18)$$

Если два корня одинаковы, возьмем их равными  $\lambda_2$ ; тогда  $K'_{22} = K'_{33}$ ,  $K'_{23} = 0$ , и квадратичная форма принимает вид

$$\lambda_1 x_1'^2 + \lambda_2 (x_2'^2 + x_3'^2) = \text{const}. \quad (19)$$

Это уравнение определяет поверхность вращения, и любая прямая в плоскости  $x'_2$  и  $x'_3$  является главной осью.

Если все три корня равны между собой, то квадратичная форма определяет сферу.

В обоих этих особых случаях можно указать три перпендикулярных направления, удовлетворяющих (4), но теперь это

можно сделать неограниченным числом способов, тогда как в случае различных корней это можно сделать лишь единственным способом.

**3.081. Тензор инерции абсолютно твердого тела.** Рассмотрим абсолютно твердое тело, движущееся с угловой скоростью  $\omega$  и имеющее одну неподвижную точку  $O$ . Если частица  $P(x_i)$  имеет массу  $m$ , то

$$\frac{d}{dt} Sm(x_k \dot{x}_m - x_m \dot{x}_k) = Sm(x_k \ddot{x}_m - x_m \ddot{x}_k) = S(x_k X_m - x_m X_k), \quad (1)$$

что является соотношением между антисимметричными тензорами. Его также можно записать в векторном виде

$$\frac{d}{dt} Sm \epsilon_{ikm} x_k \dot{x}_m = S \epsilon_{ikm} x_k X_m, \quad (2)$$

или

$$\frac{d}{dt} Sm(\mathbf{x} \times \dot{\mathbf{x}}) = S(\mathbf{x} \times \mathbf{X}). \quad (3)$$

Выражение  $Sm(\mathbf{x} \times \dot{\mathbf{x}})$  называется *моментом импульса* данного тела относительно точки  $O$  и обозначается  $\mathbf{h}(O)$ . Далее, так как точка  $O$  неподвижна и  $\omega$  — угловая скорость, то

$$\dot{\mathbf{x}} = \omega \times \mathbf{x}, \quad (4)$$

т. е.

$$\dot{x}_i = \epsilon_{ikm} \omega_k x_m \quad (5)$$

и

$$\begin{aligned} h_i(O) &= Sm \epsilon_{ikm} x_k \epsilon_{mps} \omega_p x_s = \\ &= Sm(\delta_{ip} \delta_{ks} - \delta_{is} \delta_{kp}) x_k x_s \omega_p = \\ &= Sm(x_s^2 \omega_i - x_i x_p \omega_p) = I_{ik} \omega_k, \end{aligned} \quad (6)$$

где  $I_{ik}$  — симметричный тензор, равный

$$Sm(r^2 \delta_{ik} - x_i x_k). \quad (7)$$

В аффинорном обозначении

$$\mathbf{h}(O) = \mathbf{I} \cdot \omega. \quad (8)$$

Величина  $I_{ik}$  называется *тензором инерции* данного тела относительно точки  $O$ . Переписанная в полной форме, она имеет вид

$$\begin{pmatrix} Sm(x_2^2 + x_3^2) & -Smx_1x_2 & -Smx_1x_3 \\ -Smx_1x_2 & Sm(x_3^2 + x_1^2) & -Smx_2x_3 \\ -Smx_1x_3 & -Smx_2x_3 & Sm(x_1^2 + x_2^2) \end{pmatrix}. \quad (9)$$

Диагональные компоненты являются моментами инерции относительно осей, а недиагональные компоненты — центробежными моментами инерции, *умноженными на  $-1$* . Так как  $I_{ik}$  — симметричный тензор, то можно найти такие координатные оси, что центробежные моменты инерции обратятся в нуль и тензор примет вид

$$\begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}. \quad (10)$$

Величины  $A, B, C$  являются главными моментами инерции относительно точки  $O$ . Легко показать, что:

1) момент инерции относительно прямой, имеющей направляющие косинусы  $n_i$ , равен

$$I_{ik}n_in_k = n_i I_{ik} n_k = \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n}; \quad (11)$$

2) центробежный момент инерции относительно двух взаимно перпендикулярных прямых, имеющих направляющие косинусы  $n_i$  и  $n'_i$ , равен

$$-n_i I_{ik} n'_k = -\mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n}'; \quad (12)$$

3) если центром масс тела является точка  $G$  с координатами  $\bar{x}_i$  относительно точки  $O$  и  $I_{ik}(O)$  и  $I_{ik}(G)$  — тензоры инерции относительно точек  $O$  и  $G$  соответственно, а  $Sm = M$ , то

$$I_{ik}(O) = I_{ik}(G) + M(\bar{x}_m^2 \delta_{ik} - \bar{x}_i \bar{x}_k); \quad (13)$$

4) кинетическая энергия движущегося тела, имеющего неподвижную точку  $O$ , равна

$$\frac{1}{2} I_{ik}(O) \omega_i \omega_k = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{I}(O) \cdot \boldsymbol{\omega}; \quad (14)$$

5) кинетическая энергия тела, движущегося произвольным образом, равна

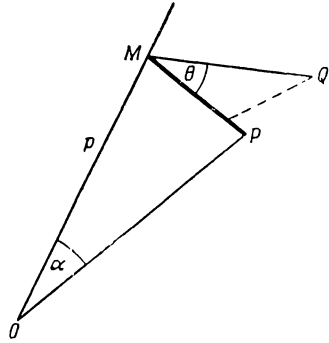
$$\frac{1}{2} MV^2 + \frac{1}{2} I_{ik}(G) \omega_i \omega_k = \frac{1}{2} M \mathbf{V} \cdot \mathbf{V} + \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{I}(G) \cdot \boldsymbol{\omega}, \quad (15)$$

где  $V$  — скорость центра масс.

Так как  $I_{ik} \omega_k$  — вектор, то его компоненты относительно любых осей можно написать сразу. В частности, если за координатные оси взять главные оси инерции, то компоненты момента импульса относительно их равны  $(A\omega_1, B\omega_2, C\omega_3)$ . Для абсолютно твердого тела  $A, B, C$  не зависят от времени, поэтому удобно пользоваться движущимися координатными осями.

**3.09. Ограниченное вращение абсолютно твердого тела.** Мы показали\*), что ограниченное вращение твердого тела относительно данной точки не может быть представлено при помощи вектора в направлении оси вращения. Теперь покажем, как такое вращение можно представить посредством тензора.

За начало координат возьмем точку  $O$  тела, а  $O123$  — система отсчета. Пусть  $P(x_i)$  — некоторая точка этого тела. Тело поворачивается на угол  $\theta$  вокруг прямой, проходящей через  $O$  и имеющей направляющие косинусы  $n_i$ . Точка  $P$  движется к  $Q(y_i)$ . Пусть  $M$  — проекция  $P$  на ось вращения, тогда  $M$  определяется соотношением  $n_k x_k n_i = p n_i$ . Вращательное перемещение  $P$  складывается из перемещения на отрезок  $(1 - \cos \theta) PM$  по направлению к  $M$  и на отрезок  $PM \sin \theta$  по перпендикуляру к плоскости  $OPM$ . Если  $\theta$  отсчитывается в правостороннем направлении, то последнее перемещение совпадает по направлению с векторным произведением  $\mathbf{n} \times \mathbf{x}$ . Определим его величину. Обозначим угол  $MOP$  через  $\alpha$ , тогда модуль  $\mathbf{n} \times \mathbf{x}$  равен  $OP \sin \alpha = PM$ . Следовательно, вторая часть перемещения равна  $\sin \theta (\mathbf{n} \times \mathbf{x})$  и



Р и с. 14.

$$\begin{aligned} y_i - x_i &= -(1 - \cos \theta)(x_i - p n_i) + \sin \theta (\mathbf{n} \times \mathbf{x})_i, \\ y_i &= \cos \theta x_i + (1 - \cos \theta) n_i n_k x_k + \sin \theta \epsilon_{ikm} n_k x_m = \\ &= [\cos \theta \delta_{ik} + (1 - \cos \theta) n_i n_k - \sin \theta \epsilon_{ikm} n_m] x_k. \end{aligned} \quad (1)$$

Очевидно, что величина в квадратных скобках является тензором второго порядка, который обозначим  $R_{ik}$ . Он не является ни симметричным, ни антисимметричным.

Если тело последовательно совершает повороты, представляемые тензорами  $R_{ik}^{(1)}, R_{ik}^{(2)}, R_{ik}^{(3)}, \dots, R_{ik}^{(n)}$ , то конечное положение точки  $P$  определяется выражением

$$x_i^{(n)} = R_{ik_{n-1}}^{(n)} R_{k_{n-1}k_{n-2}}^{(n-1)} \dots R_{k_1k}^{(1)} x_k.$$

В 2.03 мы видели, что порядок следования поворотов существен.

\*) См. 2.03, стр. 119, часть текста, набранная петитом.

**3.091.** Если  $\theta$  мало, то с точностью до членов первого порядка  $R_{ik}$  равно  $\delta_{ik} - \theta \epsilon_{ikm} n_m$  и

$$\mathbf{y} - \mathbf{x} = \theta \times \mathbf{x}, \quad (2)$$

где

$$\theta = \theta \mathbf{n}. \quad (3)$$

Результирующий поворот, состоящий из двух последовательных малых поворотов  $\theta$  и  $\theta'$ , равен

$$(\theta + \theta') \times \mathbf{x}, \quad (4)$$

и в этом смысле малый поворот может быть представлен вектором. Этот же результат был получен в гл. 2, но там было невозможно получить члены порядка  $\theta^2$ , которыми мы пренебрегли при выводе.

**3.092. Тензорная запись угловой скорости.** В гл. 2 было показано, что имеется вектор  $\boldsymbol{\omega}$ , представляющий угловую скорость абсолютно твердого тела, и что скорости  $\mathbf{v}_P$  и  $\mathbf{v}_Q$  двух точек связаны друг с другом соотношением

$$\mathbf{v}_Q = \mathbf{v}_P + \boldsymbol{\omega} \times \mathbf{PQ}. \quad (1)$$

Если  $\mathbf{PQ}$  имеет компоненты  $x_i$ , то, записав компоненты  $\mathbf{v}_P$  и  $\mathbf{v}_Q$ , как  $v_i^P$  и  $v_i^Q$ , при помощи **3.072** (13) получим, что если

$$\Omega_{ik} = \epsilon_{ikm} \omega_m, \quad (2)$$

то

$$v_i^Q - v_i^P = -\Omega_{ik} x_k. \quad (3)$$

Выражение (3) в некотором смысле более обще, чем (1). В трехмерном случае поворот около  $O3$  есть то же самое, что и поворот от  $O1$  к  $O2$ . В случае любого числа измерений ( $\geq 2$ ) можно говорить о повороте от  $O1$  к  $O2$ , но лишь в случае трех измерений такой поворот происходит *вокруг* некоторой фиксированной оси. Мы рассмотрим это в следующей главе.

**3.10. Общее движение жидкости.** Если на частицы системы не наложена связь, требующая неизменности расстояний между ними, то движение не может продолжительное время определяться скоростью отдельной точки и угловой скоростью.

Пусть  $x_i$  — радиус-вектор некоторой частицы  $P$  жидкости, и пусть  $v_i$  — вектор ее скорости в данный момент времени. В таком случае  $v_i$  является функцией как  $x_i$ , так и  $t$ . Скорость  $v_i + \delta v_i$  в соседней точке  $Q(x_i + \delta x_i)$  равна

$$v_i + \delta v_i = v_i + \frac{\partial v_i}{\partial x_k} \delta x_k + O(\delta x_k)^2. \quad (1)$$



С точностью до членов первого порядка по  $\delta x_k$  имеем

$$\delta v_i = \frac{1}{2} \left( \frac{\partial v_k}{\partial x_i} + \frac{\partial v_i}{\partial x_k} \right) \delta x_k - \frac{1}{2} \left( \frac{\partial v_k}{\partial x_i} - \frac{\partial v_i}{\partial x_k} \right) \delta x_k = e_{ik} \delta x_k - \xi_{ik} \delta x_k. \quad (2)$$

Тогда  $e_{ik}$  — симметричный тензор, а  $\xi_{ik}$  — антисимметричный тензор, причем оба они имеют размерность  $1/t$ . Часть  $\delta v_i$ , зависящая от  $\xi_{ik}$ , есть не что иное, как перемещение, возникающее при вращении абсолютно твердого тела, с компонентами  $(\xi_{23}, \xi_{31}, \xi_{12})$ . Какое движение описывает вторая часть, мы увидим несколько позже. Рассмотрим скорость изменения величины  $\frac{1}{2} P Q^2$ . Она равна

$$\delta x_i \delta v_i = (e_{ik} \delta x_k - \xi_{ik} \delta x_k) \delta x_i. \quad (3)$$

Часть, зависящая от  $\xi_{ik}$ , обращается в нуль, потому что  $\xi_{ik}$  — антисимметричный тензор и все члены сокращаются. Поэтому

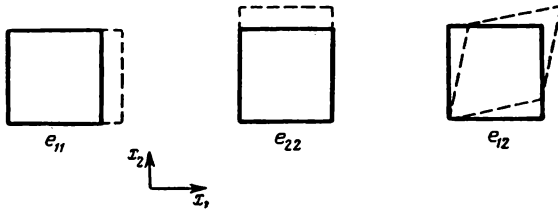


Рис. 15.

изменение расстояния между соседними частицами зависит исключительно от  $e_{ik}$ . Любая из этих величин может быть отличной от нуля, в то время как все остальные равны нулю. Этого, конечно, нельзя требовать для величин, которые не обращаются в нуль из-за условия симметрии при  $i \neq k$ . Таким образом, если  $\delta v_1 = e \delta x_1$ ,  $\delta v_2 = \delta v_3 = 0$ , то  $e_{11} = e$ , а все остальные равны нулю. Если  $\delta v_i = (e \delta x_2, e \delta x_1, 0)$ , то  $e_{12} = e_{21} = e$ , а все остальные — нули. Аналогичные соотношения по симметрии получим для других компонент. Соответствующие изменения в плоскости  $\delta x_3 = 0$  показаны на рис. 15.

Следовательно, тензор  $e_{ik}$  характеризует скорость изменения размеров и формы элемента жидкости в точке  $P$ . Его называют *тензором скорости деформации*. Этот тензор имеет три главные оси, и описанные изменения можно представить как расширения вдоль этих трех осей. Если главные значения тензора равны, то скорости расширения одинаковы во всех направлениях, т. е. деформация около точки  $P$  является симметричным расширением или сжатием.

В некотором смысле  $\xi_{ik}$  представляет локальную угловую скорость. Это утверждение требует оговорок, потому что  $e_{ik}$ , очевидно, содержит в себе угловые скорости, хотя они имеют разные значения для различных частей элемента, и без дополнительных ограничений понятие угловой скорости элемента, окружающего точку  $P$ , имеет неопределенный смысл. Рассмотрим малый элемент жидкости с центром масс в точке  $P$  и представим малое *абсолютно твердое* тело с точно таким же распределением плотности и тем же моментом импульса относительно точки  $P$ . В таком случае покажем, что если главные оси тензора инерции  $I_{ik}(P)$  совпадают с главными осями тензора  $e_{ik}$ , то угловая скорость абсолютно твердого тела равна

$$\omega_i = \frac{1}{2} e_{ikm} \xi_{km} = \frac{1}{2} (\text{rot } \mathbf{v})_i.$$

Если  $h_i$  — момент импульса рассматриваемого элемента, то

$$h_i = I_{ik} \omega_k = S m e_{ikm} \delta x_k (v_m + \delta v_m) = e_{ikm} v_m S m \delta x_k + S m e_{ikm} e_{mp} \delta x_i \delta x_p - S m e_{ikm} \xi_{mp} \delta x_k \delta x_p. \quad (4)$$

Далее, так как  $P$  — центр масс выбранного элемента, то  $S m \delta x_k = 0$ . Кроме того,

$$S m \delta x_k \delta x_p = -I_{kp} + \delta_{kp} S m (\delta x_s)^2. \quad (5)$$

Предполагая теперь, что  $I_{ik}$  и  $e_{ik}$  имеют совпадающие главные оси, т. е. можно взять эти оси так, что  $I_{ik} = 0$ ,  $e_{ik} = 0$ , если  $i \neq k$ . Тогда

$$e_{ikm} e_{mp} I_{kp} = 0, \quad e_{ikm} e_{mp} \delta_{kp} = 0, \quad (6)$$

потому что  $e_{mp} I_{kp}$  и  $e_{mp} \delta_{kp}$  обращаются в нуль при  $m \neq k$ , а  $e_{ikm} = 0$  при  $m = k$ . Следовательно, второй член в правой части (4) равен

$$[-I_{kp} + S m (\delta x_s)^2 \delta_{kp}] e_{ikm} e_{mp} = 0. \quad (7)$$

Если запишем  $\xi_{mn} = e_{smn} \xi_s$ ,  $\xi_i = e_{ikm} \xi_{km}/2$  [ср. 3.072 (10) и (12)], то последний член в (4) преобразуется к виду

$$-e_{ikm} e_{smn} \xi_s \delta x_k \delta x_p = [(\delta x_p)^2 \delta_{ip} - \delta x_i \delta x_k] \xi_i = I_{ik} \xi_k,$$

что является моментом импульса абсолютно твердого тела, заполняющего этот элемент и имеющего тензор инерции  $I_{ik}$  и угловую скорость  $\xi_k$ . Следовательно,

$$I_{ik} \omega_k = I_{ik} \xi_k, \quad (8)$$

что дает  $\omega_k = \xi_k$  при условии  $\|I_{ik}\| \neq 0$ .

Обратно, если  $\omega_k = \xi_k$ , то

$$S m e_{ikm} e_{mp} \delta x_k \delta x_p = 0. \quad (9)$$

Перейдем к главным осям тензора  $I_{ik}$ . Тогда  $S m \delta x_k \delta x_p = 0$  при  $k \neq p$ . Например, если  $i = 1$ , то в (9) отличны от нуля только члены с  $k = p = 2$ ,  $m = 3$  и  $k = p = 3$ ,  $m = 2$ , что дает

$$e_{32} S m (\delta x_2)^2 - e_{23} S m (\delta x_3)^2 = 0. \quad (10)$$

Следовательно, либо

$$e_{23} = 0, \quad \text{либо} \quad I_{22} = I_{33}. \quad (11)$$

Аналогично получим, что либо

$$e_{31} = 0, \quad \text{либо} \quad I_{33} = I_{11}, \quad e_{12} = 0 \quad \text{либо} \quad I_{11} = I_{22}. \quad (12)$$

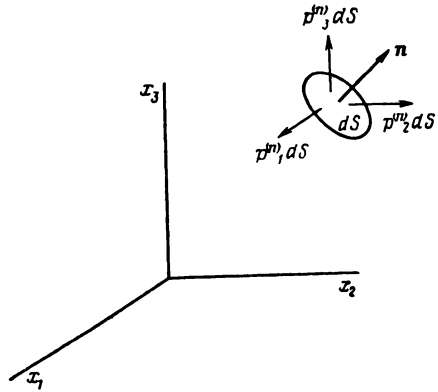
В таком случае, если главные оси элемента определены ( $I_{11} \neq I_{22} \neq I_{33}$ ), то  $e_{ik} = 0$  при  $i \neq k$  и, следовательно, главные оси  $e_{ik}$  и  $I_{ik}$  совпадают.

**3.101. Упругая деформация.** Анализ *смещений* в упругом твердом теле почти идентичен с анализом *скоростей* в жидкости. Пусть частица  $P(x_i)$  в момент времени  $t$  уже получила малое смещение  $u_i$ , так что ее первоначальные координаты были  $x_i - u_i$ . В таком случае если  $u_i + \delta u_i$  — смещение в точке  $Q(x_i + \delta x_i)$ , то точно тем же способом получим

$$\begin{aligned} \delta u_i &= \frac{\partial u_i}{\partial x_k} \delta x_k = \frac{1}{2} \left( \frac{\partial u_k}{\partial x_i} + \frac{\partial v_l}{\partial x_k} \right) \delta x_k - \frac{1}{2} \left( \frac{\partial u_k}{\partial x_i} - \frac{\partial v_l}{\partial x_k} \right) \delta x_k = \\ &= e_{ik} \delta x_k - \xi_{ik} \delta x_k, \end{aligned}$$

где  $e_{ik}$  и  $\xi_{ik}$  — симметричный и антисимметричный тензоры соответственно. Их можно интерпретировать как заданные изменения величины и формы элемента и его поворот. Так как в жидкости  $v_i$  обозначало скорость, то  $e_{ik}$  и  $\xi_{ik}$  соответствуют скоростям изменения величин, определенных здесь для упругого твердого тела.

**3.102. Напряжение.** Внутри вещества, будь то твердое тело, жидкость или газ, обычно имеют место взаимодействия между частицами. Общие свойства этих взаимодействий можно обнаружить, если рассмотреть, как можно приложить внешние силы к твердому телу, имеющему, скажем, одну закрепленную грань. Они могут быть приложены к любой части поверхности, и тело можно либо сжимать, либо растягивать перпендикулярно этой поверхности. Можно также приложить тангенциальное силовое воздействие, подобное трению. Понятие напряженного состояния обобщает представление о силе, действующей через поверхность на все элементы поверхности и даже на внутренние элементы. Если  $dS$  — малый элемент поверхности, нормаль которого имеет направление  $\mathbf{n}$ , то говорят о взаимодействиях *через*  $dS$ , представляющих силы, действующие между частицами, находящимися на противоположных сторонах этого элемента поверхности. Компоненты силы зависят как от величины, так и от направления  $dS$  и, следовательно, записываются в виде  $p_i^{(n)} dS$ . В частности, если  $\mathbf{n}$  совпадает с направлением  $x_k$ , то эту силу обозначим  $p_{ki} dS$  и назовем  $p_{ki}$



Р и с. 16.

компонентами напряжения, которые, следовательно, являются силами, действующими на единицу площади. Знак определим так. Пусть  $p_{ji} dS$  — сила противодействия, действующая на вещество, расположенное со стороны меньших  $x_k$ , и направленная в сторону увеличения  $i$ -й координаты. Силу, действующую на вещество, расположенное со стороны больших  $x_k$ , и направленную в сторону увеличения  $i$ -й координаты, обозначим  $-p_{ji} dS$ .

Компоненты напряжения обладают двумя замечательными свойствами. Они образуют симметричный тензор и связаны простым линейным соотношением с тензором скоростей деформации для жидкости и с тензором деформации упругого твердого тела при условии, что деформации малы. Мы будем предполагать, что они имеют непрерывные производные.

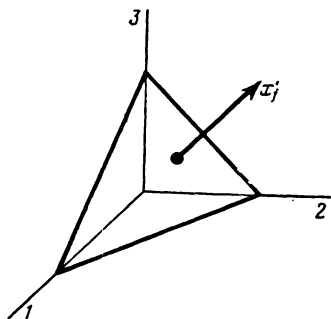


Рис. 17.

**3.103.** Покажем вначале, что компоненты напряжения образуют тензор. Пусть плоскость  $x'_j = \text{const}$  отсекает малые, порядка  $a$ , отрезки на осях  $x_i$ . Рассмотрим силы, дей-

ствующие на малый тетраэдр между этой плоскостью и осями  $x_i$ . Пусть  $dS$  — площадь плоскости  $x'_j$ , образующей основание данного тетраэдра. Тогда  $p_{ji} dS$  — силы, действующие на тетраэдр через  $dS$ . Величина силы, действующей на грань  $x_k = \text{const}$ , равна  $p_{ji}$ , умноженной на площадь этой грани, а площадь равна  $l_{kj} dS$ , где  $l_{kj}$  — косинус угла между  $x_k$  и  $x'_j$ . Однако эта сила действует на вещество с положительной стороны  $x_k = 0$  и, следовательно, должна быть взята со знаком минус.

Вещество внутри тетраэдра, вообще говоря, может подвергаться воздействию внешних сил, таких, как сила тяжести, которые будут порядка  $a^3$ , когда  $a$  мало. Они называются *объемными силами*. Это вещество к тому же может двигаться с ускорением, которое мы всегда предполагаем конечным. В таком случае скорость изменения импульса также порядка  $a^3$ , и из условия, что скорость изменения импульса элемента объема равна полной силе, следует

$$(p_{ji} - l_{kj} p_{ki}) dS = O(a^3).$$

Это соотношение связывает силы, действующие в направлении  $x_i$ . Их можно разложить и в других направлениях  $x'_i$ , что дает

$$(l_{il} p_{ij} - l_{il} l_{kj} p_{ki}) dS = O(a^3).$$

Но  $l_{il}p_{lj}dS$  — компонента силы, действующей на плоскость  $x'_j = \text{const}$  в направлении  $x'_i$ , и, следовательно, равна  $p'_{ji}dS$ , где  $p'_{ji}$  — компоненты напряжений относительно новых осей. Кроме того,  $dS$  порядка  $a^2$ , и, следовательно, при  $a$ , стремящемся к нулю, имеем

$$p'_{ji} = l_{il}l_{kj}p_{ki} = l_{ij}l_{kl}p_{ik}.$$

Таким образом,  $p_{ik}$  является тензором.

**3.104.** Теперь возьмем малый параллелепипед с центром в  $x_i$  и со сторонами  $\delta x_i$  и рассмотрим моменты сил относительно прямой, проходящей через центр параллелепипеда параллельно оси  $x_3$ . Вначале пренебрежем изменением компонент напряжения в рассматриваемой области. На грань  $x_2 + \frac{1}{2}\delta x_2$  параллельно  $x_1$  действует сила  $p_{21}\delta x_1\delta x_3$ . Ее момент равен  $-p_{21}\delta x_1\delta x_3(\frac{1}{2}\delta x_2)$ . Сила, действующая на грань  $x_2 - \frac{1}{2}\delta x_2$  имеет такую же величину, но направлена в противоположную сторону. Так как она приложена к противоположной грани, то ее момент тот же самый. Силы, параллельные  $x_2$  и действующие на грани, перпендикулярные оси  $x_1$ , имеют момент  $p_{12}dx_1dx_2dx_3$ . Очевидно, что все силы, образованные из других компонент тензора напряжений и действующие на эти же грани, дают момент, равный нулю. Следовательно, если напряжение однородно, то момент сил будет равен

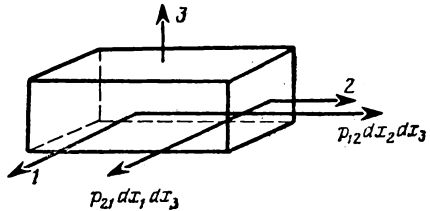


Рис. 18.

$$(p_{12} - p_{21})\delta x_1\delta x_2\delta x_3.$$

Легко показать, что изменение момента, обусловленное неоднородностью напряжения, при максимально допустимой неоднородности имеет порядок  $a^4$ , где  $\delta x_1, \delta x_2, \delta x_3$  порядка  $a$ . Полная величина объемных сил имеет порядок  $a^3$ , а их момент — порядок  $a^4$ . Аналогично полная скорость изменения момента импульса, пока ускорение конечно, имеет порядок по крайней мере  $a^4$ . Следовательно,

$$(p_{12} - p_{21})\delta x_1\delta x_2\delta x_3 = O(a^4).$$

Взяв достаточно малую область, видим, что

$$p_{12} = p_{21}.$$

Из симметрии следует, что

$$p_{ik} = p_{ki}.$$

Таким образом,  $p_{ik}$  — симметричный тензор.

**3.105. Уравнения движения.** Снова рассмотрим малый параллелепипед. Пусть  $\rho$  — плотность, а  $f_i$  — ускорение частицы вещества, находящейся в данный момент в точке  $x_i$ . Пусть  $X_i$  — объемная сила, действующая на единицу массы. Тогда  $\iiint \rho f_i dx_1 dx_2 dx_3$  по объему рассматриваемого элемента равен полной силе, действующей на этот элемент. Вклад объемной силы равен  $\iiint \rho X_i dx_1 dx_2 dx_3$ . Рассмотрим, какой вклад вносится за счет неоднородности напряжения. Грань  $x_1 + \frac{1}{2} \delta x_1$  дает  $p_{1i} \delta x_2 \delta x_3$ , где  $p_{1i}$  — заданные средние значения на этой грани. Противоположная грань дает  $-p_{1i} \delta x_2 \delta x_3$ , где  $p_{1i}$  — заданное среднее значение на *первой* грани. Если компоненты напряжения дифференцируемы, что обычно имеет место, то вклад обеих граней равен

$$\frac{\partial p_{1i}}{\partial x_1} \delta x_1 \delta x_2 \delta x_3 + O(a^4).$$

От всех шести граней имеем

$$\frac{\partial p_{ki}}{\partial x_k} \delta \tau + O(a^4).$$

Взяв достаточно малый параллелепипед, получим уравнения движения

$$\rho f_i = \rho X_i + \frac{\partial p_{ki}}{\partial x_k}.$$

В приведенном доказательстве ничего не предполагается о свойствах вещества, за исключением того что действие и противодействие между соседними частями равны и противоположно направлены, все ускорения ограничены и компоненты тензора напряжения дифференцируемы. Это одинаково справедливо для твердых тел и жидкостей. Различия между агрегатными состояниями вещества сказываются тогда, когда рассматривается связь между напряжением и деформацией.

**3.106. Соотношение напряжение — деформация. 3.1061. Упругое твердое тело.** Достаточно очевидно, что при простом перемещении или повороте упругого твердого тела в новое положение равновесия не происходит изменения напряжений, следовательно, компоненты тензора напряжения не зависят от

поворота тела. Основное соотношение выражается законом Гука, который в наиболее общем виде формулируется так: до тех пор пока компоненты смещений малы, напряжение линейно связано с ними. Это верно для большинства анизотропных кристаллов. Обычная теория упругости справедлива для изотропных твердых тел. В этой теории предполагается, что имеет место гораздо более частное соотношение. Один из возможных способов состоит в предположении, что главные оси тензора напряжений и тензора деформаций всегда совпадают. Когда растягивающее напряжение приложено вдоль однородного стержня, то этот стержень удлиняется в продольном направлении и сжимается в поперечном. Изменения размеров одинаковых элементов во всех поперечных направлениях будут равны. Если ось  $x_1$  взята вдоль оси стержня, то это можно записать так:

$$Ee_{11} = p_{11}, \quad Ee_{22} = Ee_{33} = -\sigma p_{11}. \quad (1)$$

Величина  $E$  называется модулем Юнга, а  $\sigma$  — коэффициент Пуассона. Обе они являются константами данного вещества. В рассмотренном случае все компоненты тензора напряжения, кроме  $p_{11}$ , равны нулю и три компоненты тензора деформации  $e_{21}$ ,  $e_{31}$ ,  $e_{12}$  также равны нулю. Если затем рассмотреть напряжения  $p_{22}$ ,  $p_{33}$ , то, поскольку соотношение напряжение — деформация линейно, можно добавить соответствующие деформации и получить в более общем виде

$$Ee_{11} = p_{11} - \sigma(p_{22} + p_{33}), \quad e_{23} = 0, \quad p_{23} = 0. \quad (2)$$

Аналогично получаются симметричные соотношения. Первое из этих соотношений можно записать

$$Ee_{11} = (1 + \sigma)p_{11} - \sigma(p_{11} + p_{22} + p_{33}), \quad (3)$$

а вся система в симметричной форме имеет вид

$$Ee_{i\bar{i}} = (1 + \sigma)p_{i\bar{i}} - \sigma p_{m\bar{m}}\delta_{i\bar{i}}. \quad (4)$$

Эта система уравнений справедлива для выбранных координатных осей, которые являются главными осями тензора напряжения и тензора деформации. Если теперь перейдем к любой другой прямоугольной системе координат, то каждый член выражается в соответствии с правилом преобразования тензоров второго порядка и тогда получим

$$Ee'_{\mu\bar{\mu}} = (1 + \sigma)p'_{\mu\bar{\mu}} - \sigma p'_{n\bar{n}}\delta_{\mu\bar{\mu}} \quad (5)$$

и

$$p'_{n\bar{n}} = p_{m\bar{m}}. \quad (6)$$

Следовательно, соотношение (4) имеет такой вид не только в системе координат, совпадающей с главными осями, но и в любой другой прямоугольной системе координат. Практическая ценность этого результата состоит в том, что в большинстве задач теории упругости главные оси тензора напряжения не во всех точках направлены одинаково и требуется система уравнений, вид которой сохраняется в любых осях. Этим также удобно пользоваться, если известны напряжения, а по ним надо найти деформации.

С другой стороны, во многих задачах напряжения бывают не известны, и уравнения движения нужно рассматривать как дифференциальные уравнения для определения смещений. В таком случае надо выразить напряжения через смещения, а следовательно, через деформации. Это можно сделать следующим способом. Прежде всего произведем свертывание уравнений (4), что дает

$$E e_{mm} = (1 + \sigma) p_{mm} - 3\sigma p_{mm} = (1 - 2\sigma) p_{mm}, \quad (7)$$

$$(1 + \sigma) p_{ik} = E e_{ik} + \frac{\sigma E}{1 - 2\sigma} e_{mm} \delta_{ik}, \quad (8)$$

$$p_{ik} = \lambda e_{mm} \delta_{ik} + 2\mu e_{ik}, \quad (9)$$

где

$$\lambda = \frac{\sigma E}{(1 + \sigma)(1 - 2\sigma)}, \quad \mu = \frac{E}{2(1 + \sigma)}. \quad (10)$$

Величины  $\lambda$  и  $\mu$  известны как постоянные Ляме. Из них  $\lambda$  не имеет специального названия, а  $\mu$  называется *жесткостью* \*). Очевидно, если брусок закрепить на плоскости  $x_2 = 0$ , а к противоположной его грани приложить касательное напряжение  $p_{21}$ , то брусок перекосится. Предположим, что смещение равно  $(\eta x_2, 0, 0)$ . В таком случае  $\eta$  — малый угол и является мерой сдвига. Все компоненты тензора деформации обращаются в нуль, за исключением  $e_{12}$ ,  $e_{21}$ , которые равны  $1/2\eta$ ,  $-1/2\eta$ . Тогда из (9) получим

$$p_{21} = \mu \eta. \quad (11)$$

Таким образом,  $\mu$  является отношением сдвигового напряжения к сдвигу и определяет сопротивление вещества перекашиванию.

Другая важная постоянная проявляется в случае сферически симметричной деформации, т. е. когда  $e_{11} = e_{22} = e_{33}$  и  $e_{21} = e_{31} = e_{12} = 0$ . Это означает, что относительные линейные размеры вещества изменяются одинаково во всех направлениях.

\*) Или модулем сдвига. — Прим. ред.



Соответствующее напряжение называется *гидростатическим*. В этом случае из (9) следует

$$p_{11} = 3\lambda e_{11} + 2\mu e_{11} = 3k e_{11}, \quad (12)$$

где

$$k = \lambda + \frac{2}{3}\mu. \quad (13)$$

Величина  $k$  называется *модулем объемного сжатия*, потому что относительное изменение объема в первом порядке равно  $3e_{11}$ , а  $k$  является отношением симметричного напряжения к этому изменению объема. Эту величину также называют *несжимаемостью*, а  $1/k$  — *сжимаемостью*.

Выражая  $E$  и  $\sigma$  из (10) через  $\lambda$  и  $\mu$ , получим

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}, \quad \sigma = \frac{\lambda}{2(\lambda + \mu)}. \quad (14)$$

Упругие постоянные  $E$ ,  $k$ ,  $\mu$  наиболее легко определяются экспериментально при непосредственных измерениях. Все они имеют размерность напряжения.

Другой метод формулировки теории упругости состоит в том, что если имеет место линейное соотношение общего вида

$$p_{ik} = c_{iklm} e_{lm},$$

справедливое для всех систем координат, то  $c_{iklm}$  — тензор четвертого порядка. Если к тому же его компоненты имеют одинаковые значения во всех системах координат, то он изотропный, и, следовательно, согласно 3.031, можно написать

$$c_{iklm} = \lambda \delta_{ik} \delta_{lm} + \mu (\delta_{il} \delta_{km} + \delta_{lp} \delta_{im}) + \nu (\delta_{il} \delta_{kp} - \delta_{lp} \delta_{km}),$$

где  $\lambda$ ,  $\mu$ ,  $\nu$  — скаляры. Тогда

$$\begin{aligned} p_{ik} &= \lambda \delta_{ik} e_{mm} + \mu (e_{il} + e_{kl}) + \nu (e_{ik} - e_{ki}) = \\ &= \lambda \delta_{ik} e_{mm} + 2\mu e_{ik}, \end{aligned}$$

так как  $e_{ik}$  симметричен.

Этот способ имеет то преимущество, что позволяет при помощи соответствующих модификаций метода 3.031 обнаружить свойства симметрии тензоров четвертого порядка для различных типов кристаллов. Следовательно, этот метод можно использовать для получения связи между напряжением и деформацией для кристаллов.

**3.1062. Жидкость.** В жидкости среднее напряжение, равное  $1/3 p_{mm}$ , почти всегда отрицательно и обозначается  $-p$ . Величина  $p$  называется *давлением*. (В противоположность тому, что говорится в некоторых учебниках, жидкость, тщательно освобожденная от растворенных газов, может выдерживать значительные растягивающие напряжения. Но верно и то, что это растяжение редко осуществляется практически.) В классической жидкости тензор напряжения имеет простой вид  $-p \delta_{ik}$ .

Это является хорошим приближением для многих задач, относящихся к реальным жидкостям. Отклонение напряжения от этой величины линейно зависит от скорости деформации, и если эту скорость обозначить  $e_{ik}$ , то для реальной жидкости, так же как и для изотропного твердого тела, будет справедливо утверждение, что главные оси тензора напряжения совпадают с главными осями тензора  $e_{ik}$ . Таким образом, искомое соотношение можно записать в виде

$$p_{ik} + p\delta_{ik} = \lambda' e_{mm}\delta_{ik} + 2\mu' e_{ik}. \quad (1)$$

Согласно определению  $p$ , тензор, стоящий слева, должен обращаться в нуль при свертывании. Следовательно,

$$(3\lambda' + 2\mu') e_{mm} = 0 \quad (2)$$

и

$$p_{ik} = -p\delta_{ik} + 2\mu' \left( e_{ik} - \frac{1}{3} e_{mm}\delta_{ik} \right). \quad (3)$$

Величина  $\mu'$  называется *вязкостью*. Ее размерность совпадает с размерностью напряжения, умноженного на время. Функция, стоящая после коэффициента  $2\mu'$ , описывает отклонение тензора скорости деформации от сферической симметрии.

Давление  $p$  обладает важными свойствами. Оно почти не зависит от скорости деформации, хотя теоретически может содержать малый член, пропорциональный  $e_{mm}$ . Однако он настолько мал, что не имеет практического значения. Следовательно, давление можно рассматривать как функцию только плотности и температуры согласно общим законам теплового расширения и сжимаемости для жидкости или газа.

**3.1063. Ускорение.** В левой части уравнений движения имеется член с ускорением вида  $\rho f_i$ . Необходимо его выразить для жидкости через производные от скорости, а для твердого тела — через производные от смещения. Наш вывод уравнений движения использовал параллелепипед, фиксированный в пространстве. Можно взять элемент объема, движущийся вместе с веществом, но он, вообще говоря, не должен оставаться прямоугольным и разложение сил при этом должно быть очень сложным. Однако ускорение относится к конкретной частице вещества.

Чтобы это стало ясным, удобно временно воспользоваться методом Лагранжа при описании движения. Предположим, что частица, имеющая координату  $x_i$  в момент времени  $t$ , имела координату  $a_i$  в момент  $t_0$ . Тогда движение каждой частицы описывается уравнением вида

$$x_i = g_i(a_1, a_2, a_3, t), \quad (1)$$

где для данной частицы  $a_i$  не зависит от  $t$ . В таком случае скорость и ускорение частицы равны

$$v_i = \left( \frac{\partial x_i}{\partial t} \right)_{a_i}, \quad f_i = \left( \frac{\partial^2 x_i}{\partial t^2} \right)_{a_i}, \quad (2)$$

где индекс обозначает, что  $a_k$  остается постоянной при дифференцировании.

В обычном эйлеровом методе описания движения скорость частицы рассматривается как функция ее пространственного положения в момент времени  $t$ , а не в момент  $t_0$ . Следовательно, если за время  $\delta t$  частица переместилась от  $x_i$  до  $x_i + \delta x_i$ , то ее скорость  $v_i$  будет зависеть от  $t + \delta t$  и  $x_i + \delta x_i$ . Отсюда ускорение частицы равно

$$\lim_{\delta t \rightarrow 0} \frac{v_i(t + \delta t, x_i + \delta x_i) - v_i(t, x_i)}{\delta t} = \lim \left[ \left( \frac{\partial v_i}{\partial t} \right)_{x_i} + \frac{\partial v_i}{\partial x_k} \left( \frac{\delta x_k}{\delta t} \right) \right]. \quad (3)$$

В пределе  $\delta x_k / \delta t$  равна собственной скорости частицы  $v_k$ , поэтому

$$f_i = \left( \frac{\partial v_i}{\partial t} \right)_{x_i} + v_k \frac{\partial v_i}{\partial x_k}. \quad (4)$$

Оператор  $\partial/\partial t + v_k \partial/\partial x_k$ , который определяет производную по времени от любой величины, связанной с определенной частицей (для которой  $a_i$  постоянно в методе Лагранжа), в английской литературе обычно обозначают  $D/Dt$ . Он превращается в оператор частной производной  $\partial/\partial t$ , если постоянным считается  $a_i$  вместо  $x_i$ . Когда же, как в методе Эйлера, вообще не используется величина  $a_i$ , то, по-видимому, нет достаточных причин не рассматривать этот оператор как обыкновенную полную производную и обозначать ее  $d/dt$ . В действительности обозначение  $D/Dt$  является пережитком тех времен, когда  $d/dt$  использовалось для обозначения частного дифференцирования.

Таким образом, уравнение движения жидкости имеет вид

$$\rho \left( \frac{\partial v_i}{\partial t} + v_k \frac{\partial v_i}{\partial x_k} \right) = \rho X_i + \frac{\partial p_{,i}}{\partial x_k}, \quad (5)$$

где компоненты тензора напряжений выражаются через тензор скорости деформации и давление согласно 3.1062 (3).

Для твердого тела условия несколько иные, поскольку в определении смещения в явном виде содержится начальное положение частицы. Действительно,

$$a_i = x_i - u_i. \quad (6)$$

Однако почти во всех задачах теории упругости смещения малы, и если пренебречь их квадратами, то  $d/dt$  можно заменить

на  $\partial/\partial t$ . Квадратами деформаций также пренебрегают, когда принимается, что напряжения линейно зависят от деформации. Поэтому, не уменьшая общности результата, пренебрежем ими и в ускорении \*).

**3.11. Тензор электромагнитных напряжений.** Вначале рассмотрим электрические силы. Пусть  $K$  — диэлектрическая проницаемость, которая предполагается постоянной,  $\mathbf{E}$  — напряженность электрического поля и  $\rho$  — плотность электрического заряда, тогда электрическая сила, действующая на единицу объема, равна

$$\mathbf{X} = \rho \mathbf{E}. \quad (1)$$

Кроме того,

$$4\pi\rho = K \operatorname{div} \mathbf{E}. \quad (2)$$

Тогда

$$\begin{aligned} 4\pi X_i &= K E_i \frac{\partial E_k}{\partial x_k} = K \frac{\partial}{\partial x_k} (E_i E_k) - K \frac{\partial E_i}{\partial x_k} E_k = \\ &= K \frac{\partial}{\partial x_k} (E_i E_k) - K \frac{\partial E_k}{\partial x_i} E_k. \end{aligned} \quad (3)$$

Здесь было использовано, что  $\mathbf{E}$  равняется градиенту потенциала. Далее получим

$$4\pi X_i = K \frac{\partial}{\partial x_k} \left( E_i E_k - \frac{1}{2} E_m^2 \delta_{ik} \right). \quad (4)$$

Следовательно, механическую силу можно получить из тензора напряжения

$$p_{ik} = \frac{K}{4\pi} \left( E_i E_k - \frac{1}{2} E_m^2 \delta_{ik} \right). \quad (5)$$

Теперь рассмотрим силу, обусловленную магнитным полем  $\mathbf{H}$ , которая действует на среду, проводящую электрический ток плотности  $\mathbf{j}$ . Магнитную проницаемость  $\mu$  считаем постоянной. В таком случае имеем

$$\mathbf{X} = \mu \frac{\mathbf{j}}{c} \times \mathbf{H}, \quad \operatorname{rot} \mathbf{H} = \frac{4\pi \mathbf{j}}{c}, \quad (6)$$

$$\begin{aligned} 4\pi X_i &= \mu \epsilon_{ikm} (\operatorname{rot} \mathbf{H})_k H_m = \mu \epsilon_{ikm} \epsilon_{kps} \frac{\partial H_s}{\partial x_p} H_m = \\ &= \mu (\delta_{mp} \delta_{is} - \delta_{ms} \delta_{ip}) \frac{\partial H_s}{\partial x_p} H_m = \mu \left( \frac{\partial H_i}{\partial x_m} - \frac{\partial H_m}{\partial x_i} \right) H_m. \end{aligned} \quad (7)$$

Однако

$$\frac{\partial}{\partial x_m} (H_i H_m) = \frac{\partial H_i}{\partial x_m} H_m + H_i \frac{\partial H_m}{\partial x_m} \quad (8)$$

---

\*) Наиболее полно члены второго порядка рассмотрены в работе [2].

и  $\partial H_m / \partial x_m = 0$ . Следовательно,

$$4\pi X_i = \mu \frac{\partial}{\partial x_m} \left( H_i H_m - \frac{1}{2} H_k^2 \delta_{im} \right), \quad (9)$$

и  $X_i$  можно получать из тензора напряжения

$$p_{ik} = \frac{\mu}{4\pi} \left( H_i H_k - \frac{1}{2} H_m^2 \delta_{ik} \right). \quad (10)$$

О дополнительных членах, появляющихся при непостоянных  $K$  и  $\mu$ , см. [3].

**3.12. Скорость изменения вектора во вращающейся системе координат.** Нам уже известно правило преобразования векторов, в котором содержится зависимость только от углов поворота координатных осей. До сих пор у нас не было случая рассмотреть, что произойдет, если эти углы сами изменяются со временем. До тех пор пока преобразование от одной координатной системы к другой является чисто алгебраическим, никаких затруднений не возникает. Все тождества, зависящие от определяемой компоненты вектора в заданном направлении, остаются справедливыми, даже если направляющие косинусы сами изменяются. Это справедливо при условии, что все величины берутся в один и тот же момент времени. Однако если требуется дифференцировать по времени, то приходится рассматривать различные моменты времени, при этом необходимо обратить особое внимание на изменение направляющих косинусов. Рассмотрим очень простой случай. Пусть частица движется по кругу с постоянной угловой скоростью, так что

$$x_1 = a \cos \omega t, \quad x_2 = a \sin \omega t, \quad x_3 = 0.$$

Отсюда получим компоненты скорости  $(-\omega a \sin \omega t, \omega a \cos \omega t, 0)$  и компоненты ускорения  $(-\omega^2 x_1, -\omega^2 x_2, 0)$ . Теперь возьмем вращающуюся систему координат, у которой  $x'_3$  совпадает с  $x_3$ , а  $x'_1$  образует с  $x_1$  угол  $\omega t$  и, следовательно, все время направлена на частицу. В таком случае в системе  $O 1' 2' 3'$  координаты частицы все время равны  $(a, 0, 0)$ , а скорости их изменения будут  $(0, 0, 0)$ . Однако компоненты скорости *относительно* системы  $O 123$  вдоль этих же осей равны  $(0, \omega a, 0)$ , а компоненты ускорения вдоль этих же осей равны  $(-\omega^2 a, 0, 0)$ . *Преобразования скоростей изменения координат относительно неподвижной системы к вращающейся системе координат не происходит по правилу преобразования векторов.* Можно сказать, что операции дифференцирования по времени и разложения

по осям в заданном направлении *коммутируют*, если только это направление фиксировано \*).

Простейшие уравнения движения  $m\ddot{x}_i = X_i$  справедливы в инерциальной системе координат. Если вместо этой системы используется вращающаяся система координат  $x'_j$ , то силу  $X_i$  можно представить через составляющие вдоль этих осей по правилу разложения векторов, а уравнения движения примут вид

$$ml_{ij}\ddot{x}_i = l_{ij}X_i = X'_j. \quad (1)$$

Однако здесь левая часть не равна  $m\ddot{x}'_j$ . При рассмотрении движения абсолютно твердых тел обычно бывает удобно записывать уравнения движения относительно вращающейся системы координат. Эти координатные оси часто жестко скрепляют с телом, и, следовательно, надо выразить  $l_{ij}\ddot{x}_i$  через  $x'_j$  и его производные. Аналогично если  $A_i$  — компоненты смещения, скорости или момента импульса относительно инерциальной системы координат, то необходимо выразить  $l_{ij}dA_i/dt$  через  $l_{ij}A_i$  и их производные. Обозначим  $l_{ij}A_i$  через  $A'_j$ . Если  $B_i = A_i$ , то соотношение

$$l_{ij}B_i = l_{ij}A_i \quad (2)$$

справедливо даже тогда, когда координатные оси  $x'_j$  вращаются. Справедливо и обратное утверждение. В таком случае имеем

$$A_i = l_{ij}A'_j, \quad (3)$$

$$\frac{dA_i}{dt} = l_{ij} \frac{dA'_j}{dt} + A'_j \frac{dl_{ij}}{dt}. \quad (4)$$

Возьмем точку на оси  $x'_j$  на постоянном расстоянии  $c$  от начала координат. Ее координаты в системе  $x_i$  равны  $cl_{ij}$ , а компоненты ее скорости  $c dl_{ij}/dt$ . Координатные оси  $x'_j$  образуют жесткий каркас, и, следовательно, скорость точки, жестко скрепленной с ними и имеющей координаты  $x_i$ , равна  $-\Theta_{ik}x_k$ ,

---

\*) Со свойством некоммутативности операторов мы встретимся еще неоднократно. Простейший пример таких операций: умножение на  $x$  и дифференцирование по  $x$ :

$$\frac{d}{dx} x f(x) = x \frac{d}{dx} f(x) + f(x),$$

что не равно

$$x \frac{d}{dx} f(x).$$

где  $\Theta_{ik}$  — скорость вращения осей, как и в 3.092. Следовательно, компонента скорости точки вдоль оси  $x_i$  равна  $-c\Theta_{ik}l_{kj}$ , а

$$\frac{dl_{ij}}{dt} = -\Theta_{ik}l_{kj}, \quad (5)$$

$$\frac{dA_i}{dt} = l_{ij} \frac{dA'_j}{dt} - \Theta_{ik}l_{kj}A'_j, \quad (6)$$

$$l_{il} \frac{dA_i}{dt} = l_{il}l_{ij} \frac{dA'_j}{dt} - \Theta_{ik}l_{kj}l_{il}A'_j. \quad (7)$$

Но  $l_{il}l_{ij} = \delta_{lj}$ , а  $l_{il}l_{kj}\Theta_{ik}$ , равное  $\Theta'_{lj}$ , которая определяется посредством преобразования  $\Theta_{ik}$  к осям  $x'_j$ ,  $x'_i$  согласно правилу преобразования тензоров второго порядка. Это позволяет для искомых компонент написать

$$l_{il} \frac{dA_i}{dt} = \frac{dA'_l}{dt} - \Theta'_{lj}A'_j = \left(\frac{d\mathbf{A}}{dt}\right)_l + (\boldsymbol{\theta} \times \mathbf{A})_l, \quad (8)$$

$$l_{il} \frac{dA_i}{dt} = (\dot{A}'_l - A'_2\theta_3 + A'_3\theta_2, \dot{A}'_2 - A'_3\theta_1 + A'_1\theta_3, \dot{A}'_3 - A'_1\theta_2 + A'_2\theta_1). \quad (9)$$

Эти выражения являются компонентами такого вектора вдоль осей  $x'_l$ , у которого компоненты вдоль осей  $x_i$  равны  $dA_i/dt$ . Легко проверить, что если взять третью систему координат, то компоненты относительно нее полчатся одинаковыми, независимо от того, будем ли производить преобразования непосредственно из системы  $i$  или через систему  $j$ . Эти компоненты, следовательно, удовлетворяют правилу последовательного преобразования для векторов.

Считать ли в отдельности члены  $d\mathbf{A}'/dt$  и  $(\boldsymbol{\theta} \times \mathbf{A})$  векторами, это вопрос определения, и в связи с этим появляются два возможных пути. Согласно нашему исходному определению, правило преобразования служит единственным критерием для установления, являются ли совокупности трех величин компонентами вектора. В этом правиле ничего не говорится о постоянстве направляющих косинусов при преобразовании. Если мы сохраним это определение для случая, когда  $l_{il}$  — функции времени, и так как

$$\frac{dA'_l}{dt} \neq l_{il} \frac{dA_i}{dt},$$

то  $d\mathbf{A}/dt$  не является вектором, если системы координат вращаются относительно друг друга. Однако  $d\mathbf{A}/dt + (\boldsymbol{\theta} \times \mathbf{A})$  в этом смысле вектор, а следовательно,  $(\boldsymbol{\theta} \times \mathbf{A})$  — не вектор. Тот факт, что для компактности записи это выражение представлено

в виде векторного произведения, не превращает эту величину в вектор. Проверкой является правило преобразования, и прежде чем сказать, что совокупность величин является компонентами вектора, необходимо показать, что это правило выполняется. В данном случае это правило не имеет места.

Другой путь состоит в ограничении сферы действия правила преобразования лишь системами координат, не вращающимися относительно друг друга. Предположим для простоты, что  $\mathbf{A}$  — смещение из начала координат. Тогда  $d\mathbf{A}/dt$  по определению является скоростью, компоненты которой отнесены к покоящейся системе координат. Однако можно представить бесконечный набор закрепленных систем координат с общим началом, тогда в любой момент времени движущаяся система координат совпадает с одной из них, и тем самым эти системы отождествляются. Произведем разложение относительно этой координатной системы и получим, что компоненты скорости относительно не равны  $l_{ij}dA_i/dt$ . Эти компоненты являются компонентами вектора. Величина  $\boldsymbol{\theta} \times \mathbf{A}$  — скорость точки, жестко скрепленной с движущейся системой координат, относительно любой неподвижной координатной системы. Такое движение возможно, и в этом смысле  $\boldsymbol{\theta} \times \mathbf{A}$  вектор. В таком случае  $l_{ij}dA'_j/dt$  — компоненты вектора относительно системы координат  $x_i$ , но это *не тот же вектор*, что  $d\mathbf{A}/dt$ . Это скорость точки, движущейся относительно системы координат  $x_i$  и равной *части* скорости  $d\mathbf{A}/dt$ , которая не выражается через  $\boldsymbol{\theta} \times \mathbf{A}$ .

При данной интерпретации соотношение (8) следует рассматривать как выражение для компонент скорости изменения  $\mathbf{A}$  не вдоль движущихся координатных осей, а вдоль неподвижных осей, совпадающих в данный момент с движущимися.

Любая из этих интерпретаций пригодна. Важно понять, что эти интерпретации не тождественны и ими нельзя пользоваться вместе, иначе неизбежны ошибки \*).

Угловую скорость  $\Theta'_{ij}$  иногда называют угловой скоростью вращающихся осей координат „относительно самих себя“ без какого-либо объяснения, почему такая угловая скорость не равна нулю. Рассматривая происхождение данной величины, видим, что это *угловая скорость осей координат относительно неподвижных осей, в данный момент совпадающих с ними*. Если известна угловая скорость относительно любой неподвижной системы координат, то  $\Theta'_{ij}$  или ее векторное представление определяются посредством разложения по осям.

---

\*) Подобный вопрос возникает в так называемом „ковариантном дифференцировании“ в общей теории относительности [4,5].



**3.13. Приложения к механике. Углы Эйлера.** Положение абсолютно твердого тела, имеющего одну закрепленную точку  $O$  относительно неподвижной системы координат  $Oxyz$ , можно определить следующим образом. Вначале выберем какую-либо реперную линию, проходящую в данном теле, и будем рассматривать ее в качестве оси  $O3$  прямоугольной системы координат, жестко скрепленной с телом. Полярные углы этой линии обозначают  $\theta$ ,  $\lambda$ . Положение тела теперь известно, за исключением возможного поворота вокруг оси  $O3$ . Оно будет полностью определено, если известен угол между какой-либо реперной плоскостью, скрепленной с телом и проходящей через ось  $O3$ , и плоскостью  $zO3$ . Назовем этот угол  $\chi$ . Три угла  $\theta$ ,  $\lambda$ ,  $\chi$  — углы Эйлера.

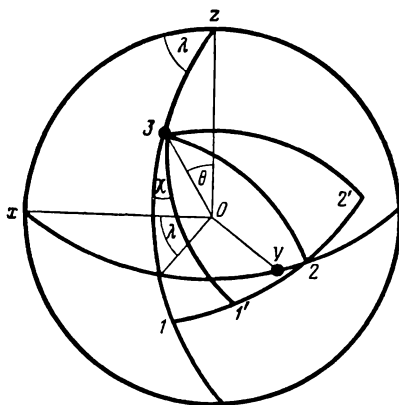


Рис. 19.

Выразим угловую скорость твердого тела через скорости изменения этих трех углов. Для этого воспользуемся промежуточной системой отсчета  $O123$ , у которой ось  $O1$  лежит в плоскости  $zO3$  и перпендикулярна  $O3$ . В таком случае ось  $O2$  перпендикулярна плоскости  $zO3$ . Все системы отсчета берутся правосторонними. Система координат  $O123$ , вообще говоря, не закреплена ни в пространстве, ни в теле. Однако если плоскость  $O31'$  скреплена с телом и образует угол  $\chi$  с плоскостью  $O31$ , то третью ось  $O2'$  можно взять перпендикулярно к плоскости  $O31'$  и она также будет скреплена с телом.

Очевидно, что скорость вращения тела определяется скоростями вращения  $\dot{\lambda}$  вокруг  $Oz$ ,  $\dot{\theta}$  вокруг  $O2$  и  $\dot{\chi}$  вокруг  $O3$ . Они могут быть разложены вдоль любых удобных направлений. Их компоненты относительно системы координат  $O123$ , очевидно, равны  $(-\sin \theta \dot{\lambda}, \dot{\theta}, \dot{\chi} + \cos \theta \dot{\lambda})$ . Снова разлагая теперь относительно координат  $O1'2'3'$ , получим

$$(-\sin \theta \dot{\lambda} \cos \chi + \dot{\theta} \sin \chi, \sin \theta \dot{\lambda} \sin \chi + \dot{\theta} \cos \chi, \dot{\chi} + \cos \theta \dot{\lambda}),$$

что является компонентами угловой скорости тела относительно осей, скрепленных с телом.

Во многих задачах твердое тело имеет ось симметрии, которую можно принять за ось  $O3$ . В таком случае удобно пользоваться системой координат  $O123$  вместо системы  $O1'2'3'$ , так

как при этом получаются более простые соотношения между компонентами угловой скорости и углами Эйлера. Отличие состоит в том, что необходимо рассматривать вращение как твердого тела, так и системы отсчета, когда последняя не закреплена в пространстве. Обозначим угловую скорость системы отсчета через  $\theta(\theta_1, \theta_2, \theta_3)$ ; тогда если эта система отсчета скреплена с телом, то

$$\theta = \omega.$$

Но система координат  $O123$  не имеет компоненты угловой скорости  $\dot{\chi}$ , поскольку ось  $O1$  всегда остается в плоскости  $zO3$ . Следовательно, для этой системы имеем

$$(\theta_1, \theta_2, \theta_3) = (-\sin \theta \dot{\lambda}, \dot{\theta}, \cos \theta \dot{\lambda}),$$

тогда как компоненты угловой скорости тела равны

$$(\omega_1, \omega_2, \omega_3) = (-\sin \theta \dot{\lambda}, \dot{\theta}, \dot{\chi} + \cos \theta \dot{\lambda}).$$

Соотношения  $\theta_1 = \omega_1$ ,  $\theta_2 = \omega_2$ , очевидно, обозначают, что ось  $O3$  имеет ту же угловую скорость, что и частицы тела, расположенные вдоль нее, т. е. ось  $O3$  закреплена в твердом теле. Соотношение  $\theta_3 \neq \omega_3$  напоминает о том, что оси  $O1$  и  $O2$  не скреплены с телом.

**3.131.** Для иллюстрации рассмотренного метода вычислим компоненты ускорения частицы в сферических координатах  $(r, \theta, \lambda)$ . Выберем ось  $O3$  по направлению к частице, оси  $O1$  и  $O2$  выберем так же, как в 3.13. Компоненты угловой скорости этих осей равны  $(-\sin \theta \dot{\lambda}, \dot{\theta}, \cos \theta \dot{\lambda})$ . Так как частица все время находится на оси  $O3$ , то компоненты ее скорости равны

$$(0, 0, \dot{r}) + (\theta_1, \theta_2, \theta_3) \times (0, 0, r) = (r\dot{\theta}, r \sin \theta \dot{\lambda}, \dot{r}),$$

в чем можно убедиться путем проверки. В таком случае компоненты ускорения равны

$$(a_1, a_2, a_3) = \frac{d}{dt} (r\dot{\theta}, r \sin \theta \dot{\lambda}, \dot{r}) + (\theta_1, \theta_2, \theta_3) \times (r\dot{\theta}, r \sin \theta \dot{\lambda}, \dot{r}),$$

$$a_1 = \frac{d}{dt} (r \dot{\theta}) + \dot{r} \dot{\theta} - r \sin \theta \cos \theta \dot{\lambda}^2 = r(\ddot{\theta} - \sin \theta \cos \theta \dot{\lambda}^2) + 2\dot{r}\dot{\theta},$$

$$a_2 = \frac{d}{dt} (r \sin \theta \dot{\lambda}) + \cos \theta \dot{\lambda} r \dot{\theta} + \sin \theta \dot{\lambda} \dot{r} = \frac{1}{r \sin \theta} \frac{d}{dt} (r^2 \sin^2 \theta \dot{\lambda}),$$

$$a_3 = \ddot{r} - r \sin^2 \theta \dot{\lambda}^2 - r \dot{\theta}^2 = \ddot{r} - r(\dot{\theta}^2 + \sin^2 \theta \dot{\lambda}^2).$$

**3.14. Динамические уравнения Эйлера.** Пусть оси системы координат  $O1'2'3'$  направлены вдоль главных осей инерции абсолютно твердого движущегося тела, имеющего неподвижную

точку  $O$ . Если  $A, B, C$  — главные моменты инерции, а  $\omega$  — угловая скорость данного тела относительно тех же осей, то момент импульса относительно точки  $O$  равен  $\mathbf{h}(O) = (A\omega_1, B\omega_2, C\omega_3)$ . Скорость изменения момента импульса относительно точки  $O$ , следовательно, будет

$$\frac{d\mathbf{h}}{dt} + \omega \times \mathbf{h}$$

с компонентами  $(A\dot{\omega}_1 - (B - C)\omega_2\omega_3$  и т. д.) относительно мгновенных положений главных осей инерции.

**3.141. Движение волчка.** Выберем ось симметрии за координатную ось  $OZ$ . Воспользуемся, как и в 3.13, системой координат  $O123$ , тогда угловая скорость волчка равна  $(-\dot{\lambda} \sin \theta, \dot{\theta}, \dot{\chi} + \dot{\lambda} \cos \theta)$ , а угловая скорость системы координат будет  $(-\dot{\lambda} \sin \theta, \dot{\theta}, \dot{\lambda} \cos \theta)$ . Поскольку моменты инерции равны  $A, A, C$ , то для компонент момента импульса имеем  $[-A\dot{\lambda} \sin \theta, A\dot{\theta}, C(\dot{\chi} + \dot{\lambda} \cos \theta)]$ .

Если  $\mathbf{N}$  — момент внешних сил, то

$$\frac{d\mathbf{h}}{dt} + \theta \times \mathbf{h} = \mathbf{N}. \quad (1)$$

Если волчок движется в поле силы тяжести, а его центр масс имеет координаты  $(0, 0, h)$ , то

$$\mathbf{N} = (0, Mgh \sin \theta, 0). \quad (2)$$

Третья компонента уравнений (1) непосредственно дает

$$\dot{\chi} + \dot{\lambda} \cos \theta = \text{const} = n. \quad (3)$$

Для второй компоненты получим

$$A\ddot{\theta} + Cn\dot{\lambda} \sin \theta - A\dot{\lambda}^2 \sin \theta \cos \theta = Mgh \sin \theta, \quad (4)$$

так что условие установившейся прецессии при  $\theta = \alpha$  и  $\dot{\lambda} = \Omega$  имеет вид

$$A\Omega^2 \cos \alpha - Cn\Omega + Mgh = 0.$$

Для описания полного движения используем два других первых интеграла.

1) Момент импульса относительно вертикали постоянен, что дает

$$A \sin^2 \theta \dot{\lambda} + Cn \cos \theta = \text{const}.$$

2) Полная энергия постоянна, откуда

$$A(\dot{\lambda}^2 \sin^2 \theta + \dot{\theta}^2) + C(\dot{\lambda} \cos \theta + \dot{\chi})^2 + 2Mgh \cos \theta = \text{const}.$$

Если известны выражения для кинетической и потенциальной энергии, то уравнения движения можно получить, воспользовавшись уравнениями Лагранжа 10.07.

**3.15. Двухмерные тензоры.** Если возможны только такие повороты координатных осей, когда ось  $x_3$  не смещается, то изменения осей будут более ограничены, чем в случае произвольного поворота. Определяемые при этом дополнительные наборы функций позволяют произвести необходимые преобразования при таких изменениях. Действительно, если  $u_i$  ( $i = 1, 2$ ) является двухмерным вектором, то его компоненты в направлении  $\alpha$  равны  $u_1 \cos \alpha + u_2 \sin \alpha$ . Далее пусть имеется вектор  $\mathbf{v}$ , компоненты которого в направлении  $\alpha$  равны компонентам  $\mathbf{u}$  в направлении  $\pi/2 + \alpha$ . Это дает

$$v_1 \cos \alpha + v_2 \sin \alpha = u_1 \cos \left( \frac{\pi}{2} + \alpha \right) + u_2 \sin \left( \frac{\pi}{2} + \alpha \right) = -u_1 \sin \alpha + u_2 \cos \alpha, \quad (1)$$

и следовательно, если

$$v_1 = u_2, \quad v_2 = -u_1, \quad (2)$$

то  $v_1$  и  $v_2$  — компоненты вектора, равные компонентам вектора  $\mathbf{u}$  относительно осей координат, повернутых на прямой угол. В частности,  $(x_2, -x_1)$  — компоненты вектора.

Однако производная вектора является тензором второго порядка. Взяв в качестве этого вектора  $(x_2, -x_1)$ , получим

$$\eta_{ik} = \frac{\partial v_k}{\partial x_i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (3)$$

что является двухмерным тензором. Таким образом, не только  $\delta_{ik}$  является изотропным двухмерным тензором второго порядка, но и любая линейная комбинация

$$\lambda \delta_{ik} + \mu \eta_{ik}, \quad (4)$$

где  $\lambda$  и  $\mu$  — скаляры, также является таким тензором.

Теперь рассмотрим тензор четвертого порядка  $\eta_{im}\eta_{kp}$ . Все компоненты с  $i = m$  или  $k = p$  равны нулю, а для остальных имеем

$$\eta_{12}\eta_{12} = \eta_{21}\eta_{21} = 1, \quad \eta_{12}\eta_{21} = -1. \quad (5)$$

Далее,  $\eta_{im}\eta_{kp} \frac{\partial^2 \varphi}{\partial x_m \partial x_p}$  — тензор второго порядка, если  $\varphi$  — скаляр.

Но если  $i = k = 1$ , то отличен от нуля лишь один член при  $m = p = 2$ , который равен  $\partial^2 \varphi / \partial x_2^2$ . Если  $i = 1, k = 2$ , то необходимо

взять  $m = 2$ ,  $p = 1$ , что дает  $-\partial^2\varphi/\partial x_1\partial x_2$ . Продолжая таким же образом, получим

$$\begin{pmatrix} \frac{\partial^2\varphi}{\partial x_2^2} & -\frac{\partial^2\varphi}{\partial x_1\partial x_2} \\ -\frac{\partial^2\varphi}{\partial x_1\partial x_2} & \frac{\partial^2\varphi}{\partial x_1^2} \end{pmatrix}, \quad (6)$$

т. е. компоненты двумерного тензора второго порядка. Этот тензор находит применение в теории упругости, в частности при рассмотрении изгиба тонких пластин и распределения напряжения между параллельными плоскостями.

**3.16. Параллакс.** Тензорные методы иногда бывают полезны при получении формул сферической астрономии. Рассмотрим параллакс Луны или планеты. Возьмем за начало координат центр Земли. Ось 3 направим на полюс, ось 2 — в плоскости меридиана наблюдателя. Пусть координаты планеты равны  $x_i = r l_i$ , а наблюдателя  $\xi_i = a \lambda_i$ . Расстояние от наблюдателя до планеты, равное  $R$ , определяется соотношениями

$$R^2 = (x_i - \xi_i)^2 = r^2 - 2ra\lambda_k l_k + a^2, \quad (1)$$

$$R = r - a\lambda_k l_k \quad (2)$$

с точностью до членов первого порядка по  $a/r$ . В таком случае направляющие косинусы линии, проходящей от наблюдателя к планете  $l'_i$ , равны

$$l'_i = \frac{x_i - \xi_i}{R} = \left(l_i - \frac{a}{r}\lambda_i\right)\left(1 + \frac{a}{r}\lambda_k l_k\right) = l_i + P(-\lambda_i + l_i\lambda_k l_k), \quad (3)$$

где  $P = a/r$  — горизонтальный параллакс. Теперь выразим направляющие косинусы через угловые координаты

$$\begin{aligned} l_1 &= \cos \delta \sin h, & l_2 &= \cos \delta \cos h, & l_3 &= \sin \delta, \\ \lambda_1 &= 0, & \lambda_2 &= \cos \varphi, & \lambda_3 &= \sin \varphi, \end{aligned} \quad (4)$$

и

$$\lambda_k l_k = \cos \varphi \cos \delta \cos h + \sin \varphi \sin \delta. \quad (5)$$

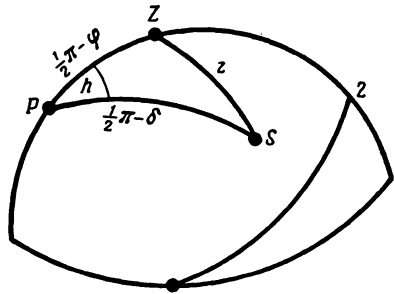


Рис. 20.

Для параллакса в склонении  $\omega_\delta$  имеем

$$l'_3 - l_3 = \cos \omega_\delta = P [-\sin \varphi + \sin \delta (\cos \varphi \cos \delta \cos h + \sin \varphi \sin \delta)] = \\ = P (\sin \delta \cos \delta \cos \varphi \cos h - \sin \varphi \cos^2 \delta), \quad (6)$$

откуда

$$\omega_\delta = P (\sin \delta \cos \varphi \cos h - \sin \varphi \cos \delta). \quad (7)$$

Если  $h'$  — наблюдаемый часовой угол, то

$$\operatorname{tg} h' = \frac{l'_1}{l'_2}, \quad \operatorname{tg} h = \frac{l_1}{l_2}, \quad \lg \operatorname{tg} h = \lg l_1 - \lg l_2, \quad (8)$$

$$\frac{\sec^2 h}{\operatorname{tg} h} (h' - h) = \frac{l'_1 - l_1}{l_1} - \frac{l'_2 - l_2}{l_2} = P \left( -\frac{\lambda_1}{l_1} + \frac{\lambda_2}{l_2} \right) = P \frac{\cos \varphi}{\cos \delta \cos h}, \quad (9)$$

$$h' - h = P \cos \varphi \sin h \sec \delta \quad (10)$$

и параллакс в прямом восхождении  $\omega_\alpha = -(h' - h)$ .

### ПРИМЕРЫ

1. Показать, что

$$(\mathbf{K} \cdot \mathbf{A})(\mathbf{K} \cdot \mathbf{B}) \times (\mathbf{K} \cdot \mathbf{C}) = \|\mathbf{K}\| \mathbf{A} \cdot \mathbf{B} \times \mathbf{C}.$$

2. Показать, что

$$\operatorname{rot}(\varphi \mathbf{A}) = \varphi \operatorname{rot} \mathbf{A} + \operatorname{grad} \varphi \times \mathbf{A}.$$

3. Показать, что

$$[\operatorname{rot}(\mathbf{A} \times \mathbf{B})]_l = A_l \operatorname{div} \mathbf{B} - B_l \operatorname{div} \mathbf{A} + B_k \frac{\partial A_l}{\partial x_k} - A_k \frac{\partial B_l}{\partial x_k}.$$

4. Если  $K_{ik}$  — тензор, то доказать, что

$$\left| \begin{matrix} K_{22}K_{23} \\ K_{32}K_{33} \end{matrix} \right| + \left| \begin{matrix} K_{33}K_{31} \\ K_{13}K_{11} \end{matrix} \right| + \left| \begin{matrix} K_{11}K_{12} \\ K_{21}K_{22} \end{matrix} \right| = \frac{1}{2} K_{ii}K_{kk} - \frac{1}{2} K_{ik}K_{ki}$$

и является скаляром. Установить связь этого скаляра с инвариантами корней уравнения  $\|K_{ik} - \lambda \delta_{ik}\| = 0$ .

5. Если тензор напряжения имеет вид  $2A_i A_k - \delta_{ik} A^2$ , где  $A_i$  — радиус-вектор,  $A^2$  обозначает  $A_1^2 + A_2^2 + A_3^2$ , а  $\delta_{ik}$  равен единице, если  $i = k$ , и равен нулю во всех остальных случаях, проверить, что в любой точке направление  $A_i$  является главной осью тензора напряжений. Показать, что напряжение в этой точке можно представить как растяжение  $A^2$  вдоль направления  $A_i$  и что давление  $A^2$  перпендикулярно этому направлению.

6. Тело деформировано внутренними напряжениями так, что частица, координаты которой в деформированном состоянии  $(x, y, z)$  относительно неподвижной прямоугольной системы координат, испытала смещение с компонентами

$$3\kappa(2x - y + z), \quad -3\kappa(x + y), \quad \kappa(3x + 5z),$$

где  $\kappa$  — малая постоянная величина. Показать, что никакая малая часть тела в окрестности точки  $(x, y, z)$  не подвергается вращению и что одно из главных растяжений равно  $3\kappa$ . Определить два других главных растяжения.

(Prelim., 1941.)

7. Цепочка из трех однородных взаимно перпендикулярных стержней, каждый массы  $m$  и длиной  $2a$ . Показать, что уравнение нормальной квадратичной формы тензора инерции относительно центра масс, отнесенное к координатным осям, параллельным этим стержням, можно записать в виде

$$\frac{2}{3} ma^2 (5x^2 + 5y^2 + 3z^2 - 3yz + 3zx + xy) = K.$$

Доказать, что один из главных моментов инерции относительно центра масс равен  $ma^2$ , и найти остальные главные моменты инерции. (Prelim., 1940.)

8. Два одинаковых однородных конуса высотой  $h$  соединены своими основаниями радиуса  $a$ . Определить тензор инерции относительно общего центра оснований и дать отношение главных моментов инерции. (I. С., 1943.)

9. Однородная твердая правильная призма массы  $M$  и длиной  $2H$ , основаниями которой служат правильные треугольники со стороной  $a$ . Определить тензор инерции этой призмы относительно центра одного из ее оснований. Определить главные моменты инерции, когда 1)  $a = H$ ; 2)  $a = 10H$ . Во втором случае показать, что момент инерции относительно любого из ребер основания равен  $\frac{83}{6} MH^2$ . (I. С., 1940.)

10. Движение относительно Земли. Радиус-вектор частицы относительно точки  $O$ , расположенной на поверхности Земли, обозначим через  $\mathbf{r}$ . Скорость и ускорение этой частицы относительно Земли в точке  $O$  равны  $\dot{\mathbf{r}}$  и  $\ddot{\mathbf{r}}$ . Если частица движется под действием земного притяжения и силы  $\mathbf{F}$ , приходящейся на каждую единицу ее массы, то доказать, что

$$\ddot{\mathbf{r}} + 2\boldsymbol{\omega} \times \dot{\mathbf{r}} = \mathbf{g}(\mathbf{r}) + \mathbf{F},$$

где  $\boldsymbol{\omega}$  — угловая скорость вращения Земли, а  $\mathbf{g}(\mathbf{r})$  — ускорение свободно падающей относительно системы координат с началом в точке  $O$  частицы, покоящейся в то же время относительно системы координат, расположенной в точке  $\mathbf{r}$ . Если частица брошена со скоростью  $\mathbf{V}$  из точки  $O$  в момент  $t=0$ , то показать, что с точностью до первого порядка по  $\boldsymbol{\omega}$  ее радиус-вектор  $\mathbf{r}$  в момент времени  $t$  определяется соотношением

$$\mathbf{r} = \mathbf{V}t + \frac{1}{2} \mathbf{g}t^2 - t^2\boldsymbol{\omega} \times \mathbf{V} - \frac{1}{3} t^3\boldsymbol{\omega} \times \mathbf{g},$$

где  $\mathbf{g} = \mathbf{g}(O)$ .

11. Маятник Фуко. Малые колебания. Начало координат возьмем в точке подвеса. Пусть  $\mathbf{i}$  — единичный вектор в направлении к грузу массы  $m$ . Тогда, если  $T$  — натяжение нити, то с точностью до первого порядка по  $\boldsymbol{\omega}$  получим

$$\ddot{\mathbf{r}} + 2\boldsymbol{\omega} \times \dot{\mathbf{r}} = \mathbf{g} - \frac{T}{m} \mathbf{i}.$$

Полагая  $\mathbf{g} = g\mathbf{z}$ , а  $\mathbf{i} = \mathbf{z} + \boldsymbol{\rho}$ , показать, что с точностью до первого порядка по  $\boldsymbol{\rho}$  будем иметь

$$\ddot{\boldsymbol{\rho}} + 2\boldsymbol{\omega} \sin \lambda \mathbf{z} \times \dot{\boldsymbol{\rho}} + \frac{g}{l} \boldsymbol{\rho} = 0,$$

где  $\lambda$  — широта точки  $O$ , а  $l$  — длина нити.

Взяв компоненты этого уравнения для двух взаимно перпендикулярных направлений в плоскости, перпендикулярной  $\mathbf{z}$ , и воспользовавшись методом 2.12, показать, что плоскость колебаний поворачивается вокруг вертикали, если смотреть на нее сверху вниз, с угловой скоростью  $\boldsymbol{\omega} \sin \lambda$ .

## ЛИТЕРАТУРА

1. *Jeffreys H.*, Cartesian Tensors, ch. 7.
2. *Murnaghan F. D.*, Amer. J. Math., **59**, 235—260 (1937).
3. *Abraham G.*, *Becker A.*, Classical Electricity and Magnetism, p. 104, 146.  
(Русский перевод с оригинального немецкого издания: *Абрагам Г.*, Электродинамика. М.—Л., ОНТИ, 1937; *Беккер*, Электронная теория, М.—Л., ОНТИ, 1938.)
4. *Eddington A. S.*, Mathematical Theory of Relativity, Cambridge, 1923.
5. *McConnell A. J.*, Absolute Differential Calculus.

## ПРИЛОЖЕНИЕ К ГЛАВЕ 3

**3.03а.** Утверждение, что только изотропные тензоры второго и третьего порядка равны скалярам, умноженным на  $\delta_{ik}$  и  $\epsilon_{ikm}$  соответственно, легко доказать методом **3.031**.



## МАТРИЦЫ

Все, что не Белгрейвская площадь,  
есть Стренд и Пикадили \*).

*W. S. Gilbert „Utopia Limited“*

**4.01. Введение: определения.** При рассмотрении тензоров второго порядка в трех измерениях мы использовали сокращенное обозначение  $K_{ik}$  для совокупности девяти величин

$$\begin{pmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{pmatrix}.$$

Если они расположены в указанном порядке, то первый индекс относится к строке, а второй — к столбцу. Такая совокупность образует тензор, если каждый индекс относится к одной из осей выбранной системы координат и при повороте осей эти величины преобразуются определенным образом. До сих пор мы рассматривали только прямоугольные оси, однако это частный случай более общей тензорной алгебры. Обобщение идет по трем направлениям: а) на тензоры любого порядка в  $n$  измерениях (особенно в четырех), отнесенные по-прежнему к прямоугольным осям; б) на тензоры любого порядка в криволинейных координатах; в) на изучение алгебры квадратных таблиц величин, которые формально имеют много общего с тензорами второго порядка, но не ограничены в интерпретации связью с системой координат.

Принимая алгебраическую точку зрения, мы говорим о таблице величин

$$\begin{pmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ K_{m1} & K_{m2} & \dots & K_{mn} \end{pmatrix}$$

---

\*) Англичанину это говорило бы: „Все, что не строчка или не столбец, есть квадратная матрица“. — *Прим. ред.*

как о матрице порядка \*)  $m \times n$ . Она имеет  $m$  строк и  $n$  столбцов. Мы можем записать ее как  $(K_{ik})$ , причем первый индекс указывает строку, а второй — столбец. Квадратная матрица порядка  $n \times n$  является частным случаем. Другими частными случаями являются матрица с единственным столбцом и  $m$  строками (порядка  $m \times 1$ ) и матрица с единственной строкой и  $n$  столбцами (порядка  $1 \times n$ ). Эти три типа матриц имеют много приложений в физике.

Мы будем обозначать матрицу единым символом, который заменяет систему  $m \times n$  величин, или *элементов* матрицы. Для такого объекта, как матрица, все, что подразумевается под сложением, вычитанием, умножением и делением, есть дело определения. Использование жирного шрифта указывает на то, что один или более индексов опущены, как в векторных или диадных обозначениях.

**Сложение.** Сумма двух матриц **a** и **b** записывается как **a + b** и обозначает матрицу с элементами  $a_{ik} + b_{ik}$ .

**Вычитание.** Матрица **-a** определяется как матрица с элементами  $(-a_{ik})$ , а **a - b** определяется как матрица с элементами  $a_{ik} - b_{ik}$ . Чтобы сложение и вычитание имели смысл, матрицы должны иметь одинаковое число строк и одинаковое число столбцов.

Очевидно, что при сложении матриц справедливы как ассоциативный закон

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (1)$$

так и коммутативный закон

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}. \quad (2)$$

**Умножение.** Закон умножения состоит в том, что **ab** есть матрица с элементами  $(\mathbf{ab})_{ik}$ , определяемыми равенством

$$(\mathbf{ab})_{ik} = a_{ij}b_{jk}, \quad (3)$$

где подразумевается соглашение о суммировании. Чтобы это действие имело смысл,  $j$  должно пробегать одинаковое число значений в обоих сомножителях и поэтому число столбцов в **a** должно быть равно числу строк в **b**. Тогда произведение есть матрица с числом строк, как у **a**, и с числом столбцов, как у **b**. В частности, если **a** — матрица  $n \times n$ , то **b** должна иметь  $n$  строк, хотя может быть матрицей-столбцом, и тогда произ-

---

\*) Таким образом, о приведенном выше тензоре порядка 2 следует говорить как о матрице порядка  $3 \times 3$ . Слово „порядок“ имеет довольно различные значения применительно к тензорам и матрицам. Тензор порядка выше 2 не может быть записан в виде матрицы.

ведение будет также матрицей-столбцом. Если  $\mathbf{b}$  — также матрица  $n \times n$ , то  $\mathbf{ab}$  будет матрицей  $n \times n$ . С другой стороны, если  $\mathbf{a}$  — матрица-столбец,  $\mathbf{b}$  должна быть матрицей-строкой, и  $\mathbf{ab}$  — матрица, чьи элементы равны  $a_i b_k$ . Если  $\mathbf{a}$  есть матрица-строка ( $1 \times n$ ), то  $\mathbf{b}$  должна иметь  $n$  строк. Если  $\mathbf{b}$  — квадратная матрица, тогда произведение — матрица-строка; если  $\mathbf{b}$  имеет только один столбец, произведение имеет одну строку и один столбец, т. е. является скалярной величиной. Таким образом, для нас важны следующие случаи:

$\mathbf{a}$		$\mathbf{b}$		$\mathbf{ab}$		В индексах
Строки	Столбцы	Строки	Столбцы	Строки	Столбцы	
$n$	$n$	$n$	$n$	$n$	$n$	$a_{ij}b_{jk}$
$n$	$n$	$n$	1	$n$	1	$a_{ij}b_j$
$n$	1	1	$n$	$n$	$n$	$a_jb_k$
1	$n$	$n$	$n$	1	$n$	$a_jb_{jk}$
1	$n$	$n$	1	1	1	$a_jb_j$

где один индекс использован для матрицы с одной строкой или столбцом и два индекса — для квадратных матриц. В третьем случае не применяются соглашения о суммировании, и он называется *внешним произведением*  $\mathbf{a}$  на  $\mathbf{b}$ . Остальные случаи называются *внутренними произведениями*.

В работах по алгебре матрицы-строки и матрицы-столбцы часто называются векторами. В отличие от использования слова *вектор* в физике в алгебре вектор — это совокупность  $n$  элементов, не имеющих отношения к какому-либо закону преобразования. В физике в определение вектора входят как элементы, так и предписание определяющего их некоторого закона преобразования. Таким образом, в физике мы говорим о векторе, имеющем различные компоненты в различных системах координат, как о едином объекте; алгебраисты называли бы эти представления различными векторами. Мы будем избегать такой терминологии.

Умножение не подчиняется, вообще говоря, коммутативному закону, даже если  $\mathbf{a}$  и  $\mathbf{b}$  являются квадратными матрицами, поскольку  $\mathbf{ba}$  должно быть определено как матрица с элементами  $b_{ij}a_{jk}$ , а они будут равны  $a_{ij}b_{jk}$  только в особых случаях. В общем случае

$$\mathbf{ab} \neq \mathbf{ba}.$$

Говорят, что пары матриц, для которых удовлетворяется равенство  $\mathbf{ab} = \mathbf{ba}$ , *коммутируют*, а те, для которых удовлетворяется равенство  $\mathbf{ab} = -\mathbf{ba}$ , *антикоммутируют*.

По правилу умножения множители всегда располагаются так, что повторяющиеся индексы *соседствуют*. Справедливо, что  $a_{ij}b_{jk} = b_{jk}a_{ij}$ , но последнее нельзя свести к  $\mathbf{ba}$ , потому что индексы  $j$  не являются соседними. Подробная запись индексов не может привести к путанице, однако при опускании индек-

сов путаница будет, если не иметь определенного правила о том, где должны быть повторяющиеся индексы. Это правило достаточно, чтобы отличить  $\mathbf{ba}$  от  $\mathbf{ab}$  при сохранении порядка множителей в явном выражении для произведения.

Ассоциативный закон

$$(\mathbf{ab})\mathbf{c} = \mathbf{a}(\mathbf{bc}) \quad (4)$$

и дистрибутивный закон

$$\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac} \quad (5)$$

справедливы при условии, что сохранен порядок сомножителей и операции имеют смысл. Эти законы легко проверяются явным выписыванием элементов. Для первого закона

$$[(\mathbf{ab})\mathbf{c}]_{il} = (a_{ij}b_{jk})c_{kl} = a_{ij}(b_{jk}c_{kl}) = [\mathbf{a}(\mathbf{bc})]_{il}. \quad (6)$$

Следовательно, это произведение можно записать без скобок в виде  $\mathbf{abc}$ , так как положение скобок не существенно. Отсюда следует, что все положительные степени данной матрицы коммутируют, ибо  $\mathbf{a}^2\mathbf{a} = \mathbf{aa}^2$ , и по индукции  $\mathbf{a}^m\mathbf{a}^n = \mathbf{a}^n\mathbf{a}^m$ .

Единичную матрицу будем обозначать  $\mathbf{I}^*$ . Ее компоненты равны  $\delta_{ik}$ , где

$$\delta_{ik} = 1 \quad (i = k), \quad \delta_{ik} = 0 \quad (i \neq k). \quad (7)$$

Очевидно, что

$$\mathbf{Ia} = \mathbf{aI} = \mathbf{a}. \quad (8)$$

Единичная матрица часто обозначается просто  $\mathbf{I}$ .

Нулевая матрица — это матрица, все элементы которой нули. Произведение двух матриц может быть нулевой матрицей, если каждый из сомножителей не является нулевой матрицей. Так, например,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ненулевой элемент первой матрицы умножается только на элементы первой строки второй матрицы, которые оба равны нулю. Но если  $\mathbf{AB} = 0$  при любой  $\mathbf{B}$ , то  $\mathbf{A} = 0$ . Аналогично если  $\mathbf{AB} = 0$  при любой  $\mathbf{A}$ , то  $\mathbf{B} = 0$ .

Для матрицы  $\mathbf{a}$  транспонированная матрица есть матрица, образованная из  $\mathbf{a}$  заменой строк на столбцы. Мы будем обозначать ее  $\tilde{\mathbf{a}}$ , а ее элементы  $\tilde{a}_{ik}$ . Тогда

$$\tilde{a}_{ik} = a_{ki}. \quad (9)$$

---

\*) Мы вводим несколько необычное обозначение единичной матрицы  $\mathbf{I}$ , чтобы отличить ее от обозначения  $\mathbf{I}$ . — Прим. ред.

Поскольку

$$(\mathbf{ab})_{ik} = a_{ij}b_{jk} = \tilde{a}_{ji}\tilde{b}_{kj} = \tilde{b}_{kj}\tilde{a}_{ji} = (\mathbf{ba})_{ki}, \quad (10)$$

то, следовательно, транспонированная матрица произведения  $\mathbf{ab}$ , обозначаемая  $\tilde{\mathbf{ab}}$ , равна произведению  $\tilde{\mathbf{ba}}$  в указанном порядке.

Заметим, что *можно* было бы определить  $(\mathbf{ab})_{ik}$  как  $a_{ij}b_{kj}$ ; такая матрица действительно часто встречается. Но тогда бы мы имели

$$(\mathbf{ab} \cdot \mathbf{c})_{ik} = (\mathbf{ab})_{ij}c_{kj} = a_{il}b_{jl}c_{kj}, \quad (\mathbf{a} \cdot \mathbf{bc})_{ik} = a_{il}(\mathbf{bc})_{kl} = a_{il}b_{kj}c_{lj},$$

что не то же самое. *Правило суммирования по соседним индексам необходимо для того, чтобы ассоциативный закон умножения оставался справедливым.* Поэтому мы пишем

$$a_{ij}b_{kj} \text{ не как } (\mathbf{ab})_{ik}, \text{ а как } a_{ij}\tilde{b}_{jk} = (\tilde{\mathbf{a}}\mathbf{b})_{ik}.$$

Начинающие иногда находят умножение матриц более простым, если они сначала транспонируют вторую матрицу и умножают строку на строку.

Если  $\mathbf{a}$  есть матрица-столбец, то  $\tilde{\mathbf{a}}$  — матрица-строка, и наоборот. Если  $\mathbf{b}$  — квадратная матрица, то имеем

$$(\mathbf{ba})_i = b_{ij}a_j = a_jb_{ji} = \tilde{a}_j\tilde{b}_{ji} = (\tilde{\mathbf{a}}\tilde{\mathbf{b}})_i, \quad (11)$$

т. е. применимо то же правило, что и для квадратных матриц.

*Комплексно сопряженная матрица.* Элементы матрицы могут быть действительными или комплексными. Матрица, элементы которой являются комплексно сопряженными для элементов матрицы  $\mathbf{a}$ , обозначается  $\mathbf{a}^*$ , а ее элементы  $a_{ik}^*$ .

*Транспонированная матрица комплексно сопряженной матрицы*  $\tilde{\mathbf{a}}^*$  обозначается  $\mathbf{a}^+$ , а ее элементы равны

$$a_{ik}^+ = a_{ki}^*. \quad (12)$$

*Свойства симметрии.* Говорят, что матрица является *симметричной*, если она не изменяется при замене строк столбцами, т. е.

$$a_{ik} = a_{ki}, \quad (13)$$

или

$$\mathbf{a} = \tilde{\mathbf{a}}. \quad (14)$$

Матрица является *антисимметричной*, или *кососимметричной*, если она меняет знак при замене строк на столбцы, т. е.

$$a_{ik} = -a_{ki}, \quad (15)$$

$$\mathbf{a} = -\tilde{\mathbf{a}}. \quad (16)$$

О комплексной матрице говорят, что она *эрмитова*, если она равна своей транспонированной комплексно сопряженной, т. е.

$$\mathbf{a} = \mathbf{a}^+, \quad (17)$$

и *антиэрмитова*, если

$$\mathbf{a} = -\mathbf{a}^+. \quad (18)$$

Для действительных матриц равенства (17) и (18) сводятся к (14) и (15) соответственно. Если  $\mathbf{a}$  эрмитова, то  $i\mathbf{a}$  антиэрмитова.

*Диагональной* является матрица, все элементы которой равны нулю, за исключением расположенных на главной диагонали, т. е.  $a_{11}, a_{22}, \dots, a_{nn}$ . Все пары диагональных матриц коммутативны.

*Присоединенная и обратная матрицы.* Если мы рассмотрим определитель, образованный из элементов квадратной матрицы  $\mathbf{a}$

$$\|a_{ik}\| \equiv \det \mathbf{a}, \quad (19)$$

то каждый элемент  $a_{ik}$  имеет алгебраическое дополнение, которое мы будем обозначать  $A_{ik}$ , причем

$$a_{ik}A_{jk} = (\det \mathbf{a})\delta_{ij}. \quad (20)$$

Для сохранения правила о суммировании по соседним индексам мы перепишем (20), образовав *присоединенную* матрицу  $\mathbf{a}^*$ :

$$(\text{adj } \mathbf{a})_{ik} = A_{ki}. \quad (21)$$

Если  $\mathbf{a}$  симметрична или эрмитова, то такова же и  $\text{adj } \mathbf{a}$ . Далее, при условии, что  $\det \mathbf{a}$  не равен нулю, имеем

$$\frac{a_{ij}A_{kj}}{\det \mathbf{a}} = \delta_{ik} = \frac{A_{ji}a_{jk}}{\det \mathbf{a}}. \quad (22)$$

Поэтому если определим *обратную матрицу*  $\mathbf{a}^{-1}$  матрицы  $\mathbf{a}$  посредством

$$(\mathbf{a}^{-1})_{ik} = \frac{A_{ki}}{\det \mathbf{a}}, \quad (23)$$

то будем иметь

$$a_{ij}(\mathbf{a}^{-1})_{jk} = \delta_{ik} = (\mathbf{a}^{-1})_{ij}a_{jk} \quad (24)$$

или

$$\mathbf{a}\mathbf{a}^{-1} = \mathbf{a}^{-1}\mathbf{a} = \mathbf{1}. \quad (25)$$

Произведение матрицы на свою обратную есть единичная матрица независимо от порядка выполнения умножения; при условии, что обратная матрица существует, порядок не имеет значения. Если, однако,  $\det \mathbf{a}$  равен нулю, то  $\mathbf{a}$  не имеет обратной матрицы и говорят, что она *особая*.

**Деление.** Деление на неособую матрицу можно теперь определить как умножение на обратную ей, однако частное зависит от порядка, как и произведение. Значит,  $\mathbf{a}^{-1}\mathbf{b}$  не то же самое, что  $\mathbf{ba}^{-1}$ .

$$\mathbf{a} \mathbf{b} \mathbf{b}^{-1} \mathbf{a}^{-1} = \mathbf{a} \mathbf{I} \mathbf{a}^{-1} = \mathbf{I}, \quad (26)$$
$$\begin{aligned} \mathbf{a}_{ij} \mathbf{b}_{jk} (\text{adj } \mathbf{b})_{kl} (\text{adj } \mathbf{a})_{lm} &= a_{ij} (\det \mathbf{b}) \delta_{il} (\text{adj } \mathbf{a})_{lm} = \\ &= a_{ij} (\text{adj } \mathbf{a})_{jm} (\det \mathbf{b}) = (\det \mathbf{a}) (\det \mathbf{b}) \delta_{im}, \\ (\mathbf{a} \mathbf{b})_{ik} (\mathbf{b}^{-1} \mathbf{a}^{-1})_{km} &= \mathbf{I}_{im}. \end{aligned} \quad (27)$$
$$\mathbf{a}\mathbf{a}^\dagger = \mathbf{a}^\dagger\mathbf{a} = \mathbf{1}, \quad (28)$$
$$\mathbf{a}^{-1} \mathbf{a} \mathbf{a}^{\dagger} = \mathbf{a}^{-1} \mathbf{I}, \text{ откуда } \mathbf{a}^{\dagger} = \mathbf{a}^{-1}, \quad (29)$$
$$\tilde{\mathbf{a}} = \mathbf{a}^{-1}, \quad (30)$$

Если унитарная матрица к тому же эрмитова, то  $\mathbf{a}^+ = \mathbf{a}$  и

$$\mathbf{a} = \mathbf{a}^{-1}, \quad (31)$$

$$\mathbf{a}\mathbf{a} = \mathbf{a}^2 = \mathbf{1}. \quad (32)$$
[illegible]

может быть записана в сокращенном виде

$$a_{ij}x_j = y_i, \quad (2)$$

где  $i$  и  $j$  пробегает значения от 1 до  $n$ . Если  $x_i$  мыслить как матрицу-столбец, то (2) можно записать в виде

$$ax = y, \quad (3)$$

где  $y$  есть также матрица-столбец. Мы предположим здесь, что  $a$  — неособая матрица и обозначим алгебраическое дополнение  $a_{ik}$  через  $A_{ik}$ .

Теперь, поскольку

$$A_{ik}a_{ij} = (\det a) \delta_{kj}, \quad (4)$$

можно умножить  $i$ -е уравнение (2) на  $A_{ik}$  и просуммировать. Сумма будет равна

$$(\det a) \delta_{kj} x_j = A_{ik} y_i, \quad (5)$$

т. е., если  $\det a \neq 0$ ,

$$x_k = \frac{A_{ik}}{\det a} y_i = (a^{-1})_{ki} y_i \quad (6)$$

или

$$x = a^{-1}y, \quad (7)$$

что можно было бы получить из (3) непосредственно, умножая обе стороны слева на  $a^{-1}$ .

Каждое из выражений (5–7) дает решение уравнений (1) в компактном виде\*).

Если матрица  $a$  унитарная, то (6) и (7) принимают соответственно вид

$$x_k = a_{ki}^+ y_i, \quad (8)$$

$$x = a^+ y. \quad (9)$$

Если  $a$  ортогональная, то

$$x_k = a_{ik} y_i, \quad (10)$$

$$x = \tilde{a} y. \quad (11)$$

**4.021. Умножение определителей.** Мы предполагаем, что читатель уже знаком с правилом перемножения определителей, однако следующее доказательство представляет интерес, потому что оно непосредственно связывает это правило с существованием решений однородных линейных уравнений.

---

\*) На практике численное решение этим методом не является наиболее простым. Практически удобный метод мы укажем в 9.16.



Пусть  $a_{ij}$ ,  $b_{ij}$  — элементы двух определителей. Тогда условие  $\|a_{ij}\| = 0$  является необходимым и достаточным для того, чтобы уравнения

$$a_{ij}x_j = 0$$

имели решение  $x_j$ , отличное от нулевого. Умножим их на  $b_{ik}$  и просуммируем. Тогда уравнения

$$b_{ik}a_{ij}x_j = 0$$

имеют решение, отличное от нулевого, и поэтому если

$$\|a_{ij}\| = 0, \text{ то и } \|b_{ik}a_{ij}\| = 0.$$

Аналогично если  $\|b_{ij}\| = 0$ , то

$$\|a_{ik}b_{ij}\| = 0 \quad \text{и} \quad \|b_{ik}a_{ij}\| = \|a_{ik}b_{ij}\|,$$

так как левая часть этого равенства получается транспонированием правой. Следовательно,  $\|b_{ik}a_{ij}\|$  содержит  $\|a_{ij}\|$  и  $\|b_{ik}\|$  в качестве сомножителей, а сравнивая коэффициенты при  $a_{11}b_{11}a_{22}b_{22} \dots a_{nn}b_{nn}$ , видим, что коэффициент пропорциональности равен 1.

В матричных обозначениях

$$a_{ij}b_{ik} = \tilde{a}_{ji}b_{ik} = (\tilde{\mathbf{a}}\mathbf{b})_{jk} \quad \text{и} \quad \det(\tilde{\mathbf{a}}\mathbf{b}) = \det \tilde{\mathbf{a}} \det \mathbf{b} = \det \mathbf{a} \det \mathbf{b}.$$

**4.03. Преобразования.** Если имеем систему соотношений

$$\mathbf{y} = \mathbf{a}\mathbf{x}, \tag{1}$$

где  $\mathbf{a}$  — неособая квадратная матрица, то можно заключить, как в предыдущем разделе, что

$$\mathbf{x} = \mathbf{a}^{-1}\mathbf{y}. \tag{2}$$

Если  $x_i$  являются переменными, то можно сказать, что матрица  $\mathbf{a}$  осуществляет преобразование переменных. В частности, важны случаи, когда  $\mathbf{a}$  ортогональная или симметричная.

**4.031. Ортогональные преобразования \*).** Для преобразования прямоугольных осей мы имеем правила

$$x'_j = l_{ij}x_i \tag{1}$$

и

$$x_i = l_{ij}x'_j. \tag{2}$$

---

\*) См. замечание в конце гл. 4, стр. 276.

Последнее выражение уже имеет вид матричного произведения, причем  $x_i$  и  $x'_j$  являются матрицами-столбцами. Поэтому можно записать

$$\mathbf{x} = \mathbf{l}\mathbf{x}'. \quad (3)$$

Но тогда можно предположить, что эти уравнения разрешены относительно  $x'_j$ , а это решение имеет вид

$$\mathbf{x}' = \mathbf{l}^{-1}\mathbf{x} \quad (4)$$

или

$$x'_j = (\mathbf{l}^{-1})_{ji} x_i. \quad (5)$$

Сравнивая это с (1) и замечая, что эти выражения должны быть эквивалентны для всех значений  $x_i$ , имеем

$$(\mathbf{l}^{-1})_{ji} = l_{ij}, \quad (6)$$

и поэтому  $\mathbf{l}$  — ортогональная матрица. Так как

$$\mathbf{l}^{-1} = \tilde{\mathbf{l}}, \quad (7)$$

(1) можно переписать также в виде

$$\tilde{\mathbf{x}}' = \tilde{\mathbf{x}}\tilde{\mathbf{l}}. \quad (8)$$

Мы должны писать  $\tilde{\mathbf{x}}$ , а не  $\mathbf{x}$ , потому что матрица-столбец не может стоять слева; поэтому нужно взять транспонированную от  $\mathbf{x}$ , что дает матрицу-строку. Таким образом, поворот осей можно представить как ортогональное преобразование с единичным определителем. Далее,

$$\tilde{\mathbf{x}}\mathbf{x} = x_i x_i \quad (9)$$

и

$$x'_j x'_j = \tilde{\mathbf{x}}\mathbf{l}\mathbf{l}^{-1}\mathbf{x} = \tilde{\mathbf{x}}\mathbf{l}\mathbf{x} = \tilde{\mathbf{x}}\mathbf{x} = x_i x_i. \quad (10)$$

Это подтверждает, что форма  $x_i x_i$  является инвариантной при ортогональных преобразованиях.

Обратно, если

$$x_i = l_{ij} x'_j \quad (11)$$

и

$$x_i x_i = x'_j x'_j \quad (12)$$

для всех значений  $x'_j$ , то имеем

$$(l_{ij} x'_j)(l_{il} x'_l) = \delta_{jl} x'_j x'_l, \quad (13)$$

и поэтому, поскольку  $l_{ij} l_{il}$  — симметричная матрица,

$$l_{ij} l_{il} = \delta_{jl}, \quad \mathbf{l}\tilde{\mathbf{l}} = \mathbf{I}. \quad (14)$$

Следовательно,  $\mathbf{l}$  — ортогональная матрица.

В свете этого (6) представляет  $n^2$  соотношений между  $n^2$  неизвестными, и мы можем ожидать, что существует лишь конечное число решений и они даже могут быть комплексными. Но каково бы ни было  $l$ , матрица  $l_{ij}l'_{ij}$  симметрична по  $j$  и  $l$ , и поэтому для ортогональности достаточно только  $\frac{1}{2}n(n+1)$  независимых условий.

Таким образом,  $l_{ij}$  можно, по-видимому, подчинить  $\frac{1}{2}n(n-1)$  дополнительным условиям, и это число действительно оказывается правильным.

Допустим теперь, что мы сохраняем оси  $x_i$  неподвижными и вращаем некоторое жесткое тело вокруг начала координат. Вообразим систему осей  $x'_j$ , первоначально совпадающую с осями  $x_i$  и фиксированную в теле, а потому вращающуюся вместе с ним. Тогда координаты любой частицы тела по отношению к осям  $x'_j$  не изменятся. Нам требуется найти новые координаты этой частицы  $y_i$  по отношению к осям  $x_i$ . Если направляющие косинусы по-прежнему обозначить  $l_{ij}$ , то будем иметь

$$y_i = l_{ij}x'_j = l_{ij}x_j, \quad y = lx, \quad (15)$$

и аналогично любой вектор  $u$  при вращении переходит в  $v$ , где

$$v = lu. \quad (16)$$

**4.032. Однопараметрическое представление ортогональной матрицы порядка  $2 \times 2$ .** Пусть теперь  $l$  обозначает ортогональную матрицу

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \quad (17)$$

В силу ортогональности

$$l\tilde{l} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \begin{pmatrix} \alpha^2 + \beta^2 & \alpha\gamma + \beta\delta \\ \gamma\alpha + \delta\beta & \gamma^2 + \delta^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (18)$$

Тогда мы можем выбрать такое  $\lambda$ , что

$$\alpha = \cos \lambda, \quad \beta = \sin \lambda, \quad (19)$$

и затем

$$\frac{\gamma}{\delta} = -\frac{\beta}{\alpha} = -\operatorname{tg} \lambda, \quad (20)$$

что вместе с

$$\gamma^2 + \delta^2 = 1 \quad (21)$$

дает

$$\gamma = -\sin \lambda, \quad \delta = \cos \lambda, \quad (22)$$

или

$$\gamma = \sin \lambda, \quad \delta = -\cos \lambda. \quad (23)$$

Равенства (22) дают возможность получить представление

$$I = \begin{pmatrix} \cos \lambda & \sin \lambda \\ -\sin \lambda & \cos \lambda \end{pmatrix}, \quad (24)$$

которое содержит лишь одну константу. При  $\lambda = 0$  получаем  $I = I$ . Равенства (23) дают

$$I = \begin{pmatrix} \cos \lambda & \sin \lambda \\ \sin \lambda & -\cos \lambda \end{pmatrix}. \quad (25)$$

При  $\lambda = 0$  получим

$$I = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Если рассматривать  $x_1, x_2$  как прямоугольные координаты в двух измерениях и подставить (24) в (15), то получим поворот на угол  $-\lambda$ . Соотношения (23) при  $\lambda = 0$  оставляют  $x_1$  прежним, но меняют знак  $x_2$ . В общем случае (25) представляет *отражение*.

**4.033. Общее ортогональное преобразование.** В  $n$  измерениях, если взять

$$I = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 & 0 & \dots \\ -\sin \alpha & \cos \alpha & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \quad \tilde{I} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 & 0 & \dots \\ \sin \alpha & \cos \alpha & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix},$$

то сразу обнаружим, что  $\tilde{I}I = I$  и  $I$  ортогональна с детерминантом, равным  $+1$ . Это преобразование можно рассматривать как вращение в плоскости осей  $x_1$  и  $x_2$  при неизменных координатах остальных  $n-2$  измерений. Мы не можем, как в трех измерениях, говорить о вращении *относительно* какой-либо оси, но можем продолжать говорить, что вращение параллельно плоскости. Такое вращение, произвольное по величине, можно осуществить для любой плоскости, содержащей две оси; следо-

вательно, в  $n$  измерениях возможны  $\frac{1}{2}n(n-1)$  независимых вращений.

Антисимметричная матрица имеет то же число независимых компонент, и можно было бы показать, что элементы ортогональной матрицы всегда выражаются через элементы антисимметричной, однако доказательство довольно длинное \*).

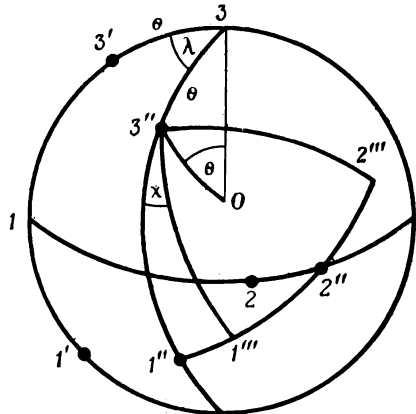
#### 4.034. Произвольный поворот абсолютно твердого тела.

В трех измерениях матрица, осуществляющая поворот вправо на угол  $\alpha$  относительно оси с направляющими косинусами  $n_i$  в силу 3.09 есть

$$\begin{pmatrix} \cos \alpha + n_1^2 (1 - \cos \alpha) & n_1 n_2 (1 - \cos \alpha) - n_3 \sin \alpha & n_1 n_3 (1 - \cos \alpha) + n_2 \sin \alpha \\ n_1 n_2 (1 - \cos \alpha) + n_3 \sin \alpha & \cos \alpha + n_2^2 (1 - \cos \alpha) & n_2 n_3 (1 - \cos \alpha) - n_1 \sin \alpha \\ n_3 n_1 (1 - \cos \alpha) - n_2 \sin \alpha & n_3 n_2 (1 - \cos \alpha) + n_1 \sin \alpha & \cos \alpha + n_3^2 (1 - \cos \alpha) \end{pmatrix}.$$

Пусть  $O123$  будет система осей, а  $x_i$  — координаты точки тела относительно этой системы. Пусть тело будет повернуто в положение, характеризуемое углами Эйлера, которые мы будем обозначать  $\theta$ ,  $\lambda$ ,  $\chi$ . Нам нужно узнать окончательное положение частицы, которая до вращения имела координаты  $x_i$ . Допустим, что тело сначала повернулось вправо на угол  $\theta$  относительно  $O2$ . Частицы, первоначально находившиеся на осях, займут положение  $1'2'3'$ . Матрица поворота есть

$$I_1 = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix},$$



Р и с. 21.

и частица из положения  $x$  переместится в  $y = I_1 x$ . Осуществим теперь поворот на  $\lambda$  вокруг  $O3$  (а не  $O3'$ ). Частицы из  $1'2'3'$  переместятся в  $1''2''3''$ , и произвольная частица займет

\*) См. ([1], стр. 162—167) или пример 10, стр. 275.

положение  $\mathbf{z}$ , где

$$\mathbf{z} = \mathbf{I}_2 \mathbf{I}_1 \mathbf{x}, \quad \mathbf{I}_2 = \begin{pmatrix} \cos \lambda & -\sin \lambda & 0 \\ \sin \lambda & \cos \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{I}_2 \mathbf{I}_1 = \begin{pmatrix} \cos \lambda \cos \theta & -\sin \lambda & \cos \lambda \sin \theta \\ \sin \lambda \cos \theta & \cos \lambda & \sin \lambda \sin \theta \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}.$$

Эта важная матрица дает координаты частицы, положение которой относительно осей  $O1''2''3''$  известно, и ее можно считать основой большинства формул сферической астрономии. Однако для абсолютно твердого тела положение частиц не определяется без третьего угла;  $\theta$  и  $\lambda$  могут быть выбраны для установления положения линии частиц в теле, но частицы, находившиеся на  $O1$ , не обязательно окажутся на  $O1''$ . Поэтому мы должны рассмотреть еще поворот на  $\chi$  вокруг  $O3''$ . Можно попытаться найти матрицу этого поворота и затем умножить ее на  $\mathbf{I}_2 \mathbf{I}_1$ . Но это сложный путь и его можно избежать. Смещение от  $O123$  до  $O1'''2'''3'''$  есть смещение абсолютно твердого тела. Если бы повороты  $\lambda$  и  $\theta$  не были произведены, плоскость  $3''1'''$  была бы все равно наклонена под углом  $\chi$  к плоскости  $31$ . Следовательно, позволительно осуществить *сначала* поворот на  $\chi$  вокруг  $O3$ , а затем сделать поворот  $\mathbf{I}_2 \mathbf{I}_1$ . Матрица первого поворота есть

$$\mathbf{I}_3 = \begin{pmatrix} \cos \chi & -\sin \chi & 0 \\ \sin \chi & \cos \chi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

и окончательный поворот задается матрицей

$$\mathbf{I}_2 \mathbf{I}_1 \mathbf{I}_3 =$$

$$= \begin{pmatrix} \cos \lambda \cos \theta \cos \chi - \sin \lambda \sin \chi & -\cos \lambda \cos \theta \sin \chi - \sin \lambda \cos \chi & \cos \lambda \sin \theta \\ \sin \lambda \cos \theta \cos \chi + \cos \lambda \sin \chi & -\sin \lambda \cos \theta \sin \chi + \cos \lambda \cos \chi & \sin \lambda \sin \theta \\ -\sin \theta \cos \chi & \sin \theta \sin \chi & \cos \theta \end{pmatrix}.$$

Так, например, точка, первоначально занимавшая положение  $(a, 0, 0)$ , окажется в положении

$$a (\cos \lambda \cos \theta \cos \chi - \sin \lambda \sin \chi, \sin \lambda \cos \theta \cos \chi + \cos \lambda \sin \chi, -\sin \theta \cos \chi).$$

**4.04. Симметричные матрицы.** Мы уже знаем, что действительная симметричная матрица порядка  $3 \times 3$  имеет три взаимно-ортогональные главные оси. Если эти оси взять в качестве

системы отсчета  $x'_j$ , то их направляющие косинусы  $l_{ij}$  относительно первоначальных осей  $x_i$  уже найдены в 3.08, причем

$$x_i = l_{ij}x'_j, \quad \mathbf{x} = \mathbf{l}\mathbf{x}'. \quad (1)$$

Тогда

$$K_{ik}x_ix_k = l_{ij}l_{kl}K_{ik}x'_jx'_l = K'_{jl}x'_jx'_k, \quad (2)$$

где

$$K'_{jl} = l_{ij}l_{kl}K_{ik} = \tilde{l}_{ji}K_{ik}l_{kl}, \quad (3)$$

$$\mathbf{K}' = \tilde{\mathbf{l}}\mathbf{K}\mathbf{l}. \quad (4)$$

Но, сравнивая с 3.08 (12), видим, что если  $j$  и  $l$  различны, то  $K'_{jl} = 0$ , а если они одинаковы и равны, скажем 1, то

$$K'_{11} = \lambda_1 l_{i1}l_{i1} = \lambda_1. \quad (5)$$

Таким образом,

$$\mathbf{K}' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad (6)$$

и преобразование (3) привело  $\mathbf{K}$  к диагональному виду.

Этот результат можно обобщить. Действительную симметричную матрицу любого порядка можно привести к диагональной форме ортогональным преобразованием вида (4). Дальнейшее обобщение состоит в том, что эрмитова матрица любого порядка может быть приведена к диагональной форме унитарным преобразованием вида  $\mathbf{K}' = \mathbf{I}^\dagger \mathbf{K} \mathbf{I}$ .

При рассмотрении произвольного движения жидкости мы видели, что часть движения в малой окрестности точки представляется симметричной матрицей и что эта часть характеризует изменение расстояний между частицами. В связи с этим рассмотрим преобразование, осуществляемое симметричной матрицей  $\mathbf{K}$ . Возьмем

$$y_i = K_{ij}x_j = x_j K_{ji}, \quad (7)$$

$$\mathbf{y} = \mathbf{K}\mathbf{x}, \quad \tilde{\mathbf{y}} = \tilde{\mathbf{x}}\mathbf{K}. \quad (8)$$

Тогда

$$\tilde{\mathbf{y}}\mathbf{y} = \tilde{\mathbf{x}}\mathbf{K}\mathbf{K}\mathbf{x}. \quad (9)$$

Для симметричной матрицы произведение  $\mathbf{K}\mathbf{K}$ , вообще говоря, не равно  $\mathbf{I}$ , и

$$\tilde{\mathbf{y}}\mathbf{y} \neq \tilde{\mathbf{x}}\mathbf{x}.$$

Однако в главных осях  $\mathbf{K}$  будем иметь

$$\mathbf{y} = \mathbf{l}\mathbf{y}', \quad \mathbf{x} = \mathbf{l}\mathbf{x}', \quad \tilde{\mathbf{y}} = \tilde{\mathbf{y}}'\tilde{\mathbf{l}}, \quad \tilde{\mathbf{x}} = \tilde{\mathbf{x}}'\tilde{\mathbf{l}}, \quad (10)$$

$$\mathbf{y}' = \tilde{\mathbf{l}}\mathbf{y} = \tilde{\mathbf{l}}\mathbf{K}\mathbf{l}\mathbf{x}' = \mathbf{K}'\mathbf{x}' \quad (11)$$

и

$$y'_1 = \lambda_1 x'_1, \quad y'_2 = \lambda_2 x'_2, \quad y'_3 = \lambda_3 x'_3, \quad (12)$$

так как  $\mathbf{K}'$  диагонального вида. Следовательно, такое перемещение среды изменяет длины трех перпендикулярных линий в теле, но их направления остаются неизменными; все отрезки, параллельные каждой из этих линий, изменяются в одинаковом отношении. Такое движение называется *чистой деформацией*.

Мы займемся общим приведением эрмитовой матрицы  $n \times n$  к диагональному виду; частным случаем этого преобразования является приведение симметричной матрицы. Эта задача возникает во многих областях математической физики. Однако, прежде чем перейти к этому, мы должны рассмотреть решение систем линейных однородных уравнений и некоторые свойства определителей.

#### 4.05. Ранг матрицы. Однородные линейные уравнения.

Обычно система  $n$  линейных уравнений относительно  $n$  неизвестных с  $n$  постоянными в правой части имеет единственное решение. Но если взять пару уравнений

$$x + y = 1, \quad 2x + 2y = 3,$$

то они решения не имеют. Это обстоятельство связано со свойствами коэффициентов. В то время как пара уравнений

$$x + y = 0, \quad x - y = 0$$

не имеет решения, отличного от  $x = y = 0$ , пара

$$x + y = 0, \quad 2x + 2y = 0$$

имеет бесчисленное множество решений, а именно решением будет любая пара чисел, таких, что  $y = -x$ . В первом случае матрица коэффициентов неособая, в последнем случае особая. Это является общим правилом: если правые части все равны нулю, то необходимым и достаточным условием существования решения, отличного от нуля, является равенство нулю определителя коэффициентов; если правые части не все равны нулю, а определитель из коэффициентов равен нулю, то либо уравнения несовместны и решения не существует, либо они совместны и существует бесчисленное множество решений.

Более того, возможно, что  $n$  уравнений могут иметь двухмерный бесконечный набор решений или даже такой, в котором любое число неизвестных, меньшее чем  $n$ , можно задать



произвольно. Чтобы увидеть, отчего это возникает, удобно ввести понятие *ранга* матрицы или определителя. Рассмотрим систему  $n$  уравнений

$$ax = a_{ik}x_k = 0. \quad (1)$$

Очевидно, одно решение  $x_k = 0$  ( $k = 1, 2, \dots, n$ ) всегда существует.

Определитель, полученный из  $\|a_{ik}\|$  вычеркиванием  $m$  строк и  $m$  столбцов, называется *минором* порядка  $n - m$ ; сам  $\|a_{ik}\|$  можно назвать минором порядка  $n$ , а любой отдельно взятый элемент — минором порядка 1. Если вычеркнуто  $m$  строк и  $m$  столбцов, таких, что номера строк  $i$  совпадают с номерами столбцов  $k$ , то образовавшийся минор называется *главным*. Главный минор расположен симметрично относительно главной диагонали. Главный минор в левом верхнем углу называется *ведущим* (leading) *минором*. Так, в определителе четвертого по-

$$\begin{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix} \\ \begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix} & \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} \end{vmatrix}$$

рядка минор, обведенный пунктирной линией, есть простой минор второго порядка. Минор, обведенный штриховой линией, является главным минором, а сплошной линией — ведущим минором второго порядка.

Может оказаться, что все миноры порядка  $r + 1$  равны нулю (а потому и все миноры порядка выше, чем  $r + 1$ ), а некоторые из миноров порядка  $r$  отличны от нуля. Тогда говорят, что матрица и определитель имеют *ранг*  $r$ . Число  $r$  есть наибольшее, для которого можно сказать: „не все миноры порядка  $r$  равны нулю“. В частности, определитель порядка  $n$ , равный нулю, в то время как не все его первые миноры (миноры порядка  $n - 1$ ) равны нулю, имеет ранг  $n - 1$ . Если сам определитель  $\|a_{ik}\|$  отличен от нуля, то матрица и определитель имеют ранг  $n$  \*).

\*) Можно распространить понятие ранга на матрицы порядка  $m \times n$ . Например,

$$x + y + z = 0, \quad 2x + 2y + 2z = 1$$

несовместны для всех конечных  $x, y, z$ . Мы не будем останавливаться на таких случаях, поскольку они для физики мало интересны, а аналитически весьма сложны.

Сразу очевидно, что если ранг  $\mathbf{a}$  не меньше, чем  $n$ , то уравнения (1) не имеют решения, отличного от  $\mathbf{x} = 0$ . В самом деле, если матрица  $\mathbf{a}$  ранга  $n$ , то она имеет обратную  $\mathbf{a}^{-1}$  и из

$$\mathbf{a}^{-1}\mathbf{a}\mathbf{x} = 0 \quad (2)$$

следует

$$\mathbf{x} = 0. \quad (3)$$

Обратно, предположим, что  $\mathbf{a}$  ранга  $n - 1$ . Это означает, что  $\|a_{ik}\| = 0$ , но не все алгебраические дополнения  $A_{ik}$  равны нулю. Для определенности положим, что ведущий первый минор  $A_{nn}$  отличен от нуля. Это всегда можно достигнуть перегруппировкой уравнений и перенумерацией неизвестных. Тогда решение уравнений (1) можно получить, положив  $x_n$  равным произвольной постоянной  $b$ . Действительно, мы можем решить первые  $n - 1$  уравнений относительно  $x_1, x_2, \dots, x_{n-1}$  обычным путем, так как определитель коэффициентов при этих неизвестных равен  $A_{nn}$  и по предположению отличен от нуля. Решение есть

$$x_k = \frac{A_{nk}b}{A_{nn}} \quad (k = 1, 2, \dots, n). \quad (4)$$

Мы должны убедиться, что это решение удовлетворяет  $n$ -му уравнению. Подставляя, получим

$$a_{nk}x_k = \frac{b}{A_{nn}} \sum_k a_{nk}A_{nk} = \frac{b}{A_{nn}} \|a_{ik}\|, \quad (5)$$

так как сумма есть разложение  $\|a_{ik}\|$  по элементам последней строки. Но это выражение равно нулю, потому что  $\|a_{ik}\| = 0$ , а  $A_{nn} \neq 0$ . Следовательно,  $n$ -е уравнение также удовлетворяется. Таким образом, отношения  $x_i$  единственны, но сами  $x_i$  определены с точностью до произвольного множителя.

Если алгебраическое дополнение  $A_{pq}$ , полученное вычеркиванием  $p$ -й строки и  $q$ -го столбца, не равно нулю, то мы можем аналогично, полагая  $x_q = c$ , получить решение

$$x_k = \frac{A_{pk}c}{A_{pq}}, \quad (6)$$

но, чтобы (4) и (6) были совместными, для всех  $k, p, q$  должно быть

$$A_{pk}A_{nq} - A_{nk}A_{pq} = 0. \quad (7)$$

Обращение этого выражения в нуль при  $\|a_{ik}\|$ , равном нулю, есть частный случай теоремы Якоби, которую мы докажем позже. Далее можно видеть, что если определитель и один пер-

вый минор равны нулю, то миноры всех элементов либо той же самой строки, либо того же столбца равны нулю.

Предположим теперь, что матрица  $a$  имеет ранг  $r$ . Для определенности допустим, что уравнения сгруппированы так, что ведущий минор порядка  $r$  отличен от нуля. Здесь придется отказаться от соглашения о суммировании, так как оно производится не от 1 до  $n$ . Тогда уравнения можно переписать в виде

$$\sum_{k=1}^r a_{ik}x_k = - \sum_{\kappa=r+1}^n a_{i\kappa}x_{\kappa} \quad (i=1, 2, \dots, r), \quad (8)$$

$$\sum_{k=1}^r a_{ik}x_k = 0 \quad (i=r+1, \dots, n). \quad (9)$$

Для любой совокупности значений  $x_{\kappa}$  (при  $\kappa$  от  $r+1$  до  $n$ ) первая система имеет единственное решение. Действительно, матрица коэффициентов левой части (8) неособая и имеет обратную. Обозначим ее определитель через  $\alpha$ , а алгебраическое дополнение  $a_{ik}$  в этом определителе через  $\alpha_{ik}$ . Тогда для  $k$  от 1 до  $r$  уравнения (8) удовлетворяются тогда и только тогда, когда

$$x_k = - \sum_{i=1}^r \frac{\alpha_{ik}}{\alpha} \sum_{\kappa=r+1}^n a_{i\kappa}x_{\kappa}. \quad (10)$$

Подставим (10) в любое из уравнений (9). Для  $i > r$  получим

$$\begin{aligned} \sum_{k=1}^r a_{ik}x_k &= - \sum_{k=1}^r \sum_{p=1}^r \frac{\alpha_{pk}}{\alpha} a_{ik} \sum_{\kappa=r+1}^n a_{p\kappa}x_{\kappa} + \sum_{\kappa=r+1}^n a_{i\kappa}x_{\kappa} = \\ &= \sum_{\kappa=r+1}^n \frac{x_{\kappa}}{\alpha} \left( \alpha a_{i\kappa} - \sum_{p=1}^r \sum_{k=1}^r a_{ik} \alpha_{pk} a_{p\kappa} \right). \end{aligned} \quad (11)$$

Разложим следующий минор порядка  $r+1$  по элементам последней строки и последнего столбца:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1\kappa} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2\kappa} \\ \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{r\kappa} \\ a_{i1} & a_{i2} & \dots & a_{ir} & a_{i\kappa} \end{vmatrix}; \quad (12)$$

коэффициент при  $a_{i\kappa}$  в этом разложении равен  $\alpha$ . Коэффициент при  $a_{ik}a_{p\kappa}$  равен коэффициенту при  $a_{i\kappa}a_{pk}$  с обратным знаком, т. е.  $-\alpha_{pk}$ . Следовательно, коэффициент при  $x_{\kappa}/\alpha$  в (11) есть

минор порядка  $r + 1$  исходного определителя, равный нулю по предположению. Таким образом, если матрица  $\mathbf{a}$  имеет ранг  $r$ , меньший чем  $n$ , то  $n - r$  неизвестных можно задать произвольно, а остальные будут однородными линейными функциями от них.

**4.06. Детерминантные (determinantal) уравнения.** Часто нам нужно будет рассматривать систему  $n$  уравнений

$$\lambda a_{ik} x_k = b_{ik} x_k, \quad (1)$$

в которой  $\lambda$  должно быть определено таким образом, чтобы существовало решение с некоторым числом  $x_k$ , отличным от нуля. Это требование приводит к тому, чтобы  $\lambda$  было равно некоторому корню уравнения

$$\|\lambda a_{ik} - b_{ik}\| = 0. \quad (2)$$

Обычно оно будет иметь  $n$  различных корней. Для каждого корня  $\lambda_j$  можно найти некоторую совокупность значений  $x_i$ , скажем  $l_{ij}$ , удовлетворяющую уравнениям (1); любое кратное этой совокупности также будет решением. Для некоторых распространенных типов матриц, в особенности для симметричных и унитарных, решение обладает особыми свойствами, которые мы изучим позже.

Если  $a_{ik} = \delta_{ik}$ , то левая часть (1) сводится к  $\lambda x_i$ . Тогда (1) можно записать в матричном виде

$$\lambda \mathbf{x} = \mathbf{b} \mathbf{x}, \quad (3)$$

причем  $\lambda$  называют *характеристическими числами* или *собственными значениями* матрицы  $\mathbf{b}$  \*). Определяющее их уравнение

$$\|b_{ik} - \lambda \delta_{ik}\| = 0 \quad (4)$$

называется *характеристическим уравнением* матрицы. Совокупность  $x_i$ , скажем  $l_{ij}$ , удовлетворяющая (3) для конкретного значения  $\lambda$ , скажем  $\lambda_j$ , можно назвать *характеристическим решением*, или *собственным вектором*.

Если взять произвольную совокупность чисел  $X_i$ , то всегда возможно найти совокупность  $\xi_j$ , такую, что

$$X_i = l_{ij} \xi_j. \quad (5)$$

Для этого матрица  $l_{ij}$  не должна быть особой. Это легко доказывается, если все  $\lambda_j$  различны. Действительно, если  $l_{ij}$

---

\*) Как  $\lambda$ , так и  $\lambda^{-1}$  встречаются иногда в практических задачах. Гильберт и Курант [2] называют  $\lambda$ , определяемые в (3), характеристическими числами, а  $1/\lambda$  — собственными значениями.

была бы особой, мы могли бы взять ненулевые значения  $\xi_j$ , такие, что

$$\sum_j l_{kj} \xi_j = 0 \quad (6)$$

для всех  $k$ .

Пусть  $r$  будет наименьшее число величин  $\xi_j$ , входящих в любую такую систему уравнений. Мы всегда можем сделать так, чтобы  $j$  пробегало от 1 до  $r$ . Тогда для всех  $i$

$$0 = (b_{ik} - \lambda_i \delta_{ik}) \sum_{j=1}^r l_{kj} \xi_j = \sum (\lambda_j - \lambda_i) l_{ij} \xi_j. \quad (7)$$

Но если мы возьмем  $i$  совпадающим с каким-либо  $j$  из (6), соответствующий множитель  $\lambda_j - \lambda_i = 0$ , и мы получим систему соотношений для всех  $i$  между по крайней мере  $r-1$  величинами  $l_{ij} \xi_j$ , поскольку по предположению  $\lambda_j$  все различны и в (7) все члены не могут обратиться в нуль. Таким образом, предположение о том, что  $l_{ij}$  особая, приводит к противоречию.

Когда некоторые корни совпадают,  $l_{ij}$  обычно также является неособой, хотя и не всегда. Рассмотрим матрицы  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  и  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ . Каждая из них имеет двойной корень 1. Первая имеет два собственных вектора (1, 0), (0, 1). Вторая имеет единственный собственный вектор (1, 0).

Рассмотрим матрицу  $\lambda_{ij}$ , у которой *диагональные* элементы равны  $\lambda_j$ , а все остальные — нули. Тогда (3) можно записать как

$$b_{ik} l_{kj} = l_{ik} \lambda_{kj} \quad (8)$$

для всех  $i, j$ . Если  $\mathbf{l}$  неособая, то, умножая слева на  $\mathbf{l}^{-1}$ , получим

$$\mathbf{l}^{-1} \mathbf{b} \mathbf{l} = \boldsymbol{\lambda}. \quad (9)$$

Таким образом, преобразование вида (9) приводит  $\mathbf{b}$  к диагональному виду. Применение его к единичной матрице дает  $\mathbf{l}^{-1} \mathbf{l}$ , что равно  $\mathbf{I}$ . Следовательно, такое преобразование оставляет единичную матрицу без изменения; оно называется преобразованием *подобия* (collineatory), а две матрицы, связанные таким преобразованием, называются *подобными*. Не обязательно нужно оговаривать, что  $\mathbf{l}$  неособая; это само собой разумеется из того, что в формуле фигурирует  $\mathbf{l}^{-1}$ . Если бы  $\mathbf{l}$  была особой, то не было бы такой формулы.

**4.061. Важность диагональных матриц.** Все диагональные матрицы коммутируют между собой. Если диагональные элементы такой матрицы  $\boldsymbol{\lambda}$  равны  $\lambda_r$ , то диагональные элементы

матрицы  $\lambda^n$  будут  $\lambda_r^n$ , а все остальные — нулями. Если ни одно из  $\lambda_r$  не равно нулю, то диагональные элементы  $\lambda^{-1}$  равны  $\lambda_r^{-1}$ , а все остальные — нули. Если  $\mathbf{I}^{-1}\mathbf{a}\mathbf{I} = \lambda$ , то  $\mathbf{a}^n = (\mathbf{I}\lambda\mathbf{I}^{-1})^n = \mathbf{I}\lambda^n\mathbf{I}^{-1}$ ; все промежуточные произведения  $\mathbf{I}\mathbf{I}^{-1}$  выпадают, когда это выражение выписывается в развернутом виде.

Этот результат непосредственно связан с поведением динамических систем, описываемых линейными уравнениями движения. Если координаты и скорости заданы в момент времени 0, то в момент времени  $\tau$  они являются линейными функциями своих начальных значений. Если обозначим их значения в момент 0 и  $\tau$  через  $X_i$  и  $Y_i$  ( $i$  принимает значения от 1 до  $2n$ , где  $n$  — число степеней свободы), то

$$Y_i = a_{ik}X_k, \quad \mathbf{Y} = \mathbf{a}\mathbf{X}.$$

Теперь если 4.06 (5) удовлетворяется характеристическими решениями матричных уравнений  $\mathbf{a}\mathbf{x} = \lambda\mathbf{x}$ , то можно написать

$$X_k = l_{kj}\xi_j, \quad \mathbf{X} = \mathbf{l}\xi, \quad \xi = \mathbf{I}^{-1}\mathbf{X}$$

и

$$Y_i = a_{ik}l_{kj}\xi_j = l_{ik}\lambda_{kj}\xi_j, \quad \mathbf{Y} = \mathbf{l}\lambda\mathbf{I}^{-1}\mathbf{X}.$$

Если теперь  $Z_i$  — значения в момент времени  $2\tau$ , то их находят по формуле  $\mathbf{Z} = \mathbf{a}\mathbf{Y} = \mathbf{a}^2\mathbf{X} = \mathbf{l}\lambda^2\mathbf{I}^{-1}\mathbf{X}$ . Процесс можно продолжить для любого времени, кратного  $\tau$ , и поэтому, если  $\mathbf{l}$  неособая, решение найдено. Очевидно,  $\lambda_{ji}$  ( $j = l$ ) являются временными множителями типа  $\exp(\gamma\tau)$  в решении динамических уравнений, полученных обычным методом.

Поскольку, как правило, корни различны,  $\mathbf{l}$  обычно неособая; в некоторых важных частных случаях  $\mathbf{l}$  является неособой, если даже корни кратные, и метод по-прежнему применим.

**4.062. Матрица удовлетворяет своему характеристическому уравнению.** Если уравнение

$$\|a_{ik} - \lambda\delta_{ik}\| = 0 \quad (1)$$

разложить по степеням  $\lambda$ , получим уравнение  $n$ -й степени относительно  $\lambda$ :

$$D(\lambda) = \alpha_n + \alpha_{n-1}\lambda + \alpha_{n-2}\lambda^2 + \dots + (-1)^n\lambda^n = 0. \quad (2)$$

Теорема, сформулированная в заголовке, означает, что если вместо  $\lambda$  подставить матрицу  $\mathbf{a}$  в каждый член уравнения и воспользоваться правилами матричного умножения и сложения, то получится матрица  $\mathbf{D}(\mathbf{a})$ , каждый элемент которой равен 0.

Предположим сначала, что уравнение (2) имеет только простые корни. Для каждого корня  $\lambda_j$  существует совокупность  $x_{ij}$ , не равных одновременно

менно нулю и удовлетворяющих уравнениям

$$\lambda_j x_{ij} = a_{ik} x_{kj}, \quad a x_j = x_j \lambda_j, \quad (3)$$

$$a^r x_j = x_j \lambda_j^r, \quad (4)$$

так как  $\lambda_j$  есть число. Возьмем тогда

$$y_i = \beta_j x_{ij}, \quad y = \beta_j x_j, \quad (5)$$

где  $\beta_j$  —  $n$  произвольных множителей. Тогда

$$\begin{aligned} D(a)y &= (\alpha_n \mathbf{1} + \alpha_{n-1} a + \alpha_{n-2} a^2 + \dots + (-1)^n a^n) \beta_j x_j = \\ &= \sum_j (\alpha_n + \alpha_{n-1} \lambda_j + \alpha_{n-2} \lambda_j^2 + \dots + (-1)^n \lambda_j^n) \beta_j x_j = \\ &= 0, \end{aligned} \quad (6)$$

так как все  $\lambda_j$  удовлетворяют (2). Но матрица  $x_{ij}$  неособая, а  $\beta_i$  можно выбрать так, чтобы сделать  $y_i$  каким угодно. Поэтому

$$D(a) = 0.$$

Если уравнение  $D(\lambda) = 0$  имеет кратные корни, то можно предположить, что элементы изменены таким образом, что корни стали различными, и затем применить (7). Если теперь непрерывно стремиться элементы к их исходным значениям, то каждый элемент  $D(a)$  также будет непрерывно стремиться к своему исходному значению. Но он всегда равен нулю, и, следовательно, исходное значение каждого элемента должно быть нулем.

Существуют более короткие доказательства. Это доказательство выбрано для иллюстрации пользы диагональных матриц.

**4.063. Блочные матрицы.** Предположим, что элементы матрицы сгруппированы в блоки (квадратные и прямоугольные). Сами эти блоки можно рассматривать как элементы матрицы более низкого порядка. Так, из матрицы

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \text{ можно получить } \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad (1)$$

где

$$\mathbf{A}_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathbf{A}_{12} = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}, \quad \mathbf{A}_{21} = (a_{31}, \quad a_{32}), \quad \mathbf{A}_{22} = a_{33}. \quad (2)$$

Мы будем говорить о матрице в правой части (1), как о матрице в *блочном виде*, и обозначать ее  $\mathbf{A}$ . О матрице в левой части будем говорить, как о матрице в *развернутом виде*. Две матрицы  $\mathbf{A}$  и  $\mathbf{B}$  в блочном виде могут быть перемножены по обычному правилу, что дает новую матрицу  $\mathbf{AB} = \mathbf{C}$  в блочном виде

$$C_{ik} = \sum_j \mathbf{A}_{ij} \mathbf{B}_{jk} \quad (3)$$

при условии, что 1) число блоков-столбцов  $\mathbf{A}$  равно числу блоков-строк  $\mathbf{B}$  и 2) число столбцов матрицы  $\mathbf{A}_{ij}$  равно числу строк матрицы  $\mathbf{B}_{jk}$ . Например, если  $\mathbf{A}$  и  $\mathbf{B}$  построены следующим образом:

$$\mathbf{A} = \begin{pmatrix} n_1 & n_2 \\ \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{matrix} m_1 \\ m_2 \end{matrix}, \quad \mathbf{B} = \begin{pmatrix} p_1 & p_2 & p_3 \\ \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{pmatrix} \begin{matrix} n_1 \\ n_2 \end{matrix}. \quad (4)$$

Здесь приписанные по краям числа  $m, n, p$  указывают количество строк и столбцов компонент, так что  $\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{21}, \mathbf{A}_{22}$  имеют соответственно порядки  $m_1 \times n_1, m_1 \times n_2, m_2 \times n_1, m_2 \times n_2$ , а  $\mathbf{B}_{11}, \mathbf{B}_{12}, \dots$  имеют порядки  $n_1 \times p_1, n_1 \times p_2, \dots$ . Произведение матриц есть

$$\mathbf{C} = \begin{pmatrix} p_1 & p_2 & p_3 \\ \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{C}_{23} \end{pmatrix} \begin{matrix} m_1 \\ m_2 \end{matrix}, \quad (5)$$

где  $\mathbf{C}_{11}, \mathbf{C}_{12}, \dots$  имеют порядки  $m_1 \times p_1, m_1 \times p_2, \dots$ .

Если  $\mathbf{a}$  и  $\mathbf{b}$  — простые матрицы и  $\mathbf{ab} = \mathbf{c}$ , тогда если  $\mathbf{A}$  и  $\mathbf{B}$  — блочные формы  $\mathbf{a}$  и  $\mathbf{b}$ , то  $\mathbf{C}$  есть блочная форма  $\mathbf{c}$ . Другими словами, если мы образуем  $\mathbf{C}$ , то можем получить  $\mathbf{c}$ , выписывая полностью все элементы матричных элементов  $\mathbf{C}$  и убирая внутренние скобки. Мы докажем это для случая, когда  $\mathbf{A}_{11}, \mathbf{B}_{11}$  порядка  $n_1 \times n_1$ ;  $\mathbf{A}_{12}, \mathbf{B}_{12} — n_1 \times n_2$ ;  $\mathbf{A}_{21}, \mathbf{B}_{21} — n_2 \times n_1$  и  $\mathbf{A}_{22}, \mathbf{B}_{22} — n_2 \times n_2$ . Удобно в  $\mathbf{A}_{12}, \mathbf{A}_{22}$  нумеровать столбцы от  $n_1 + 1$  до  $n_1 + n_2$ , а в  $\mathbf{A}_{21}, \mathbf{A}_{22}$  нумеровать строки от  $n_1 + 1$  до  $n_1 + n_2$ . Тогда

$$\left. \begin{aligned} a_{ik} &= (\mathbf{A}_{11})_{ik} & (i \leq n_1, \quad k \leq n_1), \\ a_{ik} &= (\mathbf{A}_{12})_{ik} & (i \leq n_1, \quad n_1 + 1 \leq k \leq n_1 + n_2), \\ a_{ik} &= (\mathbf{A}_{21})_{ik} & (n_1 + 1 \leq i \leq n_1 + n_2, \quad k \leq n_1), \\ a_{ik} &= (\mathbf{A}_{22})_{ik} & (n_1 + 1 \leq i \leq n_1 + n_2, \quad n_1 + 1 \leq k \leq n_1 + n_2). \end{aligned} \right\} \quad (6)$$

То же правило применяется к  $\mathbf{b}$ . Теперь

$$c_{ik} = (\mathbf{ab})_{ik} = \left( \sum_{j=1}^{n_1} + \sum_{j=n_1+1}^{n_1+n_2} \right) a_{ij} b_{jk}. \quad (7)$$

Для  $i, k \leq n_1$  это равно

$$\sum_{j=1}^{n_1} (\mathbf{A}_{11})_{ij} (\mathbf{B}_{11})_{jk} + \sum_{j=n_1+1}^{n_1+n_2} (\mathbf{A}_{12})_{ij} (\mathbf{B}_{21})_{jk} = (\mathbf{A}_{11} \mathbf{B}_{11} + \mathbf{A}_{12} \mathbf{B}_{21})_{ik} = (\mathbf{C})_{ik}. \quad (8)$$

Для остальных частей произведения доказательство аналогично.



Мы будем говорить, что матрица имеет *квазидиагональную форму*, если она в блочной форме и все матричные элементы, отличные от диагональных, являются нулями. Развернутая матрица, полученная расписыванием блоков в полном виде и убираанием скобок, будет тогда обладать тем свойством, что все ее ненулевые элементы находятся в непересекающихся квадратах, симметричных относительно главной диагонали\*). Диагональная матрица является частным случаем квазидиагональной матрицы.

Если **A** и **B** имеют квазидиагональную форму, то ясно, что произведение **AB** также будет иметь квазидиагональную форму.

**4.064.** Если **A** есть квазидиагональная матрица с диагональными элементами  $\mathbf{A}_{rs}$  ( $s=r$ ) и ее развернутая матрица есть **a**, то необходимым и достаточным условием приводимости **a** к диагональному виду преобразованием подобия будет условие приводимости к диагональному виду каждого из  $\mathbf{A}_{rs}$ ; собственные векторы **a** получаются из собственных векторов  $\mathbf{A}_{rs}$  добавлением нулевых компонент.

Остановимся на случае, когда **A** порядка  $2 \times 2$ , поскольку случай  $n \times n$  получится повторением рассуждений. Утверждение, что  $\mathbf{A}_{11}$  приводима к диагональной форме, означает, что существует неособая **L**, такая, что  $\mathbf{L}_1^{-1} \mathbf{A}_{11} \mathbf{L}_1 = \lambda_1$ , где  $\lambda_1$  диагональная. Тогда если  $\mathbf{A}_{22}$  аналогичным образом приводится посредством  $\mathbf{L}_2$  к  $\lambda_2$ , то

$$\begin{pmatrix} \mathbf{L}_1^{-1} & 0 \\ 0 & \mathbf{L}_2^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{L}_1 & 0 \\ 0 & \mathbf{L}_2 \end{pmatrix} = \\ = \begin{pmatrix} \mathbf{L}_1^{-1} \mathbf{A}_{11} \mathbf{L}_1 & 0 \\ 0 & \mathbf{L}_2^{-1} \mathbf{A}_{22} \mathbf{L}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (1)$$

Развернутая матрица, образованная из  $\begin{pmatrix} \mathbf{L}_1 & 0 \\ 0 & \mathbf{L}_2 \end{pmatrix}$ , является неособой, потому что  $\mathbf{L}_1$  и  $\mathbf{L}_2$  неособые, и

$$\begin{pmatrix} \mathbf{L}_1^{-1} & 0 \\ 0 & \mathbf{L}_2^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{L}_1 & 0 \\ 0 & \mathbf{L}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}. \quad (2)$$

Следовательно, условие теоремы достаточно.

---

\*) Это утверждение справедливо только для квадратных матриц, а предыдущее определение пригодно и для прямоугольных матриц (см. Гантмахер, Теория матриц, ГТИ, 1953). — Прим. перев.

Обратно, пусть  $\mathbf{a}$  приводится к диагональной форме преобразованием подобия. Пусть  $\mathbf{A}_{11}$  и  $\mathbf{A}_{22}$  порядка  $n_1 \times n_1$  и  $n_2 \times n_2$  соответственно. Утверждение, что  $\mathbf{a}$  приводима к диагональной форме преобразованием  $\mathbf{I}^{-1}\mathbf{a}\mathbf{I}$ , означает, что существуют  $n_1 + n_2$  линейно независимых совокупностей  $l_{i(k)}$  ( $i$  принимает значения от 1 до  $n_1 + n_2$ ), таких, что для каждой совокупности существует  $\lambda_k$ , удовлетворяющее

$$a_{ij}l_{j(k)} = \lambda_{(k)}l_{i(k)}, \quad (3)$$

где скобки в индексах означают, что суммирование по  $k$  не производится. Не все  $\lambda_k$  обязательно должны быть различными, но утверждение, что  $\mathbf{I}^{-1}\mathbf{a}\mathbf{I}$  является диагональной, подразумевает, что даже если некоторые из  $\lambda_k$  равны, то соответствующие решения все же можно выбрать так, чтобы они были линейно независимыми.

Заметим сначала, что

$$\det(\mathbf{a} - \lambda\mathbf{I}) = \det(\mathbf{A}_{11} - \lambda\mathbf{I}) \det(\mathbf{A}_{22} - \lambda\mathbf{I}). \quad (4)$$

Поэтому каждое собственное значение  $\mathbf{a}$  является собственным значением либо  $\mathbf{A}_{11}$ , либо  $\mathbf{A}_{22}$ ; и если  $\lambda$  есть  $p$ -кратное собственное значение  $\mathbf{A}_{11}$  и  $q$ -кратное собственное значение  $\mathbf{A}_{22}$ , то оно есть  $(p+q)$ -кратное собственное значение.

Мы должны показать, что при заданных условиях  $\mathbf{A}_{11}$  и  $\mathbf{A}_{22}$  имеют  $n_1$  и  $n_2$  линейно независимых собственных векторов соответственно. (И, конечно, ни одна из них не может иметь больше.) Пронумеруем строки и столбцы в  $\mathbf{A}_{11}$  от 1 до  $n_1$ , а в  $\mathbf{A}_{22}$  от  $n_1 + 1$  до  $n_1 + n_2$ . Если  $L_{i(k)1}$  есть собственный вектор  $\mathbf{A}_{11}$ , так что

$$(\mathbf{A}_{11})_{ij}L_{j(k)1} = \lambda L_{i(k)1}, \quad (5)$$

тогда если

$$l_{i(k)1} = \begin{cases} L_{i(k)1} & (i \leq n_1), \\ 0 & (n_1 + 1 \leq i \leq n_1 + n_2), \end{cases} \quad (6)$$

то получим для  $i \leq n_1$

$$a_{ij}l_{j(k)1} = (\mathbf{A}_{11})_{ij}L_{j(k)1} = \lambda L_{i(k)1} = \lambda l_{i(k)1}, \quad (7)$$

а для  $i > n_1$

$$a_{ij}l_{j(k)1} = 0L_{j(k)1} + (\mathbf{A}_{22})_{ij}0 = \lambda l_{i(k)1}. \quad (8)$$

Следовательно,  $l_{i(k)1}$  — собственный вектор  $\mathbf{a}$  с тем же самым  $\lambda$ . Аналогично если  $L_{i(k)2}$  — собственный вектор  $\mathbf{A}_{22}$  при  $i \geq n_1 + 1$ , то имеется собственный вектор  $l_{i(k)2}$  матрицы  $\mathbf{a}$ , который получается, если положить

$$l_{i(k)2} = 0 \quad (i \leq n_1); \quad l_{i(k)2} = L_{i(k)2} \quad (n_1 + 1 \leq i \leq n_1 + n_2). \quad (9)$$

Далее, если собственные векторы  $\mathbf{A}_{11}$  и  $\mathbf{A}_{22}$  линейно независимы, то полученные по ним  $l_{i(k)}$  также линейно независимы.

Предположим теперь, что имеется  $m_1$  линейно независимых собственных векторов  $\mathbf{A}_{11}$  и  $m_2$  собственных векторов  $\mathbf{A}_{22}$ , причем

$$m_1 \leq n_1, \quad m_2 \leq n_2, \quad \text{но } m_1 + m_2 < n_1 + n_2.$$

Тогда тем же методом можно получить меньше чем  $n_1 + n_2$  собственных векторов  $\mathbf{a}$ . Следовательно, существует по крайней мере один собственный вектор  $\mathbf{a}$ , скажем  $l_{i(m_1+m_2+1)}$ , удовлетворяющий при некотором  $\lambda$  уравнению

$$a_{ij} l_{j(m_1+m_2+1)} = \lambda l_{i(m_1+m_2+1)}. \quad (10)$$

Но это эквивалентно двум системам уравнений

$$\left. \begin{aligned} (\mathbf{A}_{11})_{ij} l_{j(m_1+m_2+1)} &= \lambda l_{i(m_1+m_2+1)} & (i \leq n_1), \\ (\mathbf{A}_{22})_{ij} l_{j(m_1+m_2+1)} &= \lambda l_{i(m_1+m_2+1)} & (n_1 + 1 \leq i \leq n_1 + n_2). \end{aligned} \right\} \quad (11)$$

По предположению не все  $l_i$  равны нулю. Следовательно, первая система дает собственный вектор  $\mathbf{A}_{11}$ , а вторая — собственный вектор  $\mathbf{A}_{22}$  или и то и другое вместе. В любом случае можно записать

$$l_{i(m_1+m_2+1)} = \alpha l_{i(m_1+m_2+1)1} + \beta l_{i(m_1+m_2+1)2}. \quad (12)$$

Но по предположению мы уже нашли все собственные векторы  $\mathbf{A}_{11}$  и  $\mathbf{A}_{22}$ . Следовательно,  $l_{i(m_1+m_2+1)1}$  линейно выражается через  $m_1$  известных собственных векторов  $\mathbf{A}_{11}$ , а  $l_{i(m_1+m_2+1)2}$  — через  $m_2$  собственных векторов  $\mathbf{A}_{22}$ . Но тогда  $l_{i(m_1+m_2+1)}$  линейно выражается через уже полученные векторы, значит, это не новое независимое решение. Следовательно,  $m_1 = n_1$ ,  $m_2 = n_2$  и теорема доказана. В частности, если  $\lambda$  есть  $p$ -кратное собственное значение  $\mathbf{A}_{11}$  и  $q$ -кратное собственное значение  $\mathbf{A}_{22}$ , то этот процесс дает  $p + q$  независимых собственных векторов  $\mathbf{a}$ , соответствующих этому  $\lambda$ .

Если  $l_{i(k)}$  — любой собственный вектор  $\mathbf{a}$ , то метод, использованный при получении (II), дает собственный вектор  $\mathbf{A}_{11}$ , либо  $\mathbf{A}_{22}$ , либо обоих вместе. Если соответствующее  $\lambda$  является собственным значением  $\mathbf{A}_{11}$ , но не  $\mathbf{A}_{22}$ , то получится собственный вектор для  $\mathbf{A}_{11}$  и нулевой вектор для  $\mathbf{A}_{22}$ .

Заметим также, что

$$\begin{pmatrix} \mathbf{L}_1 & 0 \\ 0 & \mathbf{L}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{L}_1 & 0 \\ 0 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{L}_2 \end{pmatrix}, \quad (13)$$

и, следовательно, приведение  $\mathbf{a}$  к диагональной форме может быть выполнено последовательным преобразованием блоков к диагональной форме.

**4.065.** Если каждую из двух матриц можно при помощи преобразования подобия привести к диагональному виду, то для того, чтобы обе матрицы можно было привести одним и тем же преобразованием, необходимо и достаточно, чтобы эти матрицы коммутировали.

Если

$$I^{-1}aI = \lambda \quad \text{и} \quad I^{-1}bI = \mu,$$

где  $\lambda$  и  $\mu$  — диагональные и, следовательно, коммутирующие матрицы, то

$$ab = I\lambda I^{-1}I\mu I^{-1} = I\lambda\mu I^{-1} = I\mu\lambda I^{-1} = ba.$$

Необходимость доказана.

Обратно, если

$$I^{-1}aI = \lambda, \quad I^{-1}bI = c,$$

где  $\lambda$  — диагональная матрица, а  $c$  — не обязательно диагональная, и если  $ab = ba$ ,

$$\lambda c = I^{-1}aII^{-1}bI = I^{-1}abI = I^{-1}baI = c\lambda.$$

Далее,

$$(\lambda c)_{ik} = \lambda_{ij}c_{jk} = \lambda_{ii}c_{ik},$$

причем здесь суммирование по  $i$  не подразумевается. Аналогично

$$(c\lambda)_{ik} = c_{ik}\lambda_{kk},$$

где не подразумевается суммирование по  $k$ . Следовательно, либо  $c_{ik} = 0$ , либо  $\lambda_{ii} = \lambda_{kk}$ . Если все диагональные элементы матрицы  $\lambda$  различны, то все недиагональные элементы матрицы  $c$  равны нулю и  $c$  имеет диагональный вид. Если некоторые диагональные элементы  $\lambda$  равны между собой, то  $c$  можно представить в виде квазидиагональной формы, причем диагональ каждого блока соответствует совокупности равных между собой диагональных элементов матрицы  $\lambda$ . Но по предположению матрицу  $c$  можно привести к диагональной форме. Следовательно, согласно **4.064**, приведение можно произвести последовательным преобразованием к диагональной форме каждого блока. Такие преобразования не влияют на соответствующие элементы матрицы  $\lambda$ , так как они образуют матрицу, кратную единичной.

Нижеследующее доказательство достаточности нам сообщил Холл. В этом доказательстве не используются свойства блочных матриц. Если  $a$  и  $b$  — матрицы порядка  $n \times n$ , то утверждение, что их можно привести

к диагональной форме эквивалентно утверждению, что каждая из матриц имеет  $n$  линейно независимых собственных векторов. Нам надо показать, что если  $\mathbf{ab} = \mathbf{ba}$ , то  $\mathbf{a}$  и  $\mathbf{b}$  имеют общую систему  $n$  линейно независимых собственных векторов. Любой вектор можно представить в виде линейной комбинации собственных векторов матрицы  $\mathbf{a}$ , так как они линейно независимы. В частности, так можно представить любой собственный вектор  $\mathbf{x}_i$  матрицы  $\mathbf{b}$ , соответствующий собственному значению  $\mu$ .

Пусть это представление имеет вид

$$\mathbf{x}_i = \sum_{k=1}^r l_{i(k)}.$$

Можем принять, что  $\mathbf{a}l_{(k)} = \lambda_k l_{(k)}$  с различными  $\lambda_{(k)}$ , ибо, если два собственных вектора матрицы  $\mathbf{a}$  соответствуют одному и тому же собственному значению  $\lambda$ , то и любая линейная комбинация их представляет собой собственный вектор, соответствующий тому же  $\lambda$ .

Тогда

$$a_{ij} b_{jk} \mathbf{x}_k = a_{ij} \mu \mathbf{x}_j = \mu \sum_{k=1}^r \lambda_{(k)} l_{i(k)},$$

$$b_{ij} a_{jk} \mathbf{x}_k = b_{ij} \sum_{k=1}^r \lambda_{(k)} l_{j(k)},$$

и из  $\mathbf{ab} = \mathbf{ba}$  следует, что  $\sum_{k=1}^r \lambda_{(k)} l_{i(k)}$  также представляет собой собственный вектор матрицы  $\mathbf{b}$ , соответствующий собственному значению  $\mu$ . Отсюда следует, что то же самое можно сказать о том, что

$$\lambda_i \mathbf{x}_i - \sum_{k=1}^r \lambda_{(k)} l_{i(k)} = (\lambda_1 - \lambda_2) l_{i(2)} + (\lambda_1 - \lambda_3) l_{i(3)} + \dots$$

Повторяя рассуждения, можно показать, что  $(\lambda_1 - \lambda_r)(\lambda_2 - \lambda_r) \dots (\lambda_{r-1} - \lambda_r) l_{i(r)}$  — собственный вектор матрицы  $\mathbf{b}$  и поэтому  $l_{i(r)}$  — один из общих собственных векторов. Отсюда следует, что поскольку номер  $k$  несуществен, каждое  $l_{i(k)}$  представляет собой собственный вектор и  $\mathbf{b}$ , и  $\mathbf{a}$ . Следовательно, каждый собственный вектор матрицы  $\mathbf{b}$  можно представить в виде линейной комбинации собственных векторов, общих для  $\mathbf{a}$  и  $\mathbf{b}$ . Поэтому в таком виде можно представить любой из этих  $n$  векторов.

**4.066.** Теоремы 4.064 и 4.065 имеют большое значение в квантовой механике для частного случая, когда матрица  $\mathbf{a}$  эрмитова. В 4.083 будет показано, что если  $\mathbf{a}$  эрмитова матрица, то ее всегда можно привести к диагональной форме при помощи преобразования подобия с унитарной матрицей  $\mathbf{I}$ . Далее, если  $\mathbf{a}$  эрмитова матрица; имеющая вид квазидиаго-

нальной матрицы, то каждый диагональный блок также представляет собой эрмитову матрицу и приводится к диагональному виду аналогичным образом. Поэтому нам нужна не общая теорема из 4.064, а лишь свойство, выраженное формулой 4.064 (13).

**4.067.** Рассмотрим следующий численный пример, встречающийся при определении собственных значений момента количества движения для некоторых конфигураций электронов в атомах [3].

Возьмем три матрицы:

$$A = \frac{1}{4} \begin{pmatrix} 3 & 0 & & 0 \\ 0 & 3 & & \\ \hline & & 7 & 4 & 4 \\ 0 & & 4 & 7 & 4 \\ & & 4 & 4 & 7 \end{pmatrix},$$

$$B = \begin{pmatrix} 4 & -2 & & & 0 \\ -2 & 4 & & & \\ \hline & & 2 & -2 & 0 \\ 0 & & -2 & 4 & -2 \\ & & 0 & -2 & 2 \end{pmatrix},$$

$$C = \frac{1}{2} \begin{pmatrix} -1 & 0 & \sqrt{2} & -\sqrt{2} & 0 \\ 0 & -1 & 0 & \sqrt{2} & -\sqrt{2} \\ \hline \sqrt{2} & 0 & & & \\ -\sqrt{2} & \sqrt{2} & & 0 & \\ 0 & -\sqrt{2} & & & \end{pmatrix}.$$

Все три матрицы симметричны и все коммутируют. Собственные значения матрицы **A** следующие:  $3/4$  (четыре раза),  $15/4$ . Собственные значения матрицы **B** равны 2, 2, 6, 6, 0. Поскольку матрицы симметричны, каждую из них можно привести к диагональному виду ортогональным преобразованием (см. 4.081). Начиная с **B**, находим ортогональную матрицу

$$I = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & -\sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 & 1 & \sqrt{2} \\ 0 & 0 & 0 & -2 & \sqrt{2} \\ 0 & -\sqrt{3} & 0 & 1 & \sqrt{2} \end{pmatrix},$$

которая дает

$$I^{-1}AI = \frac{1}{4} \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 15 \end{pmatrix},$$

$$I^{-1}BI = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$I^{-1}CI = \frac{1}{2} \begin{pmatrix} -1 & \sqrt{2} & & & \\ \sqrt{2} & 0 & & 0 & 0 \\ \hline & & -1 & \sqrt{6} & \\ & 0 & \sqrt{6} & 0 & 0 \\ \hline & & & & 0 \end{pmatrix}.$$

Первые две матрицы диагональны, а третья имеет блочно диагональную форму. Теперь отдельные блоки можно привести к диагональному виду, не затрагивая  $I^{-1}AI$  и  $I^{-1}BI$ . Матрица этого второго преобразования

$$m = \frac{1}{\sqrt{15}} \begin{pmatrix} \sqrt{5} & -\sqrt{10} & & & \\ \sqrt{10} & \sqrt{5} & & 0 & 0 \\ \hline & & \sqrt{6} & -3 & \\ & & 3 & -\sqrt{6} & \\ \hline 0 & & & & \sqrt{15} \end{pmatrix}$$

дает

$$(Im)^{-1} C (Im) = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

**4.07. Миноры матрицы алгебраических дополнений. Теорема Якоби.** Пусть  $a$  — матрица порядка  $n$ . Обозначим ее определитель буквой  $D$ . Если  $M$  — минор порядка  $k$ , то минор порядка  $(n-k)$ , получающийся вычеркиванием из  $D$  всех строк и столбцов, содержащих элементы из  $M$ , называется *дополнением* к  $M$ . В частности, если  $M$  — отдельный элемент  $a_{ik}$ , то

дополнительный минор есть соответствующий первый минор. Но в этом случае часто бывает удобнее пользоваться *минором со знаком* или *алгебраическим дополнением*  $A_{ik}$ . Удобно также следующим образом определить алгебраическое дополнение минора  $M$ . Если  $M$  — минор порядка  $r$ , в котором представлены строки  $i_1, i_2, \dots, i_r$  и столбцы  $k_1, k_2, \dots, k_r$  определителя  $D$ , то алгебраическое дополнение  $M$  определяется равенством

$$\text{Алгебраическое дополнение } M = (-1)^{i_1+i_2+\dots+i_r+k_1+k_2+\dots+k_r} \times \\ \times (\text{дополнение } M).$$

Для главного минора  $i_1 = k_1, \dots, i_r = k_r$  и алгебраическое дополнение совпадает с дополнительным минором. По определению алгебраическое дополнение самого определителя считается равным 1.

Матрицу с элементами  $A_{ik}$  будем называть матрицей алгебраических дополнений. Присоединенная матрица [см. 4.01 (21)] получается из матрицы алгебраических дополнений транспонированием. Рассмотрим определитель

$$\Delta = \begin{vmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & \dots & \dots & A_{nn} \end{vmatrix}. \quad (1)$$

Можно доказать следующую теорему о соответствующих минорах определителей  $D$  и  $\Delta$ . Если  $M$  и  $M'$  — соответствующие миноры  $D$  и  $\Delta$  порядка  $r$ , то

$$M' = D^{r-1} \times (\text{алгебраическое дополнение } M). \quad (2)$$

В частности, для  $r = n$

$$\Delta = D^{n-1}. \quad (3)$$

Пусть  $r = n - 1$ . Тогда если  $\alpha_{ik}$  — алгебраическое дополнение элемента  $A_{ik}$  в  $\Delta$ , то

$$\alpha_{ik} = D^{n-2} a_{ik}. \quad (4)$$

Пусть  $r = 2$  и  $M$  и  $M'$  получаются вычеркиванием всех строк, кроме  $i$  и  $m$ , и всех столбцов, кроме  $k$  и  $p$ . Тогда

$$\begin{vmatrix} A_{ik} & A_{ip} \\ A_{mk} & A_{mp} \end{vmatrix} = D \times (\text{алгебраическое дополнение } M). \quad (5)$$



Эту общую теорему докажем сначала для частного случая, когда  $M$  и  $M'$  — ведущие миноры и  $D \neq 0$ . Тогда можно написать

$$M' = \begin{vmatrix} A_{11} & A_{21} & \dots & A_{r1} & 0 & 0 & \dots & 0 \\ A_{12} & \dots & \dots & A_{r2} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_{1r} & \dots & \dots & A_{rr} & 0 & 0 & \dots & 0 \\ A_{1,r+1} & \dots & \dots & A_{r,r+1} & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_{1,n} & \dots & \dots & A_{r,n} & 0 & 0 & \dots & 1 \end{vmatrix}. \quad (6)$$

Умножая  $M'$  на

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & a_{nn} \end{vmatrix}, \quad (7)$$

получим

$$DM' = \begin{vmatrix} D & 0 & \dots & 0 & a_{1,r+1} & \dots & a_{1,n} \\ 0 & D & \dots & 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & D & a_{r,r+1} & \dots & a_{r,n} \\ 0 & \dots & \dots & 0 & a_{r+1,r+1} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & a_{n,r+1} & \dots & a_{nn} \end{vmatrix} = \\ = D^r \times (\text{дополнение } M \text{ в } D), \quad (8)$$

что доказывает теорему. Следует отметить, что при нашем определении дополнения самого определителя  $D$  доказательство сохраняется и для  $r = n$ .

Если  $D = 0$ , то из 4.05 (7) следует, что миноры элементов каждых двух строк (или столбцов)  $D$  пропорциональны, и поэтому  $M' = 0$  при  $r \geq 2$  и теорема по-прежнему справедлива. Если  $D = 0$  и  $r = 1$ , то  $D^{r-1}$  не определено. Но  $M' = A_{11}$  — дополнению  $M$  в  $D$ .

Это доказательство распространяется на общий случай, когда  $M$  не является ведущим минором и достигается с помощью такого преобразования обоих определителей, чтобы минор  $M$  оказался ведущим. При этом нужно учитывать изменение знака при преобразовании определителя. Подробности доказательства можно найти в книге [4].

**4.08. Квадратичные и эрмитовы формы.** Функция  $n$  величин  $x_i$  вида  $a_{ik}x_ix_k$ , где  $a_{ik}$  действительны, называется *квадратичной формой*. Если  $y_i$  — другой набор величин, то выражение  $a_{ik}x_iy_k$  называется *билинейной формой*. В матричных обозначениях эти формы можно соответственно записать в виде  $\tilde{x}ax$  и  $\tilde{x}ay$ . В первом случае  $a_{ik}$  можно считать элементами симметричной матрицы. Действительно, если  $a_{12}x_1x_2$  — один член формы, а  $a_{21}x_2x_1$  — другой, то их сумма остается прежней при замене  $a_{12}$  и  $a_{21}$  на  $\frac{1}{2}(a_{12} + a_{21})$ . И аналогично для любой пары значений  $i$  и  $k$ .

Заменой переменных квадратичную форму всегда можно привести к сумме квадратов (не обязательно с положительными коэффициентами). Простой и очень полезный способ состоит в том, чтобы принять за новую переменную

$$\xi_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \dots + \frac{a_{1n}}{a_{11}}x_n. \quad (1)$$

Тогда при вычитании  $a_{11}\xi_1^2$  из  $a_{ik}x_ix_k$  все члены, содержащие  $x_1$ , пропадают. Затем, вводя новую переменную  $\xi_2$ , можно аналогичным способом исключить члены, содержащие  $x_2$ . В общем случае квадратичная форма будет приведена к сумме  $n$  квадратов. Тогда

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \dots \equiv a_{11}\xi_1^2 + \beta_2\xi_2^2 + \dots + \beta_n\xi_n^2, \quad (2)$$

где каждая переменная  $\xi_r$  представляется в виде суммы, начинающейся с  $x_r$ . Изложенный способ не пригоден в том случае, когда исходная квадратичная форма не содержит членов с квадратами — не с чего начать. Но члены с квадратами можно ввести простой заменой переменных. Пусть  $a_{ik} = 0$  при  $i \neq k$ , но  $a_{12} \neq 0$ . Положим  $x_2 = x_1 + \xi_2$ . Тогда  $x_1x_2 = x_1^2 + x_1\xi_2$  и дальше можно пользоваться описанным способом.

Отметим следующие свойства [1]:

1) Произведение коэффициентов правой части (2), взятых до члена с  $\beta_r$ , равно определителю, составленному из коэффициентов левой части, взятых до члена с  $x_r^2$ .

2) Квадратичную форму можно привести к сумме квадратов неограниченным числом способов. Один из них был только что изложен. Но независимо от способа приведения число положительных, отрицательных и нулевых коэффициентов остается одним и тем же для всех невырожденных преобразований, т. е. таких преобразований, при которых новые переменные взаимно однозначно связаны со старыми.

3) Форма  $a_{ik}x_i x_k$  называется *положительно определенной*, если она всегда положительно и обращается в нуль только при всех  $x_i$ , равных нулю. Для того чтобы квадратичная форма была положительно определенной, необходимо и достаточно выполнение таких условий:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \quad \dots, \quad \|a_{ik}\| > 0. \quad (3)$$

Квадратичная форма может быть существенно положительной, но не положительно определенной. Например,  $(x - 2y)^2 = 0$  при  $x = 2$ ,  $y = 1$ . Для таких форм  $\|a_{ik}\| = 0$ . Форма  $a_{ik}x_i x_k$  называется *отрицательно определенной*, если  $(-a_{ik}x_i x_k)$  — положительно определенная форма.

4) Если все определители (3) порядка больше  $r$  равны нулю, а определитель порядка  $r$  отличен от нуля, то квадратичную форму можно свести к сумме  $r$  квадратов. При этом говорят, что ранг формы равен  $r$ . Ранг матрицы  $a_{ik}$  тогда также равен  $r$ .

Если  $a_{ik}$  — эрмитова матрица, то  $a_{ik}x_i x_k^*$  называется *эрмитовой формой*. Эрмитова форма действительна, так как для любого члена  $a_{ik}x_i x_k^*$  можно найти член, получающийся перестановкой  $i$  и  $k$ , т. е.  $a_{ki}x_k x_i^*$ . Но для эрмитовой матрицы  $a_{ki} = a_{ik}^*$  и поэтому выписанные члены комплексно сопряженные, а их сумма действительна. Аналогично  $a_{ik}x_i y_k^*$  называется *эрмитовой билинейной формой*. Теория эрмитовых форм почти полностью совпадает с теорией квадратичных форм. Единственное отличие состоит в том, что (2) заменяется на

$$a_{ik}x_i x_k^* = a_{11}\xi_1\xi_1^* + \beta_2\xi_2\xi_2^* + \dots + \beta_n\xi_n\xi_n^*, \quad (4)$$

где все коэффициенты в правой части — действительные числа. Эрмитова форма будет положительно определенной при тех же условиях (3). Все определители в (3) действительны, так как они не изменяются при транспонировании, а такое преобразование приводит к замене всюду  $i$  на  $-i$ , и если бы какой-нибудь определитель был комплексным, то его мнимая часть должна была бы изменить знак.

**4.081. Приведение пары эрмитовых форм.** Вообще говоря, любую пару действительных квадратичных форм  $a_{ik}x_i x_k$ ,  $b_{ik}x_i x_k$  можно заменой переменных одновременно привести к сумме квадратов. Это утверждение всегда верно, когда одна из квадратичных форм положительно определенная. При тех же

условиях любую пару эрмитовых форм можно одновременно привести к виду (4). В обоих случаях ход рассуждений по существу одинаков. Здесь будет рассматриваться случай эрмитовой формы, так как он, с одной стороны, включает в себя случай положительно определенной квадратичной формы и, с другой стороны, имеет приложение в квантовой теории. Симметричные  $a_{ik}$  встречаются почти в каждой области физики. В этой книге был уже разобран простейший случай приведения симметричного тензора второго порядка при  $n=3$  к главным осям, что равносильно выбору  $a_{ik} = \delta_{ik}$ . Но и без этого условия можно две формы от трех переменных линейным преобразованием одновременно привести к сумме квадратов. Геометрически это значит, что любые две концентрические поверхности второго порядка имеют общую систему трех взаимно сопряженных диаметров. Случай  $n$  переменных возникает в теории малых колебаний динамических систем. Ясно, что если кинетическую и потенциальную энергии можно одновременно привести к виду

$$2T = m_1 \dot{\xi}_1^2 + m_2 \dot{\xi}_2^2 + \dots + m_n \dot{\xi}_n^2,$$

$$2V = m_1 \lambda_1 \xi_1^2 + m_2 \lambda_2 \xi_2^2 + \dots + m_n \lambda_n \xi_n^2,$$

то каждая переменная  $\xi_r$  будет удовлетворять дифференциальному уравнению

$$\ddot{\xi}_r = -\lambda_r \xi_r$$

и изменяться во времени независимо от остальных переменных. Поэтому условие устойчивости заключается в требовании положительности всех  $\lambda_r$ . Форма  $T$  положительно определенная, все  $m_r$  положительны, и поэтому для устойчивой системы форма  $V$  также должна быть положительно определенной.

Удобно начать с системы  $n$  уравнений

$$\lambda a_{ik} x_k = b_{ik} x_k. \quad (1)$$

Тогда общий метод решения приводит к уравнению для  $\lambda$

$$\|\lambda a_{ik} - b_{ik}\| = 0. \quad (2)$$

Для любого действительного  $\lambda$  матрица  $\|\lambda a_{ik} - b_{ik}\|$  эрмитова и поэтому ее определитель действительный. Следовательно, все коэффициенты в разложении определителя по степеням  $\lambda$  действительны. Будем считать, что форма  $a_{ik} x_i x_k^*$  положительно определенная.

Пусть  $\lambda_1$  и  $\lambda_2$  — два корня уравнения (2), и пусть  $x_{i1}$  и  $x_{i2}$  — соответствующие нетривиальные решения. Тогда

$$\lambda_1 a_{ik} x_{k1} = b_{ik} x_{k1}, \quad (3)$$

$$\lambda_2 a_{ik} x_{k2} = b_{ik} x_{k2}. \quad (4)$$

Умножим (3) на  $x_{i2}^*$  и сложим; умножим (4) на  $x_{i1}^*$  и тоже сложим. Тогда

$$\lambda_1 a_{ik} x_{k1} x_{i2}^* = b_{ik} x_{k1} x_{i2}^*, \quad (5)$$

$$\lambda_2 a_{ik} x_{k2} x_{i1}^* = b_{ik} x_{k2} x_{i1}^*. \quad (6)$$

При помощи комплексного сопряжения из (5) получим

$$\lambda_1^* a_{ik}^* x_{k1}^* x_{i2} = b_{ik}^* x_{k1}^* x_{i2}. \quad (7)$$

Перестановка немых индексов дает

$$a_{ik}^* x_{k1}^* x_{i2} = a_{ki} x_{i2} x_{k1}^* = a_{ik} x_{k2} x_{i1}^*. \quad (8)$$

Аналогично

$$b_{ik}^* x_{k1}^* x_{i2} = b_{ik} x_{k2} x_{i1}^*. \quad (9)$$

Из сравнения (8) и (9) с (6) следует

$$(\lambda_2 - \lambda_1^*) a_{ik} x_{k2} x_{i1}^* = 0. \quad (10)$$

Примем сначала, что сравниваемые решения совпадают. Тогда выражение  $a_{ik} x_{k1} x_{i1}^*$  действительно. Если бы оно было равно нулю, то по (5) было бы равно нулю и  $b_{ik} x_{k1} x_{i1}^*$ . Это невозможно, если одно из этих выражений представляет собой положительно определенную форму. Отсюда следует, что  $\lambda_1 = \lambda_1^*$  и поэтому  $\lambda_1$  — действительное число. Следовательно, все корни (2) действительны и для действительных форм отношения  $x_{i1}$  также действительны.

Пусть теперь  $\lambda_1$  и  $\lambda_2$  различны. Тогда, поскольку они действительны,  $\lambda_2 \neq \lambda_1$  и, следовательно,

$$a_{ik} x_{k2} x_{i1}^* = 0, \quad b_{ik} x_{k2} x_{i1}^* = 0. \quad (11)$$

Это свойство решений будем называть ортогональностью относительно матриц **a** и **b** или **a**-ортогональностью.

В каждом простом корне (2) определитель имеет производную по  $\lambda$ , отличную от нуля. Но эта производная представляет собой линейную функцию первых миноров определителя, и, следовательно, не все первые миноры равны нулю. Отсюда следует, что в простом корне уравнения (2) ранг матрицы

$$c_{ik} = \lambda a_{ik} - b_{ik}$$

равен  $n - 1$ . В  $r$ -кратном корне производная порядка  $r$  по  $\lambda$  отлична от нуля и является линейной функцией миноров порядка  $r$ . Поэтому ранг  $c_{ik}$  не меньше  $n - r$ . Мы покажем, что на самом деле ранг равен  $n - r$ . Но для этого требуется сравнительно сложное доказательство теоремы о разложении корней.

Рассмотрим сначала случай, когда все корни простые. Тогда для каждого корня, например для  $\lambda_1$ , отношения  $x_i$  определяются однозначно. Можно написать

$$x_{i1} = l_{i1} \xi_1. \quad (12)$$

Очевидно, что

$$\lambda_1 a_{ik} l_{k1} = b_{ik} l_{k1}, \quad (13)$$

$$\lambda_1 a_{ik} l_{k1}^* l_{i1}^* = b_{ik} l_{k1}^* l_{i1}^*. \quad (14)$$

Выражение  $a_{ik} l_{k1} l_{i1}^*$  действительно и отлично от нуля. Кроме того, для различных  $\lambda_1$  и  $\lambda_2$  из (11) следует

$$a_{ik} l_{k1} l_{i2}^* = 0, \quad b_{ik} l_{k1} l_{i2}^* = 0. \quad (15)$$

Если обозначить различные решения индексом  $j$ , то  $l_{ij}$  можно рассматривать как матрицу преобразования

$$l_{ij}^* a_{ik} l_{kl} = (\mathbf{I}^+ \mathbf{a})_{jl}. \quad (16)$$

Из (15) следует, что все недиагональные элементы  $\mathbf{I}^+ \mathbf{a}$  и  $\mathbf{I}^+ \mathbf{b}$  равны нулю. Следовательно, это преобразование приводит обе матрицы к диагональному виду. Если обозначить диагональные элементы матриц через  $A_{jl}$  и  $B_{jl}$  ( $j = l$ ), то по (14) отношение  $A_{jl}/B_{jl}$  равно соответствующему значению  $\lambda$ . Кроме того,

$$\|l_{ji}^*\| \|a_{ik}\| \|l_{kl}\| = \|l_{ji}^* a_{ik} l_{kl}\|. \quad (17)$$

Этот определитель представляет собой произведение диагональных элементов, каждый из которых отличен от нуля. Следовательно,  $\mathbf{I}$  и  $\mathbf{I}^+$  — неособые матрицы. Если теперь в общем случае положим

$$X_i = l_{ij} \xi_j, \quad (18)$$

то эти уравнения можно разрешить относительно  $\xi_j$  при любых заданных значениях  $X_i$ . Тогда

$$a_{ik} X_k X_i^* = a_{ik} l_{kl} \xi_l l_{ij}^* \xi_j^* = A_{il} \xi_l \xi_j^*, \quad (19)$$

$$b_{ik} X_k X_i^* = B_{il} \xi_l \xi_j^*. \quad (20)$$

Таким образом, в обеих формах отсутствуют все члены с  $j \neq l$ .

Важный частный случай  $a_{ik} = \delta_{ik}$ , так что исходные уравнения сводятся к

$$\lambda x_i = b_{ik} x_k.$$

В этом случае  $a_{ik} X_i X_k^* = |X_i|^2 + \dots + |X_n|^2$  и представляет собой положительно определенную форму, и матрица

$$A_{jl} = l_{ji}^+ \delta_{ik} l_{kl} = l_{ji}^+ l_{il}$$

должна быть диагональной. Но для каждого корня  $\lambda_j$  величины  $l_{ij}$  определены с точностью до постоянного множителя. Выберем его так, чтобы

$$l_{ij}^* l_{il} = 1 \quad (j = l).$$

Всегда  $l_{ij}^* l_{il} = 0$  при  $j \neq l$ . Отсюда

$$A_{jl} = \delta_{ik},$$

и это преобразование не изменяет единичную матрицу, значит,  $l$  — унитарная матрица. Следовательно, любую эрмитову матрицу  $\mathbf{b}$ , не имеющую кратных собственных значений, можно привести к диагональному виду при помощи преобразования  $\mathbf{l}^+ \mathbf{b} \mathbf{l}$ , где  $\mathbf{l}$  — унитарная матрица. Преобразование  $\mathbf{X} = \mathbf{l} \xi$  сводит  $b_{ik} X_i X_k^*$  к  $\sum \lambda_j X_j X_j^*$ . Если  $\mathbf{b}$  — вещественная симметричная матрица, то  $\mathbf{l}$  — вещественная унитарная, т. е. ортогональная, матрица.

**4.082. Теорема о разделении корней.** Эта теорема напоминает теорему Штурма в теории многочленов. Как и прежде, будем считать  $\mathbf{a}$  и  $\mathbf{b}$  эрмитовыми матрицами порядка  $n \times n$  и  $\mathbf{x}^+ \mathbf{a} \mathbf{x}$  положительно определенной формой. Запишем

$$\lambda a_{ik} - b_{ik} = c_{ik}, \quad \lambda \mathbf{a} - \mathbf{b} = \mathbf{c}. \quad (1)$$

Обозначим определитель  $\|c_{ik}\|$  через  $D$ . Было показано, что если  $\lambda$  — любой корень уравнения  $D = 0$ , то ранг  $\mathbf{c}$  не может быть больше чем  $n - 1$ , и что если  $\lambda$  —  $r$ -кратный корень, то ранг  $\mathbf{c}$  не может быть меньше  $n - r$ . Покажем теперь, что ранг действительно равен  $n - r$ .

Обозначим ведущие миноры определителя  $D$  порядка  $n$ ,  $n - 1$ , ..., 1, 0 через

$$D_n (= D), \quad D_{n-1}, \dots, D_1 (= c_{11}), \quad D_0 (= 1).$$

Примем сначала, что  $D$  не имеет кратных корней и не существует соседних  $D_r$ ,  $D_{r-1}$ , обращающихся в нуль для одного и того же значения  $\lambda$ . Проследим при любом  $\lambda$  перемену знаков миноров в последовательности от  $D_n$  до  $D_0$ .

Из теоремы Якоби следует, что если  $C_{ik}$  — алгебраическое дополнение элемента  $c_{ik}$  в определителе  $D$ , то

$$DD_{n-2} = \begin{vmatrix} C_{n-1, n-1} & C_{n-1, n} \\ C_{n, n-1} & C_{nn} \end{vmatrix} = D_{n-1}C_{n-1, n-1} - |C_{n-1, n}|^2. \quad (2)$$

Аналогично если  $C_{ik}^s$  — алгебраическое дополнение элемента  $c_{ik}$  в  $D_s$ , то

$$D_s D_{s-2} = D_{s-1} C_{s-1, s-1}^s - |C_{s-1, s}^s|^2. \quad (3)$$

Если  $\lambda$  — корень  $D_{s-1}$ , то, поскольку для этого значения  $\lambda$  ни  $D_s$ , ни  $D_{s-2}$  не обращаются в нуль, произведение  $D_s D_{s-2}$

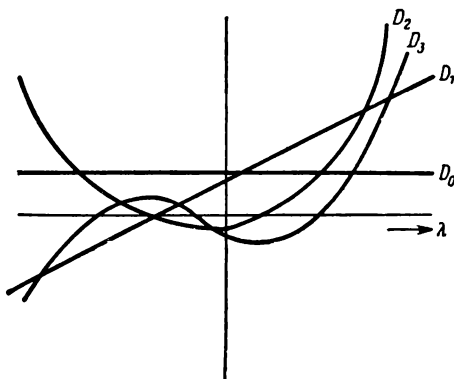


Рис. 22.

не равно нулю и должно быть отрицательным. Следовательно,  $D_s$  и  $D_{s-2}$  имеют противоположные знаки при  $D_{s-1} = 0$ . Далее, в силу того что  $a_{ik}x_i x_k^*$  — положительно определенная форма, при  $\lambda \rightarrow +\infty$  все  $D_s$  положительны. При  $\lambda \rightarrow -\infty$  они попеременно положительны и отрицательны. Следовательно, при возрастании  $\lambda$  от  $-\infty$  до  $+\infty$  учитывается  $n$  перемен знаков в последовательности миноров. Однако можно пока-

зать, что число перемен знаков изменяется только, когда при изменении  $\lambda$  меняется знак самого  $D$ . Действительно, когда  $\lambda$  достигает нуля минора  $D_{s-1}$ , миноры  $D_s$  и  $D_{s-2}$  имеют противоположные знаки и происходит по крайней мере одна перемена знака в последовательности  $D_s, D_{s-1}, D_{s-2}$ , а число перемен знаков в последовательности от  $D_n$  до  $D_0$  не меняется. Но  $D_0 = 1$  вовсе не меняет знак. Следовательно, все  $n$  перемен знаков теряются в начале последовательности, т. е. когда  $\lambda$  проходит через  $n$  действительных нулей определителя  $D$ .

На рис. 22 показано относительное расположение нулей  $D_n, D_{n-1}, \dots$ . График  $D_0$  — прямая, параллельная оси  $\lambda$ . График  $D_1$  — прямая линия, проходящая из  $-\infty$  при  $\lambda = -\infty$  в  $+\infty$  при  $\lambda = +\infty$ . Уравнение  $D_1 = 0$  имеет один действительный корень.

При  $\lambda \rightarrow \pm\infty$  минор  $D_2 \rightarrow +\infty$  и отрицателен при  $D_1 = 0$ . Поэтому уравнение  $D_2 = 0$  имеет два действительных корня. Один корень меньше, а другой больше корня уравнения  $D_1 = 0$ .



Минор  $D_3 \rightarrow -\infty$  при  $\lambda \rightarrow -\infty$  и  $\rightarrow +\infty$  при  $\lambda \rightarrow +\infty$ . При  $D_2 = 0$  и  $D_1 < 0$  должно быть  $D_3 > 0$ ; при  $D_2 = 0$  и  $D_1 > 0$  должно быть  $D_3 < 0$ . Поэтому корни  $D_2 = 0$  разделяют корни  $D_3 = 0$ . Аналогично корни  $D_{s+1} = 0$  разделены корнями  $D_s = 0$ . Это доказывает теорему о разделении корней в случае, когда все корни уравнения  $D_n = 0$  различны.

Если  $D_n$  имеет  $r$ -кратный корень  $\lambda = \lambda_p$ , то заменой  $b_{ik}$  близкими значениями  $b'_{ik}$  можно изменить это уравнение так, что все корни станут различными. Если  $D_s$  и  $D_{s-1}$  обращаются в нуль при одном и том же  $\lambda$ , то  $D_s$  можно изменить, оставив прежним  $D_{s-1}$ , например путем изменения  $b_{1s}$ . Следовательно, если поставленное условие не выполняется, то можно добиться его выполнения, задавая малые возмущения элементам матрицы  $b$ .

Рассмотрим группу  $r$  корней, которые сливаются в  $\lambda_p$  при  $b'_{ik} \rightarrow b_{ik}$ . Между этими корнями лежат  $r-1$  корней  $D_{n-1}$ ,  $r-2$  корней  $D_{n-2}$ , ..., один корень  $D_{n-r+1}$  и ни одного корня  $D_{n-r}$ .

Перестановкой строк и столбцов можно переместить в левый верхний угол любой элемент матрицы, причем матрица останется эрмитовой. Отсюда следует, что если  $\lambda_p$  — корень  $D = 0$  кратности  $r$ , то все главные миноры порядка от  $n$  до  $n-r+1$  обращаются в нуль при  $\lambda = \lambda_p$ .

Главные миноры порядка  $n-r$  не могут обращаться в нуль одновременно, так как тогда  $\lambda_p$  было бы корнем уравнения  $D = 0$  кратности  $r+1$ .

Было показано, что если определитель и один из его первых миноров равны нулю, то равны нулю либо первые миноры всех элементов соответствующей строки, либо первые миноры всех элементов соответствующего столбца. Возьмем теперь какой-либо главный минор порядка  $n-r+2$ . Этот минор и любой его первый главный минор обращаются в нуль при  $\lambda = \lambda_p$ , причем последний является главным минором порядка  $n-r+1$  определителя  $D$ . Пусть вычеркнуты строка и столбец с номерами  $i$ . Тогда равны нулю либо все миноры элементов  $c_{ik}$ , либо все миноры элементов  $c_{ki}$ . Но эти миноры комплексно сопряженные, и поэтому обе системы миноров обращаются в нуль одновременно. Следовательно, все миноры порядка  $n-r+1$  в  $D$  равны нулю и при  $\lambda = \lambda_p$  ранг матрицы  $c_{ik}$  равен  $n-r$ , что и требовалось доказать.

Кроме того,  $\lambda = \lambda_p$  по крайней мере простой корень всех миноров порядка  $n-r+1$  и выше. Но производная любого минора есть линейная функция миноров порядка, меньшего на единицу. Следовательно, миноры порядка  $n-r+2$  имеют двукратные корни, а миноры порядка  $n-1$ , т. е. первые миноры

определителя  $D$ , имеют корни кратности по крайней мере  $r - 1$ . Мы снова встретимся с этим результатом при рассмотрении малых колебаний операционным методом \*).

**4.083. Расширение свойства ортогональности.** Мы доказали ортогональность  $l_{ij}$ , соответствующих любой паре различных корней детерминантного уравнения. Если  $\lambda = \lambda_j$  — корень кратности  $r$ , то в силу того, что при  $\lambda = \lambda_j$  ранг матрицы  $c_{ik}$  равен  $n - r$ , можно выбрать  $r$  линейно независимых наборов отношений величин  $l_{ij}$ . Пусть два из них  $l_{i1}$  и  $l_{i2}$ . Тогда

$$\lambda_j a_{ik} l_{k1} = b_{ik} l_{k1}, \quad \lambda_j a_{ik} l_{k2} = b_{ik} l_{k2}, \quad (1)$$

и поэтому

$$\lambda_j a_{ik} l_{k2} l_{i1}^* = b_{ik} l_{k2} l_{i1}^*. \quad (2)$$

Эти выражения могут оказаться равными нулю. Если это так, то решения по-прежнему а-ортогональны. Если этого нет, то рассмотрим

$$m_{i2} = l_{i2} - \theta l_{i1}, \quad (3)$$

где  $\theta$  не зависит от  $i$ . Не все  $m_{i2}$  равны нулю, так как  $l_{i1}$  и  $l_{i2}$  не пропорциональны. Тогда для любого  $\theta$

$$\lambda_j a_{ik} m_{k2} = b_{ik} m_{k2}, \quad (4)$$

$$a_{ik} m_{k2} l_{i1}^* = a_{ik} l_{k2} l_{i1}^* - \theta a_{ik} l_{k1} l_{i1}^*. \quad (5)$$

Соответствующим выбором  $\theta$  можно обратить это выражение в нуль. Тогда из (2) следует, что  $b_{ik} m_{k2} l_{i1}^*$  также обращается в нуль.

Если  $\lambda_j$  — корень кратности  $r$ , то для него можно взять один из наборов  $l_{ij}$  и затем остальные сделать а-ортогональными ему. Для этого из каждого набора следует вычесть первый, умноженный на соответствующую величину. Выбирая из этих преобразованных наборов один, можно сделать остальные а-ортогональными ему. Продолжая действовать подобным образом, получим в результате  $r$  наборов, взаимно а-ортогональных. Таким образом, даже если уравнение для  $\lambda$  имеет кратные корни, все же можно выбрать  $n$  наборов  $l_{ij}$ , обладающих свойствами ортогональности, и по-прежнему  $a_{ik}$  и  $b_{ik}$  можно привести к диагональному виду (в действительности

---

\* Впервые доказано Вейерштрассом [5]. Имеется несколько других доказательств: одно в книге Лэмба [6], а другое у Бромвича [7]. Второе доказательство представляет особый интерес, так как в нем непосредственно находят нормальные координаты и квадратичные формы сводят к сумме квадратов последовательной подстановкой. В настоящей книге использован способ Рауса, как наиболее легкий при изучении гироскопических систем.

неограниченным числом способов). Если  $a_{ik} = \delta_{ik}$ , то, как и прежде, матрицу  $I$  можно считать унитарной.

Если требуется ввести нормальные координаты, соответствующие частному решению, то можно действовать следующим образом. Пусть  $x_i = l_{i1}\xi_1$  есть решение. Возьмем произвольный набор  $x_i$  и определим  $\xi_1$  так, чтобы выражение  $a_{ik}(x_k - l_{k1}\xi_1) \times \times (x_i^* - l_{i1}^*\xi_1^*)$  было стационарным относительно малых изменений  $\xi_1$ . Можно изменять и действительную и мнимую части  $\xi_1$ , однако производные по действительной и по мнимой частям  $\xi_1$  обращаются в нуль тогда и только тогда, когда обращаются в нуль производные по  $\xi_1$  и  $\xi_1^*$  как по независимым переменным. Отсюда

$$l_{i1}^* a_{ik} (x_k - l_{k1}\xi_1) = 0, \quad l_{k1} a_{ik} (x_i^* - l_{i1}^*\xi_1^*) = 0.$$

Второе равенство сводится к первому перестановкой  $i$  и  $k$  и операцией комплексного сопряжения. Следовательно,

$$A_{i1}\xi_1 = l_{i1}^* a_{ik} x_k.$$

Если, скажем,  $x_k = l_{k2}\xi_2$ , то  $l_{i1}^* a_{ik} x_k = 0$  в силу ортогональности. Следовательно, для любого собственного вектора можно непосредственно по  $x_i$  без определения всех собственных векторов найти то  $\xi$ , которое отлично от нуля.

Возьмем в качестве примера

$$a = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & 1 \\ 1 & 1 & 9 \end{pmatrix}, \quad b = \begin{pmatrix} 1 & 3 & 12 \\ 3 & 1 & -8 \\ 12 & -8 & -2 \end{pmatrix}.$$

Форма  $\tilde{h}ax$  положительно определена и корни уравнения  $\det(b - \lambda a) = 0$  равны 2, 2, -3.

При  $\lambda = 2$  все три уравнения сводятся к одному

$$x_1 - x_2 - 2x_3 = 0,$$

и в качестве решений можно взять (1, 1, 0) и (0, -2, 1). Они дают

$$\tilde{l}_1 a l_1 = 4, \quad \tilde{l}_1 a l_2 = -2.$$

Следовательно, решение, ортогональное первому решению, имеет вид  $\tilde{m}_2 = = 2 \left( \tilde{l}_2 + \frac{1}{2} \tilde{l}_1 \right) = (1, -3, 2)$ . Для  $\lambda = -3$  решение  $\tilde{l}_3 = (-3, 1, 2)$ .

Тогда можно положить

$$x_1 = \xi_1 + \xi_2 - 3\xi_3, \quad x_2 = \xi_1 - 3\xi_2 + \xi_3, \quad x_3 = 2\xi_2 + 2\xi_3.$$

Кроме того,

$$\tilde{m}_2 a m_2 = 64, \quad \tilde{l}_3 a l_3 = 64,$$

откуда

$$\tilde{h}ax = 4\xi_1^2 + 64\xi_2^2 + 64\xi_3^2 \quad \text{и} \quad \tilde{h}bx = 8\xi_1^2 + 128\xi_2^2 - 192\xi_3^2.$$

Чтобы проиллюстрировать, как обстоит дело в том случае, когда ни одна из форм, соответствующих матрицам, не является положительно определенной, возьмем формы  $\frac{1}{2}(x_1^2 - x_2^2)$  и  $x_1 x_2$ . Соответствующие им уравнения имеют вид

$$\lambda x_1 = x_2, \quad -\lambda x_2 = x_1,$$

откуда  $\lambda^2 = -1$  и  $\lambda = \pm i$ . Поэтому можно взять

$$x_1 = \xi_1 + \xi_2, \quad x_2 = i(\xi_1 - \xi_2),$$

что приводит к

$$\frac{1}{2}(x_1^2 - x_2^2) = \xi_1^2 + \xi_2^2, \quad xy = i(\xi_1^2 - \xi_2^2).$$

Таким образом, формы приведены к сумме квадратов, но не все коэффициенты получились действительными. Здесь  $a_{ik}l_{kl}^*l_{il}^* = 0$ .

Эрмитову матрицу всегда можно привести к диагональному виду при помощи преобразования  $\mathbf{l}^{-1}\mathbf{a}\mathbf{l}$ , где  $\mathbf{l}$  — унитарная матрица. Следовательно, для эрмитовых матриц всегда удовлетворяется установленное в 4.065 условие того, что матрицы можно привести к диагональному виду при помощи преобразования подобия.

Специфику случая кратных корней можно геометрически проиллюстрировать на примере задачи об определении главных осей эллипсоида вращения. Любой диаметр, лежащий в плоскости круговой симметрии, представляет собой главную ось. Если произвольно выбрать в этой плоскости два направления  $\mathbf{l}_1$  и  $\mathbf{l}_2$ , то в общем случае они не окажутся перпендикулярными. Но их можно использовать для построения вектора  $\mathbf{l}_2 - \theta \mathbf{l}_1$ , лежащего в той же плоскости и перпендикулярного  $\mathbf{l}_1$ .

**4.084. Свойство стационарности  $\lambda$ . Принцип Релея.** Для любой совокупности  $x_i$  можно задать величину  $\lambda$  в виде

$$\lambda = \frac{b_{ik}x_kx_i^*}{a_{ik}x_kx_i^*} = \frac{\sum \lambda_j A_{jj} \xi_j \xi_j^*}{\sum A_{jj} \xi_j \xi_j^*}. \quad (1)$$

Эта величина всегда заключена между наименьшим и наибольшим из  $\lambda_j$ , так как все  $A_{jj} > 0$ . Если все  $\xi_j$ , кроме одного, равны нулю, то отношение (1) становится равным соответствующему  $\lambda_j$ . Если остальные  $\xi_j$  малы, то отклонения  $\lambda$  от  $\lambda_j$  второго порядка малости. Следовательно, эта дробь при  $\lambda = \lambda_j$  стационарна относительно малых изменений всех  $\xi_j$  и, следовательно,  $x_i$ . В этом состоит *принцип Релея*.

Обратно, отношение (1) не имеет других стационарных значений. Поскольку

$$\lambda a_{ik}x_kx_i^* = b_{ik}x_kx_i^* \quad (2)$$

и если  $\delta\lambda$  второго порядка малости, то с точностью до первого порядка малости для любых действительных и мнимых вариаций  $x_i$  имеем

$$\lambda\delta(a_{ik}x_kx_i^*) = \delta(b_{ik}x_kx_i^*), \quad (3)$$

откуда

$$\lambda a_{ik}x_k = b_{ik}x_k, \quad \lambda a_{ik}x_i^* = b_{ik}x_i^*. \quad (4)$$

Второе соотношение (4) сводится к первому перестановкой  $i$  и  $k$  и операцией комплексного сопряжения.

Принцип Релея часто полезен при численном анализе. С его помощью без составления и решения детерминантного уравнения можно грубо определить отношения значений  $x_i$  в отдельном решении, обычно с наименьшим  $\lambda$ . Если полученные значения отношений  $x_i$  подставить в (1) и вычислить  $\lambda$ , то получается значение с ошибкой второго порядка малости. Может оказаться, что такая точность достаточна. В противном случае полученную величину можно использовать для уточнения приближенных значений отношений  $x_i$ . Если все  $\lambda_j$  одного знака, то получающееся приближенное значение  $\lambda$  по абсолютной величине больше, чем наименьшее из  $\lambda_j$ . Поэтому непосредственное использование этого принципа не может дать заниженной оценки наименьшего корня \*).

**4.09. Малые колебания.** Если отклонение координат динамической системы от их значения в положении равновесия обозначить через  $x_1 \dots x_n$ , то кинетическая энергия представляет собой положительно определенную квадратичную форму и записывается в виде  $T = \frac{1}{2} a_{ik}\dot{x}_i\dot{x}_k$ . Силовую функцию \*\*) можно записать в виде  $W = \frac{1}{2} b_{ik}x_ix_k$ . Тогда уравнения Лагранжа будут

$$a_{ik}\ddot{x}_k = b_{ik}x_k. \quad (1)$$

Ищем решение в виде

$$x_i \sim e^{\gamma t}, \quad (2)$$

где  $\gamma$  постоянная. Тогда

$$\gamma^2 a_{ik}x_k = b_{ik}x_k. \quad (3)$$

\*) Полное изложение дано Темплом и Бикли [8].

\*\*) В большинстве задач динамики несколько удобнее использовать силовую функцию, а не потенциальную энергию  $V = -W$ . Исключение представляет тот случай, когда  $W$  — отрицательно определенная форма, а следовательно,  $V$  — положительно определенная.

Получилась изученная выше система уравнений, в которой  $\lambda$  заменена на  $\gamma^2$ . Следовательно, все значения  $\gamma^2$  действительны. Если среди них имеются положительные, то система неустойчива. Условие устойчивости заключается в том, чтобы все значения  $\gamma^2$  были отрицательны. Посмотрим, что это означает для  $T$  и  $W$ . Обращаясь к 4.082, мы видим, что все перемены знаков в последовательности определителей должны быть утрачены при возрастании  $\gamma^2$  от  $-\infty$  до 0. Поэтому все определители последовательности

$$\| -b_{ik} \| \dots \begin{vmatrix} -b_{11} & -b_{12} & -b_{13} \\ -b_{21} & -b_{22} & -b_{23} \\ -b_{31} & -b_{32} & -b_{33} \end{vmatrix}, \quad \begin{vmatrix} -b_{11} & -b_{12} \\ -b_{21} & -b_{22} \end{vmatrix}, \quad -b_{11} \quad (4)$$

положительны, а это есть условие того, что  $W$  — отрицательно определенная форма или, наоборот, потенциальная энергия  $V = -W$  — положительно определенная форма.

Очевидно, что подстановка

$$x_i = l_{ij} \xi_j, \quad (5)$$

которая приводит  $a_{ik} x_i x_k$  к сумме квадратов  $\xi_j$ , будет приводить одновременно  $a_{ik} \dot{x}_i \dot{x}_k$  к сумме квадратов  $\dot{\xi}_j$ . Из общих свойств квадратичных форм известно, что  $T$  и  $W$  даже в том случае, когда определитель системы уравнений имеет кратные нули, можно одновременно привести к сумме квадратов (для устойчивой системы все коэффициенты при квадратах в силовой функции отрицательны). Тогда метод Лагранжа дает для  $\xi_j$  независимые дифференциальные уравнения по времени, и поэтому каждое  $\xi_j$  изменяется независимо от остальных. Такие переменные называются *нормальными координатами* системы. На практике такой способ отыскания решения оказывается трудоемким, и в тех случаях, когда требуется найти не только периоды собственных колебаний, более прост операционный метод, который будет изложен позже. Тем не менее метод нормальных координат объясняет свойство, которое обнаруживается операционным методом в том случае, когда уравнение частот имеет кратные корни. Сейчас нам нужно лишь показать, что в этом случае происходит с нормальными координатами. Рассмотрим свободное колебание маятника в произвольном горизонтальном направлении. Если  $x_1$  и  $x_2$  — горизонтальные прямоугольные координаты, то в общепринятых обозначениях с точностью до квадратичных членов получим

$$2T = m(\dot{x}_1^2 + \dot{x}_2^2), \quad 2W = -\frac{mg}{l}(x_1^2 + x_2^2).$$

Каждое из выражений уже имеет вид суммы квадратов, и, следовательно,  $x_1$  и  $x_2$  — нормальные координаты. Но можно было бы взять

$$x_1 = \xi_1 \cos \alpha - \xi_2 \sin \alpha, \quad x_2 = \xi_1 \sin \alpha + \xi_2 \cos \alpha,$$

и тогда

$$2T = m(\dot{\xi}_1^2 + \dot{\xi}_2^2), \quad 2W = -\frac{mg}{l}(\xi_1^2 + \xi_2^2),$$

так что  $\xi_1$ ,  $\xi_2$  — также нормальные координаты. Совпадение собственных периодов приводит к тому, что нормальные координаты можно выбирать неограниченным числом разных способов.

#### 4.091. Малые колебания около установившегося движения.

В таких задачах, впервые систематически исследованных Раусом, уравнения движения имеют вид

$$a_{ik}\ddot{x}_k + g_{ik}\dot{x}_k - b_{ik}x_k = 0, \quad (1)$$

где матрицы  $a_{ik}$  и  $b_{ik}$  симметричны, а матрица  $g_{ik}$  антисимметрична. Все величины действительны, и  $a_{ik}x_ix_k$  — положительно определенная квадратичная форма. Достаточное условие устойчивости получается умножением (1) на  $\dot{x}_i$  и суммированием по  $i$ . Члены с  $g_{ik}$  пропадают, и получается

$$\frac{d}{dt}(a_{ik}\dot{x}_i\dot{x}_k - b_{ik}x_ix_k) = 0. \quad (2)$$

Первый член в скобках является положительно определенным, и, следовательно, если  $b_{ik}x_ix_k$  — отрицательно определенная форма, то ее абсолютная величина не может превзойти значение, заданное начальными условиями. Поэтому при этих условиях система устойчива. Для колебаний около положения равновесия это условие является также и необходимым. Однако для колебаний около установившегося движения оно не является необходимым. Иначе, например, волчок не мог бы стоять. Способ решения для рассматриваемых систем очевиден. Тем не менее обсудим его основные черты.

Если, как и прежде, принять  $x_i \sim e^{\gamma t}$ , то получим уравнение совместности

$$D = \|a_{ik}\gamma^2 + g_{ik}\gamma - b_{ik}\| = 0. \quad (3)$$

Если обозначить общий элемент определителя через  $c_{ik}$ , то матрица  $c_{ik}$  эрмитова при чисто мнимых  $\gamma$ . Если же  $\gamma$  действительно, то матрица действительна и не симметрична. Поэтому предшествующие рассуждения нуждаются в некоторых изменениях. Далее, общий элемент содержит  $\gamma$  в трех степенях:

0, 1 и 2. Следовательно, теория малых колебаний около стационарного движения в двух отношениях сложнее теории, которую теперь можно назвать теорией простых эрмитовых матриц, члены которых содержат  $\lambda$  в степенях 0 и 1.

Замена  $\gamma$  на  $-\gamma$  с последующим транспонированием не меняет  $D$ . Следовательно,  $\gamma$  входит в равные и противоположные пары. При мнимом  $\gamma$  отношения  $x_i$ , как правило, оказываются комплексными и в каждом простом колебании фазы координат различны. Сведение к нормальным координатам невозможно.

Для формулирования условия устойчивости образуем, как и раньше, миноры  $D_n (=D)$ ,  $D_{n-1}$ , ...,  $D_0$ . По-прежнему имеем

$$D_n D_{n-2} = D_{n-1} C_{n-1, n-1}^n - C_{n-1, n} C_{n, n-1}. \quad (4)$$

Если  $\gamma$  чисто мнимое, то

$$C_{n-1, n} C_{n, n-1} = |C_{n-1, n}|^2 > 0. \quad (5)$$

Как и раньше, при возрастании  $\gamma^2$  от  $-\infty$  до 0 изменения знака могут утрачиваться только в начале последовательности миноров. Но если  $\gamma^2$  положительно, то  $C_{n-1, n}$  и  $C_{n, n-1}$  уже не будут комплексно сопряженными и их произведение не обязательно положительно. Поэтому теорема о разделении корней остается верной, лишь пока  $\gamma^2$  отрицательно. Если, как это и должно быть,  $a_{ik} \dot{x}_i \dot{x}_k$  — положительно определенная форма, и если, кроме того,  $b_{ik} x_i x_k$  — отрицательно определенная форма, то точно так, как и прежде, должно быть  $n$  действительных отрицательных значений  $\gamma^2$  и система устойчива. В этом случае устойчивость была найдена непосредственно из уравнений движения. Но если  $b_{ik} x_i x_k$  не является отрицательно определенной формой, то это не значит, что система неустойчива. Здесь в отличие от случая малых колебаний около положения равновесия изменение знака может быть *приобретено* в начале последовательности миноров. Это не происходило раньше, так как, во всяком случае,  $n$  перемен знака терялось при изменении  $\gamma^2$  от  $-\infty$  до  $+\infty$  и если бы на какой-то стадии приобретались перемены знака, то они должны бы были компенсироваться дополнительными потерями, иначе мы получили бы алгебраическое уравнение  $n$ -й степени с более чем  $n$  корнями. Изменение знака может приобретаться в случае гироскопических систем. Следовательно, для гироскопических систем, у которых  $b_{ik} x_i x_k$  не является отрицательно определенной формой, нельзя установить факт неустойчивости без подробного исследования уравнения частот.

Если система неустойчива, то  $\gamma^2$  может и не быть действительным. Теорема о разделении корней не верна для положи-



тельных  $\gamma^2$ , и поэтому перемена знака может быть потеряна тогда, когда  $\gamma^2$  проходит через нуль одного из промежуточных членов последовательности миноров. Таким образом, может быть утрачено четное число перемен знаков вне связи с действительными корнями  $D$ . Определитель  $D$  должен поэтому иметь комплексные корни.

Поскольку для отрицательно определенной формы  $b_{ik}x_ix_k$  теорема о разделении корней верна, следовательно, как и прежде, если  $D$  имеет корень порядка  $r$  для отрицательного  $\gamma^2$ , то для этого значения  $\gamma^2$  ранг матрицы  $c_{ik}$  равен  $n - r$  и каждый из первых миноров определителя  $D$  имеет корень порядка  $r - 1$ . Все это, однако, не верно, если  $b_{ik}x_ix_k$  не является отрицательно определенной.

**4.092.** Проиллюстрируем изложенное на примере вертикального волчка. В качестве координат  $x_1$  и  $x_2$  выберем косинусы углов, образуемых осью волчка с каждой из двух взаимно перпендикулярных горизонтальных осей. В обычных обозначениях уравнения движения имеют вид

$$A\ddot{x}_1 + Cn\dot{x}_2 - Mghx_1 = 0, \quad A\ddot{x}_2 - Cn\dot{x}_1 - Mghx_2 = 0. \quad (1)$$

Так что форма

$$b_{ik}x_ix_k = Mgh(x_1^2 + x_2^2) \quad (2)$$

является *положительно* определенной.

Положим, что  $x_1$  и  $x_2$  пропорциональны  $e^{\gamma t}$ . Тогда

$$\left. \begin{aligned} (A\gamma^2 - Mgh)x_1 &= -Cn\gamma x_2, & (A\gamma^2 - Mgh)x_2 &= Cn\gamma x_1, \\ D_2 = (A\gamma^2 - Mgh)^2 + C^2n^2\gamma^2 &= 0, & A\gamma^2 \pm iCn\gamma - Mgh &= 0, \\ \gamma &= \pm \frac{1}{2} \frac{iCn}{A} \pm \frac{i}{2} \left( \frac{C^2n^2}{A^2} - \frac{4Mgh}{A} \right)^{1/2}. \end{aligned} \right\} \quad (3)$$

Значения  $\gamma$  чисто мнимые при  $C^2n^2 > 4AMgh$ . При  $C^2n^2 < 4AMgh$  два корня имеют положительные действительные части и волчок неустойчив. Но при этом  $\gamma^2$  комплексно, тогда как в случае неустойчивого положения равновесия оно должно быть действительным и положительным. Движение конца оси волчка в первом приближении происходит по логарифмической спирали.

Если  $C^2n^2 = 4AMgh$ , то корни оказываются попарно равными. Выбрав  $\gamma = -\frac{1}{2} iCn/A$ , можно показать, что матрица коэффициентов равна  $2Mgh \begin{pmatrix} -2 & -2i \\ 2i & -2 \end{pmatrix}$  и имеет ранг, равный 1. Следовательно, для этого  $\gamma$  существует только одно решение

вида  $e^{\gamma t}$ , и то же самое решение для  $\gamma$  с другим знаком. Остальные решения имеют вид  $te^{\gamma t}$ . Такие решения вообще не встречаются в теории колебаний около положения равновесия.

Легко видеть, что  $D_1 = A\gamma^2 - Mgh$ ,  $D_0 = 1$ . Знаки определителей изменяются следующим образом:

$\gamma^2$	$D_2$	$D_1$	$D_0$
$-\infty$	+	-	+
0	+	-	+
$\infty$	+	+	+

что не дает возможности обнаружить нули  $D_2$  для отрицательных  $\gamma^2$ .

Если, однако, рассматривать подвешенный волчок, то форма  $b_{ik}x_ix_k$  оказывается отрицательно определенной. Нужно только изменить знак  $g$ . В этом случае не может быть кратных корней. Знаки в таблице распределяются следующим образом:

$\gamma^2$	$D_2$	$D_1$	$D_0$
$-\infty$	+	-	+
0	+	+	+
$\infty$	+	+	+

Две переменные знака утрачиваются между  $-\infty$  и 0.

**4.093.** Для гироскопических систем мнимые значения  $\gamma$  обладают свойством стационарности. Если выбрать  $\gamma$  таким, что

$$\gamma^2 = \frac{b_{ik}x_kx_i^* - \gamma g_{ik}x_kx_i^*}{a_{ik}x_kx_i^*},$$

то видно, что уравнения **4.091** (1) эквивалентны утверждению, что  $\gamma$  стационарно относительно действительных и мнимых вариаций  $x_i$ . Оказывается, что выражение  $g_{ik}x_kx_i^*$  чисто мнимое для любых комплексных  $x_i, x_k$ . Пусть оно отлично от нуля. Положим для каждого решения, имеющего форму  $x_i = l_ie^{\gamma t}$ ,

$$a_{ik}x_kx_i^* = P, \quad g_{ik}x_kx_i^* = iQ, \quad b_{ik}x_kx_i^* = R.$$

Тогда  $2P\gamma = -iQ \pm (-Q^2 + 4RP)^{1/2}$ .

Если  $Q^2 < 4RP$ , то оба корня комплексны и  $|\gamma^2| = R/P$ . Таким образом, если система неустойчива в отсутствие гироскопических членов, то действительная часть  $\gamma$  меньше, чем она была бы для тех же  $x_i$ , но при условии, что все  $g_{ik}$  равны

нулю. Но по принципу Релея действительная часть  $\gamma$  меньше, чем для реальных значений  $x_i$  без учета гироскопических членов, при которых достигается наибольшее  $\gamma$ . Следовательно, гироскопический эффект стремится уменьшить неустойчивость.

Если  $Q^2 > 4RP$ , то оба корня мнимы и система устойчива. Если  $R$  — отрицательно определенная форма, то система устойчива при всех  $Q$ . Если при колебаниях с  $g_{ik} = 0$  имеется два корня  $\pm i\sigma$ , то гироскопический эффект увеличивает частоту колебаний для одного корня и уменьшает для другого.

**4.10. Собственные значения унитарных и ортогональных матриц.** Пусть  $a_{ik}$  — унитарная матрица  $n$ -го порядка. Тогда

$$a_{ik}a_{im}^* = \delta_{km}. \quad (1)$$

Рассмотрим уравнения

$$a_{ik}x_k = \lambda x_i. \quad (2)$$

Вообще говоря, корни детерминантного уравнения

$$\|a_{ik} - \lambda\delta_{ik}\| = 0 \quad (3)$$

комплексны. Комплексное сопряжение (2) имеет вид

$$a_{im}^*x_m^* = \lambda^*x_i^*. \quad (4)$$

Перемножим (2) и (4). Тогда

$$a_{ik}a_{im}^*x_kx_m^* = \lambda\lambda^*x_ix_i^*. \quad (5)$$

С учетом (1) формула (5) сводится к

$$x_kx_k^* = \lambda\lambda^*x_ix_i^*, \quad (6)$$

и поэтому

$$\lambda\lambda^* = 1. \quad (7)$$

Следовательно, все собственные числа унитарной матрицы по модулю равны 1.

Пусть теперь  $\lambda_1$  и  $\lambda_2$  — два различных корня уравнения (3) и  $x_{i1}$ ,  $x_{i2}$  — соответствующие им  $x_i$ . Поступая, как и раньше, мы имеем

$$a_{ik}a_{im}^*x_{k1}x_{m2}^* = \lambda_1\lambda_2^*x_{i1}x_{i2}^* \quad (8)$$

и

$$a_{ik}a_{im}^*x_{k1}x_{m2}^* = x_{k1}x_{k2}^*. \quad (9)$$

Отсюда либо

$$\lambda_1\lambda_2^* = 1, \quad (10)$$

либо

$$x_{i1}x_{i2}^* = 0. \quad (11)$$

В первом случае  $\lambda_1 = \lambda_2$ , что уже было рассмотрено. Следовательно, решения обладают свойством комплексной ортогональности, подобным тому, которое было установлено для системы эрмитовых уравнений.

Унитарную матрицу  $\mathbf{a}$  всегда можно привести к диагональному виду при помощи преобразования  $\mathbf{l}^{-1}\mathbf{a}\mathbf{l}$ , где  $\mathbf{l}$  — унитарная матрица [9] \*). Это свойство сохраняется, в частности, для действительной унитарной матрицы, однако матрица  $\mathbf{l}$  в общем случае остается комплексной.

**4.101.** Рассмотрим теперь наиболее общую унитарную матрицу порядка  $2 \times 2$ . Как и в 4.032, напомним

$$\mathbf{l} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (1)$$

Тогда

$$\Pi^+ = \begin{pmatrix} aa^* + bb^* & ac^* + bd^* \\ ca^* + db^* & cc^* + dd^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2)$$

Можно взять

$$\begin{aligned} a &= \cos \alpha e^{i\beta}, & b &= -\sin \alpha e^{i\gamma}, \\ a^* &= \cos \alpha e^{-i\beta}, & b^* &= -\sin \alpha e^{-i\gamma} \end{aligned} \quad (3)$$

с действительными  $\alpha$ ,  $\beta$  и  $\gamma$ . Тогда имеем равносильные соотношения

$$\begin{aligned} \frac{c}{d} &= -\frac{b^*}{a^*} = \operatorname{tg} \alpha e^{i(\beta-\gamma)}, \\ \frac{c^*}{d^*} &= -\frac{b}{a} = \operatorname{tg} \alpha e^{-i(\beta-\gamma)}. \end{aligned} \quad (4)$$

Далее,

$$\frac{cc^*}{dd^*} = \operatorname{tg}^2 \alpha, \quad dd^* = \cos^2 \alpha, \quad cc^* = \sin^2 \alpha, \quad (5)$$

и все условия удовлетворяются при

$$d = \cos \alpha e^{i\delta}, \quad c = \sin \alpha e^{i(\beta-\gamma+\delta)}. \quad (6)$$

Положим

$$\beta + \delta = 2\varepsilon, \quad \beta - \delta = 2\eta. \quad (7)$$

---

\*) См. также пример 10 на стр. 275.

Тогда

$$I = \begin{pmatrix} \cos \alpha e^{i(\varepsilon+\eta)} & -\sin \alpha e^{i\gamma} \\ \sin \alpha e^{i(2\varepsilon-\gamma)} & \cos \alpha e^{i(\varepsilon-\eta)} \end{pmatrix} = e^{i\varepsilon} \begin{pmatrix} \cos \alpha e^{i\eta} & -\sin \alpha e^{i(\gamma-\varepsilon)} \\ \sin \alpha e^{i(\varepsilon-\gamma)} & \cos \alpha e^{-i\eta} \end{pmatrix} =$$

$$= e^{i\varepsilon} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{pmatrix}, \quad (8)$$

где

$$\theta = \frac{1}{2}(\eta + \gamma - \varepsilon), \quad \psi = \frac{1}{2}(\eta - \gamma + \varepsilon). \quad (9)$$

Таким образом, в то время как наиболее общую действительную ортогональную матрицу порядка  $2 \times 2$  можно выразить через один угол, для представления наиболее общей унитарной матрицы требуется четыре угла.

**4.102. Связь произвольного вращения в трех измерениях с унитарным преобразованием порядка  $2 \times 2$ .** Пусть  $A_1 = (A_1, B_1, C_1)$  и  $A_2 = (A_2, B_2, C_2)$  — два действительных, равных по модулю взаимно перпендикулярных вектора, так что

$$A_1^2 + B_1^2 + C_1^2 = A_2^2 + B_2^2 + C_2^2; \quad A_1 A_2 + B_1 B_2 + C_1 C_2 = 0. \quad (1)$$

Очевидно, из действительных величин  $A_1, B_1, C_1, A_2, B_2, C_2$  четыре можно задать независимо. При ортогональных преобразованиях каждый из векторов сохраняет длину и векторы остаются взаимно перпендикулярными. Если преобразование таково, что этим свойством обладают все пары равных по длине взаимно перпендикулярных действительных векторов, то преобразование является ортогональным.

Образует „комплексный вектор“  $(a, b, c)$ , где

$$a = A_1 + iA_2, \quad b = B_1 + iB_2, \quad c = C_1 + iC_2. \quad (2)$$

Тогда, согласно (1),

$$a^2 + b^2 + c^2 = 0,$$

откуда

$$(ib - a)(ib + a) = c^2. \quad (3)$$

Вектор  $(a, b, c)$  можно назвать комплексным нулевым вектором. Существуют такие  $x_1, x_2$ , что

$$\left. \begin{aligned} ib - a &= x_1^2, \\ c &= x_1 x_2, \\ ib + a &= x_2^2, \end{aligned} \right\} \quad \left. \begin{aligned} a &= \frac{1}{2}(x_2^2 - x_1^2), \\ b &= \frac{1}{2i}(x_1^2 + x_2^2), \\ c &= x_1 x_2. \end{aligned} \right\} \quad (4)$$

Тогда для заданного  $(a, b, c)$  можно с точностью до знака найти  $x_1$  и  $x_2$  и, наоборот, для каждого заданных (комплексных)  $x_1$  и  $x_2$  уравнения (4) позволяют приравниванием действительных и мнимых частей найти вектор  $(a, b, c)$ , удовлетворяющий (3) и (1). Кроме того,

$$|a^2| + |b^2| + |c^2| = \frac{1}{2} (x_1 x_1^* + x_2 x_2^*)^2, \quad (5)$$

и, следовательно,

$$|a^2| + |b^2| + |c^2| = (A_1^2 + B_1^2 + C_1^2) + (A_2^2 + B_2^2 + C_2^2).$$

Если  $x_1, x_2$  линейным преобразованием переводятся в такие  $y_1$  и  $y_2$ , что

$$x_1 x_1^* + x_2 x_2^* = y_1 y_1^* + y_2 y_2^*, \quad (6)$$

то  $a', b', c'$ , определяемые из (4) после замены  $x_1, x_2$  на  $y_1, y_2$ , линейно зависят от  $x_1^2, x_2^2, x_1 x_2$  и, следовательно, от  $(a, b, c)$ . Кроме того,

$$|a'^2| + |b'^2| + |c'^2| = |a^2| + |b^2| + |c^2|. \quad (7)$$

Отсюда следует, что векторы  $\mathbf{A}'_1, \mathbf{A}'_2$ , построенные с помощью  $a', b', c'$ , равны по величине  $\mathbf{A}_1$  и  $\mathbf{A}_2$  и взаимно перпендикулярны. Но, вообще говоря, вектор  $\mathbf{A}'_1$  зависит как от  $\mathbf{A}_1$ , так и от  $\mathbf{A}_2$ . Требуется найти действительное преобразование, которое будет соответствовать повороту  $\mathbf{A}_1$  в  $\mathbf{A}'_1$  и  $\mathbf{A}_2$  в  $\mathbf{A}'_2$ , где  $\mathbf{A}'_1$  не зависит от  $\mathbf{A}_2$  и  $\mathbf{A}'_2$  не зависит от  $\mathbf{A}_1$ . Это накладывает дополнительное ограничение, заключающееся в том, что действительные и мнимые части  $a, b, c$  должны преобразовываться независимо. Будем искать, при каких условиях преобразование (6) обладает этим свойством.

Наиболее общее унитарное преобразование двух переменных можно представить в виде

$$y_1 = (\alpha x_1 + \beta x_2) e^{i\epsilon}, \quad y_2 = (-\beta^* x_1 + \alpha^* x_2) e^{i\epsilon}, \quad (8)$$

где  $\epsilon$  действительно и  $\alpha\alpha^* + \beta\beta^* = 1$ . Тогда если

$$a' = \frac{1}{2} (y_2^2 - y_1^2), \quad b' = \frac{1}{2i} (y_1^2 + y_2^2), \quad c' = y_1 y_2, \quad (9)$$

то можно найти, что

$$\left. \begin{aligned} a'e^{-2i\varepsilon} &= \frac{1}{2}(\alpha^2 + \alpha^{*2} - \beta^2 - \beta^{*2})a + \\ &\quad + \frac{1}{2i}(\alpha^2 - \alpha^{*2} + \beta^2 - \beta^{*2})b - (\alpha\beta + \alpha^*\beta^*)c, \\ b'e^{-2i\varepsilon} &= \frac{1}{2i}(-\alpha^2 + \alpha^{*2} + \beta^2 - \beta^{*2})a + \\ &\quad + \frac{1}{2}(\alpha^2 + \alpha^{*2} + \beta^2 + \beta^{*2})b + \frac{1}{i}(\alpha\beta - \alpha^*\beta^*)c, \\ c'e^{-2i\varepsilon} &= (\alpha\beta^* + \alpha^*\beta)a + \frac{1}{i}(\alpha\beta^* - \alpha^*\beta)b + (\alpha\alpha^* - \beta\beta^*)c. \end{aligned} \right\} \quad (10)$$

Все коэффициенты в правой части действительны. Поэтому, для того чтобы действительные и мнимые части  $a$ ,  $b$  и  $c$  преобразовались независимо,  $e^{-2i\varepsilon}$  должно быть действительным и, следовательно, равным  $\pm 1$ . В каждом случае преобразование ортогонально. Сразу видно, что при  $\alpha = 1$ ,  $\beta = 0$  определитель коэффициентов правой части равен 1. При любых  $\alpha$  и  $\beta$  можно принять  $\varepsilon = 0$ . Тогда преобразование, будучи ортогональным, должно иметь определитель, равный  $\pm 1$ . Но определитель всегда будет равен 1, так как он непрерывно зависит от  $\alpha$  и от  $\beta$ . Следовательно, если  $e^{-2i\varepsilon} = 1$ , то преобразование представляет собой вращение. Если  $e^{-2i\varepsilon} = -1$ , то преобразование представляет собой вращение с отражением. [При обильной затрате бумаги можно показать, что определитель равен  $(\alpha\alpha^* + \beta\beta^*)^3$ .] Следовательно, наиболее общее вращение можно представить формулами (8) и (10), положив в них  $\varepsilon = 0$ .

Выберем, в частности,

$$\begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} \cos \frac{1}{2}\theta & -\sin \frac{1}{2}\theta \\ \sin \frac{1}{2}\theta & \cos \frac{1}{2}\theta \end{pmatrix}. \quad (11)$$

Тогда

$$A'_1 = A_1 \cos \theta + C_1 \sin \theta, \quad B'_1 = B_1, \quad C'_1 = -A_1 \sin \theta + C_1 \cos \theta. \quad (12)$$

Преобразование представляет собой вращение вправо вокруг оси  $y$  на угол  $\theta$ .

Если

$$\alpha = e^{-\frac{1}{2}i\psi}, \quad \beta = 0, \quad (13)$$

то

$$A'_1 = A_1 \cos \psi - B_1 \sin \psi, \quad B'_1 = A_1 \sin \psi + B_1 \cos \psi, \quad C'_1 = C_1. \quad (14)$$

что соответствует вращению вправо вокруг оси  $z$  на угол  $\psi$ .

Выбор

$$\alpha = \cos \frac{1}{2} \varphi, \quad \beta = -i \sin \frac{1}{2} \varphi, \quad (15)$$

$$A'_1 = A_1, \quad B'_1 = B_1 \cos \varphi - C_1 \sin \varphi, \quad C'_1 = B_1 \sin \varphi + C_1 \cos \varphi \quad (16)$$

соответствует вращению вправо вокруг оси  $x$  на угол  $\varphi$ .

Из операций такого типа можно построить общее унимодулярное унитарное преобразование. Оно задается матрицей

$$\begin{aligned} \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} &= \begin{pmatrix} e^{-\frac{1}{2} i \lambda} & 0 \\ 0 & e^{\frac{1}{2} i \lambda} \end{pmatrix} \begin{pmatrix} \cos \frac{1}{2} \theta & -\sin \frac{1}{2} \theta \\ \sin \frac{1}{2} \theta & \cos \frac{1}{2} \theta \end{pmatrix} \begin{pmatrix} e^{-\frac{1}{2} i x} & 0 \\ 0 & e^{\frac{1}{2} i x} \end{pmatrix} = \\ &= \begin{pmatrix} \cos \frac{1}{2} \theta e^{-\frac{1}{2} i (\lambda+x)} & -\sin \frac{1}{2} \theta e^{-\frac{1}{2} i (\lambda-x)} \\ \sin \frac{1}{2} \theta e^{\frac{1}{2} i (\lambda-x)} & \cos \frac{1}{2} \theta e^{\frac{1}{2} i (\lambda+x)} \end{pmatrix} \quad (17) \end{aligned}$$

и приводит к произвольному повороту в трехмерном пространстве, рассмотренному в 4.034. Каждому такому повороту там соответствуют два преобразования порядка  $2 \times 2$ .

**4.11. Спиновые матрицы Паули.** В пространстве  $x, y, z$  малые вращения  $\epsilon$  вокруг осей  $x, y$  и  $z$  задаются соответственно матрицами

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\epsilon \\ 0 & \epsilon & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (18)$$

Здесь отброшены все члены, содержащие  $\epsilon$  в степенях выше первой. Соответствующие унитарные преобразования задаются матрицами порядка  $2 \times 2$

$$\begin{pmatrix} 1 & -\frac{1}{2} i \epsilon \\ -\frac{1}{2} i \epsilon & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -\frac{1}{2} \epsilon \\ \frac{1}{2} \epsilon & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 - \frac{1}{2} i \epsilon & 0 \\ 0 & 1 + \frac{1}{2} i \epsilon \end{pmatrix}. \quad (19)$$

Если написать

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (20)$$

то матрицы (19) будут иметь вид

$$1 - \frac{1}{2} i \epsilon \sigma_1, \quad 1 - \frac{1}{2} i \epsilon \sigma_2, \quad 1 - \frac{1}{2} i \epsilon \sigma_3.$$



Матрицы (20) называются „матрицами Паули“ и встречаются в теории спина электрона в квантовой механике.

Можно проверить, что

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1, \quad \sigma_1\sigma_2 = -\sigma_2\sigma_1 = i\sigma_3, \quad \sigma_2\sigma_3 = -\sigma_3\sigma_2 = i\sigma_1, \\ \sigma_3\sigma_1 = -\sigma_1\sigma_3 = i\sigma_2.$$

Следовательно, матрицы  $\sigma_1$ ,  $\sigma_2$ ,  $\sigma_3$  антикоммутируют.

**4.12. Матрицы Эддингтона и Дирака порядка  $4 \times 4$ .** Рассмотрим следующие матрицы порядка  $4 \times 4$ :

$$E_1 = \begin{pmatrix} i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & i\sigma_3 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix}, \quad E_4 = \begin{pmatrix} -i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix},$$

где, например,

$$\begin{pmatrix} i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{pmatrix} = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & i & 0 \end{pmatrix},$$

т. е. элементы  $\sigma$  выписываются, как указано, и внутренние скобки убираются. По 4.063 можно производить умножение матриц порядка  $2 \times 2$ , элементы которых сами являются матрицами порядка  $2 \times 2$ , по обычному правилу, а затем результаты развертывать в матрицы порядка  $4 \times 4$ . Тогда

$$E_1^2 = \begin{pmatrix} -\sigma_1^2 & 0 \\ 0 & -\sigma_1^2 \end{pmatrix} = -I$$

и квадраты  $E_2$ ,  $E_3$ ,  $E_4$  также равны  $-I$ . Легко также показать, что эти матрицы антикоммутируют. Если мы теперь положим

$$iE_5 = E_1E_2E_3E_4 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix},$$

то будем иметь

$$E_5^2 = -E_1E_2E_3E_4E_1E_2E_3E_4 = E_1E_1E_2E_3E_4E_2E_3E_4 = \\ = -E_2E_2E_3E_4E_3E_4 = -E_3E_3E_4E_4 = -I$$

и

$$iE_1E_5 = -E_2E_3E_4, \quad iE_5E_1 = E_1E_2E_3E_4E_1 = -E_1E_1E_2E_3E_4 = E_2E_3E_4.$$

Таким образом,  $E_5$  антикоммутирует с  $E_1$  и аналогично с  $E_2$ ,  $E_3$  и  $E_4$ .

Итак, мы имеем антикоммутирующую пентаду квадратных корней из  $-I$ . В развернутой форме они записываются в виде (в порядке возрастания индексов)

$$\begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & i & 0 \end{pmatrix} \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Квадрат каждой из матриц  $E_\mu E_\nu$  ( $\mu \neq \nu$ ) равен  $-I$ . Обозначая его через  $E_{\mu\nu}$ , имеем

$$E_{12} = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad E_{13} = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$E_{14} = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix} = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix}, \quad E_{15} = \begin{pmatrix} 0 & i\sigma_3 \\ i\sigma_3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix},$$

$$E_{23} = \begin{pmatrix} 0 & -\sigma_1 \\ \sigma_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad E_{24} = \begin{pmatrix} -i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{pmatrix} = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & i & 0 \end{pmatrix},$$

$$E_{25} = \begin{pmatrix} 0 & -i\sigma_1 \\ -i\sigma_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix},$$

$$E_{34} = \begin{pmatrix} 0 & -iI \\ -iI & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix},$$

$$E_{35} = \begin{pmatrix} iU & 0 \\ 0 & -iU \end{pmatrix} = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}, \quad E_{45} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Мы имеем, таким образом, пятнадцать матриц — квадратных корней из  $-I$ , к которым мы можем добавить шестнадцатую  $E_{16} = iI$ . Можно показать, что первые пятнадцать матриц образуют шесть антикоммутирующих пентад, причем каждая матрица  $E$  принадлежит двум пентадам \*). Все эти матрицы являются антиэрмитовыми и унитарными.

Матрицы, введенные Дираком при решении релятивистского волнового уравнения, являются антикоммутирующими квадратными корнями из  $I$  и связаны с одной из пентад Эддингтона соотношениями  $\alpha_1 = iE_{25}$ ,  $\alpha_2 = iE_{5}$ ,  $\alpha_3 = -iE_{15}$ ,  $\beta = -iE_{35}$ ,  $\alpha_1\alpha_2\alpha_3\beta = iE_{45}$ .

**4.13. Косоугольные координаты.** Оси координат, которые мы до сих пор рассматривали, были прямоугольными. Однако положение точки может быть задано либо ортогональными проекциями на координатные оси ее перемещения из начала координат, либо перемещениями в любых трех некомпланарных направлениях, причем векторная сумма этих перемещений равна перемещению из начала координат в данную точку. Второй метод соответствует обычному разложению на косоугольные компоненты; мы увидим, что и первый метод имеет физическое значение. Характерным свойством прямоугольных осей является то, что эти две системы величин совпадают между собой. Посмотрим, что происходит при косоугольных осях. Допустим, что мы имеем систему прямоугольных осей  $x_i$ , и выберем три косоугольные оси  $x'_j$  с направляющими косинусами  $l_{ij}$  относительно прямоугольных осей. По причине, которая выяснится вскоре, мы отмечаем косоугольные координаты верхним индексом вместо нижнего, т. е.  $x'^i$ . Тогда прямоугольные координаты точки  $P$  будут, как прежде,  $l_{ij}x'^j$ . Так как  $l_{ij}$  при фиксированном  $j$  являются направляющими косинусами одной

---

\*) Эддингтон [10] не выписывал этих матриц полностью, но они неявно содержатся в уравнениях 3.61 на стр. 42. Антисимметричные матрицы входят с обратным знаком. Это связано с тем, что у него в обозначении элемента  $a_{ik}$ , в противоположность обычным обозначениям, индекс  $i$  обозначает номер столбца, а  $k$  — номер строки. В [11] на стр. 142 он дает эти матрицы подробно в обычных обозначениях; его  $E_{31}$  есть  $-E_{13}$ .

прямой, то

$$\sum_i l_{il}^2 = 1, \quad (1)$$

но в общем случае при  $j \neq l$

$$\sum_i l_{ij} l_{il} \neq 0, \quad (2)$$

так как это косинусы углов между направлениями  $x'^i$  и  $x'^l$ , которые не перпендикулярны. Расстояние точки от начала координат есть  $r$ , где

$$r^2 = x_i^2 = (l_{ij} x'^j)(l_{il} x'^l) = l_{ij} l_{il} x'^j x'^l \quad (3)$$

представляет собой квадратичную форму и может быть записано в виде

$$r^2 = g'_{jl} x'^j x'^l, \quad (4)$$

где  $g'_{11} = g'_{22} = g'_{33} = 1$ , но  $g'_{23}$ ,  $g'_{31}$ ,  $g'_{13}$  не равны нулю. Если и только если оси  $x'^i$  прямоугольные,  $g'_{jl} = \delta_{jl}$ .

Рассмотрим теперь ортогональную проекцию  $OP$  на направление  $x'^i$  и обозначим ее через  $x'_j$  с нижним индексом. Она равна

$$x'_j = l_{ij} x_i = l_{ij} l_{il} x'^l = g'_{jl} x'^l = \frac{\partial}{\partial x'^j} \left( \frac{1}{2} r^2 \right) \quad (5)$$

и

$$x'_j x'^j = g'_{jl} x'^j x'^l = r^2. \quad (6)$$

Таким образом, форма  $x_i x_i$ , правильная для  $r^2$  в прямоугольных осях, должна быть изменена для косоугольных осей поднятием одного индекса. Теперь мы не можем записать ее в виде  $x'^2_j$ .

Положение  $P$  можно одинаково хорошо задать как значениями  $x'_j$ , так и значениями  $x'^j$ . Но фактические значения этих трех компонент в обоих случаях различны. Первые называются *ковариантными* компонентами, а вторые — *контравариантными* компонентами перемещения  $OP$ . Обычно кажется, что было бы более естественно поменять эти прилагательные местами. Однако ковариантная компонента в направлении  $Oj$  обладает тем свойством, что при заданном положении точки  $P$  она не зависит от направлений двух других осей; это неверно для контравариантной компоненты, и в этом отношении ковариантные компоненты могут показаться более фундаментальными. С другой стороны, если мы изменим одну из контравариантных

компонент без изменения двух других, то мы сразу же узнаем, на сколько и в каком направлении переместится точка. Это не очевидно при изменении одной ковариантной компоненты. Поэтому обе системы имеют свои достоинства, и нам нужен способ преобразования одной в другую.

Мы пишем

$$\|g'_{il}\| = G', \quad (7)$$

обозначаем алгебраическое дополнение  $g'_{il}$  в  $G'$  через  $G'^{il}$  и полагаем

$$g'^{il} = \frac{G'^{il}}{G'} = g'^{il}. \quad (8)$$

Тогда  $g'^{il}$  представляет собой матрицу, обратную  $g_{il}$ , и

$$x'^i = g'^{ij} x'_j. \quad (9)$$

Эта формула и сопутствующая ей

$$x'_j = g'_{jl} x'^l \quad (10)$$

известны как формулы поднятия и опускания индексов. Мы видим, что определитель  $G'$  исключительно важен. Если обозначить угол между осями  $x'_1$  и  $x'_2$  через  $\alpha_3$  и т. д., то

$$g'_{12} = \cos \alpha_3, \quad (11)$$

$$G' = \begin{vmatrix} 1 & \cos \alpha_3 & \cos \alpha_2 \\ \cos \alpha_3 & 1 & \cos \alpha_1 \\ \cos \alpha_2 & \cos \alpha_1 & 1 \end{vmatrix} =$$

$$= 1 - \cos^2 \alpha_1 - \cos^2 \alpha_2 - \cos^2 \alpha_3 + 2 \cos \alpha_1 \cos \alpha_2 \cos \alpha_3. \quad (12)$$

Однако  $g'^{11}$ ,  $g'^{22}$ ,  $g'^{33}$ , вообще говоря, не равны единице или друг другу.

В другой записи

$$G' = \|l_{ij} l_{il}\| = \|l_{il}\| \|l_{il}\| = \|l_{ij}\|^2. \quad (13)$$

Но  $\|l_{ij}\|$  равняется объему параллелепипеда с ребрами единичной длины, построенного на осях, или, иными словами, смешанному произведению направляющих векторов. Поэтому он может быть равен нулю, только если оси компланарны. Следовательно, элемент объема равен  $G'^{1/2} dx'^1 dx'^2 dx'^3$ .

Если  $A$  — произвольный вектор с компонентами  $A^1$ ,  $A^2$ ,  $A^3$  вдоль косоугольных осей, то его прямоугольные компоненты будут

$$A_i = l_{ij} A'^j, \quad (14)$$

а его ковариантные компоненты будут равны

$$A'_i = g'_{ij} A'^j = l_{ij} A_j. \quad (15)$$

Теперь пусть  $\varphi$  — скаляр. Рассмотрим его производные по  $x'^j$ . Мы имеем

$$\frac{\partial \varphi}{\partial x'^j} = \frac{\partial x_i}{\partial x'^j} \frac{\partial \varphi}{\partial x_i} = l_{ij} \frac{\partial \varphi}{\partial x_i}, \quad (16)$$

и, следовательно,  $\partial \varphi / \partial x'^j$  представляют собой ковариантные компоненты  $\text{grad} \varphi$ . Чтобы получить контравариантные компоненты, нужно умножить на  $g'^{il}$  и свернуть.

Если **A** и **B** — два вектора, то

$$A'_i B'^i = l_{ij} A_j B'^i = A_i B_i \quad (17)$$

— их скалярное произведение. Аналогично

$$\frac{\partial A'^i}{\partial x'^j} = l_{ij} \frac{\partial A'^i}{\partial x_i} = \frac{\partial A_i}{\partial x_i} \quad (18)$$

есть дивергенция **A**.

Как и в прямоугольных осях, мы можем определить тензоры любого порядка, но правила преобразования будут различными в зависимости от того, является ли каждый индекс верхним или нижним. Мы можем свертывать по повторяющемуся индексу при условии, что он оказывается один раз снизу и один раз сверху, и результат будет тензором на 2 единицы меньшего порядка.

**4.131. Структура кристаллов: обратная решетка.** Простая трехмерная решетка определяется заданием трех векторов основных трансляций. Если любой атом взять за начало координат, то в каждой точке  $n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$  будет находиться такой же атом, причем  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  — радиусы-векторы трех соседних атомов, не находящихся в одной плоскости с началом, а  $n_1, n_2, n_3$  — положительные или отрицательные целые числа. Определенные так точки называются узлами решетки. Плоскость, проведенная через любые три узла решетки, будет содержать одинаковые атомы, расположенные повторяющимся образом. Направление этой плоскости может быть задано отрезками, отсекаемыми ею на осях, проведенных через начало координат в направлениях  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ . Пусть после деления на подходящее целое число эти отрезки будут  $\mathbf{a}_1/h_1, \mathbf{a}_2/h_2, \mathbf{a}_3/h_3$ , где  $h_1, h_2, h_3$  — целые числа, не содержащие общего множителя.

Тогда  $h_1, h_2, h_3$  называются индексами (*индексами Миллера*) этой плоскости \*), они равны для всех параллельных плоскостей.

Мы будем обозначать объем  $\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$  отдельной ячейки решетки через  $v_a$ .

Система векторов  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ , обратная системе  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ , определяется как

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{v_a}, \quad \mathbf{b}_2 = \frac{\mathbf{a}_3 \times \mathbf{a}_1}{v_a}, \quad \mathbf{b}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{v_a}. \quad (1)$$

Они удовлетворяют соотношениям

$$\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij} \quad (2)$$

и

$$\begin{aligned} v_b = \mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3) &= \frac{\mathbf{b}_1 \cdot [\mathbf{b}_2 \times (\mathbf{a}_1 \times \mathbf{a}_2)]}{v_a} = \\ &= \frac{\mathbf{b}_1 \cdot [(\mathbf{b}_2 \cdot \mathbf{a}_2) \mathbf{a}_1 - (\mathbf{b}_2 \cdot \mathbf{a}_1) \mathbf{a}_2]}{v_a} = \frac{1}{v_a}. \end{aligned} \quad (3)$$

Любой вектор  $\mathbf{A}$  можно представить как

$$\mathbf{A} = (\mathbf{A} \cdot \mathbf{b}_1) \mathbf{a}_1 + (\mathbf{A} \cdot \mathbf{b}_2) \mathbf{a}_2 + (\mathbf{A} \cdot \mathbf{b}_3) \mathbf{a}_3 = \quad (4)$$

$$= (\mathbf{A} \cdot \mathbf{a}_1) \mathbf{b}_1 + (\mathbf{A} \cdot \mathbf{a}_2) \mathbf{b}_2 + (\mathbf{A} \cdot \mathbf{a}_3) \mathbf{b}_3. \quad (5)$$

Следовательно, если построить решетку с  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$  в качестве векторов основных трансляций, то объем ее ячейки будет обратным объему ячейки исходной решетки.

Поскольку точки  $\mathbf{a}_1/h_1, \mathbf{a}_2/h_2, \mathbf{a}_3/h_3$  принадлежат некоторой плоскости решетки, направления векторов

$$\frac{\mathbf{a}_1}{h_1} - \frac{\mathbf{a}_2}{h_2}, \quad \frac{\mathbf{a}_1}{h_1} - \frac{\mathbf{a}_3}{h_3}$$

параллельны этой плоскости. Но оба эти направления перпендикулярны  $h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3$ , и это смещение в обратной решетке перпендикулярно плоскостям атомной решетки с индексами Миллера  $h_1, h_2, h_3$ . Поэтому уравнение каждой такой плоскости можно записать в виде

$$\frac{(h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3) \cdot \mathbf{x}}{|h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3|} = p, \quad (6)$$

где  $p$  — длина перпендикуляра, опущенного из начала координат на эту плоскость. Теперь если  $\mathbf{x}$  — радиус-вектор узла решетки, то он имеет вид  $n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ , и тогда

$$(h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3) \cdot (n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3) = h_1n_1 + h_2n_2 + h_3n_3. \quad (7)$$

---

\*) Этот термин был предложен Грассманом и др., но стал распространенным только благодаря книге Миллера [12].

Далее, если  $h_1, h_2, h_3$  не имеют общего множителя, мы можем выбрать  $n_1, n_2, n_3$  так, чтобы сумма в правой части (7) равнялась единице. Действительно, если  $h_1, h_2$  имеют общий множитель  $q$ , то процесс нахождения наибольшего общего делителя позволяет определить такие  $s_1, s_2$ , что  $s_1 h_1 + s_2 h_2 = q$ . Точно так же, если  $h_2, h_3$  имеют общим множителем  $r$ , мы можем найти такие  $t_1, t_2$ , что  $t_1 h_2 + t_2 h_3 = r$ . Но по предположению  $q$  и  $r$  не имеют общих множителей, следовательно, можно найти линейную комбинацию этих выражений с целыми коэффициентами, равную единице. Ее можно взять в качестве

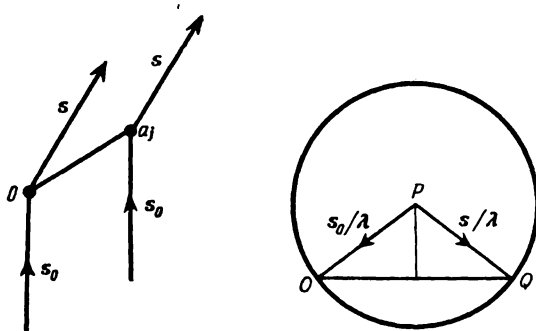


Рис. 23.

$h_1 n_1 + h_2 n_2 + h_3 n_3$ . Очевидно, выражение  $h_1 n_1 + h_2 n_2 + h_3 n_3$  можно сделать равным нулю, выбирая  $n_1 = n_2 = n_3 = 0$ . Тогда расстояние по нормали  $d_h$  между плоскостями, соответствующими этим двум значениям  $\mathbf{x}$ , т. е. промежуток между соседними кристаллографическими плоскостями с индексами Миллера  $(h_1, h_2, h_3)$ , равно  $|h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3|^{-1}$ . Рассмотрим теперь параллельный пучок рентгеновских лучей, падающих на кристалл. Допустим, что плоские волны, распространяющиеся в направлении  $\mathbf{s}_0$ , падают на атомную решетку и рассеиваются отдельными атомами. Найдем условие того, чтобы волны, распространяющиеся в направлении  $\mathbf{s}$ , усиливали друг друга. Разность хода волн, рассеянных двумя атомами, разделенными расстоянием  $\mathbf{a}_j$ , равна  $(\mathbf{s} - \mathbf{s}_0) \cdot \mathbf{a}_j$ , и условие усиления состоит в том, чтобы она была кратной длине волны  $\lambda$  для всех атомов в этой области. Отсюда

$$(\mathbf{s} - \mathbf{s}_0) \cdot \mathbf{a}_j = k_j \lambda \quad (j = 1, 2, 3), \quad (8)$$

где  $k_j$  — целые числа. Тогда из (5)

$$\mathbf{s} - \mathbf{s}_0 = \lambda (k_1 \mathbf{b}_1 + k_2 \mathbf{b}_2 + k_3 \mathbf{b}_3), \quad (9)$$



или, предполагая, что  $k_1, k_2, k_3$  имеют общий множитель  $n$ , получим

$$\mathbf{s} - \mathbf{s}_0 = n\lambda (h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3). \quad (10)$$

Выберем теперь начало координат в точке  $O$  обратной решетки. Геометрический смысл (10) состоит в том, что если  $P$  — такая точка, что  $\mathbf{PO} = \mathbf{s}_0/\lambda$  и  $\mathbf{PQ} = \mathbf{s}/\lambda$ , то  $Q$  должна быть узлом обратной решетки и все такие узлы, лежащие на сфере радиуса  $1/\lambda$  с центром в точке  $P$ , будут соответствовать отражению лучей с длиной волны  $\lambda$ . Кроме того,  $OQ$  параллельно биссектрисе внешнего угла треугольника  $OPQ$  и поэтому это отражение происходит на плоскостях с индексами Миллера  $(h_1, h_2, h_3)$ , такими, что нормаль к ним совпадает с направлением  $OQ$ .

Если мы через  $2\theta$  обозначим угол  $OPQ$ , так что  $\theta$  есть угол отражения, то будем иметь

$$\frac{n}{d_h} = n |h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3| = \frac{1}{\lambda} |\mathbf{s} - \mathbf{s}_0| = \frac{2}{\lambda} \sin \theta \quad (11)$$

или

$$n\lambda = 2d_h \sin \theta, \quad (12)$$

т. е. условие Брегга.

Далее, возводя (11) в квадрат, имеем

$$n^2 \lambda^2 |h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3|^2 = 4 \sin^2 \theta, \quad (13)$$

а также, поскольку из (10) следует

$$\mathbf{s}^2 = |n\lambda (h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3) + \mathbf{s}_0|^2, \quad (14)$$

находим

$$n\lambda = - \frac{2\mathbf{s}_0 \cdot (h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3)}{|h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3|^2}. \quad (15)$$

Дальнейшее развитие теории дано Эвальдом [13 \*).

**4.14. Криволинейные координаты.** Во многих задачах удобно задавать положение точки тремя функциями прямоугольных координат, которые не постоянны на плоскостях. Например, мы можем пользоваться сферическими координатами. В этом случае координата  $r$  постоянна на любой сфере с центром в начале координат. Если для прямоугольных координат сохранить обозначение  $x_i$ , а криволинейные обозначить через  $x'^j$ ,

\*) В особенности замечания 1 и 2.

то мы будем иметь

$$dx_i = \frac{\partial x_i}{\partial x'^1} dx'^1 + \frac{\partial x_i}{\partial x'^2} dx'^2 + \frac{\partial x_i}{\partial x'^3} dx'^3 = \frac{\partial x_i}{\partial x'^j} dx'^j, \quad (1)$$

$$dx'^j = \frac{\partial x'^j}{\partial x_i} dx_i; \quad \frac{\partial \varphi}{\partial x_i} = \frac{\partial x'^j}{\partial x_i} \frac{\partial \varphi}{\partial x'^j}, \quad (2)$$

так что правило суммирования по-прежнему применимо. Однако частные производные теперь уже не постоянны. В выражении (1)  $x_i$  нужно, конечно, рассматривать как известную функцию трех  $x'^j$ , которые могут изменяться независимо; напротив, в уравнениях (2)  $x'^j$  рассматривается как функция трех  $x_i$ . Соотношения между этими координатами теперь уже не являются линейными, однако соотношения между их малыми изменениями линейны. Линейными будут также соотношения между производными скаляра. Системы величин, преобразующихся как  $dx'^j$ , называются контравариантными, а преобразующихся как  $\partial\varphi/\partial x'^j$ , — ковариантными, как и в случае косоугольных прямолинейных осей. Если  $ds$  — расстояние между двумя близкими точками, то имеем

$$ds^2 = dx_i^2 = g'_{jl} dx'^j dx'^l, \quad (3)$$

где

$$g'_{jl} = \frac{\partial x_i}{\partial x'^j} \frac{\partial x_i}{\partial x'^l}. \quad (4)$$

Точно так же, как и в случае косоугольных прямолинейных координат, мы можем образовать обратную систему величин  $g'^{jl}$ , определяемую соотношениями

$$g'^{jl} = \frac{\partial x'^j}{\partial x^i} \frac{\partial x'^l}{\partial x^i} \quad (5)$$

и

$$\begin{aligned} g'_{jl} g'^{jn} &= \frac{\partial x^i}{\partial x'^j} \frac{\partial x^l}{\partial x'^l} \frac{\partial x'^j}{\partial x^k} \frac{\partial x'^n}{\partial x^k} = \left( \frac{\partial x^i}{\partial x'^j} \frac{\partial x'^j}{\partial x^k} \right) \left( \frac{\partial x^l}{\partial x'^l} \frac{\partial x'^n}{\partial x^k} \right) = \\ &= \delta_{ik} \frac{\partial x^i}{\partial x'^l} \frac{\partial x'^n}{\partial x^k} = \frac{\partial x'^n}{\partial x^k} \frac{\partial x^k}{\partial x'^l} = \delta_{ln}. \end{aligned} \quad (6)$$

Ковариантные и контравариантные тензоры второго порядка можно определить в соответствии с тем, как они преобразуются при дальнейших преобразованиях координат: подобно  $g'_{jl}$  или  $g'^{jl}$ . В качестве упражнения в дифференциальном

исчислении можно проверить, что если мы возьмем третью систему координат  $x''^a$ , то мы получим такую же форму  $g''_{ab}$  путем преобразования сначала от  $x^i$  к  $x'^j$ , а затем к  $x''^a$ , как если бы мы совершили преобразование прямо к  $x''^a$ .

Вся информация относительно величин перемещений при малых изменениях криволинейных координат содержится в  $g'_{il}$ . Если мы изменим  $x'^1$  без изменения  $x'^2$  и  $x'^3$ , то перемещение будет  $[V(\overline{g_{11}}) dx'^1, 0, 0]$ . Однако такие три перемещения при независимом изменении новых координат будут, вообще говоря, происходить не под прямыми углами. Углы между ними, как и в случае косоугольных координат, легко определить через  $g'_{il}$ , но это требуется редко. Обычно мы выбираем наши координаты таким образом, чтобы перемещения, соответствующие малым независимым изменениям  $x'^j$ , происходили под прямыми углами. Для этого должно выполняться условие

$$\frac{\partial x^i}{\partial x'^j} \frac{\partial x^i}{\partial x'^l} = 0 \quad (j \neq l) \quad (7)$$

или

$$g'_{jl} = 0 \quad (j \neq l). \quad (8)$$

Если эти условия выполнены, то новые координаты называют *ортогональными*. В качестве примера возьмем за  $x'^j$  сферические координаты  $r, \theta, \lambda$ . Тогда

$$\begin{aligned} x_1 &= r \sin \theta \cos \lambda, & x_2 &= r \sin \theta \sin \lambda, & x_3 &= r \cos \theta, \\ dx_1 &= \sin \theta \cos \lambda dr + r \cos \theta \cos \lambda d\theta - r \sin \theta \sin \lambda d\lambda, \\ dx_2 &= \sin \theta \sin \lambda dr + r \cos \theta \sin \lambda d\theta + r \sin \theta \cos \lambda d\lambda, \\ dx_3 &= \cos \theta dr - r \sin \theta d\theta \end{aligned}$$

$$\text{и } (dx_1)^2 + (dx_2)^2 + (dx_3)^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\lambda^2.$$

Отсюда

$$g'_{11} = 1, \quad g'_{22} = r^2, \quad g'_{33} = r^2 \sin^2 \theta.$$

Детерминант этого преобразования есть  $(g'_{11}g'_{22}g'_{33})^{1/2} = r^2 \sin \theta$ . Обратное преобразование имеет вид

$$\begin{aligned} dr &= \sin \theta \cos \lambda dx_1 + \sin \theta \sin \lambda dx_2 + \cos \theta dx_3, \\ r d\theta &= \cos \theta \cos \lambda dx_1 + \cos \theta \sin \lambda dx_2 - \sin \theta dx_3, \\ r \sin \theta d\lambda &= -\sin \lambda dx_1 + \cos \lambda dx_2. \end{aligned}$$

Матрица, обратная  $g'_{jl}$ , есть

$$g'^{il} = \frac{1}{g'_{11}g'_{22}g'_{33}} \begin{pmatrix} g'_{22}g'_{33} & 0 & 0 \\ 0 & g'_{33}g'_{11} & 0 \\ 0 & 0 & g'_{11}g'_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & r^{-2} \operatorname{cosec}^2 \theta \end{pmatrix}.$$

Здесь имеется существенное различие с любыми прямолинейными координатами. Теперь ненулевые компоненты  $g'_{jl}$  и  $g'^{il}$  не только не равны друг другу, но и имеют неодинаковую размерность. Действительно, в сферических координатах контравариантными компонентами перемещения являются  $dr$ ,  $d\theta$ ,  $d\lambda$ . Мы можем определить систему ковариантных компонент посредством

$$g'_{il} dx'^l = (dr, r^2 d\theta, r^2 \sin^2 \theta d\lambda).$$

Но ни одна из этих систем не совпадает с физическими компонентами. В качестве последних можно было бы взять компоненты перемещения относительно прямоугольных осей в этой точке, т. е.  $(dr, r d\theta, r \sin \theta d\lambda)$ . Все физические компоненты суть длины. На практике мы обычно интересуемся физическими компонентами. Мы обозначаем  $g_{11}$ ,  $g_{22}$ ,  $g_{33}$  через  $h_1^2$ ,  $h_2^2$ ,  $h_3^2$ , а перемещения, вызываемые малыми изменениями криволинейных координат, через  $ds_1$ ,  $ds_2$ ,  $ds_3$ . Тогда для физических компонент мы имеем

$$ds_1 = h_1 dx'^1, \quad ds_2 = h_2 dx'^2, \quad ds_3 = h_3 dx'^3.$$

Такие же соотношения справедливы и для компонент скорости. Если в прямоугольных координатах мы имеем векторное соотношение вида

$$\dot{x}_i = -\frac{\partial \Phi}{\partial x_i},$$

то оно непосредственно преобразуется к любой системе ортогональных осей и записанное в физических компонентах будет иметь вид

$$\dot{s}_j = \frac{\partial \Phi}{\partial s_j}, \quad h_1 \dot{x}'^1 = \frac{\partial \Phi}{h_1 \partial x'^1}$$

и т. д. Умножая или деля на соответствующие  $h$ , мы получим запись этих уравнений в виде соотношений между ковариантными или контравариантными компонентами.

В криволинейных координатах производная вектора  $A_i$  в общем случае не преобразуется как тензор второго порядка

из-за изменения  $g'_{\mu}$  от точки к точке. Это самое большое неудобство тензорного аппарата в криволинейных координатах. Его можно преодолеть соответствующей модификацией понятия производной, но это вывело бы нас далеко за рамки данной книги \*).

Правила преобразования координат будут столь же хорошо работать, даже если расстояния между близкими точками нельзя никаким выбором системы координат привести к виду  $\delta_{ik}x^i x^k$ . Например, на сфере можно выразить положение любой точки двумя переменными, но никаким образом нельзя выбрать переменные  $x_1, x_2$  так, чтобы во всех окрестностях на сфере было

$$ds^2 = dx_1^2 + dx_2^2.$$

Теория преобразований, при которых соответствующая расстоянию величина не может быть приведена к евклидовой форме без введения дополнительных измерений, составляет основу римановой геометрии и общей теории относительности.

**4.15. Электромагнитная теория.** Тензорный метод можно распространить на четыре измерения, и тогда он представляет удобный путь формулировки уравнений электромагнетизма. Если мы рассмотрим величину

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - c^2 dt^2, \quad (1)$$

где  $c$  — скорость света, то  $ds$ , взятое между двумя близкими событиями (каждое из которых задано тремя координатами положения и временем), будет одним и тем же для всех наблюдателей, даже если они находятся в равномерном относительном движении. Это математическое утверждение эквивалентно трем физическим: а) частица, движущаяся с постоянной скоростью относительно одной системы отсчета, движется с постоянной скоростью и относительно другой; б) оба наблюдателя приписывают одинаковые значения расстояниям под прямым углом к направлению их относительного движения; в) оба наблюдателя находят значение скорости света одним и тем же в любом направлении. Первые два закона взяты из ньютоновой физики, а последний — дополнительный закон, вытекающий из опытов Майкельсона — Морли. Теперь если мы напомним  $x_4 = ict$ , то  $ds^2$  сводится к сумме четырех квадратов и может рассматриваться как квадрат расстояния в евклидовой геометрии, только теперь мы должны работать в четырех измерениях.

---

\*) Полностью этот предмет освещен в [14, 15].

Так как  $ds^2$  одно и то же для всех систем отсчета, преобразование от одной системы к другой является ортогональным преобразованием в четырех измерениях.

Обозначим новую систему отсчета буквами со штрихами и выберем оси  $x_2$  и  $x'_2$ , а также  $x_3$  и  $x'_3$  под прямыми углами к направлению относительного движения. Тогда  $x_2 = x'_2$ ,  $x_3 = x'_3$  и

$$x_1^2 + x_4^2 = x_1'^2 + x_4'^2. \quad (2)$$

Это условие выполняется, если

$$x'_1 = x_1 \cos \alpha - x_4 \sin \alpha, \quad x'_4 = x_1 \sin \alpha + x_4 \cos \alpha. \quad (3)$$

Начало новой системы отсчета имеет нулевую скорость относительно этой системы. Следовательно, если мы положим  $\partial x'_1 / \partial x'_4 = 0$ , то  $dx_1 / dx_4$  будет  $v/ic$ , где  $v$  — скорость начала второй системы отсчета относительно первой. Но это дает

$$\operatorname{tg} \alpha = \frac{v}{ic}, \quad (4)$$

$$x'_1 = \beta x_1 + \frac{iv\beta}{c} x_4, \quad x'_4 = \frac{iv\beta}{c} x_1 - \beta x_4, \quad (5)$$

где

$$\beta = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (6)$$

Это преобразование, полученное Ларморовым и Лоренцем, является комплексным ортогональным преобразованием, но оно не унитарно, так как удовлетворяет соотношению  $\Pi = I$ , а не  $\Pi^\dagger = I$ . При действительных  $x_1$  и  $t$  это преобразование оставляет  $x'_1$  и  $t'$  действительными. Конечно, на него может быть наложен обычный поворот осей  $x_1$ ,  $x_2$ ,  $x_3$ , оставляющий неизменным  $x_4$ . Системы величин, определенные в каждой системе отсчета и преобразующиеся друг в друга с помощью преобразования Лармора — Лоренца, могут быть названы компонентами 4-вектора. Фундаментальным 4-вектором является сам  $x_\alpha$  ( $\alpha = 1, 2, 3, 4$ ). Но так как  $ds$  — скаляр, то  $mdx_\alpha/ds$  есть 4-вектор, где  $m$  — любой другой скаляр. Если  $m$  представляет собой собственную массу частицы, предполагаемую одинаковой во всех системах отсчета, а  $u$  — величина скорости, то получается, что  $mu_\alpha / \sqrt{1 - (u^2/c^2)}$  есть вектор, где

$$u_4 = \frac{dx_4}{dt} = ic.$$

Первые три компоненты соответствуют компонентам импульса в ньютоновой динамике, причем ньютонова масса заменяется на  $m/\sqrt{1-(u^2/c^2)}$ . Тогда, если при рассмотрении первых трех компонент отдельно от четвертой снабжать их латинскими индексами, получаем, что

$$\left( \frac{mu_a}{\sqrt{1-(u^2/c^2)}}, \quad \frac{mic}{\sqrt{1-(u^2/c^2)}} \right) \quad (7)$$

есть 4-вектор.

Кроме того, поскольку преобразование ортогонально, четырехмерный элемент объема  $dx_1 dx_2 dx_3 dx_4$  не изменяется. Следовательно, трехмерный элемент объема  $d\tau = dx_1 dx_2 dx_3$  преобразуется как  $1/dx_4$ , т. е. как

$$\frac{ds}{dx_4} = \sqrt{1 - \frac{u^2}{c^2}}. \quad (8)$$

Если определить плотность как собственную массу в единице объема, то она будет преобразовываться как  $m [1 - (u^2/c^2)]^{-1/2}$ , а импульс единицы объема — как произведение плотности и скорости. Сравнивая с (7), мы видим, что если  $\rho$  — плотность, то

$$(\rho u_a, i c \rho) \quad (9)$$

есть 4-вектор. Таким же образом, предполагая, что электрический заряд частицы одинаков во всех системах отсчета, находим, что если  $\rho$  — заряд единицы объема и  $j_a$  — плотность тока, то  $(j_a, i c \rho)$  тоже 4-вектор. Отсюда следует, что

$$\operatorname{div} \mathbf{j} + i c \frac{\partial \rho}{\partial x_4} = \operatorname{div} \mathbf{j} + \frac{\partial \rho}{\partial t} \quad (10)$$

есть скаляр и не меняется при преобразовании. Фактически он равен нулю, т. е.

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} = 0 \quad (11)$$

вследствие уравнения неразрывности.

Из пары уравнений Максвелла

$$c \operatorname{rot} \mathbf{H} - \frac{\partial \mathbf{E}}{\partial t} = 4\pi \mathbf{j}, \quad (12)$$

$$\operatorname{div} \mathbf{E} = 4\pi \rho \quad (13)$$

видно теперь, что

$$\left( c (\operatorname{rot} \mathbf{H})_a - \frac{\partial E_a}{\partial t}, \quad i c \operatorname{div} \mathbf{E} \right) \quad (14)$$

является 4-вектором. Здесь  $\mathbf{E}$  и  $\mathbf{H}$  — напряженности электрического и магнитного полей, причем первая в электростатических единицах. Очевидно, четырехмерная дивергенция (14) равна нулю.

Из второй пары уравнений Максвелла

$$\operatorname{div} \mathbf{H} = 0, \quad \operatorname{rot} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = 0 \quad (15)$$

видно, что  $\mathbf{H}$  является ротором некоторого 3-вектора  $\mathbf{A}$  и что тогда

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} \varphi, \quad (16)$$

где  $\varphi$  — некоторый скаляр. При заданном  $\mathbf{H}$  вектор  $\mathbf{A}$  определяется с точностью до градиента произвольного скаляра, влияние которого на  $\mathbf{E}$  может быть скомпенсировано соответствующим изменением  $\varphi$ . Поэтому мы можем наложить добавочное условие на  $\mathbf{A}$ ; мы потребуем, чтобы  $(\mathbf{A}, i\varphi) = A_\alpha$  было 4-вектором. Условие, что  $\partial A_\alpha / \partial x_\alpha$  должно быть скаляром, удовлетворяется, если

$$\operatorname{div} \mathbf{A} + \frac{1}{c} \frac{\partial \varphi}{\partial t} = 0. \quad (17)$$

Тогда получим

$$iE_1 = \frac{\partial A_1}{ic \partial t} - i \frac{\partial \varphi}{\partial x_1} = \frac{\partial A_1}{\partial x_4} - \frac{\partial A_4}{\partial x_1}, \quad (18)$$

$$H_1 = \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} \quad (19)$$

и аналогичные соотношения для остальных компонент. Величины  $iE_\alpha$  и  $H_\alpha$  оказываются шестью компонентами антисимметричного тензора. Введем

$$f_{\alpha\beta} = \frac{\partial A_\beta}{\partial x_\alpha} - \frac{\partial A_\alpha}{\partial x_\beta}. \quad (20)$$

Тогда

$$H_1 = f_{23}, \quad H_2 = f_{31}, \quad H_3 = f_{12}, \quad iE_1 = f_{41}, \quad iE_2 = f_{42}, \quad iE_3 = f_{43} \quad (21)$$

и тензор поля  $f_{\alpha\beta}$  принимает вид

$$\begin{pmatrix} 0 & H_3 & -H_2 & -iE_1 \\ -H_3 & 0 & H_1 & -iE_2 \\ H_2 & -H_1 & 0 & -iE_3 \\ iE_1 & iE_2 & iE_3 & 0 \end{pmatrix}. \quad (22)$$



Если мы теперь положим  $(j_\alpha, ic\rho) = s_\alpha$ , то пару уравнений (12), (13) можно будет записать в виде

$$\frac{\partial f_{\alpha\beta}}{\partial x_\beta} = \frac{4\pi}{c} s_\alpha, \quad (23)$$

а пару (15) — в виде

$$\frac{\partial f_{\alpha\beta}}{\partial x_\gamma} + \frac{\partial f_{\gamma\alpha}}{\partial x_\beta} + \frac{\partial f_{\beta\gamma}}{\partial x_\alpha} = 0. \quad (24)$$

Если все  $\alpha, \beta, \gamma$  различны, эти уравнения сводятся к (15); если два из индексов  $\alpha, \beta, \gamma$  равны, они обращаются в тождества.

*Сила Лоренца и обобщенный тензор напряжения.* Если  $\mathbf{k}$  — механическая сила на единицу объема, то

$$\mathbf{k} = \rho \mathbf{E} + \frac{1}{c} \mathbf{j} \times \mathbf{H}, \quad (25)$$

что можно записать как первые три компоненты 4-вектора

$$k_\alpha = \frac{1}{c} f_{\alpha\beta} s_\beta. \quad (26)$$

Четвертая компонента, определяемая этим соотношением, равна

$$k_4 = \frac{c}{i} \mathbf{E} \cdot \mathbf{j}, \quad (27)$$

так что  $k_4$  есть умноженная на  $c/i$  работа, совершенная полем над зарядом единицы объема в единицу времени. Теперь, используя (23), имеем

$$k_\alpha = \frac{1}{4\pi} f_{\alpha\beta} \frac{\partial f_{\beta\gamma}}{\partial x_\gamma}. \quad (28)$$

Если определим тензор  $T_{\alpha\gamma}$  соотношением

$$T_{\alpha\gamma} = \frac{1}{4\pi} \left\{ f_{\alpha\beta} f_{\beta\gamma} + \frac{1}{4} \delta_{\alpha\gamma} f_{\mu\nu} f_{\mu\nu} \right\}, \quad (29)$$

то после некоторых преобразований найдем, что

$$k_\alpha = \frac{\partial T_{\alpha\gamma}}{\partial x_\gamma}. \quad (30)$$

Тензор  $\mathbf{T}$  имеет вид

$$\begin{pmatrix} T_{11} & T_{12} & T_{13} & -iS_1/c \\ T_{21} & T_{22} & T_{23} & -iS_2/c \\ T_{31} & T_{32} & T_{33} & -iS_3/c \\ -iS_1/c & -iS_2/c & -iS_3/c & u \end{pmatrix}, \quad (31)$$

где совокупность  $3 \times 3$  величин в левом верхнем углу есть тензор напряжения Максвелла,  $\mathbf{S} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{H}$  — вектор Пойнтинга \*) и  $u = \frac{1}{8\pi} (E^2 + H^2)$  — плотность энергии.

**4.16. Цепи Маркова** (probabilities in chains). Рассмотрим систему, способную находиться в нескольких различных состояниях. В любой момент имеется некоторая вероятность  $x_i$  того, что она находится в состоянии, обозначенном индексом  $i$ . Рассмотрим вероятность того, что она будет находиться в состоянии  $i$  в следующий момент, при условии что она в первый момент находится в состоянии  $k$  (т. е. если  $x_k = 1$ ). Пусть эта вероятность равна  $a_{ik}$ . Тогда, поскольку она должна находиться во второй момент в каком-нибудь состоянии,

$$\sum_i a_{ik} = 1 \quad (1)$$

и при любом наборе значений  $x_k$  полная вероятность того, что система будет находиться во второй момент в состоянии  $i$ , равна

$$y_i = a_{ik} x_k. \quad (2)$$

Тасовка карт представляет собой наиболее известный пример. Здесь  $x_i$  будут вероятностями  $52!$  возможных раскладов карт перед тасовкой. Условия тасовки предполагают, что при любом раскладе карт перед перетасовкой после нее возможно несколько различных раскладов. При известном начальном раскладе  $k$  другие расклады будут иметь некоторые вероятности  $a_{ik}$ , которые должны в сумме давать единицу. Тогда по обычным правилам вычисления вероятностей сложных событий вероятность расклада  $i$  после перетасовки дается выражением (2).

Тот же принцип имеет место в статистической механике. Если имеются данные, что в один момент времени импульсы и координаты системы находятся в заданных конечных (ненулевых) интервалах, то последующее движение точно не определено даже в классической механике и движения будут сильно различаться в зависимости, скажем, от того, какая пара молекул встретится при следующем столкновении.

Такие вероятности могут образовывать цепи, поскольку эти процессы могут повторяться. Если сделано второе перераспре-

---

\*) В русской литературе этот вектор называют вектором Умова — Пойнтинга. — *Прим. ред.*

деление и вероятность состояния  $i$  после него есть  $z_i$ , то мы будем иметь

$$z_i = a_{im} y_m \quad (3)$$

и т. д. Последовательные вероятности получаются умножением на одну и ту же матрицу.

Это подсказывает общую трактовку. Уравнения

$$\lambda \theta_i = a_{ik} \theta_k \quad (4)$$

совместны, если

$$\|a_{ik} - \lambda \delta_{ik}\| = 0. \quad (5)$$

Допустим, что уравнение (5) имеет  $n$  корней  $\lambda_j$ , и обозначим  $\theta_i$  в соответствующих решениях через  $\theta_{ij}$ . Суммирование по  $j$  должно быть показано явно. Допустим далее, что  $x_i$  можно выразить в виде  $\sum_j \alpha_j \theta_{ij}$ . Тогда

$$y_i = \sum_k a_{ik} \sum_j \alpha_j \theta_{kj} = \sum_j \lambda_j \alpha_j \theta_{ij}$$

и результатом  $p$ -кратного применения операции (2) будет  $\sum_j \lambda_j^p \alpha_j \theta_{ij}$ , как и для класса задач динамики, рассмотренного в 4.061. Если в (5) мы сложим элементы каждого столбца, то получим  $n$  сумм, каждая из которых в силу (1) равна  $1 - \lambda$ . Следовательно,  $\lambda = 1$  всегда является корнем.

Отсюда следует, что для любой совокупности  $a_{ik}$  имеется возможный набор значений  $x_i$ , таких, что  $y_i = x_i$ . Они не обязательно должны быть равны между собой, так как мы не предполагали, что

$$\sum_k a_{ik} = 1. \quad (6)$$

Однако выполнение этого условия часто зависит просто от того, какие распределения мы объединим и какие будем рассматривать как взаимно исключающие. В задаче о тасовке карт каждое из  $a_{ik}$  равно  $1/52!$ ; как  $i$ , так и  $k$  имеют  $52!$  возможных значений, и это условие выполнено. Но если бы мы отождествили все расклады, в которых туз пик следует за тузом червей, а остальные считали бы по-прежнему различными, то все  $\sum_i a_{ik}$  были бы по-прежнему равны единице, но про  $\sum_k a_{ik}$  этого было бы утверждать нельзя, так как при данном  $i$  вероятности  $a_{ik}$  должны всегда быть больше, когда  $i$  соответствует объединенным событиям, чем когда не объединенным. Аналогично в задаче о столкновениях молекул газа

$\sum_k a_{ik}$  всегда будет больше, если  $i$  относится к области с большим объемом, чем к области с малым объемом. Но это можно устранить, выбрав все области одинакового размера. Таким образом, в фактических задачах условие (6) часто будет выполнено, а в противном случае мы можем обычно выбрать исходы таким образом, что оно станет выполняться.

Поэтому мы примем условие (6), что существенно упрощает анализ. Вернемся к (2). Пусть наибольшее и наименьшее из  $x_i$  будут  $M$  и  $m$ . Предположим, что не все  $x_i$  равны, так что  $M > m$ . Мы предположим также, что ни одно из  $a_{ik}$  не равно нулю. Тогда, используя (6), находим

$$y_i - M = \sum_k a_{ik} (x_k - M). \quad (7)$$

Члены с  $x_k = M$  не дают вклада в эту сумму, и ни один из членов суммы не положителен. Среди остальных имеется хотя бы один с  $x_k = m$ , и если  $\alpha$  — наименьшее из  $a_{ik}$  при условии, что имеется хотя бы два состояния ( $\alpha$  не превосходит  $1/2$ ), то отсюда следует

$$y_i - M < \alpha(m - M). \quad (8)$$

Аналогично

$$y_i - m > \alpha(M - m). \quad (9)$$

Тогда если  $M'$  и  $m'$  — любые два из  $y_i$ , то

$$M' < M - \alpha(M - m), \quad (10)$$

$$m' > m + \alpha(M - m) \quad (11)$$

$$|M' - m'| < (M - m)(1 - 2\alpha). \quad (12)$$

Таким образом, наибольшее колебание переменных умножается при каждом шаге на положительный множитель, меньший  $1 - 2\alpha$  и, следовательно, должно стремиться к нулю при достаточном числе попыток, каковы бы ни были начальные вероятности возможных распределений. Таким образом, вероятности всех распределений стремятся стать равными, если только ни одно из  $a_{ik}$  не равно нулю. Эту теорему можно распространить даже на случай, когда некоторые из  $a_{ik}$  равны нулю. Например, может быть так, что нельзя перейти непосредственно из состояния 1 в состояние 2, но можно перейти из состояния 1 в состояние 3 и обратно в состояние 2. Наш анализ одинаково пригоден, если мы применим его к результату после  $r$  шагов, причем  $a_{ik}$  будет теперь вероятностью того, что система достигнет состояния  $i$  за  $r$  шагов при условии, что вначале она находилась в состоянии  $k$ . Очевидно, в этом случае имеется возможность, что все  $a_{ik}$  будут положительны, хотя некоторые из них равны нулю для одного шага. Отсюда при условии, что

для некоторого конечного числа шагов  $r$  и любого состояния  $k$  имеется ненулевая вероятность достижения состояния  $i$  для каждого  $i$ , следует, что вероятности всех состояний будут стремиться к равенству независимо от начальных вероятностей.

Может случиться, что  $|M' - m'|$  в (12) равно нулю для всех пар. В этом случае однородное распределение вероятностей достигается за один шаг.

Другое доказательство, использующее теорию комплексного переменного, было дано Фреше. При данном  $\lambda_j$  пусть  $\theta_{ij}$  с наибольшим модулем таково, что  $|\theta_{ij}| = R$ . Тогда на диаграмме Аргана (см. 11.04)  $\theta_{kj}$  изображаются системой точек, лежащих внутри или на границе круга радиусом  $R$  с центром в начале координат, и каковы бы ни были  $a_{ik}$  при условии, что они неотрицательны и их сумма равна единице,  $a_{ik}\theta_{kj}$  не может лежать вне этого круга. Поэтому  $|\lambda| R \leq R$  и все

$$|\lambda_j| \leq 1. \quad (13)$$

Кроме того, для каждого  $i$  будет  $(\lambda_i - a_{ii})\theta_{ij} = \sum_{k \neq i} a_{ik}\theta_{kj}$  и

$$|\lambda_j - a_{ii}| |\theta_{ij}| \leq R \sum_{k \neq i} a_{ik} = R(1 - a_{ii}). \quad (14)$$

Но для одного из  $\theta_{ij}$  будет  $|\theta_{ij}| = R$ ; следовательно, для этого  $i$

$$|\lambda_j - a_{ii}| \leq 1 - a_{ii}. \quad (15)$$

Тогда если  $a_{ii} > 0$ , то  $\lambda_j$  лежит внутри или на круге с центром в  $a_{ii}$ , касающемся единичного круга в точке  $+1$ . Далее, такой круг с центром в наименьшем  $a_{ii}$  будет содержать все  $\lambda_j$ . Следовательно, если все  $a_{ik}$  при  $i = k$  отличны от нуля, то ни одно из  $\lambda$ , кроме  $\lambda = 1$ , не может быть по модулю равным единице. Но это условие требует просто, чтобы система, в каком бы состоянии она ни находилась, не обязательно выходила из этого состояния. В результате получается, что когда число шагов  $p$  взято достаточно большим, все  $\lambda^p$  стремятся к нулю, если  $\lambda \neq 1$ . Следовательно, если ни одно из  $a_{11}, a_{22}, \dots, a_{nn}$  не равно нулю, вероятности состояний будут стремиться к определенным пределам, которые даются решением (4) при  $\lambda = 1$ . Может случиться, что имеется более одного решения. Действительно, если бы все диагональные элементы были равны единице, то тогда  $y_i = x_i$  при всех  $i$  и вероятности вообще не изменяются. Необходимым и достаточным условием существования только одного решения с  $\lambda = 1$  является то, что матрица  $a_{ik} - \delta_{ik}$  должна иметь ранг  $n - 1$ . До некоторого момента мы можем рассмотреть все решения с  $|\lambda| = 1$  вместе. Пусть в (4)  $\theta_1$  будет  $\theta_i$

с наибольшим модулем  $R$ . Тогда если любое  $\theta_k$ , для которого  $a_{1k} \neq 0$ , таково, что  $|\theta_k| < R$ , то

$$R = |\theta_1| = |a_{1k}\theta_k| < \sum a_{ik}R = R,$$

что является противоречием. Эта аргументация может быть далее продолжена на  $i=2$ , если  $|\theta_2| = |\theta_1|$ , и т. д. Тогда, если все  $a_{ik}$  отличны от нуля, единственным решением с  $|\lambda|=1$  будет

$$\theta_1 = \theta_2 = \dots = \theta_n, \quad \lambda = 1,$$

и, следовательно, мы получили другое доказательство того, что единственные возможные предельные установившиеся значения вероятностей равны между собой. Это доказательство можно снова распространить на случай, когда некоторые из  $a_{ik}$  равны нулю; оно может оказаться недействительным только, если для всех  $|\theta_i|$ , вплоть до  $|\theta_m|$  ( $m < n$ ), окажется, что все  $a_{ik}$  исчезают при  $k > m$  и  $i \leq m$ . Но это означает, что из любого состояния с  $k > m$  невозможно перейти в состояние с  $k \leq m$ . Кроме того, в этом случае легко видеть, что на основании соотношений (1) и (6)  $a_{ik}$  должны исчезать и при  $k \leq m$ , и при  $i > m$ , т. е. обратный процесс также невозможен. Тогда будет бесконечное число возможных предельных состояний в зависимости от полных начальных вероятностей различных независимых систем. Таким образом, случай кратного корня  $\lambda = 1$  соответствует случаю, когда вероятности распадаются на две или более независимые системы и предельная вероятность будет одинаковой для каждого состояния в одной и той же системе.

Это, между прочим, показывает, что мы не *всегда* можем добиться выполнения условия (6). Ведь мы можем, конечно, иметь такую систему  $a_{ik}$ , что вероятности допускают переход из состояний с  $i \leq m$  в состояния с  $i > m$ , но не наоборот. Ясно, что в этом случае вся вероятность стремится сосредоточиться в последней системе.

Мы уже видели, что существование решения с  $|\lambda|=1$ , за исключением  $\lambda \neq 1$ , требует, чтобы некоторые  $a_{ik}$  с  $i=k$  равнялись нулю, и более того, чтобы сделать при  $|\theta_k| \leq |\theta_1|$

$$|a_{1k}\theta_k| = |\theta_1|, \quad (16)$$

должно быть некоторое  $\theta_k$ , скажем  $\theta_2$ , такое, что  $|\theta_2| = |\theta_1|$ , и все  $\theta_k$ , такие, что  $a_{1k} \neq 0$  должны быть равны  $\theta_2$ . Так как  $a_{ik}$  действительны, то  $a_{11} = 0$  (как мы уже знаем) и  $\theta_2 = \lambda\theta_1$ .

Теперь мы можем взять  $\theta_2$  вместо  $\theta_1$  и заключить, что имеется некоторое  $\theta_3$ , равное  $\lambda\theta_2$ , и для всех  $k$ , таких, что  $\theta_k \neq \theta_3$ , будет  $a_{2k} = 0$ . Но так как имеется только  $n$  значений  $k$ , этот вопрос должен замкнуться на некотором числе

шагов  $m$ , причем  $m \leq n$ , следовательно, поскольку все найденные  $\theta_k$  различны, то в каждой строке  $a_{ik}$  все элементы равны нулю, кроме одного, который поэтому должен быть равен единице. Таким образом, матрица  $a_{ik}$  содержит минор вида (при  $m = 4$ )

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

и уравнение  $\|a_{ik} - \lambda \delta_{ik}\| = 0$  удовлетворяется, если

$$1 - \lambda^m = 0. \quad (17)$$

Корни этого уравнения равны  $\exp(2\pi i r/m)$ , где  $r$  — любое целое число от 0 до  $m-1$ . Форма  $a_{ik}$  непосредственно показывает, что если система вначале находится в одном из этих  $m$  состояний, то она будет с необходимостью проходить через остальные в определенном порядке и вернется в исходное состояние за  $m$  шагов.

Если  $m = n$ , эта система должна описывать полный набор возможных состояний. При  $m < n$  не включенные в цикл состояния не зависят от состояний цикла и могут образовывать другие циклы или иметь  $|\lambda| < 1$  для всех не равных единице корней.

Следовательно, задача распадается на несколько случаев.

1. Если  $a_{ik}$  таковы, что независимо от начального состояния имеется ненулевая вероятность достижения системой любого состояния за определенное конечное число шагов, то вероятности всех состояний стремятся стать равными, когда число перераспределений делается большим. Это *эргодическая теорема* (для рассматриваемого случая. — *Ред.*).

2. Если  $a_{ik}$  таковы, что состояния образуют множества, такие, что для любого из них каждое состояние может быть достигну исходя из произвольного состояния этого множества и не может, если исходить из состояния любого другого множества, то характеристическое уравнение имеет единицу кратным корнем и вероятности в каждом множестве стремятся стать равными, но их предельные значения зависят от начальных значений.

3. В предыдущих случаях все корни характеристического уравнения либо равны единице, либо имеют модуль, меньший единицы. Если имеется не равный единице корень с модулем, равным единице, то некоторые из состояний будут образовывать цикл, такой, что каждое из них имеет детерминированное

последующее и первоначальное состояние достигается вновь не более, чем за  $n$  шагов.

Случай 1 встречается при обычной тасовке карт и в кинетической теории газов. (Приводимые обычно доказательства, принадлежащие Больцману и Гиббсу, ошибочны.) Случай 2 может встретиться при тасовке карт, если колоду разделить на две половины и тасовать каждую отдельно, а в конце сложить вместе. Очевидно, что при этом многие расположения, достижимые при тасовке полной колоды, становятся невозможными. Но вероятность того, что тузы пик и червей окажутся вместе, всегда будет зависеть от вероятности того, что они попадут с самого начала в одну половину колоды. Случай 3 в тасовке карт будет описывать ситуацию, когда тасовка состоит в перекладывании верхней карты вниз, что никогда не даст большего, чем один раз снять колоду, и никогда не будет правильной тасовкой. Но этот случай связывает, кроме того, любую детерминированную механику с цепями Маркова, так как он показывает, что для точной определенности будущих состояний при известном состоянии в некоторый момент необходимо и достаточно, чтобы временные множители в решениях имели модуль, равный единице \*).

**4.17. Интегральные уравнения.** Они бывают различных типов и обладают тем общим свойством, что неизвестная функция содержится под знаком интеграла. О них имеется обширная литература, но в данной книге они могут быть рассмотрены лишь кратко. Имеются следующие три родственных типа уравнений:

$$\int_0^a K(x, y) \varphi(y) dy = f(x), \quad (1)$$

$$\varphi(x) + \int_0^a K(x, y) \varphi(y) dy = f(x), \quad (2)$$

$$\int_0^a K(x, y) \varphi(y) dy = \lambda \varphi(x). \quad (3)$$

Здесь пределы интегрирования фиксированы,  $K(x, y)$  и  $f(x)$  — известные функции, а  $\varphi$  подлежит определению. Эти уравнения можно рассматривать как предельный случай матричных

---

\*) Приложения к статистической механике и квантовой теории см. в [16, 17].



уравнений. Ведь если мы выберем точки разбиения при  $y/a = 0, 1/n, 2/n, \dots, (n-1)/n$ , то можно положить

$$\begin{aligned} f(x_i) &= X_i, & \varphi(y_k) &= Y_k, \\ K(x_i, y_k) &= K_{ik}. \end{aligned} \quad (4)$$

Тогда эти уравнения являются пределами следующих;

$$\frac{a}{n} \sum_{k=0}^n K_{ik} Y_k = X_i, \quad (5)$$

$$Y_i + \frac{a}{n} \sum_{k=0}^n K_{ik} Y_k = X_i, \quad (6)$$

$$\frac{a}{n} \sum_{k=0}^n K_{ik} Y_k = \lambda Y_i \quad (7)$$

при  $n \rightarrow \infty$ . Последние уравнения полезны для численного решения. При использовании квадратурных формул более высокой степени точности уравнения можно решать непосредственно при числе интервалов разбиения до 12 на машине Маллока для решения систем линейных уравнений. Но они также показывают, что между этими типами интегральных уравнений и линейными алгебраическими уравнениями должно быть большое сходство. В частности, уравнения (7) будут в общем случае разрешимы только при некоторой дискретной системе значений  $\lambda$ , число которых равно  $n$ , и, следовательно, в пределе будет бесконечное число решений. Затруднения также будут при решении (5) и (6) в случаях, когда обратится в нуль определитель, составленный из коэффициентов при  $Y_k$ .

Функция  $K(x, y)$  называется *ядром* уравнения. Условие  $K_{ik} = K_{ki}$  для симметричной матрицы соответствует условию

$$K(y, x) = K(x, y). \quad (8)$$

Если оно выполняется, ядро называется симметричным. Мы можем определить эрмитово ядро условием

$$K(x, y) = K^*(y, x). \quad (9)$$

Аналогично имеем антисимметричное ядро, определяемое соотношением

$$K(x, y) = -K(y, x). \quad (10)$$

Существует аналог ортогональных матриц. Уравнение (1) может иметь решение

$$\varphi(y) = \int_0^a K(y, x) f(x) dx, \quad (11)$$

которое, как легко видеть, соответствует решению системы уравнений умножением на обратную матрицу в случае, когда  $aa = I$ .

Интегральные уравнения редко решаются в конечном виде; однако имеются обширные теории их решения методом последовательных приближений [18—20].

### ПРИМЕРЫ

1. Рассматривая матрицы  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , покажите, что две симметричные матрицы не обязательно коммутируют.

2. Докажите, что  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & -\operatorname{tg} \frac{1}{2} \theta \\ \operatorname{tg} \frac{1}{2} \theta & 1 \end{pmatrix} \begin{pmatrix} 1 & \operatorname{tg} \frac{1}{2} \theta \\ -\operatorname{tg} \frac{1}{2} \theta & 1 \end{pmatrix}^{-1}$ .

3. Одна из квадратичных форм

$$3x^2 + 2y^2 + 5z^2 + 2yz - 2zx, \quad x^2 + 2y^2 + 8yz + 12zx + 12xy$$

положительно определена. Укажите ее и найдите действительное линейное преобразование, которое приводит эти формы к виду

$$\xi^2 + \eta^2 + \zeta^2, \quad 4\xi^2 - 2\eta^2 - \zeta^2. \quad (\text{Prelim., 1943.})$$

4. Найдите действительное несингулярное линейное преобразование, которое приводит форму

$$P - \frac{2}{n-1} Q$$

к стандартному виду  $y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_{p+q}^2$ , где  $P \equiv x_1^2 + \dots + x_n^2$ ,  $Q = \sum_{i < j} x_i x_j$ .

Определите для всех действительных  $\lambda$  ранг  $r = p + q$  и сигнатуру  $s = p - q$  формы  $P + 2\lambda Q$ . (М. Т., 1935.)

5. Два однородных стержня  $AB$ ,  $BC$  шарнирно скреплены в  $B$  и опираются в  $A$ ,  $B$  и  $C$  на пружины жесткости  $\lambda$ , образуя прямую линию. Если масса каждого стержня равна  $m$ , покажите, что периоды собственных колебаний определяются соотношением:  $m\gamma^2/\lambda = 3,3 \pm \sqrt{3}$ .

6. Длинная цепочка шарнирно соединенных стержней массы  $3m$  и длины  $2a$  каждый, расположена вдоль прямой линии, и в каждом соединении имеется пружина, создающая возвращающую силу, равную произведению  $-m\lambda_0^2$  на поперечное смещение. На один конец цепочки действует сила

с периодом  $2\pi/\lambda$ , а второй конец закреплен. Покажите, что при  $\lambda_0/\sqrt{3} < \lambda < \lambda_0$  все стержни возбуждаются одинаково, но при других значениях  $\lambda$  движение сосредоточивается в окрестности возбуждаемого конца. (Это механический аналог радиотехнического частотного фильтра.)

7. На невесомой струне  $AB$  длины  $l$  расположены  $n-1$  частиц  $P_1, P_2, P_3, \dots, P_{n-1}$  массы  $m/n$ , так что  $AP_1 = P_1P_2 = \dots = P_{n-1}B$ .  $A$  и  $B$  закреплены, и вся струна находится под натяжением  $T$ .

Покажите, что в собственном колебании с периодом  $2\pi/\gamma$  амплитуды  $a_j$  малых поперечных движений частиц связаны соотношением

$$a_{j+1} - 2 \cos \alpha a_j + a_{j-1} = 0 \quad (1 \leq j \leq n-1),$$

где  $\cos \alpha = 1 - \gamma^2 ml / 2\pi^2 T$ ; найдите собственные частоты.

Переходя к пределу при  $n \rightarrow \infty$ , получите собственные частоты однородной тяжелой струны с закрепленными концами.

8. Проиллюстрируйте метод Релея на примере колебаний однородного стержня длины  $l$  с жесткостью на изгиб  $EI$ , оба конца которого закреплены и свободны от продольных усилий. Покажите, что метод Релея дает для

основной частоты  $\nu$  значение  $\frac{1}{2\pi} \sqrt{\frac{504EI}{\rho l^4}}$ , если приближенная форма стержня в любой момент времени выбирается в виде  $y = f(x)$ , где  $f(x)$  — простейший полином, удовлетворяющий граничным условиям. (М/с, III, 1936.)

9. Если  $\mathbf{a}$  и  $\mathbf{b}$  эрмитовы, покажите, что  $\mathbf{ab} + \mathbf{ba}$  и  $i(\mathbf{ab} - \mathbf{ba})$  также эрмитовы.

10. Покажите, что если  $\mathbf{H}$  — эрмитова матрица, то матрица  $\mathbf{U} = (\mathbf{H} - i\mathbf{I}) \times (\mathbf{H} + i\mathbf{I})^{-1}$  существует и унитарна, и что если  $\mathbf{U}$  — унитарная матрица, то  $\mathbf{H} = -i(\mathbf{I} + \mathbf{U})^{-1}(\mathbf{I} - \mathbf{U})$  существует и эрмитова при условии, что  $\mathbf{U}$  не имеет собственного числа, равного единице.

Получите отсюда, что любой собственный вектор  $\mathbf{U}$  является собственным вектором  $\mathbf{H}$  и наоборот и что, следовательно, существует некоторая унитарная матрица  $\mathbf{I}$ , такая, что  $\mathbf{I}^+ \mathbf{U} \mathbf{I}$  диагональная. Обобщите последний результат на случай, когда  $\mathbf{U}$  имеет собственное число, равное 1.

11. Покажите, что любая матрица  $2 \times 2$ , антикоммутирующая с матрицей Паули  $\sigma_1$ , имеет вид  $a\sigma_2 + b\sigma_3$ . Покажите далее, что существуют 8 линейно независимых матриц  $4 \times 4$ , антикоммутирующих с матрицей Эддингтона  $E_1$ .

12. Пусть кватернион определен правилом

$$\mathbf{u} = iu_0 + iu_1 + ju_2 + ku_3,$$

где

$$i^2 = j^2 = k^2 = -1, \quad ij = k, \quad ji = -k,$$

а  $u_0, u_1, u_2, u_3$  — действительные числа. Покажите, что для любых кватернионов  $\mathbf{u}, \mathbf{w}$  существует кватернион  $\mathbf{v}$  такой, что

$$\mathbf{uv} = \mathbf{w}$$

при условии, что не все  $u_0, u_1, u_2, u_3$  равны нулю.

Покажите, что матрицы Эддингтона  $E_{12}, E_{23}, E_{31}$  удовлетворяют условиям для  $i, j, k$ .

13. Треугольная матрица  $\mathbf{T}$  определяется как матрица, у которой  $T_{ik} = 0$  при  $k > i$  или  $k < i$ . Покажите, что треугольная матрица, все диагональные элементы которой отличны от нуля, имеет обратную и сопоставьте этот результат с решением системы линейных уравнений  $a_{ik}x_k = b_i$  методом последовательного исключения.

Покажите также, что если  $\mathbf{T}\mathbf{T}^+ = \mathbf{T}^+\mathbf{T}$ , то  $\mathbf{T}$  — диагональная.

14. Покажите, что при любом преобразовании вида  $\mathbf{B} = \mathbf{P}\mathbf{A}\mathbf{Q}$ , где  $\mathbf{P}$  и  $\mathbf{Q}$  неособые, ранги матриц  $\mathbf{A}$  и  $\mathbf{B}$  равны.

15. Покажите, что если  $\mathbf{H}$  — эрмитова матрица, то  $\mathbf{I}^+\mathbf{H}\mathbf{I}$  — также эрмитова, и что если  $\mathbf{A}$  антиэрмитова, то  $\mathbf{I}^+\mathbf{A}\mathbf{I}$  — антиэрмитова.

16. Пусть  $\mathbf{A}$  есть матрица  $n \times n$  и имеет  $n$  различных собственных значений, и  $\mathbf{B}$  коммутирует с  $\mathbf{A}$ . Докажите, что  $\mathbf{B}$  выражается в виде полинома от  $\mathbf{A}$  степени  $n - 1$ .

17. Пусть  $\mathbf{I}^{-1}\mathbf{A}\mathbf{I}$  — диагональная, и все ее диагональные элементы по модулю равны 1, а  $\mathbf{I}$  унитарная. Докажите, что  $\mathbf{A}$  унитарная.

## ЛИТЕРАТУРА

1. *Ferrar W. L.*, Algebra, p. 162—167.
2. *Hilbert D., Courant R.*, Methoden d. Math. Phys., **1**, 13. (Русский перевод: Гильберт Д., Курант Р., Методы математической физики, М.-Л. Гостехиздат, 1956.)
3. *Condon J., Shortley G.*, Theory of Atomic Spectra, CUP, 1935, p. 222—226. (Русский перевод: Кондон Е., Шортли Г., Теория атомных спектров, М., ИЛ, 1949.)
4. *Böcher M.*, Introduction to Higher Algebra, p. 32. (Русский перевод: Бохер М., Введение в высшую алгебру, М.-Л., Гостехиздат, 1933.)
5. *Weierstrass*, Monatsber. d. K. Acad. d. Wiss. Berlin, 1858, Werke, **1**, S. 223—246.
6. *Lamb H.*, Higher Mechanics, 1920, p. 222—226.
7. *Bromwich T. J. I. A.*, Quadratic Forms, 1906.
8. *Temple G., Bickley W. G.*, Rayleigh's Principle, Oxford, 1933.
9. *Littlewood D. E.*, The theory of Group Characters, 1940.
10. *Eddington A. S.*, Relativity Theory of Protons and Electrons, 1936.
11. *Eddington A. S.*, Fundamental Theory, 1946.
12. *Miller*, Lehrbuch der Kristallographie, 1846.
13. *Ewald P. P.*, Kristalle und Röntgenstrahlen, 1923.
14. *McConell A. J.*, Applications of Absolute Differential Calculus, 1931.
15. *Levi — Civita T.*, The Absolute Differential Calculus, 1927.
16. *Jeffreys H.*, Proc. Roy. Soc., **A 160**, 337—347 (1937).
17. *Jeffreys H.*, Phil. Mag. (7), **33**, 815—831 (1942).
18. *Schmidt E.*, Math. Ann., **63**, 433—476 (1907).
19. *Smithies F.*, Proc. Lond. Math. Soc., **43**, 255—279 (1937); **46**, 409—460 (1940), Duke Math. J., **8**, 107—133 (1941).
20. *Bückner H.*, Die praktische Behandlung von Integralgleichungen, Springer, 1952.
21. *Coexeter H. S. M.*, Introduction to Geometry, 1961.
22. *Jeffreys H.*, Mathematical Gazette, XLIX, 192—194 (1965).

## ЗАМЕЧАНИЯ ОБ ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЯХ

В гл. 2, 3 и 4 рассмотрение ортогональных преобразований в трехмерном евклидовом пространстве было почти целиком ограничено преобразованиями с определителем, равным  $+1$ , т. е. вращениями. Преобразования с определителем  $-1$  представляют собой отражения относительно некоторой плоскости

с последующим вращением вокруг  $z$  нормали к этой плоскости (*вращение с отражением*). Так как отражение в некоторой плоскости эквивалентно инверсии относительно начала координат (изменению знака всех координат) с последующим поворотом на  $\pi$  вокруг нормали к этой плоскости, заключаем, что любое ортогональное преобразование с детерминантом  $-1$  изображает также *вращение с инверсией*.

Полное обсуждение изометрических преобразований евклидова пространства дано Коксетером в [21] в гл. 3 и 7. Матрицы этих преобразований приведены в полном виде Джеффрисом в [22].

## КРАТНЫЕ ИНТЕГРАЛЫ

Один за другим или все сразу.

Джильберт „Дворцовый страж“

**5.01. Расстояние, окрестности, кривые, области.** В этой главе мы будем иметь дело с функциями двух или более переменных, скажем  $x, y, z$ . Часто (но не всегда) эти переменные будут прямоугольными координатами точки в обычном смысле. В большинстве случаев мы будем считать, что число переменных равно двум или трем, однако распространение на большее число переменных очевидно. Некоторые простые понятия евклидовой геометрии можно с пользой распространить и на случаи, когда переменные не являются прямоугольными координатами. Одним из важнейших таких понятий является *расстояние*. Расстояние  $r$  между двумя точками, заданными прямоугольными координатами  $(x, y, z)$   $(x', y', z')$ , является неотрицательной величиной, удовлетворяющей соотношению

$$r^2 = (x' - x)^2 + (y' - y)^2 + (z' - z)^2. \quad (1)$$

Если переменные имеют различные размерности, то это выражение бессмысленно, однако мы можем распространить определение и на этот случай, приняв

$$r^2 = \alpha (x' - x)^2 + \beta (y' - y)^2 + \gamma (z' - z)^2, \quad (2)$$

где  $\alpha, \beta, \gamma$  — положительные константы, выбранные так, чтобы сделать сложение возможным, а затем с помощью замены переменных свести это выражение снова к виду (1). Расстояние, определенное таким образом, вообще говоря, не будет евклидовым расстоянием; например,  $x, y, z$  могут быть сферическими координатами, и  $r$  тогда будет совершенно отлично от расстояния, выраженного с помощью этих координат. Часто  $x, y, z$  будут числами. Оправдание такого определения заключается в том, что если  $x' \rightarrow x, y' \rightarrow y, z' \rightarrow z$ , то  $r \rightarrow 0$ , и наоборот. Утверждение, что два набора значений переменных близки друг другу, равносильно утверждению, что расстояние мало.

Заметим, что если  $P, Q, R$  являются тремя точками с координатами  $x_i, y_i, z_i$  ( $i$  не обязательно изменяется от 1 до 3) то

$$y_i - z_i = (y_i - x_i) - (z_i - x_i), \quad (3)$$

$$r_{QR}^2 = r_{PQ}^2 + r_{PR}^2 - 2 \sum (y_i - x_i)(z_i - x_i). \quad (4)$$

Но по неравенству Коши — Буняковского мы имеем

$$\left[ \sum (y_i - x_i)(z_i - x_i) \right]^2 \leq \sum (y_i - x_i)^2 \sum (z_i - x_i)^2 = r_{PQ}^2 r_{PR}^2. \quad (5)$$

Поэтому

$$(r_{PQ} - r_{PR})^2 \leq r_{QR}^2 \leq (r_{PQ} + r_{PR})^2, \quad (6)$$

$$|r_{PQ} - r_{PR}| \leq r_{QR} \leq r_{PQ} + r_{PR}. \quad (7)$$

Это является обобщением известного неравенства евклидовой геометрии, однако его можно рассматривать здесь как чисто аналитическую теорему. По индукции если  $P_r(x_{ri})$  является множеством точек, то расстояние от  $P_1$  до  $P_n$  меньше суммы расстояний от  $P_1$  до  $P_2$ , от  $P_2$  до  $P_3$ , ..., от  $P_{n-1}$  до  $P_n$ .

Мы будем широко пользоваться геометрической терминологией; в частности, мы будем говорить, что данный набор значений переменных определяет *точку*. Нам нужен для случая нескольких переменных аналог понятия *интервала* для одного переменного. Для этого заметим: 1) что значения переменной (при необходимости после умножения на постоянный множитель размерности) являются действительными числами, и воспользуемся следующими фактами: 2) между двумя любыми действительными числами имеется третье число; 3) любая предельная точка множества действительных чисел также является действительным числом. Далее, 4) если  $x_1, x_2$  являются точками интервала, то любое значение между  $x_1$  и  $x_2$  также является точкой интервала — интервалы не содержат пустых промежутков внутри себя; 5) любой интервал, за исключением интервала, состоящего из множества всех конечных действительных чисел, имеет по крайней мере одну концевую точку, которая характеризуется тем, что на любом сколь угодно малом расстоянии от нее имеются действительные числа как принадлежащие к интервалу, так и не принадлежащие. Сама концевая точка может либо принадлежать, либо не принадлежать к интервалу. Условия 2 и 4 требуют только понятия упорядоченности. Нужно лишь уметь четко определять, когда  $x_2$  больше  $x_1$ , не говоря на сколько больше. Однако условия 3 и 5 действительно требуют понятия расстояния, которое мы будем обозначать с помощью  $|x_2 - x_1|$ .

Обобщение понятия интервал на случай двух или более переменных будет называться *областью*. Мы еще не готовы

определить его, но мы видели, что для случая одного переменного главную роль играет понятие расстояния, а расстояние легко определить для любого конечного числа переменных. Это дает нам возможность немедленно дать определение предельной точки в  $n$  измерениях через понятие окрестность. Окрестностью точки  $P$ , заданной координатами  $(x, y, z)$  в трех измерениях, называется множество точек  $Q(x', y', z')$ , таких, что расстояние  $PQ$ , определенное как неотрицательная величина  $r$ , получающаяся по формуле

$$r^2 = (x' - x)^2 + (y' - y)^2 + (z' - z)^2,$$

меньше, чем некоторая положительная величина  $\delta$ . (Если мы исключаем саму точку  $P$  из окрестности, то будем отмечать это.) Для любого множества точек  $R$  точка  $P$  является предельной точкой, если для любого  $\delta$  найдется точка множества  $R$ , отличная от  $P$  и отстоящая от  $P$  на расстояние меньше  $\delta$ , или, короче, если любая окрестность точки  $P$  содержит по крайней мере одну точку  $R$ , отличную от  $P$ , а следовательно, содержит бесконечно много точек  $R$  так же, как и в случае одного измерения. Множество называется замкнутым, если все предельные точки его подмножеств принадлежат самому множеству.

Для случая нескольких переменных мы могли бы, вообще говоря, потребовать, когда все, кроме одной переменной, фиксированы, чтобы оставшаяся переменная могла принимать значения только из одного интервала. Однако это немедленно привело бы к трудностям. Для двух измерений при фиксированном  $y$  это означало бы, что  $x$  должны принадлежать некоторому интервалу; таким образом, мы не смогли бы назвать областью внутреннюю часть выпуклого многоугольника. Поэтому обобщение условия 4 на случай нескольких переменных требует некоторой модификации очевидной процедуры, когда переменные изменяются поодиночке. Для случая нескольких переменных мы скажем, что если точки  $P, Q$  лежат внутри области, то они могут быть соединены кривой, лежащей целиком внутри области. Для этого правила, однако, необходимо дать определение кривой.

Понятие концевой точки заменяется понятием граничной точки; понятие внутренней точки не представляет никаких трудностей.

Поскольку мы дадим аналитическое определение кривой, то могли бы попытаться дать аналитическое определение области, сказав, что область — это множество точек, удовлетворяющих некоторому неравенству, подобно тому как интервал может быть охарактеризован неравенством  $a < x < b$ . Действительно, область часто определяют таким образом, но при этом необ-



ходима одна предосторожность. Предположим, что мы взяли неравенство  $|x| \times |x - 2| + |y| < 1/2$ . Это неравенство определяет две площадки вокруг точек  $(0, 0)$  и  $(2, 0)$ , которые не имеют общих точек. Неравенство  $|\sin x| + |y| < 1/2$  определяет бесконечное множество площадок вокруг точек  $x = n\pi$ ,  $y = 0$ , причем никакие две из них не имеют общих точек. Поэтому одиночное неравенство не всегда дает удовлетворительное описание области.

**5.011. Кривые.** Мы должны перевести на математический язык то, что подразумеваем, когда говорим, что множество точек образует кривую. Мы будем вести рассуждение для случая трех переменных, хотя оно применимо и для любого конечного числа переменных. Первое существенное замечание состоит в том, что точки появляются в определенном порядке, когда мы продвигаемся по кривой. Второе состоит в том, что на кривой нет разрывов; мы можем прийти из одной точки кривой в другую, не покидая кривой. Эти условия можно выразить следующим образом: переменные  $x, y, z$  для точек, принадлежащих кривой, являются непрерывными функциями параметра  $t$  в некотором интервале значений  $t$ . (Если интервал значений  $t$  конечен и замкнут, то  $x, y, z$  будут ограничены.) Направление возрастания  $t$  определяет порядок точек на кривой; наконец, условие, что  $x, y, z$  являются непрерывными функциями  $t$  и все значения  $t$  в данном интервале допустимы, обеспечивает отсутствие разрывов. Может показаться, что предположение о непрерывной зависимости функций  $x, y, z$  от непрерывного параметра  $t$  не является очевидным и верным для всех множеств, которые мы могли бы назвать кривыми. Но это определение дает больше общности, чем нам нужно, а не меньше. Были определены кривые, удовлетворяющие этому условию с ограниченным  $t$ , которые проходят через каждую точку замкнутой области, ограниченной единичным квадратом \*). Возможность получить полезные результаты относительно кривых возникает лишь при дальнейшем ограничении, к которому мы сейчас перейдем, а именно при условии, что кривые имеют длину; и если они ее имеют, то эта длина вдоль кривой от некоторой фиксированной до рассматриваемой точки и будет служить параметром  $t$ .

Кривая имеет кратную точку, если имеются неравные значения  $t = c, t = d$ , такие, что  $(x, y, z)_{t=c} = (x, y, z)_{t=d}$ . Кривые без кратных точек называются *простыми*.

Если  $x, y, z$  ограничены вдоль кривой и принимают одинаковые значения для обоих крайних значений  $t$ , то кривая называется *замкнутой*. Простая кривая, не являющаяся замкнутой,

\*) Впервые это сделал Пеано, ср. [7, стр. 330].

называется *дугой*. (Значение слова *замкнутый* в применении к кривым никак не связано с соответствующим значением этого слова, применяемого к интервалам и областям.) Замкнутая двухмерная кривая обычно называется *контуром*.

**5.012.** Длина кривой определяется следующим образом. Возьмем на кривой точки  $P_r(x_r, y_r, z_r)$  в порядке возрастания  $t_r$ , причем  $A$  является точкой  $(x_0, y_0, z_0)$ , а  $B$  — точкой

$$(X, Y, Z) = (x_n, y_n, z_n),$$

и определим  $h_r$  как положительную величину, удовлетворяющую соотношению

$$h_r^2 = (x_{r+1} - x_r)^2 + (y_{r+1} - y_r)^2 + (z_{r+1} - z_r)^2,$$

и пусть

$$s_n = \sum_{r=0}^{n-1} h_r.$$

Когда  $n$  бесконечно возрастает и наибольший из интервалов  $t$  стремится к нулю, сумма  $s_n$  может стремиться к некоторому пределу. Если это так и если этот предел не зависит от способа выбора  $t_r$  на каждом шаге, то этот предел и называется длиной кривой между точками  $A$  и  $B$ . Эквивалентное утверждение состоит в том, что длина является верхней гранью (если она существует) сумм  $s_n$  для всех способов выбора  $n$  и  $t_r$  и, очевидно, не изменяется, если  $t$  заменить любым параметром  $t'$ , являющимся монотонной функцией от  $t$ . Мы обозначим длину кривой от точки  $A$  до произвольной точки  $P$  через  $s$ . Данное значение  $s$  определяет некоторую точку на кривой, и длина  $s$ , если она существует, может быть использована как параметр вместо  $t$ . Кривые, имеющие длины, называются *спрямляемыми*.

**5.013.** *Необходимое и достаточное условие того, что кривая имеет конечную длину, заключается в том, что  $x, y, z$  все имеют ограниченную вариацию на кривой.* Предположим сначала, что кривая имеет конечную длину. Тогда

$$h_r \geq |x_{r+1} - x_r|, \quad s \geq s_n \geq \sum |x_{r+1} - x_r|.$$

Поэтому для любого набора точек  $P_r$  сумма  $\sum |x_{r+1} - x_r|$  не превосходит длины кривой; поэтому  $x$  имеет ограниченную вариацию на кривой. Аналогично  $y$  и  $z$  также имеют ограниченную вариацию на кривой. Обратно, пусть  $V_x, V_y, V_z$  — полные вариации переменных  $x, y, z$  на кривой все меньше  $M$ . Тогда для любого подразделения

$$\sum h_r \leq \sum [|x_{r+1} - x_r| + |y_{r+1} - y_r| + |z_{r+1} - z_r|] \leq 3M,$$

и поэтому сумма  $\sum h_r$  ограничена сверху. Но добавление новых точек в подразбиение не может уменьшить  $\sum h_r$ ; поэтому для любого процесса подразбиений сумма  $\sum h_r$  стремится к пределу. Единственность предела доказывается так же, как для интеграла Римана.

**5.014.** Таким образом, мы приходим к следующим определениям. *Внутренней точкой* множества точек в  $m$ -мерном пространстве называется такая точка  $P$  этого множества, что для некоторого  $\delta$  любая точка, отстоящая от  $P$  не дальше, чем на  $\delta$ , является точкой множества. *Внешней точкой* множества называется такая точка  $Q$ , что для некоторого  $\delta$  все точки, отстоящие от  $Q$  не дальше, чем на  $\delta$ , не являются точками множества. *Граничной точкой* множества называется такая точка  $R$  (принадлежащая множеству или же не принадлежащая), что, какое бы  $\delta$  мы ни выбрали, на расстоянии меньше  $\delta$  от  $R$  найдется по крайней мере одна точка множества и по крайней мере одна точка, не принадлежащая множеству. *Областью* называется такое множество точек, когда любую пару точек этого множества можно соединить кривой, все точки которой, за исключением, может быть, концевых точек, являются внутренними точками множества. Область называется *замкнутой*, если она содержит все свои граничные точки, и *открытой*, если все ее точки являются внутренними. Область *ограничена*, если для всех пар точек  $P, Q$  области расстояния  $PQ$  имеют верхнюю грань. Эта верхняя грань называется *диаметром*. Очевидно, что в ограниченной области значения всех координат ограничены. Так же как и для интервалов, если мы говорим, что область замкнута, то мы подразумеваем, что она ограничена.

Рассмотрим несколько примеров, большей частью в двух измерениях:

1) Окружность  $x^2 + y^2 = 1$  не является областью, поскольку если мы возьмем любой круг около некоторой ее точки, то он будет содержать как точки, принадлежащие окружности, так и не принадлежащие ей. Поэтому все точки множества являются граничными; и если мы переходим от одной точки к другой по окружности, мы никогда не пройдем через внутреннюю точку.

2) Точки, удовлетворяющие неравенству  $x^2 + y^2 < 1$ , образуют открытую область. Границей ее является окружность

$$x^2 + y^2 = 1.$$

Вокруг любой точки, лежащей внутри этой окружности, можно нарисовать круг достаточно малого радиуса, такой, что все его точки будут удовлетворять неравенству  $x^2 + y^2 < 1$ ; поэтому все

точки являются внутренними и область открытая. Ее диаметр равен 2.

3) Точки, для которых  $x^2 + y^2 \leq 1$  образуют ограниченную замкнутую область.

4) Точки, удовлетворяющие неравенствам  $0 \leq x \leq a$ ,  $0 \leq y \leq b$ , образуют ограниченную замкнутую область диаметра  $(a^2 + b^2)^{1/2}$ .

5) Точки, удовлетворяющие неравенствам  $0 \leq x \leq a$ ,  $0 < y < b$ , образуют ограниченную область, которая не является ни замкнутой, ни открытой, так как часть ее граничных точек принадлежит области, а часть нет.

6) Множество всех точек плоскости является неограниченной открытой областью.

7) Множество всех точек плоскости, за исключением точек, лежащих на прямой от  $(-1, 0)$  до  $(1, 0)$ , образует неограниченную открытую область. Прямая является границей — в данном случае внутренней границей.

8) Множество точек, удовлетворяющих неравенству  $0 < x^2 + y^2 < 1$ , является ограниченной открытой областью; начало координат и окружность образуют границу области.

9) Множество всех точек, лежащих внутри двух окружностей радиуса  $1/2$  и с центрами в  $(-1, 0)$  и в  $(1, 0)$ , не образует одну область, так как никакую точку одного круга нельзя соединить кривой с любой точкой другого, не выходя за круг.

10) Даже множество всех точек, лежащих внутри и на границе двух окружностей радиуса 1 с центрами в  $(-1, 0)$  и  $(1, 0)$ , не является одной областью. Потому что, хотя мы и можем пройти от одной точки множества к другой по точкам множества, мы не можем пройти от  $(-1, 0)$  до  $(1, 0)$  по точкам множества, не проходя через точку  $(0, 0)$ , которая является не внутренней точкой, а граничной.

11) Множество точек, определяемых неравенством  $1 \leq x^2 + y^2 \leq 4$ , является замкнутой областью, но обладает той особенностью, что мы можем пройти из одной его произвольной точки в другую произвольную точку с помощью траекторий двух видов (по часовой стрелке и против часовой стрелки), оставаясь все время внутри области, и траекторию одного вида нельзя деформировать в траекторию другого вида, не проходя вне области. Такие области называются *много связными* и играют важную роль в теории функций комплексной переменной.

12) С другой стороны, в трехмерном случае если мы возьмем область, задаваемую неравенством

$$1 \leq x^2 + y^2 + z^2 \leq 4,$$

то, хотя она имеет то же самое свойство, которое можно описать как наличие дыры в ней, тем не менее любая траектория, соединяющая две точки, может быть непрерывно деформирована в любую другую. Таким образом, имеется большое различие для областей в двухмерном и трехмерном случаях. В двухмерном случае любая простая замкнутая кривая разделяет плоскость на две области: внутри кривой и вне ее. Этот факт выглядит очевидным, но его трудно доказать (см. [10, стр. 104]), и мы будем предполагать это. Обратно, граница области в двухмерном случае может быть простой замкнутой кривой, но может состоять частично или полностью из дуг или изолированных точек, как это было в примерах (7), (8). В  $m$ -мерном случае, когда  $m > 2$ , замкнутая кривая не делит пространство на части; граница области является обычно областью в  $(m - 1)$  измерениях, но может иметь и любую меньшую размерность. Если  $m = 3$ , то граница обычно является двухмерной и задается функциями двух параметров; такие множества называются *поверхностями*.

По определению любая точка области и любая граничная точка области являются предельными точками для точек области независимо от того, является ли область открытой или замкнутой или неоткрытой, незамкнутой. Обратно, любая предельная точка множества точек области является либо точкой области, либо граничной точкой (и, следовательно, для замкнутой области также точкой области). Ибо если бы это было не так, то предельная точка была бы внешней точкой, и поэтому все точки, отстоящие от нее меньше, чем на некоторое положительное расстояние, не были бы точками области.

**5.02. Теорема Гейне—Бореля для  $m$  измерений.** Пусть  $D$ —ограниченное замкнутое множество точек; пусть  $F$ —семейство областей  $I$ , таких, что каждая точка  $P$  из  $D$  является внутренней точкой одной из областей семейства, скажем  $I_P$ ; тогда существует конечное подмножество  $F$ , такое, что те же самые соотношения верны для областей  $I$  из этого подмножества. Доказательство почти то же самое, что и для одного измерения. Мы приведем его для случая трех измерений. Поскольку  $D$  ограничено, оно лежит внутри некоторого куба  $E$  с ребром  $L$ , причем все ребра куба параллельны осям. Если все точки  $D$  лежат внутри одной области  $I$ , то доказывать нечего. Если это не так, разделим  $E$  плоскостями, параллельными осям, на 8 кубов с ребром  $L/2$ . Любой куб, не содержащий точек  $D$ , не требует дальнейшего рассмотрения. Если для каждого из кубов, которые содержат точки  $D$ , все эти точки лежат внутри одной области  $I$ , то требуемый результат доказан; если нет,

то снова разделим каждый куб, который не лежит целиком в одной области  $I$ , на 8 равных кубов.

Далее, если желаемый результат не получается после конечного числа шагов, мы получаем совокупность кубов, причем каждый предыдущий куб содержит последующий и ребро последующего в два раза меньше предыдущего; все кубы содержат по крайней мере одну точку  $D$  и никакой из них не содержится целиком в какой-нибудь области  $I$ . Это гнездо кубов определяет предельную точку  $P_0$ , которая должна быть точкой  $D$ , так как  $D$  замкнуто. Но  $P_0$  лежит внутри некоторой области  $I$ , скажем  $I_0$ ; поэтому существует такое  $\delta$ , что все точки, отстоящие от  $P_0$  не дальше, чем на  $\delta$ , лежат в  $I_0$ . Но для некоторого  $n$  диаметр куба с ребром  $2^{-n}L$  будет меньше  $\delta$ , и поэтому куб целиком будет содержаться в  $I_0$ . Таким образом, вопреки предположению, существует некоторое  $n$ , такое, что принадлежащий нашей совокупности куб с ребром  $2^{-n}L$  целиком лежит в  $I$ . Требуемый результат доказан.

Заметим, что при доказательстве не требовалось, чтобы  $D$  было областью. Оно могло бы состоять из точек  $(1/n, 0, 0)$ , где  $n$  — положительное целое число, вместе с точкой  $(0, 0, 0)$ , которая является единственной предельной точкой множества.

Заметим также, что, хотя мы делим только кубы, не содержащиеся ни в одной из областей  $I$ , теорема точно так же доказывается, если мы будем делить все кубы на каждом шаге; ибо если какая-то область  $I$  включает некоторый куб, то она включает и любую из его частей. Таким образом, нет никаких возражений против того, что все кубы берутся равными.

Так же, как и в одномерном случае, мы можем слегка ослабить условия. Если  $P$  есть граничная точка  $D$ , то достаточно, чтобы она была также граничной точкой замкнутой области  $I_P$ , такой, что все граничные точки  $D$  в окрестности  $P$  принадлежали к  $I_P$ . Ибо вокруг каждой граничной точки мы можем взять сферу и определить  $I_P$  как область, состоящую из всех точек, принадлежащих либо к  $I_P$ , либо к этой сфере (или к обеим), а для других точек  $D$  брать область  $I_P$ , совпадающую с  $I_P$ . Тогда условия теоремы верны для  $I_P$ ; и каждая граничная точка является точкой  $I_P$ , хотя не обязательно внутренней точкой.

Так же как и в одномерном случае, из теоремы немедленно вытекает следствие: *бесконечное множество точек в ограниченной замкнутой области имеет предельную точку, принадлежащую этой области.*

**5.021. Модифицированная теорема Гейне—Бореля.** Пусть  $D$  — ограниченная замкнутая область. Пусть каждая точка  $P$

области  $D$  является внутренней точкой некоторой области  $I_P$  из семейства  $F$ . Тогда  $D$  можно разделить на конечное множество областей  $G$ , таких, что каждая область  $G_P$  содержится в области  $I_P$ , соответствующей точке  $P$ , являющейся общей точкой для  $G_P$  и  $D$ . Для граничных точек  $D$  можно сделать условия менее строгими, как и в основной теореме.

Доказательство то же самое, что и для теоремы Гейне — Бореля. Но теперь мы уже не можем всегда делать кубы равными, так как если куб, покрытый некоторой областью  $I_P$ , соответствующей точке  $P$  внутри куба, делится, то только одна часть может содержать эту точку  $P$  как внутреннюю, в то время как другая часть может не покрываться никакой областью  $I_Q$ , соответствующей какой-либо точке  $Q$  этой части.

**5.03. Функции нескольких переменных.** Так же как и для одной переменной, функция определяется для всех точек некоторого множества. В частности, мы можем рассмотреть ее поведение в некоторой области.

**5.031. Непрерывность для двух и более измерений.** Мы скажем, что функция  $f(P) = f(x, y), f(x, y, z), \dots$  является непрерывной в точке  $P$ , если для любого  $\epsilon$  найдется такая окрестность  $P$ , что для всех точек  $Q$  этой окрестности выполняется неравенство  $|f(Q) - f(P)| < \epsilon$ . Ясно, что если функция непрерывна в этом смысле, то она является непрерывной функцией любой координаты при фиксированных остальных. Обратное неверно. Определим функцию

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

для всех  $x, y$ , за исключением  $x = y = 0$ , где положим  $f(0, 0) = 0$ . Она является непрерывной функцией  $x$  для любого фиксированного  $y$  и непрерывной функцией  $y$  при любом фиксированном  $x$ . Однако если  $x = r \cos \theta, y = r \sin \theta$ , то

$$f(x, y) = \cos \theta \sin \theta,$$

что может принимать все значения от  $-1/2$  до  $+1/2$  независимо от  $r$ . Поэтому в любой окрестности точки  $(0, 0)$  имеются точки, где  $f(x, y)$  отличается от  $f(0, 0)$  более чем на  $1/4$ .

Функция является непрерывной в области, если для каждой точки  $P$  из этой области и для любого  $\epsilon$  найдется некоторое  $\delta(P)$  (зависящее, возможно, от положения  $P$ ), такое, что в этой области для любой точки  $Q$ , удовлетворяющей неравенству  $r_{PQ} < \delta(P)$ , выполняется неравенство

$$|f(Q) - f(P)| < \epsilon.$$

Это определение не предполагает, что  $f$  непрерывна в граничных точках; указанное неравенство может не выполняться, если  $P$  — граничная точка и  $Q$  — внешняя точка области, поэтому мы и вставили слова „в этой области“.

Поскольку окрестность является областью, мы можем сразу вывести из теоремы Гейне — Бореля, что *если  $f(P)$  непрерывна в ограниченной замкнутой области, то для любого  $\varepsilon$  найдется  $\delta$ , такое, что если  $P$  и  $Q$  являются любыми точками области, удовлетворяющими неравенству  $r_{PQ} < \delta$ , то*

$$|f(Q) - f(P)| < \varepsilon.$$

Следующие теоремы можно доказать, используя методы, аналогичные тем, которые применимы в случае одной переменной.

*Функция, непрерывная в замкнутой области, является ограниченной в этой области и достигает своих верхней и нижней граней.*

*Если  $f(P)$  непрерывна в области и существуют точки  $A, B$  в области, такие, что*

$$f(A) < 0, \quad f(B) > 0,$$

*то для любой расположенной в области кривой, соединяющей  $A, B$ , найдется точка  $P$  на кривой, такая, что  $f(P) = 0$ .*

**5.032.** Мы определим расстояние от точки  $P$  до множества точек  $S$  как нижнюю грань расстояний между  $P$  и точками множества  $S$ . Расстояние между двумя множествами определим аналогично.

*Расстояние от точки  $P$  до множества  $S$  является непрерывной функцией координат точки  $P$ . Обозначим это расстояние через  $\delta(P)$ . Тогда для любого  $\omega$  найдется точка  $Q$  из  $S$ , такая, что*

$$\delta(P) \leq PQ \leq \delta(P) + \omega.$$

Возьмем другую точку  $P'$  и положим  $PP' = r$ . Тогда

$$PQ - r \leq P'Q \leq PQ + r.$$

Аналогично имеется точка  $Q'$ , такая, что

$$\delta(P') \leq P'Q' \leq \delta(P') + \omega$$

и

$$P'Q' \leq P'Q.$$

Поэтому

$$\delta(P') \leq P'Q' \leq P'Q \leq PQ + r \leq \delta(P) + r + \omega.$$

Аналогично

$$\delta(P) \leq \delta(P') + r + \omega.$$

Поэтому

$$|\delta(P') - \delta(P)| \leq 2r + 2\omega.$$



Возьмем  $\omega = \varepsilon/4$ ; тогда для всех  $r < \varepsilon/4$

$$|\delta(P') - \delta(P)| < \varepsilon,$$

откуда следует требуемый результат.

Отсюда же вытекает, что если  $C$  — кривая, все точки которой лежат внутри некоторой области, то расстояние от  $C$  до границы области больше нуля. Поскольку расстояния от точек  $C$  до границы образуют множество положительных чисел, то их нижняя грань есть либо положительное число, либо нуль. Но поскольку расстояние есть непрерывная функция координат кривой  $C$ , то оно принимает значение, равное своей нижней грани; так как расстояние никогда не равно нулю, то и нижняя грань не равна нулю.

**5.033.** Из любой точки  $A$  области можно перейти в любую другую точку  $B$  с помощью конечного числа перемещений, при каждом из которых изменяется только одна координата. Мы проведем доказательство для случая двух измерений. По определению области  $A$  и  $B$  могут быть связаны кривой  $C$ , целиком лежащей внутри области, и эта кривая расположена на положительном расстоянии  $\delta$  от границы. Далее, так как координаты  $(x, y)$  кривой являются непрерывными функциями параметра  $t$ , то расстояние от некоторой точки на кривой до фиксированной точки также будет непрерывной функцией  $t$  и, следовательно, ограничено. Возьмем квадрат со сторонами, параллельными осям, столь большой, что он содержит всю кривую  $C$ , и разделим его на квадраты диаметра меньше  $\delta$ . Никакой из этих квадратов, пересекающий кривую  $C$ , не пересекается с границей области. Кривая может пересекать границу квадрата много раз, даже бесконечное число раз. Но если  $t = 0$  в точке  $A$ , то значения  $t$ , соответствующие точкам на кривой, лежащим в том же квадрате, что и  $A$ , имеют верхнюю грань, определяющую точку  $P_1$ , в которой кривая последний раз выходит из квадрата, содержащего точку  $A$ , и входит в другой квадрат и т. д. Число квадратов конечно, поэтому число квадратов, пересеченных кривой, также конечно, так что можно попасть из  $A$  в  $B$  с помощью конечного числа шагов, каждый из которых проходит внутри одного из квадратов. В каждом из квадратов мы можем перейти из  $P_r$  в  $P_{r+1}$ , изменяя последовательно координаты.

Мы видим, между прочим, что любые две внутренние точки области можно соединить спрямляемой кривой.

Следствие. Если частные производные функции равны нулю всюду в области, то функция является константой во всех внутренних точках. Если при некотором смещении изменяется

одна координата  $x$  и  $\partial f / \partial x = 0$  от  $x = x_r$  до  $x_{r+1}$  при постоянных  $y, z$ , то из 1.13 следует, что  $f(x_r, y, z) = f(x_{r+1}, y, z)$  и аналогично при изменении только  $y$  или  $z$ . Поэтому мы можем показать за конечное число шагов, что  $f$  имеет то же самое значение, скажем  $a$ , в двух любых внутренних точках. Если при этом  $f(x, y, z)$  является непрерывной в замкнутой области и  $P$  — граничная точка, то для любого  $\varepsilon$  найдется внутренняя точка  $Q$ , такая, что  $|f(Q) - f(P)| < \varepsilon$  и поэтому  $|f(P) - a| < \varepsilon$ . Таким образом,  $f(x, y, z)$  также равна  $a$  и в граничных точках, если она непрерывна в замкнутой области.

Этот результат является основой для распространения на несколько измерений метода нахождения интеграла как операции, обратной дифференцированию. Он не был бы верен, если бы мы определили область таким образом, что она могла бы состоять из нескольких отдельных кусков.

**5.04. Дифференцируемость для двух и более измерений.** Производной функции одной переменной  $f(x)$  является величина

$$f'(x) = \lim_{h \rightarrow 0} \frac{1}{h} [f(x+h) - f(x)]. \quad (1)$$

Это можно записать в виде

$$f(x+h) = f(x) + hf'(x) + o(h). \quad (2)$$

В случае нескольких измерений мы хотим получить приближение к значению  $f(P)$  в окрестности точки  $P$  с ошибкой, которая также становится сколь угодно малой по сравнению с расстоянием, когда расстояние достаточно мало. Поэтому нужное нам приближение должно иметь следующий вид для двух переменных:

$$f(x+h, y+k) = f(x, y) + hf_x + kf_y + \lambda(h^2 + k^2)^{1/2}, \quad (3)$$

где  $f_x, f_y$  не зависят от  $h, k$ , а  $\lambda$  для данных  $x, y$  является функцией от  $h, k$ , стремящейся к нулю, когда  $h^2 + k^2 \rightarrow 0$ . Беря по очереди  $k=0, h=0$ , мы получим

$$\frac{\partial f}{\partial x} = f_x, \quad \frac{\partial f}{\partial y} = f_y \quad (4)$$

по определению частных производных. Теперь мы можем записать это условие следующим образом: *функция  $f(x, y)$  дифференцируема в точке  $(x, y)$ , если для любого  $\varepsilon$  найдется  $\delta$ , такое, что для  $h^2 + k^2 < \delta^2$  имеет место*

$$|f(x+h, y+k) - f(x, y) - hf_x - kf_y| \leq \varepsilon(h^2 + k^2)^{1/2}. \quad (5)$$

где  $f_x, f_y$  не зависят от  $h$  и  $k$ . Мы увидим через некоторое время, что это условие является более жестким, чем просто условие существования частных производных, и менее жестким, чем их непрерывность.

Если для каждой точки  $P(x, y)$  области соотношение вида (5) справедливо при данном  $\varepsilon$  для значений  $f$  в других точках  $Q(x+h, y+k)$  области, удовлетворяющих условию  $(h^2 + k^2)^{1/2} < \delta$ , где  $\delta$  теперь может зависеть от  $P$ , то  $f(x, y)$  называется дифференцируемой в области.

Модификации для большего числа переменных очевидны.

Это определение дифференцируемости функции нескольких переменных принадлежит Штольцу и В. Юнгу (см. [12, стр. 157—180]). Оно является основным в теории функций комплексной переменной (ср. с гл. 11) и имеет также важные физические приложения, некоторые из которых мы уже использовали в гл. 3 и которые немедленно следуют из определения. Если функция непрерывна в области, то достаточное условие для того, чтобы ее градиент в некоторой точке был вектором, состоит в том, что функция дифференцируема в этой точке. Для того чтобы градиент некоторой функции был вектор-функцией в области, достаточно, чтобы функция была дифференцируема в каждой точке области. Аналогично если  $u_i$  является векторной функцией в области, то для того, чтобы производные  $du_i/dx_k$  образовали тензор второго порядка, достаточно, чтобы компоненты  $u_i$  были дифференцируемы (см. приложение к гл. 5).

Заметим, что частные производные функции могут существовать всюду и при этом функция не будет дифференцируемой. Это имеет место, например, для функции

$$f(x, y) = \frac{xy}{x^2 + y^2}, \quad f(0, 0) = 0.$$

Здесь в точке  $(0, 0)$  мы имеем  $\partial f/\partial x = \partial f/\partial y = 0$ , но для малых  $x, y$  получаем

$$\frac{1}{r} \left[ f(x, y) - f(0, 0) - x \frac{\partial f(0, 0)}{\partial x} - y \frac{\partial f(0, 0)}{\partial y} \right] = \frac{xy}{r(x^2 + y^2)},$$

что является неограниченной функцией в любой окрестности  $(0, 0)$ .

Даже для функции

$$f(x, y) = \frac{x^2 y}{x^2 + y^2}, \quad f(0, 0) = 0$$

мы находим, что

$$\frac{1}{r} \left[ f(x, y) - f(0, 0) - x \frac{\partial f(0, 0)}{\partial x} - y \frac{\partial f(0, 0)}{\partial y} \right] = \frac{x^2 y}{r(x^2 + y^2)} = \frac{x^2 y}{r^3},$$

и это выражение принимает значения из интервала  $(-2/3\sqrt{3}, 2/3\sqrt{3})$  в любой окрестности начала координат.

**5.041.** Если частные производные функции непрерывны в окрестности некоторой точки, то функция дифференцируема в этой точке. Поскольку

$$f(x+h, y+k) - f(x, y) = [f(x+h, y+k) - f(x+h, y)] + \\ + [f(x+h, y) - f(x, y)] = k \left( \frac{\partial f}{\partial y} \right)_{x+h, y+\theta k} + h \left( \frac{\partial f}{\partial x} \right)_{x+\varphi h, y} \quad (1)$$

для некоторых  $\theta, \varphi$  между 0 и 1, производные существуют всюду в рассматриваемом интервале и непрерывны, то для любого положительного  $\omega$  найдется  $\delta$ , такое, что если

$$h^2 + k^2 < \delta^2, \quad (2)$$

$$\left| \left( \frac{\partial f}{\partial x} \right)_{x+h, y+k} - \left( \frac{\partial f}{\partial x} \right)_{x, y} \right| < \omega, \quad \left| \left( \frac{\partial f}{\partial y} \right)_{x+h, y+k} - \left( \frac{\partial f}{\partial y} \right)_{x, y} \right| < \omega, \quad (3)$$

и так как  $h^2 + \theta^2 k^2 < \delta^2$ ,  $\varphi^2 h^2 < \delta^2$ , то

$$\left| f(x+h, y+k) - f(x, y) - h \left( \frac{\partial f}{\partial x} \right)_{x, y} - k \left( \frac{\partial f}{\partial y} \right)_{x, y} \right| < \\ < \omega |h| + \omega |k| \leq \omega \sqrt{2(h^2 + k^2)}. \quad (4)$$

Это является простейшим достаточным условием дифференцируемости. Чтобы показать, что оно не является необходимым, возьмем

$$f(x, y) = x^2 \sin \frac{1}{x}; \quad f(0, 0) = 0.$$

Производные  $\partial f / \partial x$  и  $\partial f / \partial y$  существуют всюду и равны нулю в точке  $(0, 0)$ ; далее, очевидно,

$$\frac{1}{r} [f(x, y) - f(0, 0)] = \frac{x^2}{r} \sin \frac{1}{x},$$

что стремится к нулю, когда  $r$  стремится к нулю. Поэтому функция дифференцируема в точке  $(0, 0)$ . Но для  $x \neq 0$

$$\frac{\partial f}{\partial x} = -\cos \frac{1}{x} + 2x \sin \frac{1}{x},$$

что не является непрерывной функцией в окрестности точки  $(0, 0)$ .

**5.042.** Теорема о покрытии для дифференцируемых функций. Для разнобразия мы рассмотрим случай трех переменных. Если функция  $f(x, y, z)$  дифференцируема во всей ограничен-

ной замкнутой области  $D$ , то для любого  $\varepsilon$  можно найти конечное множество кубов  $G$ , таких, что для каждого  $G_r$  из этого множества найдется точка  $P_r$ , принадлежащая как  $G_r$ , так и  $D$  и такая, что для любой другой точки, принадлежащей  $G_r$  и  $D$ , выполняется соотношение

$$|f(x, y, z) - (x - x_r)f_x(P_r) - (y - y_r)f_y(P_r) - (z - z_r)f_z(P_r)| < \varepsilon [(x - x_r)^2 + (y - y_r)^2 + (z - z_r)^2]^{1/2}.$$

Неравенство вида 5.04 (5), распространенное на случай трех переменных, верно в окрестности каждой точки области  $D$ . Поэтому с помощью модифицированной теоремы Гейне—Бореля мы можем покрыть множество  $D$  конечным множеством кубов, обладающих требуемым свойством. Кубы будут, вообще говоря, неравными.

**5.05. Двойные интегралы.** Мы определим сначала двойной интеграл от функции  $f(x, y)$ , взятый по прямоугольной области  $0 \leq x \leq A$ ,  $0 \leq y \leq B$ . Предположим, что этот прямоугольник подразделяется на прямоугольники со сторонами  $h_r, k_s$ , причем диагонали всех прямоугольников не превосходят  $\delta$ ; в каждом из них мы возьмем некоторое значение функции  $f(x, y)$ , скажем  $f(\xi_r, \eta_s)$ , и образуем сумму  $\sum \sum f(\xi_r, \eta_s) h_r k_s$ . Если эта сумма стремится к единственному пределу, когда  $\delta \rightarrow 0$ , то этот предел и является двойным интегралом

$$\int_0^A \int_0^B f(x, y) dx dy. \quad (1)$$

*Необходимое и достаточное условие для существования двойного интеграла состоит в том, что функция  $f(x, y)$  должна быть ограничена и что для любых  $\omega$  и  $\alpha$  точки, где скачок функции  $f(x, y)$  превосходит  $\omega$ , могут быть заключены в конечное множество прямоугольников, полная площадь которых не превосходит  $\alpha$ .* Доказательство является прямым обобщением доказательства соответствующей теоремы для интеграла Римана. Легко показать, что спрямляемую кривую можно заключить в конечное число прямоугольников произвольно малой площади. Следовательно, разрывы вдоль конечного множества спрямляемых кривых не влияют на существование интеграла.

Двойной интеграл, скажем  $\iint_S g(x, y) dx dy$  по  $S$  внутренней части замкнутой плоской кривой  $C$ , можно определить, взяв прямоугольник  $D$ , заключающий внутри себя кривую  $C$ , и полагая  $f(x, y) = g(x, y)$  в  $S$  и  $f(x, y) = 0$  вне  $S$ . Тогда двойной

интеграл  $\int_S \int g(x, y) dx dy$  определяется как  $\int_D \int f(x, y) dx dy$  по внутренней части прямоугольника  $D$  при условии, что  $f(x, y)$  удовлетворяет в  $D$  условиям интегрируемости.

**5.051. Повторные интегралы.** Двойной интеграл, если он существует, вычисляется обычно с помощью интегрирования по каждой из переменных по очереди. Для данного  $x$  выражение  $\int_0^B f(x, y) dy$  является однократным интегралом и в то же время функцией от  $x$ . Если эту функцию проинтегрировать по  $x$ , то в результате получится повторный интеграл  $\int_0^A \left[ \int_0^B f(x, y) dy \right] dx$ ;

он обычно записывается в виде  $\int_0^A dx \int_0^B f(x, y) dy$ , но используются и другие обозначения. Наоборот, мы могли бы сначала проинтегрировать по  $x$  и затем по  $y$ , получив повторный интеграл  $\int_0^B dy \int_0^A f(x, y) dx$ .

Если двойной интеграл  $\int_0^A \int_0^B f(x, y) dx dy$  существует, то оба повторных интеграла существуют и равны двойному интегралу. Это надо понимать в довольно широком смысле. Если для некоторого  $x$ , скажем  $x = a$ , функция  $f(a, y)$  равна нулю для иррациональных  $y$  и равна единице для рациональных  $y$ , то интеграл  $\int_0^B f(a, y) dy$  не существует. Но линия интегрирования может быть заключена в прямоугольник сколь угодно малой площади, и такая нерегулярность только для одного значения  $x$  не повлияет на существование двойного интеграла. Однако мы видели при обсуждении интеграла Римана, независимо от того, существует интеграл или нет, что, если только функция является ограниченной, верхний и нижний интегралы  $H, h$  существуют. Теперь мы определим  $H(x), h(x)$  как верхний и нижний интегралы от  $f(x, y)$  по  $y$  при данном  $x$  и образуем суммы

$$\sum_r H(\xi_r) h_r, \quad \sum_r h(\xi_r) h_r \quad (x_r \leq \xi_r \leq x_{r+1}).$$

Из существования двойного интеграла 5.05 (1), который мы будем обозначать  $I$ , следует, что для любого  $\varepsilon$  найдется такое  $\delta$ , что, каково бы ни было подразбиение на прямоугольники с диагоналями, меньшими  $\delta$ , справедливы неравенства

$$I - \varepsilon \leq \sum_r \sum_s m_{rs} h_r k_s \leq \sum_r \sum_s M_{rs} h_r k_s \leq I + \varepsilon, \quad (2)$$

где  $x_{r+1} - x_r = h_r$ ,  $y_{s+1} - y_s = k_s$  и  $M_{rs}$ ,  $m_{rs}$  — верхняя и нижняя грани функции  $f(x, y)$  в прямоугольнике, указанном индексами. Возьмем прямую  $x = \text{const}$  ( $x_r \leq x \leq x_{r+1}$ ). В любом прямоугольнике верхняя и нижняя грани  $f(x, y)$  на этой прямой лежат в интервале  $(m_{rs}, M_{rs})$ ; поэтому

$$\sum_s m_{rs} k_s \leq h(x) \leq H(x) \leq \sum_s M_{rs} k_s, \quad (3)$$

и если наибольшее из  $k_s$  стремится к нулю, то

$$I - \varepsilon \leq \sum n_r h_r \leq \sum N_r h_r \leq I + \varepsilon, \quad (4)$$

где под  $n_r$  мы понимаем нижнюю грань  $h(x)$ , а под  $N_r$  — верхнюю грань  $H(x)$  в соответствующем интервале  $x$ . Устремим теперь наибольшее  $h_r$  к нулю и возьмем верхнюю грань  $I_1$  первой суммы и нижнюю грань  $I_2$  второй суммы по всем таким подразбиениям. Если  $I_1$  и  $I_2$  равны, то мы скажем, что повторный интеграл существует. Далее мы имеем

$$I - \varepsilon \leq I_1 \leq I_2 \leq I + \varepsilon, \quad (5)$$

и так как  $I_1$ ,  $I_2$  не зависят от  $\varepsilon$ , то обе эти величины должны быть равны  $I$ . Отсюда  $h(x)$ ,  $H(x)$  обе интегрируемы по  $x$  и интегралы от них равны двойному интегралу. Поэтому повтор-

ный интеграл можно представить либо как  $\int_0^A h(x) dx$ , либо как

$\int_0^A H(x) dx$ . Наше рассуждение показывает, между прочим, что если имеются точки, где  $H(x) - h(x) \geq \omega > 0$ , то эти точки можно заключить в конечное множество интервалов  $x$ , общая длина которых произвольно мала.

Обратное неверно; оба повторных интеграла могут существовать и быть равными, в то время как двойной интеграл не существует. Однако на практике такие случаи редки и в любом обычном случае с помощью непосредственной проверки функции можно быстро решить, выполнены ли условия существования двойного интеграла.

Двойной интеграл  $\int \int f(x, y) dx dy$  по всей бесконечной области  $R$  можно определить, взяв последовательность областей  $\{R_n\}$ , таких, что любая часть  $R$  входит во все  $R_n$  для значений  $n$ , больших некоторого  $m$ . Если двойные интегралы по  $R_n$  сходятся к одному и тому же пределу для всех таких последовательностей, то этот предел и можно взять как определение интеграла по  $R$ . Несобственные двойные интегралы можно определить аналогичным образом. Оказывается, однако, что если только такой же процесс не дает единственное значение, когда  $f(x, y)$  заменена на  $|f(x, y)|$ , то величина предела будет зависеть от формы областей  $R_n$  и, следовательно, не абсолютно сходящийся двойной интеграл не имеет смысла, если эти области не заданы. Для абсолютно сходящегося двойного интеграла перемена порядка предельных переходов может быть оправдана с помощью теоремы 1.111 (более детально см. в [7, гл. 5] и в [8, том I, гл. 6]).

Аналогичные утверждения справедливы для тройного и  $n$ -кратного интегралов. Используя определение площадки поверхности, которое мы дадим, мы получим, что если поверхность имеет конечную площадь, то она может быть заключена во множество параллелепипедов произвольно малого общего объема (см. приложение 5.07а).

Если  $f(x, y) = 0$  в области, не совпадающей с прямоугольником, стороны которого параллельны осям, то пределы для  $y$  в повторном интеграле, когда  $x$  зафиксирован, будут зависеть от  $x$ . Предположим, что область представляет собой четверть круга и определяется неравенствами  $y \geq 0$ ,  $x \geq 0$ ,  $x^2 + y^2 \leq a^2$ . Для данного  $x$  из этих неравенств вытекает

$$0 \leq y \leq (a^2 - x^2)^{1/2},$$

я область интегрирования по  $y$  поэтому заключена от 0 до  $(a^2 - x^2)^{1/2}$ . Далее,  $x$  может изменяться от 0 до  $a$  и интеграл по четверти круга можно записать в виде

$$\int_0^a dx \int_0^{\sqrt{a^2 - x^2}} f(x, y) dy.$$

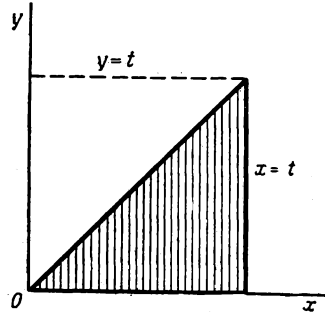
Точно так же он может быть записан и в виде

$$\int_0^a dy \int_0^{\sqrt{a^2 - y^2}} f(x, y) dx.$$



Область интегрирования всегда задается рядом неравенств, которым должны удовлетворять независимые переменные. Из этих неравенств пределы для повторных интегралов могут быть найдены либо с помощью непосредственных преобразований, как в только что рассмотренном примере, либо с помощью рисунка. Последний способ часто бывает полезным, особенно когда граница состоит из нескольких кривых, задаваемых различными уравнениями.

Повторный интеграл из конкретной задачи с определенным порядком интегрирования часто появляется, но иногда его легче всего взять с помощью обращения порядка интегрирования. Когда пределы интегрирования не являются постоянными, то наиболее простой способ определения пределов интегрирования после изменения порядка интегрирования включает в качестве промежуточного шага нахождение неравенств для двойного интеграла. Так, например, рассмотрим интеграл



Р и с. 24.

$$I = \int_0^t dx \int_0^x f(y) dy.$$

Неравенства, которые должны выполняться для  $x$ ,  $y$ , лежащих в области интегрирования, имеют вид

$$0 \leq y \leq x, \quad 0 \leq x \leq t,$$

и при данном  $y$  переменная  $x$  меняется в пределах  $y \leq x \leq t$ , в то время как без ограничений, наложенных на  $x$ , переменная  $y$  может изменяться от 0 до  $t$ . Далее, имеем

$$I = \int_0^t dy \int_y^t f(y) dx = \int_0^t (t - y) f(y) dy,$$

так что одно интегрирование выполняется немедленно. Геометрически интеграл берется по области, являющейся треугольником с вершинами в точках  $(0, 0)$ ,  $(t, 0)$ ,  $(t, t)$ , и если мы интегрируем сначала по  $x$ , то мы должны интегрировать от  $y$  до  $t$ , чтобы покрыть область возможных изменений переменной

$x$  при фиксировании  $y$  внутри треугольника (см. приложение 5.051a).

**5.052. Замена переменных.** Рассмотрим двойной интеграл

$$I = \int \int f(x, y) dx dy,$$

который берется обычно с помощью последовательного интегрирования по  $x$  и по  $y$ . Пусть  $\xi$  и  $\eta$  — две дифференцируемые функции от  $x$  и  $y$ . Пересечение кривых  $\xi = \text{const}$  и  $\eta = \text{const}$  разметит на плоскости, вообще говоря, четырехугольные фигуры, которые можно было бы использовать как элементы площади при определении двойного интеграла. Мы должны выразить эти элементы площади через  $\xi$  и  $\eta$ . Переменные  $x$  и  $y$  теперь будут рассматриваться как функции от  $\xi$ ,  $\eta$ . Выберем направления  $\xi$  и  $\eta$  так, что если мы будем обходить начало координат против часовой стрелки, переходя от оси с положительными значениями  $x$  к оси с положительными значениями  $y$ , то соответствующий переход от направления возрастания  $\xi$  к направлению возрастания  $\eta$  будет совершаться также против часовой стрелки; однако эти направления не будут, вообще говоря, перпендикулярны. Простейший подход состоит в следующем. Заметим, что когда  $\delta\xi$  и  $\delta\eta$  стремятся к нулю, элемент соответствующей площади стремится к параллелограмму, за исключением, возможно, окрестностей специальных точек, таких, как центр круга, когда  $\xi$ ,  $\eta$  являются полярными координатами. С точностью до членов первого порядка по отношению к  $\delta\xi$ ,  $\delta\eta$ , если один из углов параллелограмма имеет координаты  $(x, y)$ , то два соседних имеют координаты

$$\left(x + \frac{\partial x}{\partial \xi} \delta\xi, \quad y + \frac{\partial y}{\partial \xi} \delta\xi\right), \quad \left(x + \frac{\partial x}{\partial \eta} \delta\eta, \quad y + \frac{\partial y}{\partial \eta} \delta\eta\right).$$

Но площадь параллелограмма, три вершины которого имеют координаты  $(x_1, y_1)$ ,  $(x_2, y_2)$ ,  $(x_3, y_3)$ , равна

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix},$$

поэтому площадь рассматриваемого нами параллелограмма равна

$$\begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{vmatrix} \delta\xi \delta\eta.$$

Этот определитель называется *якобианом* \*) от функций  $x, y$  по отношению к  $\xi, \eta$  и обозначается  $\partial(x, y)/\partial(\xi, \eta)$ . Далее,

$$I = \int \int f(x, y) \frac{\partial(x, y)}{\partial(\xi, \eta)} d\xi d\eta,$$

где  $f(x, y)$  должна быть выражена как функция от  $\xi$  и  $\eta$  и пределы интегрирования таковы, что каждый элемент площади из первоначальной области появляется один и только один раз в преобразованном интеграле.

Этот путь является простейшим для получения ответа, но он имеет несколько недостатков. Трудно оценить ошибку при замене элемента, образуемого кривыми  $\xi = \text{const}$ ,  $\eta = \text{const}$ , на параллелограмм и поэтому трудно показать, что полная ошибка стремится к нулю, когда все приращения  $\delta x$ ,  $\delta y$  стремятся к нулю. Кроме того, этот способ трудно обобщать, хотя рассуждения при этом и обобщаются легко на тройные интегралы, поскольку мы имеем удобную формулу для объема произвольного параллелепипеда, однако распространение на  $n$ -кратные интегралы не очевидно. В самом деле, мы имеем

$$\int \int \int f(x, y, z) dx dy dz = \int \int \int f(x, y, z) \frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)} d\xi d\eta d\zeta,$$

где

$$\frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)} = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} & \frac{\partial z}{\partial \eta} \\ \frac{\partial x}{\partial \zeta} & \frac{\partial y}{\partial \zeta} & \frac{\partial z}{\partial \zeta} \end{vmatrix}.$$

Но четыре переменных интегрирования встречаются в общей теории относительности, и для более чем трех измерений трудно усмотреть, что мы понимаем под каким-либо обобщением объема даже параллелепипеда, если только не принять чисто аналитического определения с помощью интеграла. Далее, известные свойства площади и объема в евклидовой геометрии нельзя уже больше использовать, чтобы сократить путь при непосредственном преобразовании переменных. Переменные  $x, y, z$  могут не быть декартовыми координатами; при этом трудно усмотреть, что мы понимаем под площадью или объемом области. По всем этим причинам лучше всего проводить непосредственное преобразование переменных последовательно одно за другим. Удобно использовать индексные обозначения и рассмотреть случай тройного интеграла для того, чтобы

\*) Или функциональным определителем. — *Прим. ред.*

проиллюстрировать метод, который легко можно распространить на любое число переменных. Мы начнем с интеграла

$$I = \int \int \int f(x_1, x_2, x_3) dx_1 dx_2 dx_3 \quad (1)$$

и будем рассматривать  $x_1, x_2, x_3$  как функции от  $\xi_1, \xi_2, \xi_3$  с непрерывными первыми частными производными и такие, что имеется взаимно однозначное соответствие между точками областей  $x_i$  и  $\xi_i$ . Далее, используя теорему о среднем значении, получаем

$$\delta x_i = \frac{\partial x_i}{\partial \xi_j} \delta \xi_j, \quad (2)$$

где каждую из производных  $\partial x_i / \partial \xi_j$  надо рассматривать как вычисленную в некоторой точке (не обязательно все производные в одной точке) внутри элемента, ограниченного двумя наборами значений переменных  $\xi_j$ , отличающихся на  $\delta \xi_j$ . Выразим сначала  $x_1$  в виде функции от  $\xi_1, x_2$  и  $x_3$ . Далее, считая  $x_2$  и  $x_3$  постоянными, мы должны решить уравнения

$$\delta x_1 = \frac{\partial x_1}{\partial \xi_j} \delta \xi_j, \quad 0 = \frac{\partial x_2}{\partial \xi_j} \delta \xi_j, \quad 0 = \frac{\partial x_3}{\partial \xi_j} \delta \xi_j, \quad (3)$$

и решение будет следующим:

$$\begin{vmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_2} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_2}{\partial \xi_1} & \frac{\partial x_2}{\partial \xi_2} & \frac{\partial x_2}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_2} & \frac{\partial x_3}{\partial \xi_3} \end{vmatrix} \delta \xi_1 = \begin{vmatrix} \frac{\partial x_2}{\partial \xi_2} & \frac{\partial x_2}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_2} & \frac{\partial x_3}{\partial \xi_3} \end{vmatrix} \delta x_1. \quad (4)$$

Когда  $\delta \xi_1 \rightarrow 0$ , определители стремятся к якобианам, вычисленным в точке  $(x_1, x_2, x_3)$ ; поэтому

$$I = \int \int \int f(x_1, x_2, x_3) \left[ \frac{\partial (x_1, x_2, x_3)}{\partial (\xi_1, \xi_2, \xi_3)} / \frac{\partial (x_2, x_3)}{\partial (\xi_2, \xi_3)} \right] d\xi_1 dx_2 dx_3. \quad (5)$$

Теперь выразим  $x_2$  в виде функции от  $\xi_1, \xi_2$  и  $x_3$ ; мы находим аналогичным образом, считая  $\xi_1, x_3$  постоянными, что

$$\delta x_2 = \frac{\partial x_2}{\partial \xi_2} \delta \xi_2 + \frac{\partial x_2}{\partial \xi_3} \delta \xi_3, \quad (6)$$

$$0 = \frac{\partial x_2}{\partial \xi_2} \delta \xi_2 + \frac{\partial x_3}{\partial \xi_3} \delta \xi_3,$$

$$\frac{\partial (x_2, x_3)}{\partial (\xi_2, \xi_3)} \delta \xi_2 = \frac{\partial x_3}{\partial \xi_3} \delta x_2, \quad (7)$$

$$I = \int \int \int f(x_1, x_2, x_3) \left[ \frac{\partial (x_1, x_2, x_3)}{\partial (\xi_1, \xi_2, \xi_3)} / \frac{\partial x_3}{\partial \xi_3} \right] d\xi_1 d\xi_2 dx_3. \quad (8)$$

Наконец, выразим  $x_3$  в виде функции от  $\xi_3$ ,  $\xi_1$  и  $\xi_2$ ; затем, считая  $\xi_1$  и  $\xi_2$  постоянными, получаем

$$\delta x_3 = \frac{\partial x_3}{\partial \xi_3} \delta \xi_3 \quad (9)$$

и

$$I = \int \int \int f(x_1, x_2, x_3) \frac{\partial (x_1, x_2, x_3)}{\partial (\xi_1, \xi_2, \xi_3)} d\xi_1 d\xi_2 d\xi_3. \quad (10)$$

Последовательная замена переменных законна в силу результатов 1.032 при условии, что интегралы существуют, и величина (10) может быть отождествлена с тройным интегралом, если этот последний существует. При рассмотренных условиях  $J$  (якобиан) является непрерывной и ограниченной функцией, так что тройной интеграл обязательно существует.

Имеется, однако, некоторая трудность, когда  $J$  меняет знак в области интегрирования. Даже при замене одной переменной, скажем  $x$ , если  $dx/d\xi$  меняет знак в области интегрирования, данному значению  $x$  будут соответствовать два или более значений  $\xi$ , и если мы хотим получить правильный результат, мы должны записать интеграл в виде интеграла Стильбеса. При преобразовании повторных интегралов, если якобиан обращается в нуль в некоторой точке внутри области, в некоторой окрестности этой точки данному набору значений  $x_1, x_2, x_3$  будет соответствовать более одного набора значений  $\xi_1, \xi_2, \xi_3$ . Если это так, то лучше всего разбить область на подобласти, в каждой из которых якобиан сохраняет знак, и рассматривать каждую подобласть отдельно.

Этот подход показывает, что первый метод, основанный на том, что с элементами площади или объема обращаются как с параллелограммами или параллелепипедами, на самом деле дает правильный ответ и может быть использован непосредственно во многих случаях, где он более удобен, чем явное решение якобиана. Преобразование часто упрощается с помощью использования теоремы

$$\frac{\partial (x_1 \dots x_n)}{\partial (\eta_1 \dots \eta_n)} = \frac{\partial (x_1 \dots x_n)}{\partial (\xi_1 \dots \xi_n)} \frac{\partial (\xi_1 \dots \xi_n)}{\partial (\eta_1 \dots \eta_n)}, \quad (11)$$

так как

$$\frac{\partial x_i}{\partial \eta_k} = \frac{\partial x_i}{\partial \xi_j} \frac{\partial \xi_j}{\partial \eta_k}, \quad (12)$$

где  $x_i$  слева рассматриваются как функции от  $\eta_k$  и справа как функции от  $\xi_j$ . Но определитель выражений, стоящих в правой части (12), равен произведению двух определителей в правой части (11) согласно правилу перемножения определителей. Или же, иначе говоря, теорема должна вытекать из

непротиворечивости, ибо, если бы она была неверна, мы могли бы получить различные результаты при преобразовании интеграла непосредственно к переменным  $\eta_k$  и при преобразовании сначала к  $\xi_j$ , а затем к  $\eta_k$ .

Существует теорема, утверждающая, что если якобиан функций  $\xi_1, \dots, \xi_n$  по отношению к  $x_1, \dots, x_n$  равен нулю всюду, то одна из функций  $\xi$  определяется по остальным и эти функции, таким образом, не являются подходящей системой координат \*). Доказательство можно найти в большинстве книг по математическому анализу.

**5.053. Изменение пределов интегрирования.** Замена переменных, естественно, влечет за собой изменение пределов интегрирования. Новые пределы можно найти либо аналитически, записав область интегрирования в виде неравенств, либо же нарисовав чертеж и найдя с помощью проверки такие пределы для новых переменных, которые обеспечивали бы покрытие всей области их изменения. Никакого общего правила преобразования координат дать нельзя, но мы проиллюстрируем эти методы на примерах.

Имеется очевидное противоречие между якобианом преобразования и простым изменением порядка двух переменных в интеграле, так как  $\partial(x, y)/\partial(y, x) = -1$ . Это противоречие объясняется следующим образом: якобиан  $\partial(x, y, z)/\partial(\xi, \eta, \zeta)$  не меняется при любой циклической замене  $x, y, z$  или  $\xi, \eta, \zeta$ . Он является положительным, если при повороте направлений  $d\xi, d\eta, d\zeta$  так, что направления  $d\xi$  и  $d\eta$  лежат в плоскости  $dx$  и  $dy$ , а вращение вокруг  $dz$  от  $d\xi$  к  $d\eta$  считается положительным, направление  $d\zeta$  образует острый угол с положительным направлением  $dy$ . Таким образом, положительность якобиана означает, что  $d\xi, d\eta, d\zeta$  образуют правостороннюю систему направлений, и если  $x, y, z$  возрастают в своих интервалах интегрирования, то  $\xi, \eta, \zeta$  будут также возрастать; для каждого интегрирования нижний предел меньше верхнего. Если якобиан отрицательный, то  $d\xi, d\eta, d\zeta$  образуют левостороннюю систему направлений и нечетное число из них будет убывать в интервале своего изменения. Если мы все же хотим, чтобы нижние пределы были меньше верхних, то должны изменить знак. Поэтому, если для каждой переменной мы хотим, чтобы нижний предел был меньше верхнего, мы должны вместо якобиана взять его модуль. Это правило для двух переменных также дает правильный ответ, и так как пределы

---

\*) Так как мерность нового пространства меньше мерности исходного пространства. — *Прим. ред.*

интегрирования не меняются местами, то мы должны взять не  $\partial(x, y)/\partial(y, x)$ , а модуль этого выражения, который равен  $+1$ .

**5.054. Полярные координаты.** Рассмотрим двойной интеграл

$$I = \iint f(x, y) dx dy \quad (0 \leq x^2 + y^2 \leq a^2)$$

и преобразуем его к полярным координатам  $r, \theta$ . Тогда

$$\begin{aligned} x &= r \cos \theta, & y &= r \sin \theta, \\ \frac{\partial(x, y)}{\partial(r, \theta)} &= \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \end{aligned}$$

и

$$I = \iint f(x, y) r dr d\theta.$$

Как обычно, мы считаем, что  $r \geq 0$ ; тогда мы представляем каждую точку внутри круга один раз, взяв  $\theta$  от 0 до  $2\pi$ , и пределы для  $r$  будут равны 0 и  $a$ , а пределы для  $\theta$  равны 0 и  $2\pi$ . Можно было бы разрешить  $r$  принимать и отрицательные значения так, что точка  $(-x, -y)$  соответствовала бы точке  $(-r, \theta)$  с тем же самым  $\theta$ , что и для  $(x, y)$ ; но тогда для того, чтобы представить каждую точку один раз, угол  $\theta$  может изменяться только от 0 до  $\pi$  и пределы становятся следующими: от  $-a$  до  $a$  для  $r$  и от 0 до  $\pi$  для  $\theta$ . Каждая из систем одинаково справедлива. Начало координат является особой точкой преобразования, но она не доставляет никаких неприятностей.

**5.055.** Очевидно, что, если мы перейдем от прямоугольных координат  $x, y, z$  к цилиндрическим координатам  $\omega, \lambda, z$ , мы аналогично получим

$$\frac{\partial(x, y, z)}{\partial(\omega, \lambda, z)} = \omega.$$

Если мы теперь перейдем от цилиндрической к сферической системе координат  $(r, \lambda, \theta)$ , то мы должны положить

$$\begin{aligned} z &= r \cos \theta, & \omega &= r \sin \theta, \\ \frac{\partial(\omega, \lambda, z)}{\partial(r, \lambda, \theta)} &= \frac{\partial(\omega, z)}{\partial(r, \theta)} = -r, \end{aligned}$$

и поэтому

$$\frac{\partial(x, y, z)}{\partial(r, \lambda, \theta)} = -r\omega = -r^2 \sin \theta, \quad \frac{\partial(x, y, z)}{\partial(r, \theta, \lambda)} = r^2 \sin \theta.$$

Этот результат можно также получить непосредственно, взяв

$$x = r \sin \theta \cos \lambda, \quad y = r \sin \theta \sin \lambda, \quad z = r \cos \theta.$$

Мы представляем внутреннюю область сферы радиуса  $a$ , взяв следующие интервалы изменения: от 0 до  $a$  для  $r$ , от 0 до  $\pi$  для  $\theta$  и от 0 до  $2\pi$  для  $\lambda$ . Но любой из следующих наборов интервалов изменений был бы эквивалентен:

$$\begin{aligned} -a \leq r \leq a, & \quad 0 \leq \theta \leq \pi, & \quad 0 \leq \lambda \leq \pi, \\ 0 \leq r \leq a, & \quad -\pi \leq \theta \leq \pi, & \quad 0 \leq \lambda \leq \pi. \end{aligned}$$

Мы не должны брать интервалы изменения, например, следующими:  $0 \leq r \leq a$ ,  $-\pi \leq \theta \leq \pi$ ,  $0 \leq \lambda \leq 2\pi$ , так как при этом мы покрыли бы сферу дважды и получили бы удвоенный правильный результат.

**5.056.** В качестве иллюстрации некоторых особенностей, возникающих при обращении с двойными интегралами, рассмотрим метод, приводимый во многих учебниках для вычисления интеграла

$$I = \int_0^{\infty} e^{-x^2} dx.$$

Поступим следующим образом:

$$I^2 = \int_0^{\infty} e^{-x^2} dx \times \int_0^{\infty} e^{-y^2} dy = \quad (1)$$

$$= \int_0^{\infty} \int_0^{\infty} e^{-(x^2+y^2)} dx dy = \quad (2)$$

$$= \int_0^{\infty} \int_0^{\pi/2} e^{-r^2} r dr d\theta = \quad (3)$$

$$= \int_0^{\infty} e^{-r^2} r dr \times \int_0^{\pi/2} d\theta = \quad (4)$$

$$= \frac{1}{2} \cdot \frac{\pi}{2}; \quad (5)$$

поэтому

$$I = \frac{\sqrt{\pi}}{2}. \quad (6)$$



Переход от (1) к (2) требует обоснования; не очевидно, что произведение двух однократных интегралов можно перевести в двойной интеграл. В самом деле,  $I$  нужно понимать как предел

$$\lim_{X \rightarrow \infty} \int_0^X e^{-x^2} dx = \lim_{Y \rightarrow \infty} \int_0^Y e^{-y^2} dy, \quad (7)$$

и без потери общности мы можем принять  $X = Y$ . Тогда (1) обозначает, что

$$I^2 = \lim_{X \rightarrow \infty} \int_0^X e^{-x^2} dx \times \int_0^X e^{-y^2} dy. \quad (8)$$

Но в интеграле

$$\int_0^X \int_0^X e^{-(x^2+y^2)} dx dy \quad (9)$$

пределы конечны и подынтегральная функция непрерывна. Поэтому этот интеграл можно выразить как повторный интеграл и он равен произведению интегралов в (8). Отсюда

$$I^2 = \lim_{X \rightarrow \infty} \int_0^X \int_0^X e^{-(x^2+y^2)} dx dy. \quad (10)$$

Однако переход от этого выражения к (3) с помощью замены переменных не очевиден, поскольку область интегрирования является квадратом и выражение (3) нужно понимать в смысле предела

$$\lim_{R \rightarrow \infty} \int_0^R \int_0^{\pi/2} e^{-r^2} r dr d\theta, \quad (11)$$

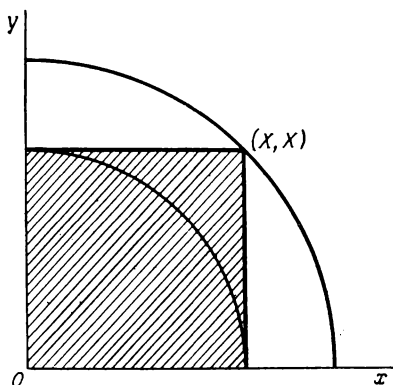


Рис. 25.

а здесь область интегрирования является четвертью круга. Обоснование заключается в том, что подынтегральная функция положительна для всех  $x, y$  и интеграл по квадрату заключен между интегралами по четверти круга с радиусом  $X$  и по четверти круга с радиусом  $X\sqrt{2}$ . Значит,

$$\int_0^X \int_0^{\pi/2} e^{-r^2} r dr d\theta < \int_0^X \int_0^X e^{-(x^2+y^2)} dx dy < \int_0^{X\sqrt{2}} \int_0^{\pi/2} e^{-r^2} r dr d\theta. \quad (12)$$

Когда мы интегрируем по  $\theta$  и затем переходим к пределу, первое и третье выражения стремятся к одному и тому же пределу; поэтому

$$I^2 = \frac{\pi}{2} \int_0^{\infty} e^{-r^2} r \, dr = \frac{\pi}{4}. \quad (13)$$

С помощью очевидного преобразования мы получаем результат, который в дальнейшем будет нам часто нужен \*):

$$\left(-\frac{1}{2}\right)! = \int_0^{\infty} x^{-1/2} e^{-x} \, dx = \sqrt{\pi}.$$

### 5.057. Многократный интеграл. Интеграл

$$I = \int \int \dots \int \exp(-W/2) \, dx_1 \dots dx_n \quad (1)$$

часто бывает нужен в теории вероятностей, где

$$W = a_{lk} x_l x_k \quad (2)$$

является положительно определенной квадратичной формой, а интегрирование ведется от  $-\infty$  до  $+\infty$  по всем переменным. Если  $W$  не является положительно определенной формой, то этот интеграл не сходится. Мы знаем, что форму  $W$  можно привести к сумме квадратов с помощью перехода к переменным  $\xi_1, \dots, \xi_n$ , причем якобиан преобразования будет равен единице (см. 4.08). Если при этом

$$W = A_{11}\xi_1^2 + \dots + A_{nn}\xi_n^2, \quad (3)$$

то  $I$  является произведением  $n$  интегралов вида

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} A_{rr} \xi_r^2\right) d\xi_r = \left(\frac{2\pi}{A_{rr}}\right)^{1/2}. \quad (4)$$

Но

$$A_{11}A_{22} \dots A_{nn} = \|a_{lk}\|. \quad (5)$$

Поэтому

$$I = \frac{(2\pi)^{n/2}}{\|a_{lk}\|^{1/2}}. \quad (6)$$

### 5.06. Интегралы вдоль спрямляемых кривых. Если

$$I = \int_A^B f(x, y, z) \, ds, \quad (1)$$

то  $s$  непрерывно возрастает, когда мы движемся по кривой, и  $I$  является интегралом Римана. Заметим, что если  $|f| < M$  и длина кривой от  $A$  до  $B$  равна  $L$ , то  $|I| < ML$ .

\*) Определение факториальной функции см. в гл. 15.

Пусть  $u, v, w$  — три функции от  $x, y, z$ , ограниченные и непрерывные на кривой; обозначим

$$J = \int_{s=s_0}^{s_1} (u \, dx + v \, dy + w \, dz). \quad (2)$$

Это выражение рассматривается как сумма трех интегралов Стильеса, так как  $x, y, z$  не обязательно являются монотонными функциями от  $s$ . Функция  $x$  непрерывна по  $s$  и имеет ограниченную вариацию. Но даже если функция  $u$  и непрерывна по  $s$ , она может не быть непрерывной по  $x$ , так как часть кривой может быть перпендикулярна к оси  $x$ , скажем при  $x=a$ , и тогда функция  $u$  может иметь разрыв в точке  $x=a$ , так как  $y$  или  $z$  могут иметь разрыв. Но, как это следует из (достаточных) условий существования интеграла Стильеса, интеграл  $J$  будет существовать, если кривая может быть разделена на конечное число дуг, на каждой из которых функции  $u, v, w$  являются либо непрерывными, либо ограниченной вариации.

Можно показать (простейшее доказательство см. в [1, стр. 205—207]), что спрямляемая кривая почти всюду имеет касательную; поэтому производные

$$l = \frac{dx}{ds}, \quad m = \frac{dy}{ds}, \quad n = \frac{dz}{ds} \quad (3)$$

существуют почти всюду. С помощью замены переменных мы можем записать

$$J = \int_{s_0}^{s_1} (lu + mv + nw) \, ds = \int l_i u_i \, ds = \int l \cdot u \, ds. \quad (4)$$

Достаточные условия, дающие возможность производить такую замену, состоят в том, что оба выражения (2) и (4) существуют и производные  $dx_i/ds$  ограничены, как это, очевидно, и есть. На практике  $u_i$  и  $l_i$  обычно непрерывны, за исключением, возможно, конечного числа точек, и после проверки становится видно, что интегралы существуют.

**5.07. Поверхностные интегралы: площадь поверхностей.** Простым замкнутым кривым и дугам соответствуют замкнутые поверхности и поверхности с границами. Замкнутая поверхность — это поверхность конечного диаметра с внутренней и внешней частью и такая, что мы можем перейти из любой точки поверхности в любую другую, не покидая поверхность; из любой внутренней точки в любую другую внутреннюю точку,

не пересекая поверхности, и аналогично из одной внешней точки в другую; но мы не можем перейти из внутренней точки ни в какую внешнюю, не пересекая поверхности. Если в дополнение к этому любую простую замкнутую кривую на поверхности можно стянуть в точку, не покидая поверхности, то поверхность называется *односвязной*. Эти условия удовлетворяются для большинства обычных поверхностей, в частности для границ твердых тел, однако можно построить аномальную поверхность, известную под названием *бутылка Клейна*, которая не имеет ни внутренней части, ни внешней, и в то же время у нее нет граничной кривой. Она имеет конечные размеры, и мы можем перейти из любой ее точки в любую другую, не покидая поверхности, но любые две точки, не лежащие на поверхности, можно соединить кривой, не пересекающей поверхность. Мы исключим такие поверхности из нашего рассмотрения. Любая прямая линия, которая хоть раз пересекает замкнутую поверхность, будет пересекать ее в четном числе точек; мы не будем рассматривать такие поверхности, как, например, поверхность, записываемая в цилиндрических координатах в виде

$$\omega = \sin^2 \frac{1}{z},$$

которая пересекается бесконечное число раз любой прямой  $\lambda = 0$ ,  $\omega = a$ , где  $0 < a < 1$ . Примером поверхности, не являющейся односвязной, является тор. Любая простая замкнутая кривая на замкнутой односвязной поверхности делит ее на две части, которые мы будем называть *шапками* (caps). Кривая будет называться *каймой* (rim) каждой из шапок. Каждая из шапок по отдельности не имеет внутренней части, но обычно обладает еще таким свойством, что простая замкнутая кривая на ней, не проходящая через границу, делит ее на две области, разделенные кривой, и любую такую внутреннюю кривую можно стянуть в точку, не покидая поверхности. Иными словами, шапка не должна иметь дыр. Мы исключим из рассмотрения *полоску Мёбиуса*, так как на ней можно нарисовать замкнутую кривую, которая не делит полоску на две взаимно недоступимые области. Полоску Мёбиуса можно сделать, взяв длинную полоску из бумаги, скрутив ее вполоборота и склеив концы вместе. Очевидно, полоска Мёбиуса имеет одну границу, и полный продольный разрез вдоль нее по первоначальной средней линии не разделяет ее на два куска. Снова мы можем сделать продольный разрез на расстоянии одной трети или меньше поперечного сечения; в результате мы обнаружим, что поверхность разделится на два куска, причем часть поверхности

у края даст полосу с двумя краями и полностью перекрученную, в то время как внутренняя часть останется в виде более узкой полосы Мёбиуса, сцепленной с первым куском. Мы видим, таким образом, что первоначальный край полосы можно преобразовать в окружность с помощью непрерывного изменения. Полоска Мёбиуса была приведена как пример односторонней поверхности \*). В последующем мы будем иметь дело только с двухсторонними поверхностями.

Изучение таких свойств фигур в пространстве, которые сохраняются при любом непрерывном преобразовании фигур, относится к разделу математики, называемому *топологией*. Такие утверждения, как „плоская замкнутая кривая имеет внутреннюю и внешнюю части“, с первого взгляда кажутся тривиальными, но в действительности они верны лишь тогда, когда определение замкнутой кривой дано совершенно четко, и даже тогда их довольно трудно доказать. С другой стороны, в теории магнитного поля, создаваемого электрическим током, обычно считается само собой разумеющимся, что замкнутый контур может быть заполнен двухсторонней поверхностью, однако кромка полосы Мёбиуса также представляет собой вполне возможную форму для такого контура и далеко не очевидно, как заполнить ее двухсторонней поверхностью, хотя это можно сделать.

Подобно тому как мы определяем длину кривой, естественно попытаться определить площадь искривленной поверхности, взяв множество точек на поверхности, соединив их так, чтобы получились треугольники, и определив площадь поверхности как предел суммы площадей треугольников, когда они делаются бесконечно малыми. (В случае плоских многоугольников с более чем тремя вершинами имеется, конечно, осложнение, так как четыре точки на искривленной поверхности не обязательно лежат в одной плоскости.) К сожалению, без дальнейших ограничений на способ выбора точек и на то, какие пары точек нужно соединить, этот способ не приводит к единственному определению площади, за исключением того случая, когда поверхность плоская. Из определения кратного интеграла естественно потребовать, чтобы стремились к нулю не просто площади треугольников, а все стороны треугольников; но даже этого недостаточно. Возьмем пример, данный Шварцем [11, стр. 309—311]. Представим себе круговой цилиндр радиусом 1 с осью, параллельной оси  $z$ , разделенный плоскими сечениями через интервал  $m$ . На окружности каждого из сечений возьмем  $n$  точек, равноотстоящих друг от друга, причем на каждом

---

\*) Более подробно см. [5, стр. 261].

сечении эти точки сдвинуты по окружности по отношению к точкам соседнего сечения на расстояние, равное половине расстояния между соседними точками по окружности сечения. Тогда эти точки определяют множество равнобедренных треугольников, вершины которых лежат на поверхности, и их стороны могут быть сделаны сколь угодно малыми, если мы возьмем достаточно малое  $m$  и достаточно большое  $n$ . Но если  $A, B$  — точки сечения, цилиндрические координаты которых  $\lambda$  отличаются на  $2\pi/n$ , то середина отрезка  $AB$  лежит внутри цилиндра на расстоянии  $1 - \cos(\pi/n)$  и ее  $z$ -координата отличается от  $z$ -координаты точки  $C$  — ближайшей точки соседнего сечения — на величину  $m$ . Поэтому плоскость треугольника наклонена к касательной плоскости в точке  $C$  на угол

$$\operatorname{arctg}\left(\frac{2}{m} \sin^2 \frac{\pi}{2n}\right).$$

Это выражение мало только тогда, когда  $mn^2$  велико; если  $m$  стремится к нулю как  $n^{-1}$ , то плоскости треугольников стремятся к касательным плоскостям, но это не так, если  $m$  стремится к нулю как  $n^{-2}$ . Далее, площадь треугольника равна

$$\sin \frac{\pi}{n} \left(m^2 + 4 \sin^4 \frac{\pi}{2n}\right)^{1/2}.$$

и площадь  $2n/m$  треугольников, покрывающих единичную длину вдоль цилиндра, равна

$$2n \sin \frac{\pi}{n} \left(1 + \frac{4}{m^2} \sin^4 \frac{\pi}{2n}\right)^{1/2}.$$

Когда  $n \rightarrow \infty$ , это выражение стремится к  $2\pi$  при условии, что  $mn^2 \rightarrow \infty$ . Таким образом, условие необходимое для того, чтобы сумма площадей треугольников стремилась к площади цилиндра, совпадает с условием, что плоскости треугольников стремятся к касательным плоскостям. Если это условие не выполнено, сумма может стремиться к любому пределу, превосходящему  $2\pi$ .

Следовательно, определение площади искривленной поверхности сложнее, чем определение длины кривой. Для кривой, если только направляющие косинусы касательной изменяются непрерывно при передвижении по кривой, направление короткой хорды должно стремиться к направлению касательной и отношение ее длины к длине дуги должно стремиться к 1. Для поверхности соответствующие приближения треугольниками требуют дополнительного условия, которое можно принять в следующем виде: все углы треугольников должны быть больше некоторого фиксированного угла  $\delta$ . Однако это условие ставит

дальнейший вопрос: для какого вида поверхностей возможен такой выбор, если стороны треугольников становятся как угодно малыми? Было показано (см. [9, стр. 100—112]), что это возможно при условии, что поверхность ограничена во всех направлениях и имеет нормаль в каждой точке, так что если  $Q_1, Q_2 \dots$  — точки поверхности, стремящиеся к  $P$ , то нормаль в точке  $Q_n$  стремится к нормали в  $P$ . Последнее условие часто формулируют так: поверхность имеет непрерывно вращающуюся нормаль. Эти условия являются до некоторой степени излишне строгими; они не выполняются, например, для куба. Но, поскольку обычная формула площади поверхности представляет собой интеграл функции от направляющих косинусов, из одного лишь существования интеграла следует, что эти условия не могут быть сильно ослаблены, если мы хотим, чтобы эта формула была верна, и они также могут быть приняты.

Если поверхность задана в виде

$$z = F(x, y), \quad (1)$$

где  $F$  имеет непрерывные частные производные  $F_x, F_y$  по  $x$  и  $y$ , то мы можем покрыть проекцию поверхности на плоскость  $x, y$  прямоугольниками со сторонами  $h, k$ , так что скачок изменений  $F_x$  и  $F_y$  в каждом прямоугольнике меньше  $\varepsilon$ . Возьмем точки, в которых прямые, проходящие через углы прямоугольников параллельно оси  $z$ , пересекают поверхность. В каждом из четырехугольников, определенных таким образом, проведем одну диагональ. Мы получим в результате множество треугольников, все вершины которых лежат на поверхности; проекция каждого на плоскость  $x, y$  имеет площадь  $hk/2$ . Далее, если вершины на плоскости  $x, y$  имеют координаты  $(x_r, y_r), (x_r + h, y_r), (x_r, y_r + k)$  и мы отметим их индексами 1, 2, 3, то проекция на плоскость  $x = \text{const}$  будет иметь площадь

$$\frac{1}{2} (y_3 - y_1)(z_2 - z_1) = \frac{1}{2} kh F_x(x_r + \theta h, y_r), \quad (2)$$

где  $0 < \theta < 1$ . Аналогично для  $(x_r + h, y_r), (x_r, y_r + k), (x_r + h, y_r + k)$  проекция будет иметь площадь

$$\frac{1}{2} kh F_x(x_r + \theta' h, y_r + k).$$

Далее, сумма площадей треугольников будет равна

$$\frac{1}{2} \sum hk [1 + (F_x)_r^2 + (F_y)_r^2]^{1/2}, \quad (3)$$

где  $F_x, F_y$  нужно брать в каждом случае в двух точках прямоугольника, лежащего в плоскости  $x, y$ . Если теперь мы

устремим  $\varepsilon$  к нулю, то эта сумма будет сходиться к интегралу

$$\iint (1 + F_x^2 + F_y^2)^{1/2} dx dy, \quad (4)$$

который существует при указанных условиях. Но следует помнить, что этот результат зависит от конкретного способа выбора треугольников, хотя бы это был и очень естественный способ. Если мы хотим получить интеграл от функции  $f(x, y, z)$  по поверхности, мы можем взять для каждого треугольника значение  $f(x, y, F(x, y))$  в некоторой точке треугольника, лежащего в плоскости  $x, y$ , и если  $f$  непрерывна, то в результате мы получим  $\iint f(x, y, z)(1 + F_x^2 + F_y^2)^{1/2} dx dy$ , что можно короче записать в виде  $\iint f dS$ .

В этом методе предполагается, что  $z$  однозначна. Если прямая, параллельная оси  $z$ , пересекает поверхность в двух точках, то необходимо рассмотреть их по отдельности; так, поверхность сферы нужно представить как две поверхности, одна из которых состоит из множества точек с меньшей координатой  $z$  при фиксированных  $x, y$ , а другая — из точек с большей координатой  $z$ . Кроме того, в этом методе предполагается, что  $F_x, F_y$  непрерывны и, следовательно, ограничены, так как интегрирование проводится по замкнутой области  $x, y$ ; но для сферы эти производные стремятся к бесконечности и интегралы нужно понимать как несобственные, исключив сначала зону около границы области и затем устремив ширину зоны к нулю. Если производные  $F_x, F_y$  имеют разрывы вдоль спрямляемой кривой или в конечном множестве изолированных точек, то поверхность можно разделить на куски, так чтобы разрывы лежали на границах кусков; при этом не возникнет никакой особой трудности. Если часть поверхности состоит из прямых, параллельных оси  $z$ , мы возьмем ее в виде  $x = G(y, z)$  или  $y = H(x, z)$ .

Такие интегралы обычно являются несобственными, и чтобы обойти трудности, связанные с бесконечными рядами несобственных интегралов, обычно требуют, чтобы поверхность не пересекалась бесконечное число раз какой-либо прямой, параллельной одной из осей. Поэтому такие ограничения на тип рассматриваемой поверхности довольно часты, но, к счастью, они выполняются для большинства поверхностей, которые мы встречаем на практике (см. приложение к гл. 5).

Поверхность можно задать параметрически с помощью уравнения  $x = x(\xi, \eta)$  и аналогичных уравнений для других координат, где  $x, y, z$  имеют непрерывные производные по  $\xi$  и  $\eta$ .



Если мы выберем  $\lambda, \mu$  столь малыми, что разности для каждой из этих производных будут всегда меньше  $\epsilon$ , когда  $|\xi_1 - \xi_2| \leq \lambda, |\eta_1 - \eta_2| \leq \mu$ , то мы можем применить аналогичные рассуждения к треугольникам, задаваемым на плоскости  $\xi, \eta$  прямоугольниками со сторонами  $\lambda, \mu$ . Мы найдем, что сумма теперь стремится к интегралу

$$\iint J d\xi d\eta, \quad (5)$$

где

$$J^2 = \left[ \frac{\partial(y, z)}{\partial(\xi, \eta)} \right]^2 + \left[ \frac{\partial(z, x)}{\partial(\xi, \eta)} \right]^2 + \left[ \frac{\partial(x, y)}{\partial(\xi, \eta)} \right]^2, \quad (6)$$

и что направляющие конусы нормали задаются соотношениями

$$J(l, m, n) = \pm \left( \frac{\partial(y, z)}{\partial(\xi, \eta)}, \frac{\partial(z, x)}{\partial(\xi, \eta)}, \frac{\partial(x, y)}{\partial(\xi, \eta)} \right), \quad (7)$$

причем знак положителен, если вращение направления  $d\xi$  к направлению  $d\eta$  положительно относительно нормали. С помощью прямой замены переменных можно показать, что этот интеграл равен интегралу (4) и сводится к нему, когда  $\xi = x, \eta = y$ , так что правило для вычисления площади инвариантно для различных параметрических представлений поверхности.

Это особенно важно, так как принятый метод триангуляции специально подбирается для каждого представления; поэтому необходимо проверить непротиворечивость.

Значительно труднее дать определение площади для поверхности общего вида. Лебег обошел трудность, отмеченную Шварцем, введя в рассмотрение нижние грани сумм площадей треугольников, приближающихся к поверхности, и предотвратив тем самым завышенную оценку площади, однако Безикович показал, что этот подход может привести к заниженной оценке. Он построил множество точек, имеющих, согласно определению Лебега, конечную площадь и в то же время положительный объем (см. [2, стр. 86—102]). Наилучшее общее определение принадлежит, по-видимому, Каратеодори (см. [4, стр. 404—426]). Все эти определения приводят к той же самой интегральной формуле для площади обычных поверхностей. Необходимые и достаточные условия для того, чтобы эта формула соответствовала определению Каратеодори, сформулированы Безиковичем (см. [3, стр. 1—7]).

В индексных обозначениях при преобразовании интеграла по поверхности мы можем интерпретировать  $l_i dS$  как

$$l_i dS = \epsilon_{ikm} \frac{\partial x_k}{\partial \xi} \frac{\partial x_m}{\partial \eta} d\xi d\eta \quad (8)$$

и

$$\varepsilon_{ips} J_i dS = \frac{\partial (x_p, x_s)}{\partial (\xi, \eta)} d\xi d\eta \quad (9)$$

снова при том ограничении, что направления  $d\xi$ ,  $d\eta$  и нормаль образуют правостороннюю систему направлений. Если они образуют левостороннюю систему, то знак нужно изменить на обратный.

**5.08. Лемма Грина.** Эта теорема, известная иначе как теорема Гаусса, теорема Остроградского (см. приложение к гл. 5) и теорема о дивергенции, утверждает, что если  $V$  — замкнутая область и  $S$  — ограничивающая ее поверхность, то

$$\int_V \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) dx dy dz = \int_S (lu + mv + nw) dS, \quad (1)$$

при условии, что тройной интеграл от  $\partial u/\partial x$ ,  $\partial v/\partial y$ ,  $\partial w/\partial z$  по области  $V$  существует и никакая прямая, параллельная какой-либо из осей, не пересекает  $S$  больше, чем некоторое фиксированное число раз;  $l$ ,  $m$ ,  $n$  являются направляющими косинусами внешней нормали к  $S$ . Следует помнить, что, как и в разделе 5.05, интеграл от  $\partial u/\partial x$  по области  $V$  означает интеграл по параллелепипеду  $D$ , включающему область  $V$  и ограниченному плоскостями  $x = X_1$ ,  $X_2$ ,  $y = Y_1$ ,  $Y_2$ ,  $z = Z_1$ ,  $Z_2$  от функции  $p(x, y, z)$ , равной  $\partial u/\partial x$  в области  $V$  и нулю вне ее; утверждение, что интеграл по области  $V$  существует, означает, что функция  $p(x, y, z)$  интегрируема в  $D$ .

Как показано в 5.051, из существования тройного интеграла вытекает, что этот интеграл можно вычислить с помощью последовательного интегрирования. Тогда

$$\int_V \int \int \frac{\partial u}{\partial x} dx dy dz = \int_{Y_1}^{Y_2} \int_{Z_1}^{Z_2} dy dz \int_{X_1}^{X_2} p(x, y, z) dx. \quad (2)$$

Для любой пары значений  $y$ ,  $z$ , которые не принимает ни одна точка из области  $V$ , функция  $p$  равна нулю для всех  $x$ . Пусть прямая, параллельная оси  $x$ , пересекает поверхность; обозначим индексы точек пересечения 1, 2, 3 ... в порядке увеличения  $x$ , так что нечетные индексы будут соответствовать точкам входа, а четные — точкам выхода. Для любых  $y$ ,  $z$  число пересечений  $2k$

ограничено. Вдоль такой прямой мы имеем

$$\int_{x_1}^{x_2} p \, dx = \int_{x_1}^{x_2} \frac{\partial u}{\partial x} \, dx + \int_{x_3}^{x_4} \frac{\partial u}{\partial x} \, dx + \dots = \sum_1^{2k} (-1)^j u_j \quad (3)$$

и

$$\iiint_V \frac{\partial u}{\partial x} \, dx \, dy \, dz = \int_{y_1}^{y_2} \int_{z_1}^{z_2} \sum_1^{2k} (-1)^j u_j \, dy \, dz = \sum_1^{2k} (-1)^j \iint u_j \, dy \, dz, \quad (4)$$

причем подынтегральное выражение равно нулю для значений  $y, z$ , которым не соответствует ни одной точки  $S$ .

Далее, из 5.07(4) имеем

$$|l| = \left[ 1 + \left( \frac{\partial x}{\partial y} \right)^2 + \left( \frac{\partial x}{\partial z} \right)^2 \right]^{-1/2} \quad (5)$$

и для любой части  $S$  получаем

$$\iint dS = \iint \frac{dy \, dz}{|l|}. \quad (6)$$

С помощью рисунка мы могли бы увидеть, что в точке выхода внешняя нормаль образует острый угол с осью  $x$  и поэтому для ее направляющего косинуса с осью  $x$  справедливо равенство  $l = |l|$ ; в точке входа мы имеем  $l = -|l|$ . Поэтому интеграл в правой части (4) равен  $\iint lu \, dS$  и берется по всем точкам пересечения. Но такие точки входа и выхода покрывают всю поверхность  $S$ , за исключением тех мест, где прямая, параллельная оси  $x$ , лежит на  $S$  для некоторого интервала значений  $x$ ; но на таких прямых  $l = 0$  и поэтому любая часть  $S$ , состоящая из таких линий, не вносит никакого вклада в интеграл.

Мы получим аналогичный результат для  $\iiint \frac{\partial v}{\partial y} \, dx \, dy \, dz$  и  $\iiint \frac{\partial w}{\partial z} \, dx \, dy \, dz$ , интегрируя сначала по  $y$  и  $z$  соответственно, а затем, складывая результаты, получим теорему (1).

Выражение  $l \, dS$  на самом деле нужно рассматривать как сокращение выражения  $\pm dy \, dz$ , причем знак выбирается в зависимости от того, являются ли точки окрестности точками выхода или входа прямой, параллельной оси  $x$ , так что правую часть (1) нужно рассматривать как сумму трех обычных двойных интегралов и не вносить сюда трудностей, возникающих при определении площади поверхности. При вычислении поверхностного интеграла, когда поверхность задается параметрически, будет

использована формула 5.07(8):

$$\iint l_i u_i dS = \iint \varepsilon_{ikm} u_i \frac{\partial x_k}{\partial \xi} \frac{\partial x_m}{\partial \eta} d\xi d\eta, \quad (7)$$

где  $l_i$ ,  $d\xi$ ,  $d\eta$  образуют правостороннюю систему направлений.

Приведенное выше доказательство имеет большую общность, чем это часто требуется. Достаточные условия, которые обычно выполняются, заключаются в следующем. 1) Поверхность ограничена и имеет всюду непрерывно вращающуюся нормаль, за исключением, возможно, конечного множества спрямляемых кривых, а также не пересекается более, чем некоторое фиксированное число раз, любой прямой линией. 2) Производные  $du/dx$ ,  $dv/dy$ ,  $dw/dz$  существуют и ограничены в области  $V$ , а также непрерывны всюду, за исключением, быть может, конечного множества поверхностей конечной площади. Нет никаких возражений против того, чтобы область  $V$  имела дыру. Она может, например, быть областью, ограниченной двумя концентрическими сферами. Тогда  $S$  состоит как из внешней, так и из внутренней сферических поверхностей, так что нормаль к внутренней границе берется по направлению от области  $V$ , т. е. к центру сфер.

Если  $u$ ,  $v$ ,  $w$  являются компонентами вектора  $u_i$ , мы можем записать нашу теорему в следующем виде:

$$\iiint_V \frac{\partial u_i}{\partial x_i} d\tau = \iint l_i u_i dS, \quad \iiint_V \operatorname{div} \mathbf{u} d\tau = \iint u_n dS, \quad (8)$$

где  $u_n$  — проекция вектора на внешнюю нормаль. Метод, использованный в приведенном доказательстве, рассматривает все члены по отдельности; поэтому не было бы никакого особого преимущества использовать в этом методе векторные или тензорные обозначения. В доказательстве, приводимом ниже (11.053), мы будем пользоваться индексными обозначениями: оно будет несколько более общим, но также и более трудным.

Обычно принято, особенно в немецких книгах, записывать интегралы только одним знаком интегрирования. Такая запись имеет тот недостаток, что, для того чтобы взять интегралы, их надо записать в виде двойных или тройных интегралов и при этом может произойти некоторая путаница: не существует переменных интегрирования  $S$  и  $\tau$ .

Если все компоненты вектора не зависят от  $z$  и мы применим нашу теорему к области, ограниченной двумя плоскостями  $z = \text{const}$  и цилиндром, образующие которого параллельны оси  $z$ , то получим двумерную форму теоремы

$$\iint \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dx dy = \iint (lu + mv) ds. \quad (9)$$

Концы не вносят в поверхностный интеграл никакого вклада, ибо там  $l = m = 0$  и значения  $n\omega$  в соответствующих точках двух концов равны по величине и противоположны по знаку; сомножитель, равный длине цилиндра, в правой и левой частях равенства сократится. Этот результат может быть легко доказан и непосредственно.

Заменяя  $u$  на  $v$  и  $v$  на  $-u$ , мы получаем

$$\iint \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy = \int (lv - mu) ds. \quad (10)$$

Но если взять за положительное направление вдоль касательной такое, при котором площадь остается слева, то направляющие косинусы касательной равны  $(-m, l) = (dx/ds, dy/ds)$ . Поэтому интеграл в правой части равен  $\int (u dx + v dy)$  и берется по границе в положительном направлении. Это является двухмерной формой *теоремы Стокса*.

**5.081. Теорема Грина.** Положим в лемме Грина

$$u_i = U \frac{\partial V}{\partial x_i}.$$

Тогда

$$\iiint \frac{\partial}{\partial x_i} \left( U \frac{\partial V}{\partial x_i} \right) d\tau = \iint l_i U \frac{\partial V}{\partial x_i} dS,$$

т. е.

$$\iiint U \nabla^2 V d\tau + \iiint \frac{\partial U}{\partial x_i} \frac{\partial V}{\partial x_i} d\tau = \iint U \frac{\partial V}{\partial n} dS,$$

где под  $\partial V / \partial n$  мы понимаем дифференцирование по направлению внешней нормали. Аналогично

$$\iiint V \nabla^2 U d\tau + \iiint \frac{\partial U}{\partial x_i} \frac{\partial V}{\partial x_i} d\tau = \iint V \frac{\partial U}{\partial n} dS,$$

и поэтому

$$\iiint (U \nabla^2 V - V \nabla^2 U) d\tau = \iint \left( U \frac{\partial V}{\partial n} - V \frac{\partial U}{\partial n} \right) dS$$

при условии, что  $U \partial V / \partial x_i$  и  $V \partial U / \partial x_i$  имеют первые производные, интегрируемые в области, т. е. вторые производные  $U$  и  $V$  существуют и интегрируемы.

**5.082. Теорема Стокса.** Пусть  $C$  — простая замкнутая кривая и  $S$  — двухсторонняя поверхность, имеющая  $C$  границей. Пусть  $x_i$  — координаты,  $u_i$  — три функции от  $x_i$  с непрерывными

первыми производными. Теорема Стокса утверждает, что

$$\int_C u_i dx_i = \int_S l_i \varepsilon_{ikm} \frac{\partial u_m}{\partial x_k} dS, \quad (1)$$

или в векторных обозначениях

$$\int_C \mathbf{u} \cdot d\mathbf{x} = \int_S \mathbf{l} \cdot \text{rot } \mathbf{u} dS, \quad (2)$$

где  $l_i$  — направляющие косинусы нормали к элементу поверхности  $dS$ . Направление, которое нужно выбрать для нормали, определяется следующим образом: если кривую  $C$  непрерывно деформировать и переместить так, что она описывается в положительном направлении в некоторой плоскости  $z = \text{const}$ , и  $S$  лежит в этой плоскости, то нормаль должна совпадать с положительным направлением оси  $z$ .

Предположим, что поверхность задается с помощью параметров  $\xi, \eta$ . Тогда с помощью двумерной теоремы Стокса получаем

$$\begin{aligned} \int \int \varepsilon_{ikm} l_i \frac{\partial u_m}{\partial x_k} dS &= \int \int \frac{\partial u_m}{\partial x_k} \left( \frac{\partial x_k}{\partial \xi} \frac{\partial x_m}{\partial \eta} - \frac{\partial x_m}{\partial \xi} \frac{\partial x_k}{\partial \eta} \right) d\xi d\eta = \\ &= \int \int \left( \frac{\partial u_m}{\partial \xi} \frac{\partial x_m}{\partial \eta} - \frac{\partial u_m}{\partial \eta} \frac{\partial x_m}{\partial \xi} \right) d\xi d\eta = \\ &= \int \int \left[ \frac{\partial}{\partial \xi} \left( u_m \frac{\partial x_m}{\partial \eta} \right) - \frac{\partial}{\partial \eta} \left( u_m \frac{\partial x_m}{\partial \xi} \right) \right] d\xi d\eta = \\ &= \int_C u_m \left( \frac{\partial x_m}{\partial \eta} d\eta + \frac{\partial x_m}{\partial \xi} d\xi \right). \end{aligned} \quad (3)$$

Но это как раз равно  $\int_C u_m dx_m$ .

При доказательстве предполагается, что  $\partial^2 x_m / \partial \xi \partial \eta$  и  $\partial^2 x_m / \partial \eta \partial \xi$  существуют и равны, это в свою очередь можно вывести из предположения, что эти производные непрерывны (см. [6, стр. 221–222]). Часто  $\mathbf{n} dS$  записывается как  $d\mathbf{S}$ , который понимают как вектор величины  $dS$ , направленный вдоль  $\mathbf{n}$ .

**5.09. Поток и циркуляция.** Если  $u_i$  — вектор,  $n_i$  — направление нормали к поверхности, то интеграл  $\int \int n_i u_i dS$ , взятый по поверхности, называется *поток*ом вектора  $u_i$  через поверхность. В гидродинамике, если  $u_i$  — скорость жидкости в некоторой точке, поток — это объем жидкости, проходящей через поверхность в единицу времени; отсюда его название. Ин-

теграл  $\int_C u_i dx_i$ , взятый по некоторому контуру, называется *циркуляцией* по этому контуру. В этом случае лемму Грина можно прочесть следующим образом: поток вектора через замкнутую поверхность равен интегралу от его дивергенции по объему, находящемуся внутри поверхности. Теорему Стокса можно прочитать так: циркуляция вектора по контуру равна потоку его ротора через шапку, заполняющую контур \*).

Если функции  $u_i$  дифференцируемы в точке  $P$  и мы берем вокруг  $P$  любую последовательность поверхностей  $S_n$  одинаковой формы и ориентации, диаметр и объем которых равны  $a_n$ ,  $V_n$ , то из 5.04 (5) следует, что

$$\frac{1}{V_n} \iint l_i u_i dS_n = \left( \frac{\partial u_i}{\partial x_i} \right)_P + \frac{1}{V_n} \iiint l_i v_i r dS_n, \quad (1)$$

где  $|v_i| \rightarrow 0$  равномерно при  $a_n \rightarrow 0$ .

Последний член стремится к нулю вместе с  $a_n$ , так как

$\iiint r dS_n = O(a_n^3) = O(V_n)$ . Поэтому

$$\lim_{a_n \rightarrow 0} \frac{1}{V_n} \iint l_i u_i dS_n = \left( \frac{\partial u_i}{\partial x_i} \right)_P. \quad (2)$$

Этот результат особенно полезен тогда, когда нам нужно выразить дивергенцию вектора с помощью криволинейных ортогональных координат.

Векторы, дивергенция которых равна нулю всюду в некоторой области, называются *соленоидальными* в этой области. Поток соленоидального вектора через любую замкнутую поверхность в этой области равен нулю; обратно, используя последний результат, получаем, что если вектор имеет нулевой поток через любую замкнутую поверхность в области, то он соленоиден в этой области. В частности, в потоке жидкости плотности  $\rho$ , где  $\rho$  может быть переменной, скорость переноса массы через поверхность равна  $\iint \rho l_i u_i dS$ . Если масса внутри каждой поверхности остается одной и той же с течением времени, то вектор  $\rho u_i$  соленоиден и  $\operatorname{div}(\rho u) = 0$ . В частном случае, когда жидкость однородна и несжимаема, так что масса внутри замкнутой поверхности в этой жидкости постоянна и  $\partial \rho / \partial x_i = 0$ , мы имеем просто  $\operatorname{div} u = 0$ . Во многих гидродинамических задачах это условие удовлетворяется. Достаточным

---

\*) В русской литературе теорему Стокса обычно формулируют так: циркуляция вектора по контуру равна потоку его ротора через любую поверхность, опирающуюся на этот контур. — *Прим. ред.*

условием, чтобы вектор был соленоидальным, будет то, что он является ротором другого вектора. Мы докажем в 6.11, что это условие также и необходимо.

Векторы, имеющие нулевую циркуляцию по любому контуру в рассматриваемой области, называются *безвихревыми* в этой области. Если  $A(x_i)$  и  $B(x_i)$  — любые две точки в области и мы соединяем их двумя различными траекториями  $L$  и  $L'$  в этой области, то имеем

$$\int_L u_i dx_i = \int_{L'} u_i dx_i. \quad (3)$$

Если траектория  $L$  от  $A$  к  $B$  вместе с траекторией  $L'$  от  $B$  к  $A$  образует замкнутый контур в области и вектор  $u_i$  безвихревой, то интеграл по этому контуру будет равен нулю. Но в этом контуре часть  $L'$  проходится в противоположном направлении, откуда и получается требуемый результат. Таким

образом, интеграл  $\int_A^B u_i dx_i$  зависит только от начальной и конечной точек, но не от пути интегрирования и его можно выразить в виде  $\varphi(B) - \varphi(A)$ , где  $\varphi$  — некоторая скалярная функция координат. Возьмем теперь точку  $B'$  с координатами  $(x_1 + \delta x_1, x_2, x_3)$ . Чтобы попасть из  $A$  в  $B'$ , мы можем пройти сначала в  $B$  и затем в  $B'$ ; в результате получим

$$\frac{1}{\delta x_1} [\varphi(B') - \varphi(B)] = \frac{1}{\delta x_1} \int_B^{B'} u_i dx_i \rightarrow u_1, \quad (4)$$

откуда

$$\frac{\partial}{\partial x_1} \varphi(B) = u_1 \quad (5)$$

и, используя симметрию, получаем

$$u_i = \frac{\partial \varphi}{\partial x_i}.$$

Отсюда следует, что безвихревой вектор является градиентом некоторого скаляра. Снова имеем

$$\frac{\partial u_i}{\partial x_k} = \frac{\partial^2 \varphi}{\partial x_i \partial x_k} = \frac{\partial u_k}{\partial x_i}, \quad (6)$$

и поэтому ротор от безвихревого вектора равен нулю. Обратно, если ротор вектора равен нулю всюду в области, то мы можем применить теорему Стокса, чтобы показать, что его циркуля-



ция по любому контуру, который можно заполнить шапкой\*), лежащей в нашей области, равна нулю, и поэтому он безвихревой и, следовательно, согласно приведенным выше рассуждениям, является градиентом некоторого скаляра.

Если  $u_i$  является и соленоидальным и безвихревым, то существует функция  $\varphi$ , и тогда, так как  $\partial u_i / \partial x_k = 0$ , мы получаем

$$\frac{\partial^2 \varphi}{\partial x_i^2} = \frac{\partial^2 \varphi}{\partial x_1^2} + \frac{\partial^2 \varphi}{\partial x_2^2} + \frac{\partial^2 \varphi}{\partial x_3^2} = 0, \quad (7)$$

что можно записать компактно в виде

$$\nabla^2 \varphi = 0. \quad (8)$$

### ПРИМЕРЫ.

1. Доказать, что если  $-\pi < \alpha < \pi$ , то

$$\int_0^\infty \int_0^\infty \exp(-x^2 - 2xy \cos \alpha - y^2) dx dy = \frac{\alpha}{2 \sin \alpha}.$$

2. Пусть  $S, T$  — фиксированные точки с координатами  $(0, 0, \pm R/2)$ , и  $P$  — переменная точка  $(x, y, z)$ ; расстояния  $PS, PT$  обозначаются через  $s, t$  соответственно. Если

$$\xi = \frac{s+t}{R}, \quad \eta = \frac{s-t}{R}$$

и  $\varphi$  — угол между плоскостью  $PST$  и плоскостью  $y=0$ , то показать, что

$$\left| \frac{\partial (\xi, \eta, \varphi)}{\partial (x, y, z)} \right| = \frac{8}{R^3 (\xi^2 - \eta^2)}.$$

Откуда доказать, что

$$\int \int \int \frac{1}{st} e^{-(s+t)/R} dx dy dz = \frac{2\pi R}{e},$$

где интеграл берется по всему пространству.

(Prelim., 1942.)

3. С помощью преобразования

$$x + iy = (u + iv)^2$$

или другим способом вычислить интеграл

$$\int \int (x^2 + y^2)^{-1/2} dx dy,$$

взятый по области, ограниченной дугами софокусных парабол

$$y^2 = 4a_r(x + a_r) \quad (r = 1, 2, 3),$$

где

$$a_1 > a_2 > 0, \quad a_3 < 0.$$

---

\*) См. примечание на стр. 319.

4. Доказать, что интеграл

$$\int \frac{dS}{p},$$

взятый по поверхности эллипсоида с полуосями  $a, b, c$ , где  $p$  — длина перпендикуляра из центра на касательную плоскость в некоторой точке эллипсоида, равен

$$\frac{4}{3} \pi abc \left( \frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} \right).$$

Вычислить

$$\int \frac{dS}{p^3}. \quad (\text{М. Т., 1940.})$$

5. Определить новые пределы следующих интегралов после изменения порядка интегрирования:

$$\int_0^1 dx \int_{x^2}^1 f(x, y) dy, \quad \int_{\frac{\pi}{2}}^{\frac{1}{2}\pi} d\theta \int_0^{b \operatorname{cosec} \theta} f(r, \theta) dr. \quad (\text{И. С., 1940.})$$

6. Записать интеграл

$$\int_0^a dz \int_0^z dy \int_0^{\sqrt{yz-y^2}} dx$$

в сферических координатах и показать, что его величина равна  $\pi a^3/24$ . (И. С., 1938).

7. Если  $K_n$  — однородный многочлен положительной целой степени  $n$  в системе координат, где выполняется условие  $\nabla^2 K_n = 0$  и  $S$  — сфера радиуса  $a$  с центром в начале координат, то доказать, что

$$\int \int \left( \frac{\partial K_n}{\partial x_i} \right)^2 dS = \frac{n(2n+1)}{a^2} \int \int K_n^2 dS.$$

8. Если  $\mathbf{A}$  и  $\mathbf{B}$  — две векторные функции от координат точки пространства, то доказать, что

$$\operatorname{div} (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \operatorname{rot} \mathbf{A} - \mathbf{A} \cdot \operatorname{rot} \mathbf{B}.$$

Если далее  $\mathbf{A}$  и  $\mathbf{B}$  — функции времени и связаны соотношениями

$$\frac{d\mathbf{A}}{dt} = \operatorname{rot} \mathbf{B}, \quad \frac{d\mathbf{B}}{dt} = -\operatorname{rot} \mathbf{A},$$

и  $\tau$  — объем, заключенный в фиксированной поверхности  $S$ , то доказать, что

$$\frac{1}{2} \frac{d}{dt} \int \int \int (\mathbf{A}^2 + \mathbf{B}^2) d\tau = - \int \int (\mathbf{A} \times \mathbf{B}) \cdot d\mathbf{S}. \quad (\text{М/с, III, 1931.})$$

9. Доказать, что

$$\operatorname{div} (\varphi \mathbf{A}) = \varphi \operatorname{div} \mathbf{A} + \mathbf{A} \cdot \operatorname{grad} \varphi.$$

Если  $\operatorname{div} \mathbf{D} = 4\pi\rho$ ,  $\mathbf{E} = -\operatorname{grad} \varphi$  и  $\mathbf{D} = K\mathbf{E}$ , где  $K$  не зависит от  $\mathbf{E}$ , то показать, что если  $\varphi$  имеет порядок  $O(1/r)$  на бесконечности, то

$$\frac{1}{2} \int \int \int \rho \varphi d\tau = \frac{1}{8\pi} \int \int \int \mathbf{E} \cdot \mathbf{D} d\tau,$$

где интегралы берутся по всему пространству.

## ЛИТЕРАТУРА

1. *Besicovitch A. S.*, J. Lond. Math. Soc., **19**, 1944.
2. *Besicovitch A. S.*, Q. J. Math. (Oxford series), **16**, 1945.
3. *Besicovitch A. S.*, Q. J. Math. (Oxford series), **20**, 1949.
4. *Carateodory C.*, Gött. Nachr., 1914.
5. *Courant R., Robbins H.*, What is Mathematics, Ind., Oxford Univ. Press, 1946. (Русский перевод: *Курант Р., Робинс Г.*, Что такое математика? М.—Л., Гостехиздат, 1948.)
6. *Gibson G. A.*, Calculus.
7. *Hobson E. W.*, Functions of a Real Variable, Camb. Univ. Press, 1907.
8. *Hobson E. W.*, Theory of Functions of a Real Variable, 1920.
9. *Kellog O. D.*, Foundations of Potential Theory, 1929.
10. *Newman M. H. A.*, Topology of Plane Sets, 1939.
11. *Schwarz H. A.*, Ges. Math. Abhand., **2**, 1890.
12. *Stolz O., Young V. H.*, Proc. Lond. Math. Soc., (2) **7** (1909).

## ПРИЛОЖЕНИЕ К ГЛАВЕ 5

## 5.04а. Если

$$f(x, y) = (x^2 + y^2)^{1/2} \frac{x^2 y}{x^4 + y^2}, \quad f(0, 0) = 0,$$

то  $\partial f / \partial x = \partial f / \partial y = 0$  в точке  $(0, 0)$ . Далее, если  $x = r \cos \theta$ ,  $y = r \sin \theta$ , то  $f/r \rightarrow 0$  при  $r \rightarrow 0$  для любого фиксированного  $\theta$ . Отсюда  $\partial f / \partial r = 0$  и градиент  $f$  по любому направлению обращается в нуль. Но  $f$  недифференцируема, так как при  $r < \delta$  мы могли бы всегда взять  $y = x^2$ , что дает  $f/r = 1/2$ . Поэтому дифференцируемость функции не является необходимым условием для того, чтобы градиент был вектором, как это было определено в 3.06. Однако свойство (5) настолько важно, что векторное свойство производной мало о чем говорит нам, если (5) не выполняется. Например, если  $f$  дифференцируема и  $r \rightarrow 0$ ,  $\theta \rightarrow \varphi$  любым способом, то

$$\frac{1}{r} [f(x, y) - f(0, 0)] \rightarrow \frac{\partial f}{\partial x} \cos \varphi + \frac{\partial f}{\partial y} \sin \varphi,$$

но это неверно для рассмотренного выше примера.

**5.051а. Дифференцирование под знаком интеграла.** Мы покажем, что при некоторых условиях имеет место равенство

$$\frac{d}{dc} \int_a^b f(x, c) dx = \int_a^b \frac{\partial f(x, c)}{\partial c} dx,$$

где  $a$  и  $b$  не зависят от  $c$ .

Предположим, что интеграл  $\int_a^b f(x, y) dx$  существует и что  $\partial f(x, y) / \partial y$  — непрерывная функция двух переменных  $x$  и  $y$ , где

$a \leq x \leq b$  и  $|y - c| \leq \delta$ ,  $\delta > 0$ . Тогда, согласно 5.051, из-за того что двойной интеграл

$$\int_a^b dx \int_c^{c+h} \frac{\partial f(x, y)}{\partial y} dy$$

существует при  $|h| < \delta$ , получаем

$$\int_a^b dx \int_c^{c+h} \frac{\partial f(x, y)}{\partial y} dy = \int_c^{c+h} dy \int_a^b \frac{\partial f(x, y)}{\partial y} dx.$$

Интеграл слева равен

$$\int_a^b [f(x, c+h) - f(x, c)] dx.$$

Интеграл в правой части равенства равен

$$h \int_a^b \left[ \frac{\partial f(x, y)}{\partial y} \right]_{y=c} dx + o(h),$$

так как однократный интеграл является непрерывной функцией  $y$ . Отсюда, разделив на  $h$  и устремив  $h$  к нулю, мы получаем

$$\frac{d}{dc} \int_a^b f(x, c) dx = \int_a^b \frac{\partial f(x, c)}{\partial c} dx.$$

Если  $a, b$  зависят от  $c$ , то нужно добавить члены  $f(b, c) \frac{db}{dc} - f(a, c) \frac{da}{dc}$ .

Распространение на случай, когда  $a, b$  стремятся к бесконечности, получается из 1.12, если  $\int_a^b \frac{\partial f}{\partial y} dx$  равномерно сходится.

Доказательство при несколько иных условиях следующее.

Пусть  $\int_a^b f(x, y) dx$  существует для всех  $y$ , таких, что  $|y - c| \leq \delta$ , и пусть  $\frac{\partial f(x, y)}{\partial y}$  существует для  $a \leq x \leq b, |y - c| \leq \delta$ . Пусть, далее,  $\partial^2 f(x, y) / \partial y^2$  существует в этой же области и ограничена верхней гранью  $M$ , но не обязательно непрерывна. Тогда

$$\frac{1}{h} \left[ \int_a^b f(x, c+h) dx - \int_a^b f(x, c) dx \right] = \int_a^b f_y(x, c+\theta h) dx,$$

где  $0 < \theta < 1$ ,  $|h| \leq \delta$ . Кроме того,

$$f_y(x, c + \theta h) - f_y(x, c) = f_{yy}(x, c + \varphi h) \theta h,$$

где  $0 < \varphi < \theta < 1$ . Отсюда

$$\left| \int_a^b f_y(x, c + \theta h) dx - \int_a^b f_y(x, c) dx \right| < M(b-a) |h|,$$

так что, переходя к пределу при  $h \rightarrow 0$ , мы получаем

$$\frac{d}{dc} \int_a^b f(x, c) dx = \int_a^b \frac{\partial f(x, c)}{\partial c} dx.$$

**5.07а.** Поверхность, удовлетворяющую условиям 5.07, можно заключить в произвольно малый объем. Так как  $l^2 + m^2 + n^2 = 1$ , то в каждой точке поверхности по крайней мере одна из величин  $l^2$ ,  $m^2$ ,  $n^2$  больше  $1/3$ . Возьмем точки, где  $n^2 \geq 1/3$ ; они образуют область или области по  $x, y$ , так как  $n$  непрерывна и

$$1 + F_x^2 + F_y^2 = \frac{1}{n^2} \leq 3.$$

Тогда

$$F_x^2 \leq 2, \quad F_y^2 \leq 2.$$

Отсюда

$$|F(x+h, y) - F(x, y)| \leq h \sqrt{2}, \quad |F(x, y+k) - F(x, y)| \leq k \sqrt{2}.$$

Возьмем  $k \leq h$ . Тогда параллелепипед со сторонами  $2h, 2k, 2h\sqrt{2}$  с центром в  $(x, y, z)$  будет заключать в себе все точки  $S$ , удовлетворяющие неравенствам  $|\xi - x| < h$ ,  $|\eta - y| < k$ , и будет перекрываться с аналогичными параллелепипедами, центры которых лежат в точках  $S$  и соответствуют соседним точкам решетки; поэтому такое множество параллелепипедов вокруг точек, соответствующих на плоскости  $x, y$  всем точкам  $(h, k)$ -решетки, будет заключать в себе всю поверхность  $S$ . Пусть амплитуды изменения  $x, y$  будут  $H, K$ . Тогда число точек решетки не превосходит

$$\left(\frac{H}{h} + 1\right) \left(\frac{K}{k} + 1\right)$$

и полный объем параллелепипедов не превосходит

$$\left(\frac{H}{h} + 1\right) \left(\frac{K}{k} + 1\right) 8h^2k \sqrt{2} = 8 \sqrt{2} (Hh + h^2)(K + k),$$

что стремится к нулю вместе с  $h$ , так как  $k \leq h$ .

Применим аналогичное рассуждение к точкам, где  $l^2$  или  $m^2$  больше  $1/3$ , и требуемый результат получится с помощью сложения.

**5.08a.** Грин вывел теорему, носящую теперь его имя (см. [1, стр. 23]), отделяя члены и интегрируя по частям. М. В. Остроградский (см. [2]) дал строгий вывод теоремы о дивергенции, но, конечно, все принципы, использованные в ней, содержатся в рассуждениях Грина.

## ТЕОРИЯ ПОТЕНЦИАЛА

Но все это движение вызвано любовью к перемещениям.

Спенсер „Королева фей“, книга 7.

**6.01.  $1/r$  как решение  $\nabla^2\varphi = 0$ .** Пусть  $x_i$  — координаты точки  $P$  и  $r$  — ее расстояние от начала координат. Тогда

$$\frac{\partial}{\partial x_i} \frac{1}{r} = \frac{d}{dr} \left( \frac{1}{r} \right) \frac{\partial r}{\partial x_i} = -\frac{x_i}{r^3}, \quad (1)$$

$$\frac{\partial^2}{\partial x_i \partial x_k} \frac{1}{r} = -\frac{\partial}{\partial x_k} \frac{x_i}{r^3} = -\frac{1}{r^3} \frac{\partial x_i}{\partial x_k} + \frac{3x_i x_k}{r^5} = \frac{3x_i x_k - r^2 \delta_{ik}}{r^5}. \quad (2)$$

Теперь положим  $k=i$  и используем правило суммирования; так как

$$x_i x_i = r^2, \quad \delta_{ii} = 3, \\ \frac{\partial^2}{\partial x_i \partial x_i} \frac{1}{r} = 0, \quad (3)$$

то  $1/r$  есть решение уравнения Лапласа  $\nabla^2\varphi = 0$  всюду, кроме начала координат.

Отсюда сразу следует, что если  $\xi_i$  — координаты другой точки  $Q$  и

$$r^2 = (x_i - \xi_i)^2$$

(подразумевается суммирование), то

$$\frac{\partial^2}{\partial x_i^2} \frac{1}{r} = 0, \quad (4)$$

кроме  $x_i = \xi_i$ . Заметим, что  $\frac{\partial}{\partial x_i} \frac{1}{r} = -\frac{\partial}{\partial \xi_i} \frac{1}{r}$ ; этот результат нам потребуется неоднократно.

Далее, если мы возьмем  $n$  точек  $Q_1 \dots Q_n$  и обозначим их координаты через  $\xi_{is}$ , а расстояния  $Q_s P$  (которые считаются различными) через  $r_s$ , то

$$\nabla^2 \sum_{s=1}^n \frac{a_s}{r_s} = 0, \quad (5)$$

где  $a_s$  — произвольные константы; равенство справедливо, если  $r_s$  не равны 0, т. е. если  $P$  не совпадает ни с одной из точек  $Q_s$ . Подразумевается, конечно, что дифференцирование производится по координатам точки  $P$ . Следовательно, с ограничением, о котором только что говорилось, любая функция вида  $\sum_{s=1}^n (a_s/r_s)$  является решением уравнения Лапласа.

Такой вид имеет гравитационный потенциал, образованный распределением весомих частиц, и электростатический потенциал от множества точечных зарядов. Следовательно, оба они удовлетворяют уравнению Лапласа. Это уравнение возникает также в гидродинамике несжимаемой жидкости. Если  $u_i$  есть скорость в точке  $P(x_i)$ , то из условия, что масса внутри любой замкнутой поверхности постоянна, следует, что  $u_i$  — соленоидальный вектор. Далее, для любого замкнутого контура в сосуде, целиком заполненном жидкостью, циркуляция  $\int u_i dx_i$  вдоль него практически равна нулю, если этот поток нигде не проходит вблизи твердой границы. Поэтому с хорошим приближением к действительности  $u_i$  является потенциальным вектором. В классической гидродинамике принимается аппроксимация

$$u_i = \frac{\partial \phi}{\partial x_i},$$

где  $\phi$  — скалярная функция координат, удовлетворяющая уравнению Лапласа и называемая *потенциалом скоростей*. Решения этого уравнения, таким образом, включают в себя целиком ту часть гидродинамики, в которой пренебрегают вязкостью и допускают, что жидкость имеет постоянную плотность независимо от давления и других осложняющих обстоятельств, например изменения температуры и состава. Ни одна реальная жидкость не удовлетворяет таким условиям, но для многих встречающихся в действительности движений жидкости, ограниченной твердыми стенками, они выполняются с точностью, сравнимой с ошибкой наблюдений. Исключение составляет часть жидкости в непосредственной близости от твердой границы; эта область, обычно тонкая, изучается в *теории пограничного слоя*, развитой в последнее время.

Если  $r$  — расстояние от  $P$  до начала, то  $r_s/r \rightarrow 1$ , когда  $r \rightarrow \infty$ . Следовательно, при достаточно большом  $r$

$$r\phi = r \sum_{s=1}^n \frac{a_s}{r_s} \rightarrow \sum_{s=1}^n a_s. \quad (6)$$



Поэтому, если  $\sum a_s \neq 0$ ,  $\varphi$  при больших  $r$  ведет себя как  $1/r$ . Если  $\sum a_s = 0$ ,  $\varphi$  убывает быстрее, чем  $1/r$ .

Рассмотрим теперь поток градиента  $1/r$  через сферу радиуса  $a$  с центром в начале координат. Направляющие косинусы внешней нормали равны  $x_i/r$ ; поэтому

$$\iint l_i \frac{\partial}{\partial x_i} \frac{1}{r} dS = - \iint \frac{x_i}{r} \frac{x_i}{r^3} dS = - \iint \frac{1}{a^2} dS.$$

Площадь сферы равна  $4\pi a^2$ ; следовательно,

$$\iint \frac{\partial}{\partial n} \left( \frac{1}{r} \right) dS = \iint l_i \frac{\partial}{\partial x_i} \frac{1}{r} dS = -4\pi. \quad (7)$$

Этот результат можно распространить на любую замкнутую поверхность, окружающую начало координат. Возьмем такую поверхность  $S$  и сферу  $\Sigma$ , достаточно большую, чтобы заключить  $S$  целиком внутри себя, и применим лемму Грина к области  $\Sigma - S$  между  $S$  и  $\Sigma$ . Применяя обозначение  $\partial/\partial v$  для дифференцирования вдоль внешней к этой области нормали, которая будет направлена вовне по отношению  $O$  на  $\Sigma$  и внутрь на  $S$ , мы получим

$$\iiint_{\Sigma-S} \frac{\partial^2}{\partial x_i^2} \left( \frac{1}{r} \right) dx_1 dx_2 dx_3 = \iint_{S, \Sigma} \frac{\partial}{\partial v} \left( \frac{1}{r} \right) dS. \quad (8)$$

Левая часть равна нулю. Вклад в правую часть от  $\Sigma$  равен  $-4\pi$ . Следовательно,

$$\iint_S \frac{\partial}{\partial v} \left( \frac{1}{r} \right) dS = 4\pi. \quad (9)$$

Так как  $dv$  взято по направлению к  $O$ , то, взяв  $dn$ , направленным от  $O$ , получим выражение

$$\iint_S \frac{\partial}{\partial n} \left( \frac{1}{r} \right) dS = -4\pi, \quad (10)$$

которое распространяет (7) на любую поверхность, заключающую начало координат. (Применяя лемму Грина, легко показать, что этот интеграл равен нулю, если  $S$  не заключает начало координат.)

В классической гидродинамике, если течение направлено по радиусу и симметрично относительно начала координат, а  $u$  — скорость на расстоянии  $r$  от начала,  $u$  зависит только от  $r$ . Объем жидкости, вытекающей из сферы радиуса  $r$  в единицу времени, равен  $4\pi r^2 u$ . Если область между двумя такими

сферами радиусов  $r_1$  и  $r_2$  заполнена жидкостью во время движения, то отсюда следует, что  $r_1^2 u_1 = r_2^2 u_2$ ; таким образом,

$$u = \frac{m}{r^2}, \quad (11)$$

где  $4\pi m$  представляет объем жидкости, вытекающей из начала координат за единицу времени. Поэтому

$$\varphi = -\frac{m}{r}, \quad (12)$$

отвлекаясь от не относящейся к делу константы. Если  $m$  положительно, тогда  $\varphi$  — потенциал скоростей *источника*, из которого вытекает объем жидкости, равный  $4\pi m$  за единицу времени. Если  $m$  отрицательно, то это же относится к *стоку*. Все это можно применить к любому множеству источников и стоков, а также и в теории гравитации и электростатике.

Сходство между этими тремя областями математической физики, к которым можно добавить электрический ток в однородном проводнике, столь велико, что удобно развивать сразу общую для них математическую теорию. С небольшими изменениями большая часть ее приложима и к магнетизму. Эта теория называется *теорией потенциала*.

Из (10) сразу следует, что если

$$\varphi = \sum_{s=1}^n \frac{a_s}{r_s}, \quad (13)$$

то

$$\iint \frac{\partial \varphi}{\partial n} dS = -4\pi \sum' a_s, \quad (14)$$

где суммирование в  $\sum'$  распространяется на  $a_s$ , которые соответствуют точкам  $Q_s$ , лежащим внутри  $S$ ; действительно, если  $Q_s$  лежит внутри  $S$ , то  $a_s/r_s$  вносит слагаемое  $-4\pi a_s$  в сумму, а если  $Q_s$  лежит вне  $S$ , то в сумму не вносится ничего. Это один из видов *теоремы Гаусса* \*).

**6.02. Непрерывные распределения.** Приложение приведенных выше результатов к непрерывным распределениям сталкивается с математическими и физическими трудностями. Если мы будем считать, что частицами являются протоны и электроны, то гравитационный и электрический потенциалы будут представлены конечными суммами того вида, который мы только что рассмотрели. На практике, однако, число частиц в любом

\*) Эту теорему в отечественной литературе называют теоремой Гаусса — Остроградского. — *Прим. ред.*

куске вещества обычного размера столь велико, что оперировать с суммами невозможно; кроме того, на малых расстояниях возникают дополнительные силы. Но именно по этой причине возможен другой способ. Интеграл является пределом конечной суммы, если число интервалов становится очень большим, в то время как общая длина остается прежней. Таким образом, сумма, взятая по большому конечному числу интервалов, является хорошим приближением к интегралу и, наоборот, интеграл хорошо аппроксимирует такую сумму. Это наводит на мысль, что вместо того, чтобы рассматривать сосредоточенные в отдельных точках вещество или заряды, можно считать, что они непрерывно распределены по объему таким образом, чтобы приближенно сохранились прежними общая масса или заряд в любой области, содержащей много частиц. Очевидно, что невозможно сохранить массу или заряд в точности прежними в любой области. Если бы это было так, мы могли бы взять замкнутую поверхность вокруг единичной частицы; масса внутри нее стремилась бы к нулю в случае замены точечной частицы непрерывным распределением массы и в то же время к конечному пределу для всякого реального распределения.

В действительности это приближение хорошо выражает только средние свойства областей, содержащих много элементов, а не свойства отдельных частиц. Свойства первого типа обычно называют *молярными* или *макроскопическими*, второго — *молекулярными* или *микроскопическими*. (Не надо думать, что эти слова выбраны потому, что в микроскопе можно увидеть отдельные молекулы!) Если, например, область, содержащая  $n$  частиц, где  $n$  велико, не слишком вытянута в одном направлении по сравнению с другим и эквивалентная масса или заряд, равные  $n \pm \sqrt{n}$ , являются непрерывными, то приближение будет достаточно хорошим. Плотность в каждой точке можно отождествить с отношением общей массы в такой области к ее объему; подобным же образом определяется плотность заряда; обе плотности будут всюду конечными.

Физическая трудность состоит в том, что твердое вещество, целиком состоящее из неподвижных частиц, связанных только такими силами, величина которых обратно пропорциональна квадрату расстояния, не было бы устойчивым и коллапсировало бы в нулевой объем. Частично ее можно преодолеть, допустив существование дополнительных отталкивающих сил, заметных только на малых расстояниях и убывающих с расстоянием быстрее, чем  $r^{-2}$ , или предположив, что частицы находятся в быстром орбитальном движении. В первом случае дополнительные силы нужно изучать отдельно; этот метод применялся главным образом Борном и Ленард-Джонсом

в теориях кристаллов и газов. Во втором случае потенциал, создаваемый данным телом, даже внешне находящимся в полном покое, будет быстро меняющейся функцией времени. Квантовая теория стремится объединить оба допущения в одной гипотезе.

Различные решения показывают, что силы, действующие между частицами, подчиняются закону обратных квадратов, если расстояние значительно больше  $10^{-8}$  см, и их средние величины на интервалах времени больше чем  $\sim 10^{-17}$  сек изменяются мало. Следовательно, математические и физические трудности можно обойти одним и тем же способом: закон обратных квадратов будет давать правильные результаты, если применять его только к изменениям среднего положения или импульса вещества внутри областей с линейными размерами, большими  $10^{-8}$  см, за интервал времени, больший; чем  $10^{-17}$  сек и если эти результаты правильны, то они будут приближенно теми же для непрерывных распределений, которые сохраняют ту же массу или заряд внутри таких областей.

В нашей формуле для  $\varphi$  мы заменим  $a_s$  на  $\rho d\xi_1 d\xi_2 d\xi_3$ , где  $\xi_i$  — точка соответствующего непрерывного распределения. Для координат точки, в которой определяется  $\varphi$ , мы продолжаем использовать обозначение  $x_i$ . Таким образом, приходим к изучению функции

$$\varphi = \gamma \iiint \frac{\rho}{R} d\xi_1 d\xi_2 d\xi_3, \quad (1)$$

где  $\rho$  теперь будет функцией координат  $\xi_1, \xi_2, \xi_3$  точки  $Q$ ;  $x_1, x_2, x_3$  — координаты точки  $P$  и  $R$  — расстояние  $QP$ , определяемое по формуле

$$R^2 = (x_i - \xi_i)^2. \quad (2)$$

Константа  $\gamma$  различна для разных физических полей. Для краткости мы будем писать  $d\xi_1 d\xi_2 d\xi_3 = d\tau$ ; таким образом,  $d\tau$  есть элемент объема.

Предположим, как и раньше, что  $\rho = 0$  в точках, находящихся от начала координат на расстоянии, большем чем данное. Тогда снова для больших  $r$  будем иметь

$$r\varphi \rightarrow \gamma \iiint \rho d\tau. \quad (3)$$

Если  $\rho = 0$  в некоторой области, то в ней  $\nabla^2 \varphi = 0$ . Действительно, дифференцируя (1) под знаком интеграла, получаем

$$\frac{\partial^2 \varphi}{\partial x_1^2} = \gamma \iiint \frac{\rho}{R^5} [3(x_1 - \xi_1)^2 - R^2] d\xi_1 d\xi_2 d\xi_3. \quad (4)$$

При достаточно малых  $R$  точка  $Q$  окажется в области, где  $\rho = 0$ . Поэтому подынтегральное выражение является непре-

рывной функцией  $x_1, x_2, x_3$  и  $R$  не обращается в нуль, а пределы интегрирования конечны. Следовательно, дифференцирование под интегралом обосновано. Сложив это и аналогичные выражения для других вторых производных, получим, как и раньше,  $\nabla^2\varphi = 0$ .

Промежуточными между случаем дискретных частиц и непрерывными распределениями являются поверхностные и линейные распределения, при которых масса или заряд на единицу площади поверхности или единицу длины линии соответственно — конечные величины. Отсюда очевидно, что потенциал удовлетворяет уравнению Лапласа всюду, кроме, может быть, такой поверхности или линии.

**6.03. Однородная сферическая оболочка.** Прежде чем развивать дальше общую теорию, мы рассмотрим несколько важных специальных случаев.

Пусть на сфере радиуса  $a$  задано распределение с постоянной поверхностной плотностью  $\sigma$ . Очевидно, что потенциал  $\varphi$  не зависит от направления и является поэтому функцией одного только  $r$ ; потенциал  $\varphi$  должен удовлетворять уравнению Лапласа везде, кроме сферической оболочки, где это уравнение не удовлетворяется. Если теперь обозначить штрихом дифференцирование по  $r$  и если  $\varphi$  — функция только  $r$ , то

$$\frac{\partial\varphi}{\partial x_i} = \frac{x_i}{r} \varphi', \quad \frac{\partial^2\varphi}{\partial x_i \partial x_k} = \frac{\delta_{ik}}{r} \varphi' - \frac{x_i x_k}{r^3} \varphi' + \frac{x_i x_k}{r^2} \varphi'', \quad (1)$$

$$\nabla^2\varphi = \frac{3}{r} \varphi' - \frac{\varphi'}{r} + \varphi'' = \varphi'' + \frac{2\varphi'}{r} = \frac{d^2}{dr^2} (r\varphi). \quad (2)$$

Поскольку  $\nabla^2\varphi = 0$  всюду, кроме сферической поверхности, то

$$\varphi = A + \frac{B}{r}, \quad (3)$$

где  $A$  и  $B$  — константы. Но при большом  $r$

$$r\varphi \rightarrow \gamma \int \int \sigma dS = 4\pi\gamma a^2\sigma. \quad (4)$$

Поэтому вне сферы  $A = 0$  и  $B = 4\pi\gamma a^2\sigma$  и

$$\varphi = \frac{4\pi\gamma a^2\sigma}{r} \quad (r > a). \quad (5) *$$

---

\*) Весьма примечательна история этой формулы. Есть основания предполагать, что Ньютон задержал на двадцать лет публикацию работы по сравнению ускорения силы тяжести от Луны с ускорением силы тяжести на земной поверхности, так как не мог доказать, что притяжение шара во внешних точках равносильно притяжению такой же массы, сосредоточенной в его центре. Но это легко доказать из уравнения Лапласа. Почему же Ньютон не открыл уравнения Лапласа? Вероятно, потому, что его метод обозначения производной не позволял ему рассматривать более одной независимой переменной и находить частные производные.

В центре  $\varphi$ , очевидно, конечно и равно  $4\pi\gamma a\sigma$ ; следовательно,  $B=0$  внутри сферы и

$$\varphi = 4\pi\gamma a\sigma \quad (r < a). \quad (6)$$

Таким образом,  $\varphi$  имеет две разные аналитические формы внутри и вне сферы. Его производная по  $r$  равна

$$\frac{\partial\varphi}{\partial r} = \begin{cases} -\frac{4\pi\gamma a^2\sigma}{r^2} & (r > a), \\ 0 & (r < a) \end{cases} \quad (7)$$

и претерпевает разрыв  $-4\pi\gamma\sigma$ , когда  $r$ , возрастая, проходит через  $a$ . Сам потенциал  $\varphi$ , однако, непрерывен при переходе через сферическую оболочку. Это справедливо и в общем случае: потенциал непрерывен при пересечении поверхностного распределения, но производная потенциала имеет разрыв, равный  $-4\pi\gamma\sigma$ .

**6.031. Однородная сфера.** Далее рассмотрим однородную сферу радиуса  $a$  с плотностью  $\rho$ . Мы можем представить, что она сложена из сферических оболочек радиуса  $\alpha$ , толщины  $d\alpha$  и поэтому с поверхностной плотностью  $\rho d\alpha$ . Тогда в точках вне сферы

$$\varphi = \frac{4\pi\gamma\rho}{r} \int_0^a \alpha^2 d\alpha = \frac{4}{3} \frac{\pi\gamma\rho a^3}{r} \quad (r > a), \quad (8)$$

как и следовало ожидать. Для внутренних точек легче получить  $\partial\varphi/\partial r$ , так как оболочки с  $\alpha > r$  не влияют на нее. Имеем

$$\frac{\partial\varphi}{\partial r} = -\gamma \int_0^r \frac{4\pi\rho\alpha^2}{r^2} d\alpha = -\frac{4}{3} \pi\gamma\rho r \quad (r < a). \quad (9)$$

В центре

$$\varphi = 4\pi\gamma\rho \int_0^a \alpha d\alpha = 2\pi\gamma\rho a^2, \quad (10)$$

и поэтому

$$\varphi = 2\pi\gamma\rho \left( a^2 - \frac{1}{3} r^2 \right) \quad (r < a). \quad (11)$$

Следовательно,  $\varphi$  и  $\partial\varphi/\partial r$  обе непрерывны на границе сферы, так как

$$\lim_{r \rightarrow a} \frac{4}{3} \frac{\pi\gamma\rho a^3}{r} = \lim_{r \rightarrow a} 2\pi\gamma\rho \left( a^2 - \frac{1}{3} r^2 \right) = \frac{4}{3} \pi\gamma\rho a^2, \quad (12)$$

$$\lim_{r \rightarrow a} \left( -\frac{4}{3} \frac{\pi\gamma\rho a^3}{r^2} \right) = \lim_{r \rightarrow a} \left( -\frac{4}{3} \pi\gamma\rho r \right) = -\frac{4}{3} \pi\gamma\rho a. \quad (13)$$

Потенциал  $\varphi$  при  $r < a$  не удовлетворяет уравнению Лапласа

$$\begin{aligned}\frac{\partial}{\partial x_i} \left( -\frac{2}{3} \pi \gamma \rho r^2 \right) &= -\frac{4}{3} \pi \gamma \rho x_i, \\ \nabla^2 \left( -\frac{2}{3} \pi \gamma \rho r^2 \right) &= -4\pi \gamma \rho.\end{aligned}\quad (14)$$

Это — уравнение Пуассона, выражающее общие свойства потенциала внутри вещества. Из предыдущего следует, что при разрывах плотности потенциал и его первая производная могут быть непрерывны, но по крайней мере одна из вторых производных будет иметь разрыв.

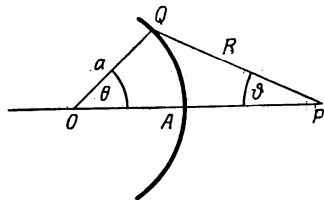


Рис. 26.

**6.032. Сферическая шапка.** Рассмотрим теперь потенциал на оси сферического сегмента, если радиус сферы равен  $a$ , а половина центрального угла сегмента равна  $\alpha$ .

Мы имеем

$$\varphi = \gamma \int \int \frac{\sigma dS}{R} = \gamma \sigma \int_0^\alpha \int_0^{2\pi} \frac{a^2 \sin \theta d\theta d\lambda}{R}, \quad (15)$$

где  $\theta$  и  $\lambda$  — сферические координаты точки  $Q$ . Но

$$R^2 = a^2 + r^2 - 2ar \cos \theta, \quad (16)$$

$$R dR = ar \sin \theta d\theta \quad (17)$$

при изменении положения  $Q$ ; следовательно,

$$\varphi = 2\pi \gamma \sigma \int_{\theta=0}^\alpha \frac{a dR}{r} = 2\pi \gamma \sigma \frac{a}{r} [(a^2 + r^2 - 2ar \cos \alpha)^{1/2} - |r - a|] \quad (18)$$

и снова  $\partial \varphi / \partial r$  делает скачок  $-4\pi \gamma \sigma$ , когда  $r$ , возрастая, проходит через  $a$ . Величина  $R$  для  $\theta = 0$ , конечно, должна быть положительной и равной  $|r - a|$  при  $r > a$  и  $r < a$ , так как она является просто расстоянием  $AP$ . При  $\alpha = \pi$ , когда сегмент становится полной сферической поверхностью, мы возвращаемся к формуле (5).

Если  $a$  становится большим при фиксированных  $r - a = x$  и  $a \sin \alpha = b$ , мы приходим к случаю кругового диска радиуса  $b$ . В этом случае

$$\varphi = 2\pi \gamma \sigma [ \sqrt{b^2 + x^2} - |x| ], \quad (19)$$

$$\frac{\partial \varphi}{\partial x} = 2\pi \gamma \sigma \left[ \frac{x}{\sqrt{b^2 + x^2}} - \frac{x}{|x|} \right]. \quad (20)$$

Производная стремится к  $2\pi\gamma\sigma$  или  $-2\pi\gamma\sigma$ , когда  $x \rightarrow 0$  со стороны отрицательных или положительных значений соответственно. Формулы (18)–(20) справедливы, конечно, только если  $P$  находится на оси симметрии. Но они показывают, как величина скачка производной по нормали сохраняет одно и то же значение, даже если величина производной по одну сторону изменяется от 0 до  $2\pi\gamma\sigma$  с изменением  $a$ .

**6.033. Линейная плотность.** В качестве примера линейной плотности рассмотрим распределение с плотностью  $\lambda$  на единицу длины на оси  $z$  от  $\zeta = -a$  до  $\zeta = +a$ . Мы имеем

$$\varphi = \gamma\lambda \int_{-a}^a \frac{d\zeta}{[x^2 + y^2 + (z - \zeta)^2]^{1/2}}. \quad (21)$$

Если положить

$$\begin{aligned} (x^2 + y^2)^{1/2} &= \omega, & z - \zeta &= \omega \operatorname{tg} \theta, & R &= [x^2 + y^2 + (z - \zeta)^2]^{1/2} = \omega \sec \theta, \\ R_1 &= [x^2 + y^2 + (z - a)^2]^{1/2} = \omega \sec \theta_1, \\ R_2 &= [x^2 + y^2 + (z + a)^2]^{1/2} = \omega \sec \theta_2, \end{aligned}$$

то это выражение станет

$$\begin{aligned} \gamma\lambda \int_{\theta_1}^{\theta_2} \sec \theta \, d\theta &= \gamma\lambda \ln \left( \frac{\sec \theta_2 + \operatorname{tg} \theta_2}{\sec \theta_1 + \operatorname{tg} \theta_1} \right) = \gamma\lambda \ln \left( \frac{R_2 + z + a}{R_1 + z - a} \right) = \\ &= \gamma\lambda \ln \frac{R_1 + R_2 + 2a}{R_1 + R_2 - 2a}, \end{aligned} \quad (22)$$

так как

$$4az = R_2^2 - R_1^2.$$

Если  $R_1 + R_2 \rightarrow 2a$ , то потенциал стремится к  $+\infty$  логарифмически; но если  $R_1$  и  $R_2$  оба велики, то предел конечен. Если  $a$  очень большое, а  $x, y, z$  невелики, то

$$\varphi \approx -2\gamma\lambda \ln \frac{\omega}{2a}. \quad (23)$$

Эта функция — решение уравнения Лапласа, так как

$$\begin{aligned} \frac{\partial}{\partial x} \ln(x^2 + y^2) &= \frac{2x}{x^2 + y^2}, \\ \frac{\partial^2}{\partial x^2} \ln(x^2 + y^2) &= \frac{2(x^2 + y^2) - 4x^2}{(x^2 + y^2)^2} = \frac{2(y^2 - x^2)}{(x^2 + y^2)^2}, \\ \frac{\partial^2}{\partial y^2} \ln(x^2 + y^2) &= \frac{2(x^2 - y^2)}{(x^2 + y^2)^2}, \quad \frac{\partial^2}{\partial z^2} \ln(x^2 + y^2) = 0. \end{aligned} \quad (24)$$

Таким образом,  $\ln \omega/2a$  является двухмерным решением уравнения Лапласа. Градиент этой функции не зависит от кон-



станты  $a$ , которая нужна только потому, что  $\ln \omega$ , строго говоря, не имеет смысла, так как  $\omega$  — длина, а не число. Мы получили простейший возможный пример другого общего результата: вблизи линии с плотностью  $\lambda$  потенциал стремится к бесконечности как  $-2\gamma\lambda \ln(\omega/b)$ , где  $\omega$  — кратчайшее расстояние до линии и  $b$  — некоторая фиксированная длина, несмотря на то что линия может быть и кривой.

**6.034.** С этим потенциалом тесно связан, как видно из теории комплексной переменной, двухмерный потенциал

$$\varphi = \operatorname{arctg} \frac{y}{x}. \quad (25)$$

Здесь

$$\begin{aligned} \frac{\partial \varphi}{\partial x} &= -\frac{y}{x^2} \frac{1}{1+(y^2/x^2)} = -\frac{y}{x^2+y^2}, & \frac{\partial^2 \varphi}{\partial x^2} &= \frac{2xy}{(x^2+y^2)^2}, \\ \frac{\partial \varphi}{\partial y} &= \frac{1}{x} \frac{1}{1+(y^2/x^2)} = \frac{x}{x^2+y^2}, & \frac{\partial^2 \varphi}{\partial y^2} &= -\frac{2xy}{(x^2+y^2)^2}, \end{aligned}$$

и

$$\nabla^2 \varphi = 0. \quad (26)$$

Величина  $\operatorname{grad} \varphi$  есть  $1/\omega$ , как и для  $\varphi = \ln \omega$ , но градиент направлен вдоль окружности  $\omega = \text{const}$ , в то время как при  $\varphi = \ln(\omega/b)$  он направлен по радиусу. Функция  $\varphi$  от  $x$  и  $y$  не является однозначной, но ее производные однозначны.

Объемные плотности хорошо приближают реальные плотности в задачах гравитации, а поверхностные — в электростатике. Линейные плотности реже возникают в теории притяжения, но частная потенциальная функция  $\ln(\omega/b)$  имеет большое значение для двухмерной гидродинамики и для изучения электрического тока в плоских пластинах. Функция  $\operatorname{arctg}(y/x)$  возникает при исследовании вихревых движений в гидродинамике и магнитном поле, создаваемом электрическим током.

**6.035. Диполь.** Большое значение для теории представляет еще один тип потенциала — диполь, или дублет. Если

$$\varphi = \gamma \left( \frac{A}{R_1} - \frac{A}{R_2} \right), \quad (27)$$

где

$$R_1^2 = (x-a)^2 + y^2 + z^2, \quad R_2^2 = (x+a)^2 + y^2 + z^2, \quad (28)$$

и мы устремим  $a \rightarrow 0$ ,  $A \rightarrow \infty$  так, чтобы  $2aA \rightarrow \mu$ , то

$$\varphi \rightarrow -\gamma\mu \frac{\partial}{\partial x} \frac{1}{r} = \gamma\mu \frac{x}{r^3}. \quad (29)$$

Это — решение уравнения Лапласа вне  $r=0$ , так как

$$\nabla^2 \varphi = -\gamma \mu \nabla^2 \frac{1}{r} = -\gamma \mu \frac{\partial}{\partial x} \nabla^2 \frac{1}{r} = 0. \quad (30)$$

Функция  $\varphi$  называется потенциалом диполя с моментом  $\mu$ ; диполь расположен в начале координат и направлен по оси  $x$ . Для диполя с моментом  $\mu$  и с направлением  $l_i$ , находящегося в точке  $\xi_i$ , будет

$$\varphi = -\gamma \mu l_i \frac{\partial}{\partial x_i} \frac{1}{R} = \gamma \mu l_i \frac{\partial}{\partial \xi_i} \frac{1}{R}, \quad (31)$$

где

$$R^2 = (x_i - \xi_i)^2. \quad (32)$$

Направление  $l_i$  называется *осью* диполя.

Поле, создаваемое диполем, хорошо отображает поле маленького намагниченного бруска и встречается во многих физических задачах. Эквипотенциальные поверхности замкнуты, симметричны относительно оси диполя и все сходятся в диполе.

**6.036. Двойной слой** представляет собой распределение на поверхности диполей, оси которых направлены вдоль внешней нормали к поверхности. В действительности ничего похожего не существует, но свойства такого слоя используются в некоторых общих теоремах. Если  $\mu dS$  есть момент элемента площади  $dS$  и  $l_i$  — направляющие косинусы нормали, то

$$\varphi = -\gamma \iint \mu l_i \frac{\partial}{\partial x_i} \left( \frac{1}{R} \right) dS = \gamma \iint \mu l_i \frac{\partial}{\partial \xi_i} \left( \frac{1}{R} \right) dS. \quad (33)$$

Это выражение просто преобразуется. Пусть  $P$  — точка с координатами  $x_i$ , а  $Q$  — с координатами  $\xi_i$  и  $QP$  образует угол  $\chi$  с нормалью в  $Q$ . Тогда

$$l_i \frac{\partial}{\partial \xi_i} \frac{1}{R} = \frac{l_i (x_i - \xi_i)}{R^3} = \frac{\cos \chi}{R^2}. \quad (34)$$

Если провести конус от  $P$  к границе  $dS$  и окружить  $P$  сферой радиуса  $\alpha$ , то площадь, вырезаемая на этой сфере, будет

$$\alpha^2 d\omega = \alpha^2 \frac{\cos \chi}{R^2} dS. \quad (35)$$

Отношение  $d\omega$  называется элементом *телесного угла*, под которым площадь  $dS$  видна из  $P$ ; суммируя  $\alpha^2 d\omega$ , получаем площадь, вырезаемую на сфере конусом, соединяющим  $P$  с границей  $S$ . Выражение для потенциала с помощью  $d\omega$  записы-

вается в замечательно простой форме

$$\varphi = \gamma \iint \mu d\omega. \quad (36)$$

Важно обратить внимание на знак  $d\omega$ , который совпадает со знаком  $\cos \chi$ . На рис. 27  $d\omega$  положителен для  $\varphi_P$ , отрицателен для  $\varphi_{P_1}$  и  $\varphi_{P_2}$ . Очевидно, что потенциал на большом расстоянии убывает как  $r^{-2}$ , а не как  $r^{-1}$ ; если функция  $\mu$  постоянна на замкнутой поверхности, то  $\varphi = 0$  во всех внешних точках и  $\varphi = -4\pi\mu$  во всех внутренних точках. Следовательно,  $\varphi$  делает скачок  $4\pi\mu$  при переходе с внутренней стороны поверхности на наружную. При пересечении однородного плоского двойного слоя  $\varphi$  изменяется от  $-2\pi\mu$  до  $2\pi\mu$ .

Рассмотрим однородный двойной слой, заполняющий полуплоскость  $y = 0$ ,  $x < 0$ , с осями диполей, направленными в сторону положительных  $y$ . Телесный угол, образуемый в точке  $P$ , составляет часть  $\theta/2\pi$  от угла, который образуется в той же точке окружающей ее сферой, поэтому он равен  $2\theta$ , где  $\theta = \text{Arctg}(y/x)$  \*). Следовательно,

$$\varphi = 2\gamma\mu\theta = 2\gamma\mu \text{Arctg} \frac{y}{x}. \quad (37)$$

Когда  $\theta \rightarrow \pi$ ,  $\varphi \rightarrow 2\pi\mu$ ; если  $\theta \rightarrow -\pi$ ,  $\varphi \rightarrow -2\pi\mu$ ; поэтому при пересечении слоя по направлению осей диполей снова появляется разрыв величиной  $4\pi\mu$ . В этом случае  $\varphi$  не стремится к 0 при  $r \rightarrow \infty$ , но это происходит потому, что слой не удовлетворяет такому условию: все вещество находится не дальше некоторого данного расстояния от начала координат.

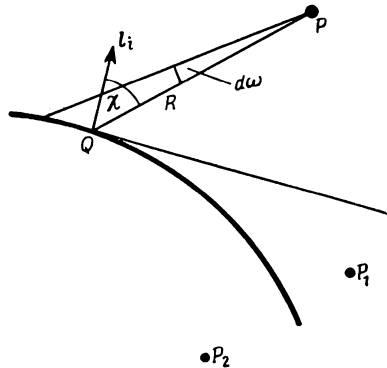


Рис. 27.

\*) Здесь

$$\text{Arctg} \frac{x}{y} = \begin{cases} \text{arctg} \frac{x}{y}, & x > 0, \\ \pi + \text{arctg} \frac{x}{y}, & x < 0, y > 0, \\ -\pi + \text{arctg} \frac{x}{y}, & x < 0, y < 0. \end{cases}$$

— Прим. перев.

Подводя итог, мы назовем обычный разрыв плотности разрывом нулевого порядка; тогда поверхностную плотность можно связать с разрывом  $-1$  порядка, а двойной слой — с разрывом  $-2$  порядка. Все они могут рассматриваться как предельные случаи непрерывных распределений. Если  $\rho = \rho_0 \operatorname{th}(x/a)$  и  $a \rightarrow 0$ , то  $\rho \rightarrow \rho_0$  для  $x > 0$  и  $\rho \rightarrow -\rho_0$  для  $x < 0$ . Если  $\rho = \frac{\sigma}{a\sqrt{\pi}} \exp\left(-\frac{x^2}{a^2}\right)$ ,

то  $\rho \rightarrow 0$  при  $a \rightarrow 0$  для любого  $x \neq 0$ , но  $\int_{-h}^h \rho dx \rightarrow \sigma$ , как бы

мало  $h$  ни было. Если  $\rho = \frac{2}{a^3\sqrt{\pi}} \mu x \exp\left(-\frac{x^2}{a^2}\right)$ , то  $\rho \rightarrow 0$  для

любого  $x \neq 0$ ,  $\int_{-h}^h \rho dx \rightarrow 0$ ,  $\int_{-h}^h \rho x dx \rightarrow \mu$ . Порядок неправиль-

ности непрерывного распределения, который обеспечивает нужные свойства в пределе, здесь возрастает: величина  $\rho_{\max} = O(1)$ ,  $O(a^{-1})$  и  $O(a^{-2})$  соответственно, где  $a$  является линейным масштабом распределения. Соответствующие порядки разрывов потенциалов, которые определяются порядком самой низкой производной, терпящей разрыв, будут 2, 1 и 0. Это правило можно было бы распространить на случай разрывов более высоких производных; в большинстве практических задач порядок разрыва потенциала на две единицы больше порядка разрыва плотности.

#### 6.04. Потенциал и поле внутри непрерывного распределения.

Мы должны теперь рассмотреть более общие случаи. Соотношение  $\nabla^2 \varphi = 0$  выполняется, если точка  $P$  расположена вне вещества. Когда  $P$  находится внутри вещества (что было бы невозможно в случае распределения частиц),  $\rho$  не обращается в нуль в  $P$  и подынтегральная функция стремится к бесконечности. Поэтому необходимо определить  $\varphi$  как несобственный интеграл, сначала исключив малую окрестность  $P$  из области интегрирования и затем взяв предел интеграла при стремлении диаметра этой малой окрестности к нулю. Для того чтобы предел существовал, мы потребуем, чтобы его значение для любой формы исключенной окрестности было бы одним и тем же.

*Л е м м а.* Интеграл  $\iiint (d\tau/R^m)$ , где  $R$  измеряется от точки  $P$  в объеме  $V$ , по которому ведется интегрирование, сходится при  $m < 3$ ; из всех областей одинакового объема интеграл достигает наибольшего значения, если область сферическая с центром в  $P$ , для  $0 < m < 3$ .

Пусть  $\tau$  — область внутри сферы радиуса  $a$ , а  $\tau'$  расположена целиком внутри  $\tau$ . Тогда

$$\begin{aligned} \int_{\tau-\tau'} \int \frac{1}{R^m} d\tau &= \int_{\tau-\tau'} \int \int R^{2-m} d\omega dR \leq \\ &\leq 4\pi \int_{\tau} R^{2-m} dR = \frac{4\pi}{3-m} a^{3-m} \quad (m < 3). \end{aligned} \quad (1)$$

При  $m < 3$  для любого положительного  $\varepsilon$  мы можем взять  $a$  таким, чтобы  $[4\pi/(3-m)] a^{3-m} < \varepsilon$ ; тогда существует сфера радиуса  $a$ , такая, что

$$\int_{\tau-\tau'} \int \frac{1}{R^m} d\tau < \varepsilon \quad (2)$$

для любой области  $\tau'$  внутри  $\tau$ . Следовательно, если исключить полость вокруг  $P$  из объема интегрирования, то интеграл стремится к единственному пределу, не зависящему от формы полости, когда ее размеры стремятся к нулю.

Предположим теперь, что объем интегрирования  $V$  заменен сферой с тем же объемом с центром  $P$  и радиусом  $b$ . Тогда

$$\int_{\text{Сфера}} \int \frac{1}{R^m} d\tau - \int_V \int \frac{1}{R^m} d\tau = \int_{\substack{\text{Часть сферы} \\ \text{вне } V}} \int \frac{1}{R^m} d\tau - \int_{\substack{\text{Часть } V \text{ вне} \\ \text{сферы}}} \int \frac{1}{R^m} d\tau. \quad (3)$$

В правой части объема, по которым производится интегрирование, одинаковы по величине в обоих интегралах: в первом  $R < b$ , а во втором  $R > b$ , кроме части границы, где  $R = b$ . Следовательно, правая часть положительна.

Всюду дальше мы будем предполагать, что функция  $\rho$  интегрируема, из чего следует, что она ограничена. Нам нужно рассмотреть два интеграла, выражающих потенциал и напряженность поля  $X_i$ :

$$\Phi = \gamma \int \int \int \frac{\rho}{R} d\tau, \quad (4)$$

$$X_i = -\gamma \int \int \int \frac{\rho}{R^3} (x_i - \xi_i) d\tau. \quad (5)$$

Оба эти интеграла несобственные. Рассмотрим сферическую область  $\tau$  около  $P$  и полость  $\tau'$  любой формы, расположенную внутри  $\tau$  и окружающую  $P$ . Рассмотрев вклад от области  $\tau - \tau'$

и обозначив верхнюю грань  $|\rho|$  в сфере через  $\rho_m$ , мы получим

$$\left| \iiint_{\tau-\tau'} \frac{\rho \, d\tau}{R} \right| \leq \rho_m \iiint_{\tau-\tau'} \frac{d\tau}{R} \quad (6)$$

и

$$\left| \iiint_{\tau-\tau'} \frac{\rho}{R^3} (x_i - \xi_i) \, d\tau \right| \leq \rho_m \iiint_{\tau-\tau'} \frac{d\tau}{R^2}. \quad (7)$$

Из леммы следует, что интегралы справа можно сделать сколь угодно малыми, и интегралы, выражающие  $\Phi$  и  $X_i$ , сходятся к пределам, которые не зависят от формы полости.

$X_i$  получается формально из  $\Phi$  дифференцированием под знаком интеграла, но для того, чтобы показать, что  $X_i$  действительно равно  $\partial\Phi/\partial x_i$ , мы должны доказать не только существование интегралов, но также и то, что, если  $P'(x_i + hl_i)$  — точка в окрестности  $P$ , тогда для любого данного положительного  $\epsilon$  можно найти  $h_0$ , такое, что для всех  $0 < h < h_0$

$$\left| \frac{\Phi(P') - \Phi(P)}{h} - l_i X_i \right| < \epsilon. \quad (8)$$

Разделим объем интегрирования на а) объем внутри сферы  $\tau_1$  с центром  $P$  и радиусом  $a$  и б) остальную область  $\tau_0$ . Обозначим вклады в интегралы от этих двух частей индексами 1 и 0 соответственно. Пусть  $P'$  — любая не совпадающая с  $P$  точка внутри  $\tau_1$ . Мы покажем, что: а) для любого заданного положительного  $\epsilon$  вклад от  $\tau_1$  в левую часть (8) можно сделать меньше чем  $\frac{1}{2}\epsilon$  для любой  $P'$ , выбрав  $a$  достаточно малым, и б) при фиксированном  $a$  можно найти  $h_0$ , такое, что вклад от  $\tau_0$  также меньше  $\frac{1}{2}\epsilon$  при всех  $h < h_0 < a$ .

а) Обозначим расстояние до  $P'$  через  $R'$ . Тогда

$$\Phi_1(P') - \Phi_1(P) = \gamma \iiint_{\tau_1} \rho \left( \frac{1}{R'} - \frac{1}{R} \right) d\tau = \gamma \iiint_{\tau_1} \rho \frac{R - R'}{RR'} d\tau. \quad (9)$$

Имеем

$$|R - R'| \leq h < a, \quad (10)$$

$$\frac{1}{RR'} \leq \frac{1}{2} \left( \frac{1}{R^2} + \frac{1}{R'^2} \right). \quad (11)$$

Следовательно,

$$\left| \frac{\Phi_1(P') - \Phi_1(P)}{h} \right| \leq \frac{\gamma \rho_m}{2} \iiint_{\tau_1} \left( \frac{1}{R^2} + \frac{1}{R'^2} \right) d\tau, \quad (12)$$

Но по второй части леммы, так как  $\int \int \int_{\tau_1} \frac{1}{R'^2} d\tau$  берется по объему шара с центром не в начале координат, он  $\leq \int \int \int_{\tau_1} \frac{d\tau}{R^2}$ , который берется по шару того же объема с центром в начале координат. Следовательно,

$$\left| \frac{\Phi_1(P') - \Phi_1(P)}{h} \right| \leq \gamma \rho_m \int \int \int \frac{d\tau}{R^2} = 4\pi \gamma \rho_m a. \quad (13)$$

Также поскольку

$$|l_i(x_i - \xi_i)| \leq R, \quad (14)$$

то

$$|l_i X_{i1}| \leq \gamma \rho_m \int \int \int \frac{d\tau}{R^2} = 4\pi \gamma \rho_m a. \quad (15)$$

Следовательно,

$$\left| \frac{\Phi_1(P') - \Phi_1(P)}{h} - l_i X_{i1} \right| \leq 8\pi \gamma \rho_m a \quad (16)$$

для всех  $P'$  в  $\tau_1$ . По любому данному положительному  $\varepsilon$  мы выберем  $a$  так, чтобы

$$8\pi \gamma \rho_m a < \frac{1}{2} \varepsilon. \quad (17)$$

б) Так как обе точки  $P$  и  $P'$  расположены вне  $\tau_0$ , то для любого заданного значения  $a$

$$\left| \frac{\Phi_0(P') - \Phi_0(P)}{h} - l_i X_{i0} \right|$$

можно сделать меньше любого данного положительного числа  $\frac{1}{2} \varepsilon$ , выбрав  $h < h_0 < a$ , причем  $h_0$  может зависеть от  $a$ .

Мы получили, что

$$\left| \frac{\Phi(P') - \Phi(P)}{h} - l_i X_i \right| \leq \left| \frac{\Phi_1(P') - \Phi_1(P)}{h} - l_i X_{i1} \right| + \left| \frac{\Phi_0(P') - \Phi_0(P)}{h} - l_i X_{i0} \right| < \varepsilon \quad (18)$$

для всех  $h < h_0$ .

Так как  $\varepsilon$  произвольно мало, то

$$\Phi(P') = \Phi(P) + l_i h X_i + \omega h, \quad (19)$$

где  $\omega \rightarrow 0$  вместе с  $h$  равномерно по  $l_i$ ; следовательно,  $\Phi$  — дифференцируемая функция, имеет производные  $X_i$  и

$$X_i = \frac{\partial \Phi}{\partial x_i} = \gamma \int \int \int \rho \frac{\partial}{\partial x_i} \left( \frac{1}{R} \right) d\tau. \quad (20)$$

Применяя лемму Грина к области между малой сферой  $\Sigma$  около  $P$  и поверхностью  $S$  всего тела, имеем

$$\left(\frac{\partial \Phi}{\partial x_i}\right)_0 = -\gamma \int \int \int_{\tau_0} \rho \frac{\partial}{\partial \xi_i} \left(\frac{1}{R}\right) d\tau. \quad (21)$$

Применим лемму Грина к этому интегралу; условия, при которых доказывалась лемма, будут выполнены, если  $\rho$  имеет интегрируемые производные. В поверхностных интегралах мы выберем направление по нормали от точки  $P$ ; тогда

$$\left(\frac{\partial \Phi}{\partial x_i}\right)_0 = \gamma \int \int \int_{\tau_0} \frac{\partial \rho}{\partial \xi_i} \frac{d\tau}{R} - \gamma \int \int_S l_i \rho \frac{dS}{R} + \gamma \int \int_{\Sigma} l_i \rho \frac{dS}{R}. \quad (22)$$

Когда радиус  $\Sigma$  стремится к 0, последний интеграл также стремится к 0, поэтому

$$\frac{\partial \Phi}{\partial x_i} = -\gamma \int \int_S l_i \rho \frac{dS}{R} + \gamma \int \int \int \frac{\partial \rho}{\partial \xi_i} \frac{d\tau}{R}, \quad (23)$$

причем последний интеграл несобственный и берется по области, лежащей внутри  $S$ . Следовательно,  $X_i$  является суммой двух потенциальных функций, одна из которых образована распределением на поверхности  $S$  с поверхностной плотностью  $-l_i \rho$ , а другая — распределением с плотностью  $\partial \rho / \partial \xi_i$  во внутренней по отношению к  $S$  области. Так как предполагалось, что  $\rho$  имеет интегрируемые производные, то  $\partial \rho / \partial \xi_i$  интегрируема и ограничена в этой области. Следовательно,  $\partial \Phi / \partial x_i$  в свою очередь дифференцируема в этой области.

Заметим, что в рассуждениях до формулы (20) предполагалось только, что  $\rho$  интегрируема, потенциал  $\Phi$  существовал и имел первые производные в местах конечных разрывов  $\rho$ , например на свободной поверхности. Принятое затем условие интегрируемости производных  $\rho$  использовалось в преобразованиях, ведущих к соотношению (23).

Заметим также, что если в теле сделать реальную полость и оставить неизменным распределение  $\rho$  вне полости, то поле в ней будет  $X_{i0}$  и поэтому произвольно близко к  $X_i$ , если полость достаточно мала.

**6.041. Теорема Гаусса.** Рассмотрим поток  $\partial \Phi / \partial x_i$  через замкнутую поверхность  $S$ . Возьмем точку  $x_i$  на  $S$ ;  $\rho$  есть функция только  $\xi_i$  и не зависит от  $x_i$ ; следовательно,

$$\begin{aligned} \int \int l_i \frac{\partial \Phi}{\partial x_i} dS &= \gamma \int \int l_i dS \int \int \int \rho \frac{\partial}{\partial x_i} \left(\frac{1}{R}\right) d\xi_1 d\xi_2 d\xi_3 = \\ &= \gamma \int \int \int \rho d\xi_1 d\xi_2 d\xi_3 \int \int l_i \frac{\partial}{\partial x_i} \left(\frac{1}{R}\right) dS, \end{aligned} \quad (24)$$



так как перестановка порядка интегрирования возможна, если  $\rho$  непрерывна и  $S$  имеет конечную площадь (в действительности и в более широких условиях). Теперь при интегрировании по  $S$  координаты  $\xi_i$  считаются константами. Если точка  $Q$  с координатами  $\xi_i$  лежит вне  $S$ , то по лемме Грина

$$\int \int l_i \frac{\partial}{\partial x_i} \left( \frac{1}{R} \right) dS = \int \int \int \nabla^2 \left( \frac{1}{R} \right) dx_1 dx_2 dx_3, \quad (25)$$

взятому по внутреннему для  $S$  пространству. Если  $Q$  находится вне  $S$ , то  $\nabla^2(1/R) = 0$  для всех точек внутри  $S$  и интеграл равен нулю. Если  $Q$  внутри  $S$ , то по 6.01 (7) имеем

$$\int \int l_i \frac{\partial}{\partial x_i} \left( \frac{1}{R} \right) dS = -4\pi. \quad (26)$$

Поэтому

$$\int \int l_i \frac{\partial \Phi}{\partial x_i} dS = -4\pi \gamma \int \int \int \rho d\tau, \quad (27)$$

где интегрирование по  $\tau$  ведется по внутреннему по отношению к  $S$  объему. Таким образом, теорема Гаусса распространяется на непрерывные распределения.

**6.042. Первый вывод уравнения Пуассона.** С помощью теоремы Гаусса можно легким, хотя и неудовлетворительным, способом получить уравнение Пуассона. Последний интеграл в (27) не изменится, если заменить  $\xi_i$  на  $x_i$  всюду, включая  $\rho$ ; первый интеграл равен

$$\int \int \int \nabla^2 \Phi dx_1 dx_2 dx_3. \quad (28)$$

Таким образом,

$$\int \int \int \nabla^2 \Phi dx_1 dx_2 dx_3 = -4\pi \gamma \int \int \int \rho dx_1 dx_2 dx_3. \quad (29)$$

Равенство справедливо для любой области, поэтому

$$\nabla^2 \Phi = -4\pi \gamma \rho \quad (30)$$

почти всюду. Это — уравнение Пуассона, которое модифицирует уравнение Лапласа для точек внутри вещества.

**6.043. Второй вывод уравнения Пуассона.** Мы вывели уравнение Пуассона способом, который не совсем удовлетворителен по трем причинам:

А. Применение леммы Грина для вывода (28) и (27) было бы вполне оправдано, если  $\partial \Phi / \partial x_i$  имела бы интегрируемые производные по всем координатам; мы доказали, что производные

$\partial^2\Phi/\partial x_i \partial x_k$  существуют при определенных условиях, но они могут быть не интегрируемы. Б. Возникает осложнение из-за точек  $Q$ , находящихся на  $S$ , но его легко устранить. Действительно, легко видеть, что в этом случае интеграл в (26) ограничен, хотя и не равен ни 0, ни  $-4\pi$ , а так как общий объем точек на  $S$  равен 0, они не влияют на правую часть в (27).

В. В случае одной переменной из равенства  $\int_a^x f(t) dt = \int_a^x g(t) dt$ , справедливого для всех  $x$  из некоторого интервала, вытекает только, что  $f(x) = g(x)$  почти всюду. Они могут различаться на некотором множестве значений  $x$ , которое можно заключить в интервалы, общая длина которых произвольно мала. Доказательство того, что  $f(x) = g(x)$  всюду, легко завершить, если известно, что  $f(x)$  и  $g(x)$  непрерывны, но в нашем случае не было доказано, что  $\nabla^2\Phi$  непрерывно, если непрерывна функция  $\rho$ .

Следующее рассуждение исходит из 6.04 (23). Если рассмотреть вклад в  $\partial\Phi/\partial x_i$  от части области, лежащей вне малой сферы  $\Sigma$  радиуса  $a$ , то из 6.04 (21) ясно, что  $\nabla^2\Phi_0 = 0$ . Вклад от внутренней части можно выразить с помощью формулы 6.04 (23), примененной к  $\Sigma$ :

$$\begin{aligned} \left( \frac{\partial^2\Phi}{\partial x_i \partial x_k} \right)_0 &= -\gamma \int_{\Sigma} l_i \rho \frac{\partial}{\partial x_k} \left( \frac{1}{R} \right) dS + \gamma \int_{\tau_i} \int \int \frac{\partial \rho}{\partial \xi_i} \frac{\partial}{\partial x_k} \left( \frac{1}{R} \right) d\tau = \\ &= \gamma \int_{\Sigma} l_i \rho \frac{\partial}{\partial \xi_k} \left( \frac{1}{R} \right) dS - \gamma \int \int \int \frac{\partial \rho}{\partial \xi_i} \frac{\partial}{\partial \xi_k} \left( \frac{1}{R} \right) d\tau. \end{aligned} \quad (31)$$

Второй интеграл стремится к 0 вместе с  $a$ . Для первого, положив  $i = k$  и просуммировав, получаем

$$\gamma \int_{\Sigma} \rho l_i \frac{\partial}{\partial \xi_i} \left( \frac{1}{R} \right) dS = -\gamma \int \int \frac{\rho}{R^2} dS. \quad (32)$$

Так как функция  $\rho$  непрерывна в  $P$ , интеграл стремится к  $4\pi\rho_P$  при  $a \rightarrow 0$ ; следовательно,

$$\nabla^2\Phi = -4\pi\rho_P. \quad (33)$$

Мы доказали это при (достаточном) условии, что  $\rho$  имеет интегрируемые производные. Это условие не является необходимым. Рассмотрим неоднородный шар радиуса  $a$ , в котором  $\rho$

есть функция  $r$ . Потенциал во внутренней точке будет

$$\begin{aligned}\varphi &= \frac{4\pi\gamma}{r} \int_0^r \rho \alpha^2 d\alpha + 4\pi\gamma \int_r^a \rho \alpha d\alpha, \\ \frac{\partial \varphi}{\partial x_i} &= \frac{x_i}{r} \frac{d\varphi}{dr} = \frac{4\pi\gamma x_i}{r} \left[ -\frac{1}{r^2} \int_0^r \rho \alpha^2 d\alpha + \frac{1}{r} \rho r^2 - \rho r \right] = \\ &= -\frac{4\pi\gamma x_i}{r^3} \int_0^r \rho \alpha^2 d\alpha, \\ \frac{\partial^2 \varphi}{\partial x_i \partial x_k} &= -4\pi\gamma \left[ \left( \frac{\delta_{ik}}{r^3} - \frac{3x_i x_k}{r^5} \right) \int_0^r \rho \alpha^2 d\alpha + \frac{x_i}{r^3} \frac{x_k}{r} \rho r^2 \right], \\ \nabla^2 \varphi &= -4\pi\gamma \rho.\end{aligned}$$

Здесь не накладывались никакие ограничения на функцию  $\rho$ , кроме непрерывности ее в рассматриваемой точке и интегрируемости на  $0 \leq r \leq a$  (см. приложение 6.043а).

**6.05. Поверхностные распределения. Поверхностная плотность.** Предположим, что на поверхности  $S$  сконцентрировано вещество с плотностью  $\sigma$  на единицу площади; тогда потенциал в  $P$  выражается так:

$$\varphi = \gamma \iint_S \frac{\sigma}{R} dS. \quad (1)$$

Если  $P$  находится на поверхности, этот интеграл надо рассматривать как несобственный, исключив внутреннюю по отношению к кривой  $C$  часть поверхности у точки  $P$  и затем стягивая  $C$  к  $P$ . Очевидно, что уравнение Лапласа удовлетворяется во всех точках, не принадлежащих  $S$ .

Возьмем точку  $O$  на  $S$  и выберем ось  $x_3$  так, чтобы в каждой точке на  $S$ , находящейся от  $O$  на расстоянии, не большем некоторого заданного, кроме, возможно, самой точки  $O$ , существовала нормаль к  $S$ , образующая с осью  $x_3$  угол, отличающийся от прямого больше, чем на некоторую фиксированную величину.

Если  $S$  — гладкая поверхность, мы можем выбрать ось  $x_3$ , совпадающей с нормалью в точке  $O$ . Если  $S$  образует в  $O$  коническую вершину (но не точку возврата), мы можем выбрать ось  $x_3$  внутри конического угла. Теперь мы выберем кривую  $C$  около  $O$  так, чтобы  $\sigma/l_3$  было ограничено внутри  $C$ ; здесь  $l_3$  — третий направляющий косинус нормали.

Если  $P$  — точка с координатами  $(0, 0, x_3)$  и если взять замкнутую кривую  $c$  на  $S$  так, чтобы  $c$  целиком лежала внутри  $C$ , то вклад в  $\varphi$  от части поверхности между  $c$  и  $C$  будет

$$\gamma \iint \frac{\sigma}{l_3} \frac{d\xi_1 d\xi_2}{[\xi_1^2 + \xi_2^2 + (\xi_3 - x_3)^2]^{1/2}}, \quad (2)$$

где интегрирование ведется по поверхности между  $c$  и  $C$ . Выберем цилиндрические координаты  $\omega, \lambda$ , тогда из

$$\left| \frac{\sigma}{l_3} \right| < K, \quad (3)$$

$$\xi_1^2 + \xi_2^2 + (\xi_3 - x_3)^2 \geq \omega^2 \quad (4)$$

следует, что (2) по модулю

$$\leq \gamma K \iint d\omega d\lambda \leq 2\pi\gamma a K, \quad (5)$$

где  $a$  — наибольшее значение  $\omega$  на  $C$ . Оно произвольно мало, поэтому и при  $x_3 = 0$  несобственный интеграл имеет одно и то же значение независимо от предельной формы  $C$ . Следовательно,  $\varphi$  является некоторой определенной величиной и в том случае, когда  $P$  лежит на  $S$ .

Выберем  $C$  так, чтобы  $4\pi\gamma a K < 1/2 \varepsilon$ . Затем возьмем  $x_3$  столь малым, чтобы вклады в  $\varphi$  в точках  $P$  и  $O$ , скажем  $\varphi$  и  $\varphi_0$ , от части  $S$ , находящейся вне  $C$ , различались бы менее чем на  $1/2 \varepsilon$ ; тогда

$$|\varphi - \varphi_0| < 4\pi\gamma a K + \frac{1}{2} \varepsilon < \varepsilon. \quad (6)$$

Следовательно,  $\varphi$  непрерывно приближается к  $\varphi_0$ , когда  $P$  стремится к  $O$  с любой стороны.

Потенциал  $\varphi$  непрерывен также и на  $S$ . Действительно, если  $P$  находится на  $S$  внутри  $C$  на расстоянии  $r$  от  $O$ , то наибольшее расстояние от  $P$  до любой точки на  $C$  будет  $\leq a + r \leq 2a$ . Следовательно, вклад в  $|\varphi - \varphi_0|$  от части поверхности, внутренней к  $C$ , будет  $\leq 6\pi\gamma a K$ , как следует из (5). Выберем  $a$  таким, чтобы этот вклад был  $\leq 1/2 \varepsilon$ , а  $r$  возьмем столь малым, чтобы вклады от части вне  $C$  различались менее чем на  $1/2 \varepsilon$ .

Тогда  $|\varphi - \varphi_0| < \varepsilon$ , следовательно, можно найти на  $S$  около  $O$  такую область, чтобы  $|\varphi - \varphi_0| < \varepsilon$  во всех точках на ней. Поэтому потенциал  $\varphi$  непрерывен всюду.

Если  $O$  лежит на краю  $S$  или же в  $O$  пересекается конечное число краевых линий области  $S$ , то этот же вывод получается, если рассмотреть каждую сторону поверхности по отдельности и просуммировать результаты.

Непрерывность справедлива и для точек возврата, если  $\sigma$  ограничена. Тем самым она установлена во всех возможных случаях для поверхностей, которые любая прямая линия пересекает лишь в конечном числе точек.

Возвращаясь к 6.04 (23), мы видим, что если объемная плотность терпит разрыв 1-го рода на некоторой поверхности,  $\partial\phi/\partial x_i$  остается непрерывной. В самом деле,  $\partial\phi/\partial x_i$  можно представить как сумму двух потенциальных функций, одна из которых образуется объемной плотностью, а другая — поверхностной. Но оба эти потенциала непрерывны, как было доказано.

Производная  $\phi$  по нормали терпит на  $S$  разрыв. Пусть  $dn$  — элемент нормали к  $S$ , направление на которой выбрано одинаковым на обеих сторонах  $S$ ; обозначим значение  $\partial/\partial n$  на стороне, где  $dn$  положительно, индексом 0, а на другой стороне — индексом 1. Мы покажем, что при соответствующих условиях

$$\left(\frac{\partial\phi}{\partial n}\right)_0 - \left(\frac{\partial\phi}{\partial n}\right)_1 = -4\pi\gamma\sigma. \quad (7)$$

Этот результат легче всего доказывается с помощью теоремы Гаусса, но рассуждения, сделанные при этом, неудовлетворительны по тем же причинам, на которые мы ссылаемся при выводе уравнения Пуассона. Лучше непосредственно изучить выражение для  $\partial\phi/\partial n$ . Мы выберем в качестве эталона для сравнения малый участок касательной плоскости около точки  $O$ . Будем считать, что начало координат находится в точке  $O$  и на касательной плоскости  $x_3 = 0$ . Предположим, что поверхность имеет конечную кривизну в точке  $O$  и вблизи нее. Проведем кривую  $C$  так, чтобы все ее точки находились на расстоянии  $a$  от оси  $x_3$ , и рассмотрим, какой вклад в  $\partial\phi/\partial x_3$  в точках на оси  $x_3$  вносится частью  $S$ , расположенной внутри  $C$ . Он равен

$$\gamma \iint \sigma \frac{\xi_3 - x_3}{R^3} dS = \gamma \iint \tau \frac{\xi_3 - x_3}{R^3} d\xi_1 d\xi_2, \quad (8)$$

где интегрирование ведется по области  $\xi_1^2 + \xi_2^2 < a^2$  и  $\tau = \sigma/l_3$ . Рассмотрим ошибку, которая получится, если считать, что  $\xi_3 = 0$ . Если производная функции существует, то

$$f(\xi_3) = f(0) + \xi_3 p, \quad (9)$$

где  $p = f'(\theta \xi_3)$ ,  $0 < \theta < 1$ . Но

$$\left| \frac{\partial}{\partial \xi_3} \frac{\xi_3 - x_3}{R^3} \right| = \left| \frac{1}{R^3} - \frac{3(\xi_3 - x_3)^2}{R^5} \right| \leq \frac{2}{R^3} \quad (10)$$

и

$$\frac{\xi_3 - x_3}{R^3} = -\frac{x_3}{R_0^3} + \frac{2A\xi_3}{R_1^3}, \quad (11)$$

где

$$R_0^2 = \xi_1^2 + \xi_2^2 + x_3^2 = \omega^2 + x_3^2, \quad (12)$$

$$|A| < 1 \quad (13)$$

и  $R_1$  находится между  $R_0$  и  $R$ . Отсюда

$$\begin{aligned} \gamma \iint_C \sigma \frac{\xi_3 - x_3}{R^3} dS = & -\gamma \int_0^a \int_0^{2\pi} \frac{\tau x_3}{R_0^3} \omega d\omega d\lambda + \\ & + \gamma \int_0^a \int_0^{2\pi} 2A\tau \frac{\xi_3}{R_1^3} \omega d\omega d\lambda. \end{aligned} \quad (14)$$

Во втором интеграле справа  $R_1 \geq \omega$ , и по допущению о конечной кривизне  $\xi_3 = O(\omega^2)$ . Следовательно, этот интеграл является малой величиной того же порядка, что и  $a$ . Поэтому для любого  $\varepsilon$  мы можем так выбрать  $a$ , чтобы второй интеграл был меньше чем  $1/8 \varepsilon$ .

Кроме того,  $\omega d\omega = R_0 dR_0$ , так как  $x_3$  — константа; поэтому

$$-\gamma \int_0^a \int_0^{2\pi} \frac{\tau x_3}{R_0^3} \omega d\omega d\lambda = -\gamma \int_{|x_3|}^{\sqrt{a^2 + x_3^2}} \int_0^{2\pi} \frac{\tau x_3}{R_0^2} dR_0 d\lambda, \quad (15)$$

$\tau$  считается непрерывной. Это равно

$$-2\pi\gamma\tau_1 \left[ \frac{x_3}{|x_3|} - \frac{x_3}{\sqrt{a^2 + x_3^2}} \right], \quad (16)$$

где  $\tau_1$  лежит между верхним и нижним пределом  $\tau$  на  $C$ . Так как  $\tau$  непрерывна, можно положить  $\tau = \sigma + \eta$ , где  $\sigma$  равно  $\tau$  в точке  $O$  и  $\eta \rightarrow 0$  вместе с  $a$ . Абсолютная величина каждого из членов в скобках  $< 1$ . Следовательно, если  $a$  достаточно мало, (15) отличается от  $-2\pi\gamma\sigma \left[ \frac{x_3}{|x_3|} - \frac{x_3}{\sqrt{a^2 + x_3^2}} \right]$  меньше чем

на  $1/8 \varepsilon$ . Затем, выбрав  $\delta$  достаточно малым по сравнению с  $a$ , можно сделать второй член в скобках сколь угодно малым, скажем  $< 1/8 \varepsilon$ , для всех  $|x_3| < \delta$ . В результате (8) будет отличаться от  $-2\pi\gamma\sigma x_3/|x_3|$  меньше чем на  $3/8 \varepsilon$ . Таким образом,

это выражение изменяется скачком, величина которого произвольно близка к  $-4\pi\sigma$ , когда  $x_3$ , возрастая, проходит через 0.

Далее, так как вклад в  $\partial\phi/\partial x_3$  от части  $S$ , находящейся вне  $C$ , непрерывен, можно выбрать  $x_3$  так, чтобы его значения в  $O$  и  $P$  различались меньше чем на  $1/8\epsilon$ . Тогда для двух значений  $x_3$ , таких, что  $|x_3| < \delta$  и имеющих разные знаки, разница между величинами  $\partial\phi/\partial x_3$  по обе стороны от  $S$  отличается от  $-4\pi\sigma$  менее чем на  $\epsilon$ .

Здесь  $a$  должно выбираться *первым* так, чтобы сделать каждое из приращений (14) и (15), возникающих от вариаций  $\xi_3$  и  $\tau$  внутри  $C$ , меньшим  $1/8\epsilon$ ; затем  $x_3$  берется таким, чтобы различия из-за второго члена в (16) и из-за  $\tau - \sigma$  были меньше  $1/8\epsilon$  и вклад от части  $S$  вне  $C$  отличался меньше чем на  $1/8\epsilon$  от его значения в  $O$ . Таким образом, мы покажем, что  $\partial\phi/\partial x_3$  стремится к пределу при  $x_3 \rightarrow 0$  с каждой из сторон, причем при нашем выборе различие между производной и ее пределом будет меньше  $1/2\epsilon$ . Наконец, разница между предельными значениями по обе стороны отличается от  $-4\pi\sigma$  менее чем на  $\epsilon$ , которое произвольно мало.

Рассуждение для тангенциальной производной проводится подобным же образом. Имеем

$$\frac{\partial\phi}{\partial x_1} = \gamma \iint \sigma \frac{\xi_1 - x_1}{R^3} dS = \gamma \iint \tau \frac{\xi_1 - x_1}{R^3} d\xi_1 d\xi_2. \quad (17)$$

Единственное существенное изменение в рассуждении появляется при рассмотрении интеграла по кругу в касательной плоскости. Подынтегральное выражение в этом случае есть четная функция  $x_3$ , и поэтому интеграл должен иметь один и тот же предел, когда  $x_3 \rightarrow 0$  с любой стороны, если этот предел вообще существует. Далее,

$$\int \tau \frac{\xi_1 - x_1}{R^3} d\xi_1 = - \int_{\xi_1 = -\sqrt{a^2 - \xi_2^2}}^{\sqrt{a^2 - \xi_2^2}} \tau d\left(\frac{1}{R}\right) = - \left[ \frac{\tau}{R} \right] + \int \frac{1}{R} \frac{\partial\tau}{\partial\xi_1} d\xi_1,$$

где оба члена являются потенциальными функциями. Следовательно, если  $\tau$  дифференцируемо, то  $\partial\phi/\partial x_1$  и  $\partial\phi/\partial x_2$  существуют во всех точках оси  $OZ$  и непрерывны.

Каждый из этих результатов доказан при разных ограничениях, налагаемых на форму поверхности и на поверхностную плотность. Потенциал  $\phi$  непрерывен, если  $\sigma$  ограничена и интегрируема и ни одна прямая не пересекает  $S$  бесконечное число раз. Градиент по нормали имеет разрыв  $-4\pi\sigma$ , если  $\sigma$

непрерывна и кривизна  $S$  конечна. Производные  $\varphi$  по касательным непрерывны, если производные  $\sigma$  интегрируемы и кривизна  $S$  конечна. Некоторые из этих условий можно немного ослабить, но случаи, когда они не выполняются, а теоремы остаются правильными, очень редки.

**6.06. Двойной слой.** Мы получили раньше

$$\varphi = \gamma \iint \mu d\omega.$$

Если снова взять точку  $O$  на поверхности и окружить ее малым контуром  $C$ , то вклад в  $\varphi$  от части  $S$ , находящейся вне  $C$ , является непрерывной функцией. Вклад же от части внутри  $C$ , если  $\mu$  непрерывна, можно сделать сколь угодно близким к  $\gamma\mu_0 \iint d\omega$ . Но это выражение увеличивается на  $4\pi\gamma\mu_0$ , когда точка, в которой рассматривается потенциал, проходит через  $O$ .

Этот результат связан с разрывом  $\partial\varphi/\partial n$  в случае поверхностной плотности. Если сместить всю поверхность на  $-dn$  параллельно  $OZ$ , то изменение  $\varphi$  с точностью порядка  $dn$  равно потенциалу, создаваемому распределением диполей с плотностью моментов  $-\sigma dn$ , оси которых параллельны оси  $x_3$  и направлены для точек, близких к этой оси, почти нормально к поверхности. Эту связь можно получить формально различными способами; например, можно показать непосредственно, что разрыв  $\partial\varphi/\partial n$  из-за поверхностной плотности  $\sigma$  такой же, как разрыв  $\varphi$  из-за распределения диполей с плотностью, равной  $-\sigma$ . Однако доказательство с помощью интегрирования по частям неудовлетворительно.

**6.07. Теорема единственности для решения уравнения Лапласа.** Пусть  $\varphi$  и  $\varphi'$  — две различные функции, удовлетворяющие уравнению Лапласа внутри замкнутой поверхности  $S$  и имеющие непрерывные первую и вторую производные. Тогда их разность  $\varphi''$  также удовлетворяет этим условиям. Но по теореме Грина

$$\begin{aligned} \iint \left( \frac{\partial \varphi''}{\partial x_i} \right)^2 d\tau &= \iint \left[ \frac{\partial}{\partial x_i} \left( \varphi'' \frac{\partial \varphi''}{\partial x_i} \right) - \varphi'' \nabla^2 \varphi'' \right] d\tau = \\ &= \iint \varphi'' \frac{\partial \varphi''}{\partial n} dS. \end{aligned} \quad (1)$$

Если  $\varphi$  и  $\varphi'$  либо  $\partial\varphi/\partial n$  и  $\partial\varphi'/\partial n$  равны во всех точках  $S$ , то интеграл равен 0. Но интеграл слева не меньше 0 и может быть равен 0, только если

$$\frac{\partial \varphi''}{\partial x_i} = 0 \quad (2)$$



во всех точках внутри  $S$ . Следовательно,  $\varphi''$  — константа, которая будет равна 0, если она равна 0 на  $S$ . Поэтому 1) если функция  $\varphi$  задана на  $S$ , то она единственным образом определена внутри  $S$ , 2) если  $\partial\varphi/\partial n$  задано на  $S$ , то  $\varphi$  определена внутри  $S$  с точностью до аддитивной константы.

**6.071.** Тот же результат получается, если  $\varphi$  и  $\varphi'$  удовлетворяют уравнению Лапласа вне замкнутой поверхности  $S$ , если только эти функции достаточно быстро убывают на больших расстояниях. Возьмем сферу большого радиуса  $A$ , охватывающую  $S$ , и применим (1) к области между  $S$  и этой сферой. Теперь  $dv$  на  $S$  будет направлено внутрь  $S$ . Пусть дано, что  $\varphi$  и  $\varphi'$  стремятся к 0 на больших расстояниях и  $\partial\varphi/\partial x_i$ ,  $\partial\varphi'/\partial x_i$  порядка  $O(1/r^2)$  при больших  $r$ . Тогда  $\varphi$  и  $\varphi'$  имеют порядок  $O(1/r)$ . Интеграл по большой сфере в этом случае — величина порядка  $O(1/A)$  и исчезающе мал при  $A \rightarrow \infty$ . Поэтому имеем

$$\iiint \left( \frac{\partial \varphi''}{\partial x_i} \right)^2 d\tau = \iint \varphi'' \frac{\partial \varphi''}{\partial v} dS,$$

где слева интегрирование ведется по всему пространству вне  $S$ . Следовательно, интеграл слева равен 0, если  $\varphi = \varphi'$  или  $\partial\varphi/\partial n = \partial\varphi'/\partial n$  во всех точках  $S$ ; тогда  $\partial\varphi''/\partial x_i = 0$  вне  $S$ . В этом случае  $\varphi'' \rightarrow 0$  на бесконечности и разность должна быть равна 0 во всех точках. Таким образом, если  $\varphi$  или  $\partial\varphi/\partial n$  дана во всех точках  $S$ , удовлетворяет уравнению Лапласа вне нее и если  $\varphi \rightarrow 0$  и  $\partial\varphi/\partial x_i = O(1/r^2)$  при  $r \rightarrow \infty$ , то  $\varphi$  однозначно определена во всех точках вне  $S$ .

Эти теоремы остаются справедливыми, если на части  $S$  задана  $\varphi$ , а на остальной части поверхности даны  $\partial\varphi/\partial n$ . Действительно, и в этом случае интеграл по поверхности равен 0.

**6.072. Теоремы о минимуме.** Пусть  $\varphi$  удовлетворяет уравнению Лапласа внутри  $S$ , а  $\varphi'$  — функция, не удовлетворяющая уравнению Лапласа, но такая, что  $\varphi' = \varphi$  во всех точках  $S$ . Положим  $\varphi' - \varphi = \varphi''$ . Тогда

$$\begin{aligned} \iiint \left[ \left( \frac{\partial \varphi'}{\partial x_i} \right)^2 - \left( \frac{\partial \varphi}{\partial x_i} \right)^2 - \left( \frac{\partial \varphi''}{\partial x_i} \right)^2 \right] d\tau &= 2 \iiint \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi''}{\partial x_i} d\tau = \\ &= 2 \iiint \left[ \frac{\partial}{\partial x_i} \left( \varphi'' \frac{\partial \varphi}{\partial x_i} \right) - \varphi'' \nabla^2 \varphi \right] d\tau = \\ &= 2 \iint \varphi'' \frac{\partial \varphi}{\partial n} dS = 0. \end{aligned} \quad (1)$$

Следовательно,

$$\iiint \left( \frac{\partial \varphi'}{\partial x_i} \right)^2 d\tau \geq \iiint \left( \frac{\partial \varphi}{\partial x_i} \right)^2 d\tau. \quad (2)$$

Таким образом,  $\iiint \left( \frac{\partial \varphi}{\partial x_i} \right)^2 d\tau$  достигает минимума, величина которого зависит от граничных условий, если  $\varphi$  — решение уравнения Лапласа. Легко распространить этот результат на область вне  $S$ , если наложить те же ограничения, что и в последней теореме, на поведение  $\varphi$  и  $\varphi'$  при больших  $r$ .

**6.073.** С этим же вопросом связана следующая теорема Кельвина. Пусть  $\varphi$  удовлетворяет уравнению  $\nabla^2 \varphi = 0$  внутри  $S$ , а  $u_i$  — соленоидальный вектор, такой, что  $l_i u_i = \partial \varphi / \partial n$  на  $S$ . Положим

$$u_i = \frac{\partial \varphi}{\partial x_i} + v_i. \quad (3)$$

Тогда  $l_i v_i = 0$  на  $S$ ;  $\frac{\partial v_i}{\partial x_i} = \frac{\partial u_i}{\partial x_i} - \nabla^2 \varphi = 0$ ,

$$\begin{aligned} \iiint \left[ u_i^2 - \left( \frac{\partial \varphi}{\partial x_i} \right)^2 - v_i^2 \right] d\tau &= 2 \iiint v_i \frac{\partial \varphi}{\partial x_i} d\tau = \\ &= 2 \iiint \left[ \frac{\partial}{\partial x_i} (\varphi v_i) - \varphi \frac{\partial v_i}{\partial x_i} \right] d\tau = \\ &= 2 \iiint \varphi l_i v_i dS = 0, \end{aligned}$$

и поэтому

$$\iiint u_i^2 d\tau \geq \iiint \left( \frac{\partial \varphi}{\partial x_i} \right)^2 d\tau. \quad (4)$$

Таким образом,  $\iiint u_i^2 d\tau$  достигает минимума на всех соленоидальных внутри  $S$  векторах  $u_i$  с заданной на  $S$  производной по нормали, если  $u_i$  — градиент решения уравнения  $\nabla^2 \varphi = 0$ .

Эти теоремы доказывают, что при определенных условиях не может быть больше одного решения. Доказать существование этого единственного решения трудно [1].

**6.074. Теорема единственности в двухмерном случае.** Теоремы, изложенные выше, легко приложить к двумерным задачам для областей, внутренних по отношению к замкнутым кривым и цилиндрам.

Для внешних задач теоремы нужно изменить, так как в двух важных случаях вектор  $u_i$  в двух измерениях является величиной порядка  $O(1/r)$ , а не  $O(1/r^2)$  при больших  $r$ . Потенциал

цилиндра с конечным зарядом на единицу длины ведет себя на больших расстояниях как  $\ln r$ . Для двух пластин, простирающихся в бесконечность, асимптоты которых наклонены к оси  $x$  под углами  $\theta_1$  и  $\theta_2$ , а потенциалы  $\varphi$  стремятся к разным предельным значениям на больших расстояниях, можно удовлетворить условиям, положив при больших  $r$ , что  $\varphi = A + B\theta$  и является решением уравнения Лапласа. Коэффициент  $B$  должен быть равен нулю в области, внешней к замкнутой кривой, для того чтобы  $\varphi$  была однозначна, но он может быть и ненулевым, если граница такова, что полный обход вокруг начала координат невозможен при больших  $r$  без пересечения границы. В обоих случаях  $u_i$  будет  $O(1/r)$ .

Но тогда  $\iint u_i^2 dS$  расходится, так как  $dS = r dr d\theta$ . Поэтому становится бессмысленным говорить, что, изменив  $u_i$ , мы увеличим этот интеграл. То же самое можно сказать, если граница, которую мы будем теперь называть  $C$ , уходит в бесконечность вдоль двух кривых с параллельными асимптотами, так как если  $\varphi$  стремится к различным пределам на них, то градиент по нормали стремится к константе и интеграл опять расходится. Теоремы о минимуме поэтому неприменимы к двумерному случаю, если не наложить дополнительных ограничений на поведение  $u_i$  на больших расстояниях.

Если, однако, мы выберем дополнительные условия в теореме единственности так, чтобы для замкнутой кривой существовали  $a$  и  $b$ , такие, что при больших  $r$

$$\varphi - a \ln \frac{r}{b} \rightarrow 0, \quad \varphi' - a \ln \frac{r}{b} \rightarrow 0,$$

$$\frac{\partial}{\partial r} \left( \varphi - a \ln \frac{r}{b} \right) = O\left(\frac{1}{r^2}\right),$$

$$\frac{\partial}{\partial r} \left( \varphi' - a \ln \frac{r}{b} \right) = O\left(\frac{1}{r^2}\right),$$

то получим

$$\varphi'' = O\left(\frac{1}{r}\right), \quad \frac{\partial \varphi''}{\partial r} = O\left(\frac{1}{r^2}\right)$$

и для достаточно больших кругов

$$\iint \varphi'' \frac{\partial \varphi''}{\partial n} ds = O\left(\frac{1}{r^2}\right).$$

Следовательно, теорема единственности остается справедливой, если подходящим образом ограничить поведение  $\varphi$  и  $\varphi'$  на больших расстояниях.

Интеграл  $\int (d\varphi/dn) ds$  по теореме Грина, очевидно, имеет одно и то же значение для  $C$  и любой замкнутой кривой.

окружающей ее. Его величина может быть включена в исходные данные во многих задачах, так как в электростатике он непосредственно связан с общим зарядом, а в гидродинамике с полной скоростью истечения, и мы можем ограничиться вариациями, сохраняющими неизменными эти величины. Но если этот интеграл конечен и не равен нулю, то производная по радиусу должна стремиться к нулю как  $1/r$  и потенциал будет вести себя как  $\ln r$  с некоторым коэффициентом.

Если  $C$  уходит в бесконечность по двум направлениям, то можно выбрать  $\varphi$  и  $\varphi'$  так, чтобы они вели себя как  $a\theta + b$  при больших  $r$ , причем  $a$  и  $b$  нужно взять таким образом, чтобы они соответствовали предельным значениям  $\varphi$  на  $C$  по этим двум направлениям. Если положить

$$\varphi - a\theta - b = o(1), \quad \varphi' - a\theta - b = o(1),$$

$$\frac{\partial}{\partial r}(\varphi - a\theta - b) = o\left(\frac{1}{r}\right), \quad \frac{\partial}{\partial r}(\varphi' - a\theta - b) = o\left(\frac{1}{r}\right),$$

тогда

$$\int \varphi'' \frac{\partial \varphi''}{\partial n} ds \rightarrow 0$$

на большой дуге и теорема единственности справедлива. К этому случаю относится и полубесконечная пластина, где предельными значениями  $\theta$  на обеих поверхностях можно считать 0 и  $2\pi$ .

Если  $C$  уходит в бесконечность вдоль двух параллельных асимптот, на которых  $\varphi$  стремится к предельным значениям, то в качестве внешней границы можно взять кривую, нормальную к этим асимптотам. Тогда для правильности теоремы единственности достаточно дополнительно условиться, что  $d\varphi''/dn \rightarrow 0$  на любой кривой, пересекающей асимптоты под прямым углом на большом расстоянии.

**6.08. Метод Релея — Ритца.** Теоремы о минимуме являются основой полезного метода численного решения, который был применен Релеем и Ритцем и обоснован Крыловым. Обычный метод состоит в том, что решается уравнение с частными производными и затем решения комбинируются так, чтобы удовлетворить граничным условиям. Но если дифференциальное уравнение эквивалентно принципу стационарности квадратичной формы, то можно выбрать функцию  $f_0$ , удовлетворяющую граничным условиям, и множество функций  $f_1, f_2, \dots$ , ничего не добавляющих к граничным значениям; тогда

$$\varphi' = f_0 + a_1 f_1 + a_2 f_2 + \dots$$

удовлетворяет граничным условиям. Если, как в рассматриваемой задаче, правильное решение минимизирует интеграл  $\iiint \left( \frac{\partial \Phi}{\partial x_i} \right)^2 d\tau$ , то можно подставить  $\Phi'$  вместо  $\Phi$ , подсчитать значение различных интегралов и затем определить  $a_1, a_2, \dots$  так, чтобы результат был минимальным.

Если функции  $f_r$  таковы, что любая дважды дифференцируемая функция может быть через них выражена, то решение теоретически является полным и может при достаточных усилиях дать правильный численный ответ там, где формальное решение дифференциального уравнения будет слишком сложным. Обобщение, сделанное Ричардсоном и Саутвеллом, не требует явного выражения для функций  $f_r$ . Метод состоит в том, что выбирается прямоугольная сетка с достаточно малым шагом и минимизируемый интеграл выражается прямо через значения в узлах сетки с использованием конечных центральных разностей. Значение функции находится затем непосредственно с помощью последовательных приближений. Метод трудоемок, но, по замечанию Саутвелла, его всегда можно осуществить на вычислительной машине. Будет ли он более трудоемок, чем табулирование решений уравнения с частными производными, зависит от конкретных особенностей задачи.

**6.09. Эквивалентный слой Грина.** Пусть  $\Phi$  удовлетворяет уравнению Лапласа внутри замкнутой поверхности и  $R$  — расстояние между точками  $Q(\xi_i)$  и  $P(x_i)$ . Если  $P$  находится внутри  $S$ , окружим ее малой сферой  $\sigma$  и применим теорему Грина к области между  $\sigma$  и  $S$ . Если же  $P$  не внутри  $S$ , применим теорему прямо к области, внутренней по отношению к  $S$ . В обоих случаях  $\nabla^2(1/R) = 0$  в рассматриваемой области, и для точки  $P$ , расположенной внутри  $S$ , имеем

$$\begin{aligned} \iint \Phi \frac{\partial}{\partial n} \left( \frac{1}{R} \right) dS + \iint \Phi \frac{\partial}{\partial \nu} \left( \frac{1}{R} \right) d\sigma = \\ = \iint \frac{1}{R} \frac{\partial \Phi}{\partial n} dS + \iint \frac{1}{R} \frac{\partial \Phi}{\partial \nu} d\sigma, \end{aligned} \quad (1)$$

где  $d\nu$  направлено к  $P$  по внешней к области нормали и поэтому равно  $-dR$ . Для  $P$ , не находящейся внутри  $S$ , интегралы по  $\sigma$  не возникают, и мы получаем просто

$$\iint \left[ \Phi \frac{\partial}{\partial n} \left( \frac{1}{R} \right) - \frac{1}{R} \frac{\partial \Phi}{\partial n} \right] dS = 0, \quad (2)$$

где  $\partial\Phi/\partial n$  — производная по нормали, если точка стремится к  $S$  изнутри.

Если  $P$  расположена внутри  $S$  и радиус  $\sigma$  равен  $a$ , причем  $a$  произвольно мало, то

$$\int \int \Phi \frac{\partial}{\partial \nu} \frac{1}{R} d\sigma \rightarrow \int \int \frac{\Phi}{a^2} a^2 d\omega = 4\pi\Phi_P, \quad (3)$$

$$\int \int \frac{1}{R} \frac{\partial \Phi}{\partial \nu} d\sigma = O \left\{ \int \int a d\omega \right\} \rightarrow 0. \quad (4)$$

Таким образом,

$$4\pi\Phi_P = \int \int \left[ \frac{1}{R} \frac{\partial \Phi}{\partial n} - \Phi \frac{\partial}{\partial n} \left( \frac{1}{R} \right) \right] dS. \quad (5)$$

Это уравнение является трехмерным аналогом теоремы 11.13 в теории функций комплексной переменной

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(t)}{t-z} dt, \quad (6)$$

где  $f(z)$  — аналитическая функция внутри  $C$  и  $z$  находится внутри  $C$ . Оно позволяет определить  $\Phi$  во всех точках внутри поверхности, если даны значения  $\Phi$  и ее нормальной производной на поверхности. Член с  $1/R$  есть потенциал от поверхностной плотности на  $S$ , а с  $\partial(1/R)/\partial n$  — потенциал двойного слоя на  $S$ .

Но плотности этих слоев нельзя установить независимо друг от друга; действительно, из формулы (2) видно, что если  $P'$  находится вне  $S$  и  $R'$  — расстояние  $QP'$ , то справедливо равенство

$$\int \int \Phi \frac{\partial}{\partial n} \left( \frac{1}{R'} \right) dS = \int \int \frac{1}{R'} \frac{\partial \Phi}{\partial n} dS \quad (7)$$

для любого такого положения  $P'$ . Этого следовало ожидать, так как мы знаем, что если решение вообще существует, то значения либо  $\Phi$ , либо  $\frac{\partial \Phi}{\partial n}$  на границе определяют его, а значит, и друг друга, кроме, возможно, аддитивной константы. Аналогом в комплексных переменных является то, что действительная или мнимая части функции  $f(z)$  на замкнутом контуре определяют одна другую с точностью до аддитивной константы, если  $f(z)$  — аналитическая функция внутри контура.

При подходящих ограничениях на  $\Phi$  на больших расстояниях этот результат можно применить для нахождения  $\Phi$  вне поверхности, если  $\Phi$  и  $\partial\Phi/\partial n$  заданы на ней. Допустим, что  $\Phi$  удовлетворяет уравнению  $\nabla^2\Phi=0$  вне  $S$  и стремится к 0 на больших расстояниях как  $1/r$ , где  $r$  — расстояние от любой фиксированной точки внутри  $S$ . Точка  $P$  находится вне  $S$ ; возьмем большую сферу  $\Sigma$ , заключающую  $S$  и  $P$  и малую

сферу  $\sigma$  около  $P$ , как и раньше, и применим теорему к области между  $S$ ,  $\sigma$  и  $\Sigma$ . Имеем

$$4\pi\varphi_P = \int \int \left[ \frac{1}{R} \frac{\partial\varphi}{\partial\nu} - \varphi \frac{\partial}{\partial\nu} \left( \frac{1}{R} \right) \right] dS + \int \int \left[ \frac{1}{R} \frac{\partial\varphi}{\partial\nu} - \varphi \frac{\partial}{\partial\nu} \left( \frac{1}{R} \right) \right] d\Sigma, \quad (8)$$

где  $d\nu$  направлено по внешней к рассматриваемой области нормали и, следовательно, внутрь на  $S$ . На  $\Sigma$  функция  $\varphi$  и  $1/R$  порядка  $1/r$ , а  $\partial\varphi/\partial\nu$  и  $\partial(1/R)/\partial\nu$  порядка  $1/r^2$ . Следовательно, интеграл по  $\Sigma$  имеет тот же порядок, что и  $\int \int d\Sigma/r^3$ , и стремится к 0 при  $r \rightarrow \infty$ . Направив  $dn$  из области  $S$ , получим

$$4\pi\varphi_P = \int \int \left[ \varphi \frac{\partial}{\partial n} \left( \frac{1}{R} \right) - \frac{1}{R} \frac{\partial\varphi}{\partial n} \right] dS. \quad (9)$$

Сравнивая с (2), мы видим, что они совместимы, только если либо  $\varphi$ , либо  $\partial\varphi/\partial n$  терпят разрыв при пересечении  $S$ ; вычитая, получаем

$$4\pi\varphi_P = \int \int \left\{ (\varphi_0 - \varphi_1) \frac{\partial}{\partial n} \left( \frac{1}{R} \right) - \frac{1}{R} \left[ \left( \frac{\partial\varphi}{\partial n} \right)_0 - \left( \frac{\partial\varphi}{\partial n} \right)_1 \right] \right\} dS, \quad (10)$$

где индексы 0 и 1 обозначают пределы при стремлении к  $S$  извне и изнутри соответственно. Если  $\varphi$  непрерывна при переходе через  $S$ , то

$$4\pi\varphi_P = - \int \int \frac{1}{R} \left[ \left( \frac{\partial\varphi}{\partial n} \right)_0 - \left( \frac{\partial\varphi}{\partial n} \right)_1 \right] dS. \quad (11)$$

Если  $\partial\varphi/\partial n$  непрерывна, то

$$4\pi\varphi_P = \int \int (\varphi_0 - \varphi_1) \frac{\partial}{\partial n} \left( \frac{1}{R} \right) dS. \quad (12)$$

Поле можно поэтому представить с помощью распределения заряда или диполей на  $S$ . Следовательно, все теоремы, полученные с помощью интегрального представления потенциала, справедливы, если только эта функция удовлетворяет уравнению Лапласа в области.

К несчастью, редко можно найти внутреннее или внешнее поле с помощью примечения этих теорем. Если дано либо  $\varphi$  либо  $\partial\varphi/\partial n$ , то в каждом уравнении по одному из членов подынтегральной функции надо найти другой, что обычно требует полного решения задачи, в процессе которого и находится потенциал в точке  $P$ . Однако это можно сделать в важных специальных случаях сферы, круга и плоскости. Случай круга рассматривается в гл. 14.

**6.091. Решение для сферы, внешняя точка.** Пусть  $\Phi$  задана на сфере радиуса  $a$ ,  $P$  — внешняя точка, находящаяся на расстоянии  $r$  и  $P'$  — точка, получающаяся из нее при помощи инверсии относительно сферы. Пусть  $Q$  — любая точка на поверхности; положим  $PQ = R$ ,  $P'Q = R'$ . Тогда

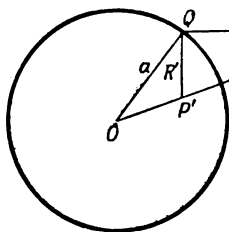


Рис. 28.

$$4\pi\Phi_P = \iint \left[ \Phi_Q \frac{\partial}{\partial n} \left( \frac{1}{R} \right) - \frac{1}{R} \left( \frac{\partial \Phi}{\partial n} \right)_Q \right] dS \quad (1)$$

$$0 = \iint \left[ \Phi_Q \frac{\partial}{\partial n} \left( \frac{1}{R'} \right) - \frac{1}{R'} \left( \frac{\partial \Phi}{\partial n} \right)_Q \right] dS, \quad (2)$$

так как  $P'$  — внутренняя точка и значение  $\partial\Phi/\partial n$  в точке  $Q$  берется по внешней нормали. Но всюду на сфере

$$\frac{R'}{R} = \frac{a}{r}. \quad (3)$$

Поэтому можно избавиться от  $(\partial\Phi/\partial n)_Q$ ; получим

$$4\pi\Phi_P = \iint \Phi_Q \left[ \frac{\partial}{\partial n} \left( \frac{1}{R} \right) - \frac{a}{r} \frac{\partial}{\partial n} \left( \frac{1}{R'} \right) \right] dS. \quad (4)$$

Обозначим угол  $POQ$  через  $\theta$ . Тогда

$$R^2 = a^2 + r^2 - 2ar \cos \theta, \quad \frac{\partial}{\partial n} \frac{1}{R} = \frac{\partial}{\partial a} \frac{1}{R} = - \frac{a - r \cos \theta}{R^3} \quad (5)$$

и подобным образом (считая  $P'$  фиксированной при дифференцировании) получаем

$$\frac{\partial}{\partial n} \frac{1}{R'} = - \frac{a - \frac{a^2}{r} \cos \theta}{R'^3}. \quad (6)$$

Следовательно, после подстановки и упрощения

$$4\pi\Phi_P = \iint \frac{\Phi}{aR^3} (r^2 - a^2) dS. \quad (7)$$

**6.092. Внутренняя точка.** Аналогично находится потенциал во внутренней точке  $P'$ , где  $OP' = r'$ :

$$4\pi\Phi_{P'} = \iint \Phi \frac{a^2 - r'^2}{aR'^3} dS.$$

Эта задача, решенная Гринном, иногда называется первой краевой задачей теории потенциала для сферы.



Вторая задача состоит в определении поля по заданному на сфере градиенту по нормали. Ее можно решить, заметив, что если  $\varphi$  — потенциальная функция, то ею же будет и  $r(\partial\varphi/\partial r)$ , значения которой заданы на поверхности. Третья задача появляется при изучении гравитационного поля, когда уровенные поверхности приблизительно сферические, но  $\partial\varphi/\partial n$  наблюдается не на сфере, а на поверхности, где  $\varphi$  постоянна. Было обнаружено, что при малых отклонениях от сферичности с точностью до первого порядка по этим отклонениям эта информация эквивалентна значениям  $\frac{\partial\varphi}{\partial r} + 2\frac{\varphi}{r}$  на сфере равного объема.

Но  $\frac{1}{r} \frac{\partial(r^2\varphi)}{\partial r}$  — потенциальная функция и она задана на сфере. Следовательно, она определена вне сферы, а соответствующие значения  $\varphi$  можно получить интегрированием по  $r$  под знаком интеграла. Задача была решена этим методом Идельсоном и Малкиным; в первоначальном решении Стокса применялись сферические гармоники (ср. 24.114).

Заметим, что если в (8) поместить  $P'$  в  $O$ , то  $r' = 0$ ,  $R' = a$  для всех  $Q$  и

$$4\pi\varphi_0 = \frac{1}{a^2} \iint \varphi dS.$$

Таким образом, *средняя величина потенциальной функции на сфере равна ее значению в центре*. Отсюда вытекает, что функция не может иметь максимум или минимум в какой-либо внутренней точке области, в которой она удовлетворяет уравнению Лапласа. Экстремальное значение обязательно достигается на границе.

**6.093. Решение для плоскости.** Взяв радиус сферы очень большим, мы получим в пределе решение соответствующей задачи Грина для плоскости. Если  $z$  — расстояние по нормали от  $P$  до плоскости, то

$$2\pi\varphi_P = \iint \frac{\varphi}{R^3} z dS. \quad (1)$$

Можно проверить, что полученное решение удовлетворяет уравнению  $\nabla^2\varphi = 0$  и стремится к нужным значениям на границе; метод [2] проверки такой же, как и в 6.091 и 6.092. Точно так же

$$2\pi\varphi_P = - \iint \left( \frac{\partial\varphi}{\partial z} \right)_Q \frac{dS}{R} \quad (2)$$

есть решение, производная которого стремится к соответствующим значениям на границе. Последний результат, конечно,

очевиден, так как поверхностная плотность, соответствующая производной  $\partial\Phi/\partial z$  на плоскости, равна  $-\frac{1}{2\pi\gamma} \frac{\partial\Phi}{\partial z}$ .

**6.10. Потенциал и поле в поляризованной среде.** Рассмотрим конечную область, полный электрический заряд которой равен нулю, а поляризация в точке  $Q$  равна  $\mathbf{P}$ ; иначе говоря, если  $d\tau$  — малый элемент объема около  $Q(\xi_i)$ , то момент диполя этого элемента будет  $\mathbf{P} d\tau$ . Рассмотрим потенциал и поле в точке  $P(x_i)$  вне этой области:

$$\begin{aligned}\Phi &= \gamma \iiint \frac{P_i(x_i - \xi_i)}{R^3} d\tau = -\gamma \iiint P_i \frac{\partial}{\partial x_i} \left( \frac{1}{R} \right) d\tau = \\ &= \gamma \iiint P_i \frac{\partial}{\partial \xi_i} \left( \frac{1}{R} \right) d\tau, \quad (1)\end{aligned}$$

$$E_i = -\frac{\partial\Phi}{\partial x_i} = -\gamma \iiint \left[ \frac{P_i}{R^3} - \frac{3P_i(x_i - \xi_i)}{R^5} \right] d\tau. \quad (2)$$

Так как

$$\iiint P_i \frac{\partial}{\partial \xi_i} \left( \frac{1}{R} \right) d\tau = \iint \frac{l_i P_i}{R} dS - \iiint \frac{1}{R} \frac{\partial P_i}{\partial \xi_i} d\tau, \quad (3)$$

то  $\Phi$  есть сумма потенциала, создаваемого зарядом с плотностью  $(-\operatorname{div} \mathbf{P})$ , распределенным в области, и потенциала заряда с поверхностной плотностью, равной  $l_i P_i$ , т. е. нормальной компоненте  $\mathbf{P}$ , распределенной по границе области.

Так же как при непрерывном распределении заряда, для определения потенциала и поля в точке  $P$ , находящейся внутри области, мы должны окружить  $P$  малой полостью и рассмотреть поведение интегралов, когда величина полости стремится к нулю \*). Однако в этом случае несобственные интегралы в (2) зависят от формы полости. Интегралы в (3) от нее не зависят. Обозначим внешнюю границу через  $S$ , границу полости через  $\Sigma$ , область внутри  $S$  через  $V$  и область внутри  $\Sigma$  через  $v$ ; отсюда область между  $S$  и  $\Sigma$  будет  $V - v$ . Тогда потенциал  $\Phi$  будет равен

$$\Phi = \gamma \iint \frac{l_i P_i}{R} dS + \lim_{v \rightarrow 0} \gamma \iint \frac{l_i P_i}{R} d\Sigma - \lim_{v \rightarrow 0} \gamma \iiint_{V-v} \frac{1}{R} \frac{\partial P_i}{\partial \xi_i} d\tau. \quad (4)$$

Если  $P_i$  ограничена, то второй интеграл стремится к нулю и

$$\Phi = \gamma \iint \frac{l_i P_i}{R} dS - \lim_{v \rightarrow 0} \gamma \iiint_{V-v} \frac{1}{R} \frac{\partial P_i}{\partial \xi_i} d\tau, \quad (5)$$

---

\*)  $\mathbf{P}$  вне полости не меняется.

т. е.  $\varphi$  состоит из потенциала, создаваемого зарядом с поверхностной плотностью  $l_i P_i$  на  $S$ , и потенциала от пространственного распределения в  $V$  с плотностью  $(-\operatorname{div} \mathbf{P})$ . Если  $\operatorname{div} \mathbf{P}$  удовлетворяет ограничениям, наложенным на  $\rho$  в 6.04, то из рассмотрения, проведенного там, известно, что (5) можно дифференцировать под знаком интеграла; тогда

$$-\frac{\partial \varphi}{\partial x_i} = \gamma \iint l_k P_k \frac{x_i - \xi_i}{R^3} dS - \lim_{v \rightarrow 0} \gamma \iiint \frac{\partial P_k}{\partial \xi_k} \frac{x_i - \xi_i}{R^3} d\tau. \quad (6)$$

Обозначим это выражение через  $E_i$  и назовем напряженностью электрического поля; она также не зависит от предельной формы полости.

Поле  $E_i$  точно так же создается зарядом с поверхностной плотностью  $l_i P_i$  на  $S$  и пространственным распределением в  $V$  с плотностью  $(-\operatorname{div} \mathbf{P})$ . Как мы увидим, оно, вообще говоря, не равно предельному значению поля в стягивающейся полости. Из уравнения Пуассона следует, что

$$\operatorname{div} \mathbf{E} = -\nabla^2 \varphi = -4\pi\gamma \operatorname{div} \mathbf{P}, \quad (7)$$

или

$$\operatorname{div} (\mathbf{E} + 4\pi\gamma \mathbf{P}) = 0. \quad (8)$$

Если кроме распределения диполей существует заряд с плотностью  $\rho$ , то в  $\varphi$  и  $\mathbf{E}$  появляются дополнительные члены и (8) надо заменить на

$$\operatorname{div} (\mathbf{E} + 4\pi\gamma \mathbf{P}) = 4\pi\gamma \rho. \quad (9)$$

Это физически осуществимо. Сделаем отверстие в куске твердого парафина, трением создадим внутри него заряд, а затем заполним отверстие снова, оставив, таким образом, заряд с некоторой плотностью внутри тела. Вектор  $\mathbf{E} + 4\pi\gamma \mathbf{P}$  называется *электрическим смещением*, или *электрической индукцией*, и обозначается через  $\mathbf{D}$ .

Рассмотрим теперь поле внутри  $v$ . Так как  $P$  — точка, внешняя для области  $V - v$ , то поле будет

$$F_i = \gamma \iint l_k P_k \frac{x_i - \xi_i}{R^3} dS + \gamma \iint l_k P_k \frac{x_i - \xi_i}{R^3} d\Sigma - \gamma \iiint_{V-v} \frac{\partial P_k}{\partial x_k} \frac{x_i - \xi_i}{R^3} d\tau, \quad (10)$$

что равно, если устремить полость к нулю,

$$E_i + \lim \gamma \iint l_k P_k \frac{x_i - \xi_i}{R^3} d\Sigma. \quad (11)$$

Возьмем полость в виде круглого цилиндра длиной  $2a$  и радиусом  $b$ , центр которого находится в  $Q$  и ось направлена вдоль  $\mathbf{P}$ ; тогда  $I_k P_k$  на основании  $A$  равно  $P$ , на основании  $B$  равно  $(-P)$  (так как рассматривается внутренняя нормаль), а на боковой поверхности равно 0, если отвлечься от влияния малых вариаций  $P$  в окрестности  $Q$ . Тогда поле в  $Q$ , создаваемое поверхностными зарядами на стенках полости, будет

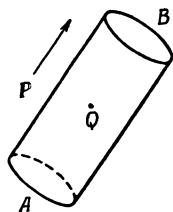


Рис. 29.

$$4\pi\gamma\mathbf{P}\left(1 - \frac{a}{(a^2 + b^2)^{1/2}}\right). \quad (12)$$

Поэтому если  $a$  и  $b$  становятся малыми, причем  $b/a \rightarrow 0$ , то этот вклад в поле стремится к 0. Если  $a/b \rightarrow 0$ , то вклад будет  $4\pi\gamma\mathbf{P}$ . Следовательно, величина  $\mathbf{E}$ , определяемая с помощью (5), представляет предел поля в иглообразной полости с осью, направленной вдоль  $\mathbf{P}$ , а  $\mathbf{D} (= \mathbf{E} + 4\pi\gamma\mathbf{P})$  — предел поля в дискообразной полости с тем же направлением оси.

В сферической полости поверхностная плотность на стенках равна  $(-P \cos \theta)$ ; она создает в центре поле  $\frac{4}{3}\pi\gamma\mathbf{P}$ .

Предположим, что существует связь

$$\mathbf{P} = \kappa\mathbf{E}, \quad (13)$$

так что

$$\mathbf{D} = \mathbf{E}(1 + 4\pi\gamma\kappa) = K\mathbf{E}, \quad (14)$$

где  $\kappa$  и также  $K$  — непрерывные скалярные функции точки. Функция  $\kappa$  называется электрической восприимчивостью, а  $K$  — диэлектрической постоянной *изотропного* вещества. Если связь между  $\mathbf{D}$  и  $\mathbf{E}$  имеет вид

$$D_i = K_{ik}E_k, \quad (15)$$

где  $K_{ik}$  — тензор, не пропорциональный  $\delta_{ik}$ , то вещество анизотропно.

Рассмотрим теперь поведение  $\mathbf{D}$  и  $\mathbf{E}$  на границе двух однородных сред с различными диэлектрическими постоянными  $K_1$  и  $K_2$ . Напряженность  $\mathbf{E}$  в любой точке создается пространственным распределением с плотностью  $\text{div } \mathbf{P}$  и поверхностным распределением с плотностью  $P_{1n} - P_{2n}$ , где  $n$  обозначает компоненту вдоль нормали, направленной из первой среды во вторую. Вследствие поверхностного распределения создается скачок нормальной компоненты  $E$ :

$$E_{2n} - E_{1n} = 4\pi\gamma(P_{1n} - P_{2n}).$$

Следовательно, нормальная компонента  $\mathbf{D}$  имеет одно и то же значение по обе стороны от границы.

Сходным образом строится соответствующая теория магнетизма. Различие состоит в том, что 1)  $\rho$  всегда равно 0, 2)  $\kappa$  и  $K$  (называемая здесь магнитной проницаемостью  $\mu$ ) могут сильно изменяться в зависимости от напряженности магнитного поля  $\mathbf{H}$ , 3) существуют постоянные магниты с постоянной намагниченностью, поэтому хотя и можно написать  $\mathbf{B} = \mathbf{H} + 4\pi\mu\mathbf{I}$ , но связи между  $\mathbf{I}$  и  $\mathbf{H}$  нет.

**6.11. Векторный потенциал.** Пусть  $\mathbf{u}$  — соленоидальный вектор. Можно показать, что существует другой вектор  $\mathbf{A}$ , такой, что  $\mathbf{u}$  является его ротором. Прежде всего удобно обозначить компоненты через  $u, v, w$  и  $A, B, C$ ; должны выполняться равенства

$$u = \frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}, \quad v = \frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}, \quad w = \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}. \quad (1)$$

Положим

$$C_1 = 0, \quad A_1 = \int_{z_0}^z v \, dz, \quad B_1 = - \int_{z_0}^z u \, dz, \quad (2)$$

где  $z_0$  не зависит от  $x, y, z$  и интегрирование ведется параллельно оси  $z$ . Таким образом, мы удовлетворим первым двум уравнениям. Но

$$\frac{\partial B_1}{\partial x} - \frac{\partial A_1}{\partial y} = - \int_{z_0}^z \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz = \int_{z_0}^z \frac{\partial w}{\partial z} dz = w - w(x, y, z_0), \quad (3)$$

так как

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (4)$$

Обычно  $w(x, y, z_0)$  не равна нулю, но она зависит только от  $x, y$  и ее можно обозначить через  $w_0$ . Пусть теперь

$$A = A_1 + A_2 \text{ и т. д.}; \quad (5)$$

тогда

$$\frac{\partial C_2}{\partial y} - \frac{\partial B_2}{\partial z} = 0, \quad \frac{\partial A_2}{\partial z} - \frac{\partial C_2}{\partial x} = 0, \quad \frac{\partial B_2}{\partial x} - \frac{\partial A_2}{\partial y} = w_0. \quad (6)$$

Положим

$$B_2 = 0, \quad C_2 = 0, \quad A_2 = - \int_{y_0}^y w_0 \, dy, \quad (7)$$

тогда будут выполняться все уравнения. Но решение все еще в значительной степени произвольно; действительно, к  $(A, B, C)$  можно было бы добавить градиент любой скалярной функции, не изменив ротора. Наоборот, можно ввести дополнительное условие, сделав дивергенцию  $(A, B, C)$  равной любой функции, в частности нулю. Для этого нужно добавить градиент такого скаляра  $\Phi$ , чтобы  $\nabla^2\Phi$  сократилась с дивергенцией уже найденного решения; при подходящих граничных условиях  $\Phi$  можно найти. Следовательно, в тензорных обозначениях мы получаем, что если

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (8)$$

то существует вектор  $A_i$ , удовлетворяющий уравнениям

$$u_i = \varepsilon_{ikm} \frac{\partial A_m}{\partial x_k}, \quad \frac{\partial A_i}{\partial x_i} = 0, \quad (9)$$

и  $A_i$  определяется единственным образом при граничных условиях того же типа, которые обеспечивают единственность решения задачи в теории потенциала.  $A$  называется *векторным потенциалом*.

**6.111.** Метод, изложенный выше, несимметричен относительно координат. Симметричное решение в том случае, когда  $\text{rot } \mathbf{u}$  задан во всем пространстве, можно найти следующим образом:

$$\varepsilon_{ikm} \frac{\partial u_m}{\partial x_k} = \omega_i, \quad \frac{\partial u_i}{\partial x_i} = 0; \quad (10)$$

предположим, что

$$u_i = \varepsilon_{ips} \frac{\partial A_s}{\partial x_p}, \quad \frac{\partial A_i}{\partial x_i} = 0. \quad (11)$$

Тогда

$$\begin{aligned} \varepsilon_{ikm} \frac{\partial u_m}{\partial x_k} &= \varepsilon_{ikm} \varepsilon_{mps} \frac{\partial^2 A_s}{\partial x_k \partial x_p} = \\ &= (\delta_{ip} \delta_{ks} - \delta_{is} \delta_{kp}) \frac{\partial^2 A_s}{\partial x_k \partial x_p} = \\ &= \frac{\partial^2 A_s}{\partial x_i \partial x_s} - \frac{\partial^2 A_i}{\partial x_k^2} = -\nabla^2 A_i. \end{aligned} \quad (12)$$

Поэтому условия удовлетворяются, если

$$A_i = \frac{1}{4\pi} \int \int \int \frac{\omega_i}{R} d\tau \quad (13)$$

во всем пространстве и если дивергенция  $A_i$  равна нулю. Применяя теорему Грина к области между малой сферой около  $x_i$  и большой сферой, получаем

$$\begin{aligned} \frac{\partial A_i}{\partial x_i} &= -\frac{1}{4\pi} \lim \int \int \int \omega_i \frac{\partial}{\partial \xi_i} \left( \frac{1}{R} \right) d\tau = \\ &= -\frac{1}{4\pi} \lim \int \int \frac{I_i \omega_i}{R} dS + \frac{1}{4\pi} \lim \int \int \int \frac{1}{R} \frac{\partial \omega_i}{\partial x_i} d\tau. \end{aligned} \quad (14)$$

Второй интеграл равен нулю, так как  $\omega_i$  — ротор вектора. Первый интеграл по внутренней сфере стремится к нулю, если  $\omega_i$  ограничена вблизи  $\xi_i = x_i$ , а по внешней сфере — если  $\omega_i$  стремится к нулю достаточно быстро на большом расстоянии. Следовательно, (13) является решением задачи. Оно не единственно, так как добавление к  $A_i$  градиента любого решения уравнения Лапласа не влияет на уравнение (11).

Этот способ не применим, если  $u_i$  — потенциальный вектор. В таком случае  $\omega_i = 0$  всюду и (13) также равно нулю. Этот способ применяется в основном, когда мы имеем нить с малым сечением, причем  $\omega_i = 0$  вне нити и велика внутри нее. Такая ситуация возникает при вихревом движении в жидкости и в магнитном поле, образованном электрическим током. В каждом из этих случаев интеграл

$$\Omega = \int_C u_i dx_i, \quad (15)$$

взятый по замкнутому контуру, равен нулю, если контур может быть краем поверхности типа шапки, не пересекающей нить или провод, и имеет одинаковое значение для всех контуров, которые могут быть краями шапок, пересекаемых нитью один раз. Рассмотрим вклад в  $A_i$  от элемента, находящегося между двумя параллельными плоскостями, разделенными расстоянием  $d\xi_i$ , и пусть  $dS$  — элемент плоскости, параллельной им; получим

$$d\tau = d\xi_i dS, \quad (16)$$

$$\int \int \omega_i dS = \int u_i d\xi_i = \Omega \quad (17)$$

по теореме Стокса; следовательно,

$$A_i = \frac{\Omega}{4\pi} \int \frac{d\xi_i}{R}, \quad (18)$$

взятому вдоль нити. Подобно этому,

$$\begin{aligned} u_l &= \varepsilon_{lkm} \frac{\Omega}{4\pi} \frac{\partial}{\partial x_k} \int \frac{1}{R} d\xi_m = \\ &= \frac{\Omega}{4\pi} \int \varepsilon_{lkm} \frac{(\xi_k - x_k) l_m}{R^3} ds, \\ \mathbf{u} &= \frac{\Omega}{4\pi} \int \frac{(\boldsymbol{\xi} - \mathbf{x}) \times d\boldsymbol{\xi}}{R^3}, \end{aligned} \quad (19)$$

где  $ds$  — элемент длины нити,  $l_m$  — направляющие косинусы касательной.

В гидродинамике  $\Omega$  — циркуляция вокруг нити. В теории электромагнетизма она равна  $4\pi J/c$ , где  $J$  — ток в электростатических единицах, а  $u_l$  — напряженность магнитного поля.

**6.112. Векторные потенциалы точечного заряда и диполя.** Возьмем диполь с осью вдоль  $z$ , так что

$$\mathbf{u} = \mu \operatorname{grad} \left( \frac{z}{r^3} \right).$$

В 6.11 (2) выберем нижний предел по  $z$  равным  $-\infty$ ; тогда

$$A_1 = -\mu \int_{-\infty}^z \frac{3uz}{r^5} dz = \frac{\mu y}{r^3},$$

$$B_1 = \mu \int_{-\infty}^z \frac{3zx}{r^5} dz = -\frac{\mu x}{r^3}$$

и из 6.11 (7)

$$A_2 = 0,$$

так как  $\omega$  стремится к 0 при  $z \rightarrow -\infty$ . Следовательно, решением будет  $\left( \frac{\mu y}{r^3}, -\frac{\mu x}{r^3}, 0 \right)$ ; дивергенция его равна нулю. Поэтому  $u_l$ , являясь градиентом потенциала  $\mu x_k/r^3$ , одновременно представляет собой ротор векторного потенциала

$$A_l = -\varepsilon_{lkm} \frac{\mu_k x_m}{r^3} = -\frac{[\mu \times \mathbf{x}]_l}{r^3}.$$

Рассмотрев подобным образом элементарный вектор  $\operatorname{grad}(1/r)$ , получим, что он является ротором вектора

$$\left[ -\frac{y}{\omega^2} \left( 1 + \frac{z}{r} \right), \quad \frac{x}{\omega^2} \left( 1 + \frac{z}{r} \right), 0 \right].$$



Часть  $(-u/\omega^2, x/\omega^2, 0)$  есть градиент скаляра  $\operatorname{arctg}(y/x)$  и ее можно отбросить; таким образом, получаем

$$\mathbf{A} = \left( -\frac{uz}{r\omega^2}, \frac{xz}{r\omega^2}, 0 \right).$$

Выражение все еще несимметрично; частичная симметрия может быть достигнута, если взять среднее этого вектора и двух других, получающихся из него циклической перестановкой координат; однако из-за этого не одна ось координат, а все три станут линиями особенностей. Обобщение при помощи вращения осей только увеличит сложность.

### ПРИМЕРЫ

1. Притяжение двух точечных масс  $m$  и  $m'$  равно  $mm'f(R)$ , где  $R$  — расстояние между ними. Показать, что направленное во внутрь ускорение, создаваемое однородной тонкой сферической оболочкой с массой  $M$  и внутренним радиусом  $a$ , действующее на внутреннюю точку на расстоянии  $x$  ( $x < a$ ) от центра, равно  $F$ , где

$$F = \frac{M}{4ax^2} \int_{a-x}^{a+x} (R^2 + x^2 - a^2) f(R) dR.$$

Если  $F=0$  для всех значений  $a$  и  $x$  при  $x < a$ , то показать, что тогда  $f(R) = A/R^2$  — единственное возможное значение ( $A = \text{const}$ ).

2. Предположим, что в любой точке тела, намагниченного до интенсивности  $I$ , магнитный потенциал такой же, как при объемном распределении магнитных полюсов плотности  $-\operatorname{div} \mathbf{I}$  с поверхностным распределением  $nI$ , где  $n$  — направляющие косинусы внешней нормали. Доказать, что а) поле внутри сферы, постоянно намагниченной с интенсивностью  $I$ , имеет однородную напряженность  $-\frac{4}{3}\pi I$ ; б) поле равно нулю в сферической полости (не обязательно концентрической) внутри этой сферы; в) если плоская пластина постоянной толщины и бесконечной протяженности постоянно намагничена с равномерной интенсивностью  $I$ , то поле в сферическом отверстии, сделанном в этой пластине, равно  $\frac{4}{3}\pi I - 4\pi(\mathbf{I} \cdot \mathbf{n})$ , где  $\mathbf{n}$  — нормаль к одной из поверхностей пластины. (М. Т., 1939.)

3. Имеем однородный куб с центром в  $O$  и массой  $M$ , а  $P$  — точка вне куба. Расстояние  $OP$  велико по сравнению с ребром куба  $a$ , направляющие косинусы  $OP$  относительно трех сходящихся граней куба равны  $l, m, n$ . Выразить гравитационный потенциал в  $P$  в виде ряда по степеням  $OP^{-1}$  и вычислить его до членов  $OP^{-5}$ . (P. elim., 1936.)

4. Если  $\Phi_1$  и  $\Phi_2$  удовлетворяют условиям а)  $\Phi$  всюду непрерывно и имеет непрерывные вторые производные всюду, за исключением некоторых поверхностей, б)  $K \frac{\partial \Phi}{\partial r}$  претерпевает разрыв определенной величины при пересечении таких поверхностей, в)  $\Phi$  стремится к нулю как  $1/r$ , а  $d\Phi/dx_i \rightarrow$

как  $1/r^2$ , когда  $r$  велико; и если, кроме того,  $\operatorname{div} [K \operatorname{grad} (\varphi_1 - \varphi_2)] = 0$ , за исключением поверхностей разрыва, то показать, что

$$\iiint K \left( \frac{\partial \varphi_1}{\partial x_l} \right)^2 d\tau - \iiint K \left( \frac{\partial \varphi_2}{\partial x_l} \right)^2 d\tau = \iiint K \left[ \frac{\partial (\varphi_1 - \varphi_2)}{\partial x_l} \right]^2 d\tau,$$

где интегралы взяты по всему объему; следовательно, при данных условиях  $\varphi$  однозначно определено, если  $K$  всюду положительно.

5. Доказать, что если  $\varphi$  удовлетворяет условиям а и б из примера 4, то условием, что

$$V = \frac{1}{8\pi} \iiint K \left( \frac{\partial \varphi}{\partial x_l} \right)^2 d\tau$$

стационарно для всех малых вариаций  $\varphi$ , будет

$$\frac{\partial}{\partial x_l} \left( K \frac{\partial \varphi}{\partial x_l} \right) = 0,$$

за исключением особых поверхностей, причем  $K \frac{\partial \varphi}{\partial n}$  непрерывно при пересечении таких поверхностей. Если  $K$  всюду положительно, то стационарное значение  $V$  осуществляет минимум. Показать также, что если  $K \frac{\partial \varphi}{\partial n}$  имеет предписанный разрыв при пересечении особых поверхностей, то  $V$  будет оставаться стационарным при условии, что  $\varphi$  не изменяется на этой поверхности.

Если  $\varphi$  — электростатический потенциал поля, содержащего диэлектрики, то дать физическую интерпретацию посылок и выводов.

6. Показать, что если  $\mathbf{a}$  — постоянный вектор, то

$$\operatorname{rot} \left( \frac{\mathbf{a} \times \mathbf{r}}{r^n} \right) = \frac{(2-n)\mathbf{a}}{r^n} + \frac{n\mathbf{r}(\mathbf{a} \cdot \mathbf{r})}{r^{n+2}},$$

и как следствие этого или каким-либо другим способом найти выражение для вектора потенциала внутри бесконечного прямого соленоида.

(М/с, III, 1936.)

7.  $\nabla^2 u_1 = 0$ ,  $\nabla^2 u_2 = 0$  в замкнутой области  $D$ ,  $u_1 = u_2$  в области  $D_1$ , которая является частью  $D$ . Доказать, что  $u_1 = u_2$  во всей области  $D$ .

8. Вывести из теоремы Стокса, что

$$\int \varphi \, ds = \int d\mathbf{S} \times \operatorname{grad} \varphi,$$

рассмотрев проекцию  $\int \varphi \, ds$  на направление касательной к контуру  $\mathbf{n}$ . Показать, что взаимный потенциал двух однородных магнитных слоев интенсивностью  $\mu$  и  $\mu'$  равен

$$-\mu\mu' \iint \frac{ds \cdot ds'}{r}.$$

## ЛИТЕРАТУРА

1. Kellogg O. D., Foundations of Potential Theory, 1929, p. 236.
2. Poincaré H., Theorie du Potential Newtonien, 1899, p. 183—191.

## ПРИЛОЖЕНИЕ К ГЛАВЕ 6

**6.043а.** Ясно, что если  $\rho$ , удовлетворяющую этим условиям, изменить в конечном числе изолированных точек, то  $\Phi$  и, следовательно,  $\nabla^2\Phi$  не изменится; но уравнение Пуассона не будет справедливо в таких точках. Интегрируемость  $\rho$  не является поэтому достаточным условием, и из приведенных рассуждений можно было бы заключить, что необходимым и достаточным условием будет непрерывность. В действительности она необходима, но недостаточна. Одно из слабых достаточных условий было дано Гельдером; оно состоит в том, что для любой точки  $Q$ , не совпадающей с  $P$ , выполняется неравенство  $|\rho_Q - \rho_P| < A r^\alpha$ , где  $A$  и  $\alpha$  фиксированы и  $\alpha$  положительно. Это условие представляет распространение условия Липшица с одного измерения на три. Одной только непрерывности недостаточно. Если плотность в сфере радиуса  $a$  равна в полярных координатах

$$\rho = -\frac{3 \cos^2 \theta - 1}{\ln(b/r)}, \quad b > a,$$

то  $\rho$  непрерывна, но имеет неограниченные производные при  $r=0$  и не удовлетворяет условию Гельдера. Можно показать, что при  $0 < r \leq a$  уравнение

$$\nabla^2\Phi = -4\pi\rho$$

имеет решение, содержащее член

$$\Phi_0 = \frac{4}{5} \pi \gamma r^2 \ln \ln \frac{b}{r} (3 \cos^2 \theta - 1),$$

но вторые производные  $\Phi_0$  не существуют при  $r=0$ . Для того чтобы теорема Пуассона была правильной для всех непрерывных  $\rho$ , предложены другие определения  $\nabla^2\Phi$ . Но и при обычном определении непрерывные распределения  $\rho$ , для которых уравнение Пуассона неправильно, встречаются очень редко.

## ОПЕРАЦИОННЫЕ МЕТОДЫ

Даже кембриджские математики заслуживают справедливости.

*Оливер Хевисайд.*

**7.1. Правила арифметики для дифференциального оператора.** В некотором смысле операции дифференцирования и взятия определенного интеграла удовлетворяют правилам арифметики. Если  $a$  и  $b$  — константы, то

$$\frac{d}{dx} f(x) + af(x) = af(x) + \frac{d}{dx} f(x), \quad (1)$$

$$\left[ \frac{d}{dx} f(x) + af(x) \right] + bf(x) = \frac{d}{dx} f(x) + [af(x) + bf(x)], \quad (2)$$

$$\frac{d}{dx} af(x) = a \frac{d}{dx} f(x), \quad (3)$$

$$\frac{d}{dx} a[bf(x)] = \frac{d}{dx} (ab) f(x), \quad (4)$$

и если мы определим оператор

$$\left( \frac{d}{dx} + a \right) f(x) = \frac{d}{dx} f(x) + af(x) = \left( a + \frac{d}{dx} \right) f(x), \quad (5)$$

что допустимо в силу (1), то мы получим следующие формы дистрибутивного закона:

$$\frac{d}{dx} (a + b) f(x) = \frac{d}{dx} af(x) + \frac{d}{dx} bf(x), \quad (6)$$

$$b \left( \frac{d}{dx} + a \right) f(x) = b \frac{d}{dx} f(x) + abf(x). \quad (7)$$

Таким образом, в любой алгебраической комбинации с константами, включающей только сложение (а следовательно, и вычитание) и умножение, с дифференциальными операторами можно обращаться так, как если бы они сами были численными константами. Функция  $f(x)$ , над которой производится операция, остается справа, и операция может быть произведена над ней в конце.

**7.011. Операция взятия определенного интеграла.** Те же самые рассуждения применимы и к операции взятия определенного интеграла. Если мы обозначим  $\int_0^t f(\xi) d\xi$  через  $Qf(t)$ , то

$$Qf(t) + af(t) = af(t) + Qf(t) \quad (1)$$

и поэтому оба выражения можно обозначить через

$$(Q + a)f(t) = (a + Q)f(t); \quad (2)$$

кроме того,

$$[Qf(t) + af(t)] + bf(t) = Qf(t) + [af(t) + bf(t)] \quad (3)$$

$$Q[af(t)] = a[Qf(t)], \quad (4)$$

$$Q[abf(t)] = Qa[bf(t)], \quad (5)$$

$$Q[(a + b)f(t)] = Qaf(t) + Qbf(t), \quad (6)$$

$$a(Q + b)f(t) = aQf(t) + abf(t). \quad (7)$$

Аналогичные соотношения можно получить, заменяя  $a$  или  $b$  на  $Q$  и интерпретируя

$$Q^2f(t) = Q[Qf(t)] = \int_0^t Qf(\tau) d\tau,$$

где по определению  $Qf(\tau)$  — та же самая функция от  $\tau$ , что и функция  $Qf(t)$  от  $t$ .

**7.012. Некоммутативное свойство дифференцирования и взятия определенного интеграла.** Указанные выше свойства дифференцирования были использованы Булем как основа известного метода получения частных интегралов для определенных типов линейных дифференциальных уравнений. Они служат также основой метода получения формул, необходимых при численной интерполяции, дифференцировании и интегрировании. Соответствующее свойство взятия определенного интеграла было отмечено, по-видимому, впервые в [1] и использовано Фуксом, Пеано, Пикаром и Бейкером для развития некоторых методов решения линейных дифференциальных уравнений. Хевисайд уделял особое внимание линейным дифференциальным уравнениям с постоянными коэффициентами, для которых из оператора интегрирования и констант можно составлять комбинации точно так же, как если бы оператор был числом. Конечно, этот оператор не коммутирует с независимыми переменными; выражение  $Q[tf(t)]$  совсем не то же самое, что

$t[Qf(t)]$ . Можно понять, почему методы Хевисайда имели успех, возвращаясь к приему, использованному Пикаром при доказательстве существования (при определенных условиях) решений дифференциальных уравнений. К сожалению, хотя Хевисайд и отмечал, что операторы дифференцирования и интегрирования можно комбинировать с константами без ограничения, он не заметил, что они не коммутируют друг с другом. В самом деле,

$$\begin{aligned}\frac{d}{dt} Qf(t) &= \frac{d}{dt} \int_0^t f(\xi) d\xi = f(t), \\ Q \frac{d}{dt} f(t) &= \int_0^t f'(\xi) d\xi = f(t) - f(0),\end{aligned}$$

что отличается от предыдущего результата на  $f(0)$ . Хевисайд получил довольно много неверных результатов, меняя порядок дифференцирования и интегрирования, и Джеффрис [2] впервые объяснил это некоммутативностью этих операций. Хевисайд также не слишком интересовался вопросами сходимости, и этот факт так возмущал математиков того времени, что они не захотели выяснить, при каких условиях предлагаемые методы могли бы быть оправданы, что было в их возможностях. Для динамических систем с конечным числом степеней свободы систематическое использование оператора взятия определенного интеграла дает гораздо более короткий способ решения, чем какой-либо другой метод. В этом случае вопрос обоснования не требует никакой более развитой математики чем та, которая была во времена Хевисайда.

**7.02. Представление \*)  $Q^n f(t)$  в виде однократного интеграла.** Если  $f(t) = 1$ , то мы получаем с помощью последовательного интегрирования

$$Q1 = t, \quad Q^2 1 = \frac{t^2}{2!}, \quad Q^3 1 = \frac{t^3}{3!}, \quad \dots, \quad Q^n 1 = \frac{t^n}{n!}. \quad (1)$$

Для произвольной функции  $f(t)$  мы имеем по определению

$$Qf(t) = \int_0^t f(\tau) d\tau, \quad (2)$$

$$Q^2 f(t) = Q \int_0^t f(\tau) d\tau = \int_0^t \left[ \int_0^\xi f(\tau) d\tau \right] d\xi. \quad (3)$$

---

\*) Под термином *представление*, или интерпретация (interpretation), автор понимает формулу, дающую результат применения оператора к функции. — Прим. ред.

Интеграл берется по заштрихованной области на рис. 30. Изменив порядок интегрирования, мы получаем

$$Q^2 f(t) = \int_0^t \left[ \int_{\tau}^t f(\tau) d\xi \right] d\tau = \int_0^t (t - \tau) f(\tau) d\tau. \quad (4)$$

Аналогично

$$\begin{aligned} Q^3 f(t) &= Q \int_0^t (t - \tau) f(\tau) d\tau = \int_0^t \left[ \int_0^{\xi} (\xi - \tau) f(\tau) d\tau \right] d\xi = \\ &= \int_0^t \left[ \int_{\tau}^t (\xi - \tau) f(\tau) d\xi \right] d\tau = \int_0^t \frac{(t - \tau)^2}{2!} f(\tau) d\tau, \end{aligned} \quad (5)$$

и далее по индукции

$$Q^n f(t) = \int_0^t \frac{(t - \tau)^{n-1}}{(n-1)!} f(\tau) d\tau. \quad (6)$$

То же самое можно доказать следующим образом. Выражение  $Q^n f(t)$  является функцией, равной нулю при  $t=0$ , а ее производная равна  $Q^{n-1} f(t)$ . Условие равенства нулю выполнено в (6). Чтобы получить второе условие, продифференцируем под знаком интеграла по  $t$ ; в результате получим

$$\int_0^t \frac{(t - \tau)^{n-2}}{(n-2)!} f(\tau) d\tau = Q^{n-1} f(t). \quad (7)$$

Подынтегральное выражение обращается в нуль при  $\tau=t$ , если  $n > 1$ ; поэтому дифференцирование пределов интегрирования ничего не дает. Отсюда если этот результат верен для  $Q^{n-1}$ , то он верен и для  $Q^n$ ; но для  $n=1$  результат справедлив; поэтому он справедлив и для всех положительных целых значений  $n$ .

Таким образом, любая целая степень оператора  $Q$ , выполняемая над некоторой функцией, дает результат, который можно выразить однократным интегралом, если только функция интегрируема.

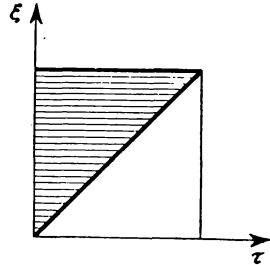


Рис. 30.

**7.03. Ряды операторов.** Эти правила позволяют нам выразить в виде однократного интеграла любую сумму конечного числа

членов вида

$$\begin{aligned}
 & (a_0 + a_1 Q + a_2 Q^2 + \dots + a_n Q^n) f(t) = \\
 & = a_0 f(t) + \sum_{r=1}^n \int_0^t a_r \frac{(t-\tau)^{r-1}}{(r-1)!} f(\tau) d\tau = \\
 & = a_0 f(t) + \int_0^t \left[ a_1 + a_2(t-\tau) + \frac{a_3(t-\tau)^2}{2!} + \dots + a_n \frac{(t-\tau)^{n-1}}{(n-1)!} \right] f(\tau) d\tau.
 \end{aligned}$$

Распространение этой формулы на случай бесконечных рядов требует специального обоснования.

**7.031. Теорема сходимости.** Если интеграл  $\int_0^T f(t) dt$  существует и если ряд

$$a_0 + a_1 z + \dots + a_n z^n + \dots \quad (1)$$

сходится для некоторого значения  $z$ , отличного от нуля, то ряд

$$(a_0 + a_1 Q + \dots + a_n Q^n + \dots) f(t) \quad (2)$$

сходится равномерно и абсолютно для  $0 \leq t \leq T$  и равен

$$a_0 f(t) + \sum_{n=1}^{\infty} \int_0^t a_n \frac{(t-\tau)^{n-1}}{(n-1)!} f(\tau) d\tau. \quad (3)$$

Два ряда (2) и (3) почленно равны и поэтому имеют одну и ту же сумму, если хотя бы один из них сходится. Но так как ряд (1) сходится для некоторого значения, скажем  $z = r$ , то для любого положительного  $\rho$ , не превосходящего  $|r|$ , мы можем найти такую величину  $M$ , что для всех  $n$  выполняются неравенства

$$|a_n| \rho^n < M. \quad (4)$$

Далее, так как  $Qf(t)$  ограничено при  $0 \leq t \leq T$ , то существует положительная величина, скажем  $C$ , такая, что

$$|Qf(t)| < C \quad (0 \leq t \leq T). \quad (5)$$

Поэтому

$$\begin{aligned}
 |Q^2 f(t)| & < \left| \int_0^t C d\tau \right| < C |t|, \\
 |Q^3 f(t)| & < \int_0^t C |t| d\tau < \frac{1}{2} C |t|^2,
 \end{aligned} \quad (6)$$



и вообще

$$|Q^n f(t)| < \frac{1}{(n-1)!} C |t|^{n-1}. \quad (7)$$

Отсюда абсолютные значения членов ряда, начиная с первого, соответственно меньше членов следующего ряда:

$$C \left( a_1 + a_2 |t| + \frac{1}{2} a_3 |t|^2 + \dots + \frac{1}{(n-1)!} |t|^{n-1} + \dots \right) < \quad (8)$$

$$< \frac{CM}{\rho} \left( 1 + \frac{|t|}{\rho} + \frac{1}{2} \frac{|t|^2}{\rho^2} + \dots + \frac{1}{(n-1)!} \frac{|t|^{n-1}}{\rho^{n-1}} + \dots \right). \quad (9)$$

Но последний ряд является экспоненциальным рядом, абсолютно сходящимся для всех  $t$ . Далее, его члены не превосходят членов ряда, получаемого заменой  $t$  на  $T$ , и этот ряд также является сходящимся рядом, члены которого положительные и не зависят от  $t$ . Поэтому наш ряд удовлетворяет  $M$ -признаку равномерной сходимости на интервале  $0 \leq t \leq T$ , т. е. на любом интервале  $t$ , где  $Qf(t)$  ограничено.

Если  $f(t)$  — непрерывная функция, то каждый член ряда является непрерывной функцией; и поскольку сумма равномерно сходящегося ряда непрерывных функций является непрерывной функцией, отсюда следует, что сумма ряда будет непрерывной функцией в интервале  $0 \leq t \leq T$ . По этой же причине суммирование можно провести под знаком интеграла; таким образом, получаем

$$\sum_{n=0}^{\infty} a_n Q^n f(t) = a_0 f(t) + \int_0^t f(\tau) \left( \sum_{n=1}^{\infty} a_n \frac{(t-\tau)^{n-1}}{(n-1)!} \right) d\tau. \quad (10)$$

Это утверждение остается справедливым, если потребовать, чтобы члены ряда (1) были ограничены, а также если интеграл  $\int_0^t f(\tau) d\tau$  существует только как несобственный интеграл вследствие неограниченности функции  $f(\tau)$ .

**7.032. Композиция (или произведение) операторов.** Пусть ряды

$$F(z) = a_0 + a_1 z + a_2 z^2 + \dots, \quad (1)$$

$$G(z) = b_0 + b_1 z + b_2 z^2 + \dots \quad (2)$$

оба сходятся для некоторого ненулевого значения  $|z|$ , скажем  $r$ . Тогда для любого положительного значения  $\rho$ , меньшего чем  $r$ , существуют величины  $M$ ,  $N$ , такие, что

$$|a_n| < \frac{M}{\rho^n}, \quad |b_n| < \frac{N}{\rho^n}. \quad (3)$$

Для любого  $z$ , такого, что  $|z| \leq \rho$ , произведение рядов

$$\begin{aligned} F(z) G(z) &= a_0 b_0 + (a_0 b_1 + a_1 b_0) z + (a_0 b_2 + a_1 b_1 + a_2 b_0) z^2 + \dots = \\ &= c_0 + c_1 z + c_2 z^2 + \dots \end{aligned} \quad (4)$$

является абсолютно сходящимся. Мы докажем, что если  $F(Q)$  и  $G(Q)$  являются операторными рядами, полученными заменой  $z$  на  $Q$ , и если  $G(Q)$  применяется к функции  $\varphi(t)$ , удовлетворяющей условиям последней теоремы, и затем  $F(Q)$  применяется к полученной функции, то окончательный результат будет такой же, как если бы мы заменили  $z$  на  $Q$  в произведении рядов (4) и применили этот оператор непосредственно к функции  $\varphi(t)$ . Мы имеем

$$F(Q) G(Q) \varphi(t) = F(Q) (b_0 + b_1 Q + b_2 Q^2 + \dots) \varphi(t). \quad (5)$$

Согласно последней теореме, ряд, представляющий функцию  $G(Q) \varphi(t)$ , является абсолютно и равномерно сходящимся и поэтому его можно интегрировать по частям. Отсюда получаем

$$\begin{aligned} Q^m (b_0 + b_1 Q + b_2 Q^2 + \dots) \varphi(t) &= \\ &= b_0 Q^m \varphi(t) + b_1 Q^{m+1} \varphi(t) + b_2 Q^{m+2} \varphi(t) + \dots \end{aligned} \quad (6)$$

и

$$F(Q) G(Q) \varphi(t) = \sum_m \sum_n a_m b_n Q^{m+n} \varphi(t), \quad (7)$$

где сначала надо провести суммирование по  $n$ . Но если  $|Q \varphi(t)| < C$  и  $m+n \geq 1$ , то

$$|a_m b_n Q^{m+n} \varphi(t)| < \frac{MNC}{(m+n-1)!} \left( \frac{r}{\rho} \right)^{m+n-1} \quad (8)$$

и слагаемые нашего ряда меньше положительных членов некоторого двойного ряда. Все члены этого ряда с одним и тем же  $m+n=k$  равны и их имеется всего  $k+1$ , так как  $n$  может изменяться от 0 до  $k$ . Поэтому их сумма равна

$$MNC \frac{(k+1)}{(k-1)!} \left( \frac{r}{\rho} \right)^{k-1} \quad (9)$$

и сумма таких членов по  $k$  сходится.

Ряд (7) абсолютно сходится и его члены можно переставить в любом порядке, что не изменит ни его сходимости, ни суммы. Мы можем, следовательно, сгруппировать все члены с одинаковым  $Q^{m+n}$ , но тогда получим

$$F(Q) G(Q) \varphi(t) = (c_0 + c_1 Q + c_2 Q^2 + \dots) \varphi(t), \quad (10)$$

что доказывает теорему.

**7.04. Линейные дифференциальные уравнения первого порядка.** Рассмотрим теперь линейное дифференциальное уравнение первого порядка

$$\frac{dx}{dt} - \alpha x = \varphi(t), \quad (1)$$

где  $\alpha$  — константа и дано, что  $x = x_0$  при  $t = 0$ . Заменяем  $t$  на  $\tau$  и проинтегрируем обе части уравнения по  $\tau$  от 0 до  $t$ . Мы получим

$$x - x_0 - \alpha Qx = Q\varphi(t), \quad (2)$$

т. е.

$$(1 - \alpha Q)x = x_0 + Q\varphi(t), \quad (3)$$

и выражение в правой части является непрерывной функцией от  $t$ , если  $\varphi(t)$  интегрируема. Теперь применим к обеим частям этого равенства ряд операторов  $(1 + \alpha Q + \alpha^2 Q^2 + \dots)$ . Согласно последней теореме, результатом последовательного применения операций  $1 - \alpha Q$  и  $1 + \alpha Q + \alpha^2 Q^2 + \dots$  к функции  $x$  будет просто  $x$ , так как если мы напомним  $z$  вместо  $Q$  в этих рядах, то получим два ряда, сходящихся для  $|z| < 1/\alpha$ , и их произведение равно 1. Поэтому

$$x = (1 + \alpha Q + \alpha^2 Q^2 + \dots)[x_0 + Q\varphi(t)]. \quad (4)$$

Это дает формальное решение, которое можно разложить в ряд по степеням  $\alpha$ . Но, с другой стороны, нам известно, что решением уравнения (1) является

$$x = x_0 e^{\alpha t} + e^{\alpha t} \int_0^t e^{-\alpha \tau} \varphi(\tau) d\tau. \quad (5)$$

Отсюда выражения в правых частях (4) и (5) равны для всех значений  $x_0$ ; поэтому

$$(1 + \alpha Q + \alpha^2 Q^2 + \dots)x_0 = e^{\alpha t} x_0, \quad (6)$$

$$(1 + \alpha Q + \alpha^2 Q^2 + \dots)Q\varphi(t) = e^{\alpha t} Q[e^{-\alpha t} \varphi(t)]. \quad (7)$$

До сих пор мы не объяснили, что означает деление на оператор. И как раз по этой причине мы теперь имеем право на это при условии, что обеспечим непротиворечивость деления. Если, например, мы придадим смысл выражению  $g(t)/(1 - \alpha Q)$ , то он должен быть таким, что применение к этому выражению оператора  $1 - \alpha Q$  давало бы  $g(t)$ . Но мы имеем

$$(1 - \alpha Q)(1 + \alpha Q + \alpha^2 Q^2 + \dots)g(t) = g(t) \quad (8)$$

согласно правилу композиции операторов для всех  $g(t)$ , таких, для которых эти операции применимы. Поэтому мы можем *определить*

$$\frac{1}{1-\alpha Q} = 1 + \alpha Q + \alpha^2 Q^2 + \dots \quad (9)$$

и записать данную выше интерпретацию в компактном виде

$$\frac{1}{1-\alpha Q} 1 = e^{\alpha t}, \quad (10a)$$

$$\frac{1}{1-\alpha Q} Q\varphi(t) = e^{\alpha t} Q[e^{-\alpha t} \varphi(t)] = \int_0^t e^{\alpha(t-\tau)} \varphi(\tau) d\tau. \quad (10b)$$

Аналогично если  $F(z)$  — произвольная функция  $z$  и ее можно разложить в степенной ряд в окрестности точки  $z=0$ , причем  $F(0) \neq 0$  и  $G(z)$  — обратный ряд, то мы можем интерпретировать  $1/F(Q)$  как  $G(Q)$ . Значит, основной интерпретацией любого оператора всегда служит его разложение по положительным степеням  $Q$ , как если бы  $Q$  было константой. Такие операторы, как  $Q^{-n}$  и  $e^{h/Q}$ , неразложимы по положительным степеням  $Q$  в указанном смысле, т. е. функции  $z^{-n}$  и  $e^{h/z}$  не могут быть разложены по положительным степеням  $z$ .

Следовательно, в настоящее время мы не можем дать им никакой интерпретации и оказывается, что в важном классе физических задач, требующих решения конечного числа линейных дифференциальных уравнений с постоянными коэффициентами, такие операторы не возникают.

Мы имеем

$$\begin{aligned} \frac{1}{(1-\alpha Q)^2} 1 &= (1 + 2\alpha Q + 3\alpha^2 Q^2 + \dots) 1 = \\ &= 1 + 2\alpha t + \frac{3}{2} \alpha^2 t^2 + \frac{4}{6} \alpha^3 t^3 + \dots, \end{aligned} \quad (11)$$

и трудно сразу увидеть, какую функцию представляет этот ряд; однако

$$\begin{aligned} \frac{Q}{(1-\alpha Q)^2} 1 &= (Q + 2\alpha Q^2 + 3\alpha^2 Q^3 + \dots + n\alpha^{n-1} Q^n \dots) 1 = \\ &= t + \alpha t^2 + \frac{1}{2} \alpha^2 t^3 + \dots + \frac{\alpha^{n-1}}{(n-1)!} t^n + \dots = t e^{\alpha t}. \end{aligned} \quad (12)$$

Теперь мы видим, что

$$\frac{1}{(1-\alpha Q)^2} = \frac{1}{1-\alpha Q} 1 + \frac{\alpha Q}{(1-\alpha Q)^2} 1 = (1 + \alpha t) e^{\alpha t}, \quad (13)$$

это совпадает с (11). Но с увеличением  $n$  представление выражения  $(1-\alpha Q)^{-n} 1$  содержит все больше и больше членов.

Иначе ведет себя выражение  $Q^{n-1}(1-\alpha Q)^{-n}1$ , так как

$$\begin{aligned} \frac{Q^{n-1}}{(1-\alpha Q)^n} 1 &= \left( Q^{n-1} + n\alpha Q^n + \frac{n(n+1)}{2!} \alpha^2 Q^{n+1} + \dots + \right. \\ &\quad \left. + \frac{n(n+1) \dots (n+r-1)}{r!} \alpha^r Q^{n+r-1} + \dots \right) 1 = \\ &= \frac{t^{n-1}}{(n-1)!} + \frac{\alpha t^n}{(n-1)!} + \frac{\alpha^2 t^{n+1}}{2! (n-1)!} + \dots = \frac{t^{n-1}}{(n-1)!} e^{\alpha t}, \quad (14) \end{aligned}$$

если мы будем интерпретировать отдельные члены с помощью 7.02(1). Аналогично с помощью разложения получаем

$$\begin{aligned} \frac{Q^n}{(1-\alpha Q)^n} \varphi(t) &= \int_0^t \varphi(\tau) \left[ \frac{(t-\tau)^{n-1}}{(n-1)!} + \frac{\alpha(t-\tau)^n}{(n-1)!} + \right. \\ &\quad \left. + \frac{\alpha^2}{2!} \frac{(t-\tau)^{n+1}}{(n-1)!} + \dots \right] d\tau = \int_0^t \varphi(\tau) \frac{(t-\tau)^{n-1}}{(n-1)!} e^{\alpha(t-\tau)} d\tau. \quad (15) \end{aligned}$$

Или же мы можем использовать (106), тогда получим

$$\begin{aligned} \frac{Q^2}{(1-\alpha Q)^2} \varphi(t) &= \frac{Q}{1-\alpha Q} \int_0^t e^{\alpha(t-\tau)} \varphi(\tau) d\tau = \\ &= \int_0^t \left[ \int_0^\tau e^{\alpha(t-\tau)} e^{\alpha(\tau-\xi)} \varphi(\xi) d\xi \right] d\tau = \\ &= \int_0^t \left[ \int_\xi^t e^{\alpha(t-\xi)} \varphi(\xi) d\tau \right] d\xi = \int_0^t \varphi(\xi) (t-\xi) e^{\alpha(t-\xi)} d\xi, \end{aligned}$$

и далее так же, как при доказательстве 7.02 (6).

**7.05. Система  $n$  линейных дифференциальных уравнений первого порядка.** Рассмотрим уравнения

$$\left. \begin{aligned} e_{11}y_1 + e_{12}y_2 + \dots + e_{1n}y_n &= S_1, \\ e_{21}y_1 + e_{22}y_2 + \dots + e_{2n}y_n &= S_2, \\ . &. . . . . \\ e_{n1}y_1 + e_{n2}y_2 + \dots + e_{nn}y_n &= S_n, \end{aligned} \right\} \quad (1)$$

где  $y_1, y_2, \dots, y_n$  — зависимые переменные,  $t$  — независимая переменная,  $S_1, \dots, S_n$  — известные интегрируемые функции  $t$  при  $0 \leq t \leq T$  и

$$e_{rs} = a_{rs} \frac{d}{dt} + b_{rs}, \quad (2)$$

где  $a_{rs}$  и  $b_{rs}$  — константы. Мы не будем в настоящее время предполагать, что  $a_{rs} = a_{sr}$ ,  $b_{rs} = b_{sr}$ , но предположим, что определитель

$$A = \|a_{rs}\| \quad (3)$$

не равен нулю. Если это не так, то, как мы сможем показать ниже, имеется некоторая неточность в задании начальных условий.

Используя правило суммирования, мы можем записать уравнения (1) в виде

$$a_{rs} \frac{dy_s}{dt} + b_{rs} y_s = S_r. \quad (4)$$

Выполним операцию  $Q$  над обеими частями каждого из уравнений, т. е. заменим временно  $t$  на вспомогательную переменную  $\tau$  и проинтегрируем по  $\tau$  от 0 до  $t$ . Мы получим

$$a_{rs}(y_s - u_s) + b_{rs} Q y_s = Q S_r, \quad (5)$$

причем  $u_s$  есть значение  $y_s$  при  $t = 0$ .

Мы перепишем это в виде

$$\hat{f}_{rs} y_s = (a_{rs} + b_{rs} Q) y_s = a_{rs} u_s + Q S_r. \quad (6)$$

Эти уравнения включают в себя как дифференциальные уравнения, так и начальные условия.

Пусть теперь  $D$  обозначает операторный определитель

$$D = \|\hat{f}_{rs}\|. \quad (7)$$

Если этот определитель расписать по правилам алгебры, мы получим полином относительно  $Q$  в общем случае степени  $n$ . Член, не содержащий  $Q$ , равен  $A$  и, согласно сделанному предположению, он отличен от нуля. Пусть  $F_{rs}$  — алгебраическое дополнение элемента  $\hat{f}_{rs}$  в этом определителе.  $F_{rs}$  также является полиномом относительно  $Q$ .

Теперь подействуем на первое уравнение в системе (1) оператором  $F_{rm}$ , на второе — оператором  $F_{2m}$  и т. д. и сложим результаты. Мы получим

$$F_{rm} \hat{f}_{rs} y_s = F_{rm} (a_{rs} u_s + Q S_r). \quad (8)$$

Но  $F_{rm} \hat{f}_{rs} = 0$ , если  $m \neq s$ , так как эта сумма является определителем с двумя одинаковыми столбцами. Если  $m = s$ , то  $F_{rm} \hat{f}_{rs} = D$ . Поэтому

$$D y_m = F_{rm} (a_{rs} u_s + Q S_r). \quad (9)$$

Выражение в правой части является ограниченной интегрируемой функцией  $t$ , поскольку такими являются функции  $S_r$ .

Точно так же функция  $D(z)$ , полученная заменой  $Q$  на число  $z$ , отлична от нуля при  $z=0$ , так как  $A \neq 0$ . Поэтому для значений  $z$ , меньших чем некоторое положительное  $\rho$ , функция  $1/D(z)$  может быть разложена в степенной ряд по  $z$ . В соответствии с правилом 7.04 мы определим  $D^{-1}$  как степенной ряд относительно  $Q$ , полученный заменой  $z$  на  $Q$  в ряде  $1/D(z)$ . Далее, подействуем оператором  $D^{-1}$  на обе части равенства (9). Мы получим

$$D^{-1}Dy_m = D^{-1}F_{rm}(a_{rs}u_s + QS_r). \quad (10)$$

Но степенные относительно  $Q$  ряды можно перемножать по правилам алгебры; поэтому выражение  $D^{-1}Dy_m$  равно просто  $y_m$ , и мы получили формальное решение

$$y_m = D^{-1}F_{rm}(a_{rs}u_s + QS_r). \quad (11)$$

Основное правило интерпретации в операционном исчислении состоит в том, что операторы должны перемножаться и интерпретироваться почленно; однако мы увидим, что с помощью правил, которые мы уже имеем, все эти операторы могут быть сведены в худшем случае к однократному интегрированию. Ряд  $D^{-1}$ , действующий на интегрируемую функцию, всегда дает сходящийся ряд; поэтому результат имеет смысл и должен быть решением, соответствующим рассматриваемым дифференциальным уравнениям и начальным условиям, если решение вообще существует. Чтобы показать, что (11) удовлетворяет начальным условиям и дифференциальным уравнениям.

Во-первых, если  $t \rightarrow 0$ , то все члены, содержащие  $Q$ , стремятся к нулю; далее,  $D^{-1} \rightarrow A^{-1}$ ,  $F_{rm} \rightarrow A_{rm}$  — алгебраическому дополнению элемента  $a_{rm}$  в  $D$ . Поэтому для  $t=0$

$$y_m = A^{-1}A_{rm}a_{rs}u_s. \quad (12)$$

Но  $A_{rm}a_{rs} = 0$  при  $m \neq s$  и  $A_{rm}a_{rms} = A$ . Отсюда

$$y_m = u_m$$

и решение удовлетворяет начальным условиям.

Во-вторых,

$$a_{rs} = f_{rs} - b_{rs}Q \quad (13)$$

и поэтому решение можно записать в виде

$$y_m = D^{-1}F_{rm}(f_{rs}u_s - b_{rs}u_sQ + QS_r). \quad (14)$$

Первый член, как и раньше, сводится к  $u_m$  после суммирования. Поэтому

$$y_m = u_m + D^{-1}F_{rm}Q(S_r - b_{rs}u_s) \quad (15)$$

и последний член содержит лишь положительные степени  $Q$ , действующие на известную функцию. Но

$$\frac{d}{dt} Qf(t) = \frac{d}{dt} \int_0^t f(\tau) d\tau = f(t). \quad (16)$$

Отсюда

$$\left(a_{vm} \frac{d}{dt} + b_{vm}\right) Qf(t) = (a_{vm} + b_{vm}Q) \dot{f}(t) = \dot{f}_{vm} f(t), \quad (17)$$

$$\left(a_{vm} \frac{d}{dt} + b_{vm}\right) y_m = b_{vm} u_m + \dot{f}_{vm} D^{-1} F_{rm} (S_r - b_{rs} u_s). \quad (18)$$

Снова  $\dot{f}_{vm} F_{rm} = 0$  при  $r \neq v$ ,  $\dot{f}_{vm} F_{rm} = D$ . Поэтому последний член приводится к  $S_v - b_{vs} u_s$ . Здесь второй член сокращается с  $b_{vm} u_m$  и окончательно

$$\left(a_{vm} \frac{d}{dt} + b_{vm}\right) y_m = S_v, \quad (19)$$

что показывает, что полученное решение удовлетворяет дифференциальным уравнениям. Этим заканчивается доказательство того, что задача имеет единственное решение, даваемое формулой (11).

Рассмотрим теперь случай  $A=0$ . Умножим уравнения (1) на  $A_{1m}$ ,  $A_{2n}$  ... и сложим. Тогда  $a_{rs} A_{rm} = 0$  даже при  $s=m$ , так как в этом случае сумма равна  $A$ . Поэтому

$$A_{rm} b_{rs} y_s = A_{rm} S_r \quad (20)$$

для всех  $t$ , и в частности для  $t=0$ . Таким образом, значения  $y_s$  при  $t=0$  не могут задаваться независимо друг от друга. Если они заданы так, что (20) удовлетворяется, то существует связь между  $y_s$  для всех значений времени и одну из переменных можно исключить; если начальные условия не удовлетворяют (20), тогда они противоречивы.

Условие  $A \neq 0$  (т. е.  $a_{rs}$  является матрицей ранга  $n$ ) эквивалентно условию, что начальные значения неизвестных независимы, и оно будет удовлетворяться в любой правильно поставленной задаче.

**7.051. Символ  $p$ .** Процесс получения операторного решения (11) из (6) совпадает с процессом решения системы алгебраических уравнений относительно  $y_s$ . Описанная выше процедура наиболее удобна для установления общих теорем,



однако действительное вычисление операторного решения облегчается с помощью изменения обозначений. Мы заменим  $Q$  на  $p^{-1}$ ; правило, согласно которому операторы должны представляться в виде

$$a_0 + a_1 Q + \dots,$$

переходит в правило, что их можно представить в виде

$$a_0 + a_1 p^{-1} + \dots,$$

где коэффициенты таковы, что ряд

$$a_0 + a_1 z + \dots$$

сходится для некоторого  $z$ , отличного от нуля. Переписывая с помощью этого обозначения и интерпретацию выражений, которую мы уже делали, мы получим

$$\begin{aligned} p^{-1}1 &= t, & p^{-2}1 &= \frac{1}{2}t^2, & p^{-n}1 &= \frac{t^n}{n!}, \\ p^{-1}f(t) &= \int_0^t f(\tau) d\tau, & p^{-n}f(t) &= \int_0^t \frac{(t-\tau)^{n-1}}{(n-1)!} f(\tau) d\tau, \\ \sum_{n=0}^{\infty} a_n p^{-n} f(t) &= a_0 f(t) + \int_0^t f(\tau) \left( \sum_{n=1}^{\infty} a_n \frac{(t-\tau)^{n-1}}{(n-1)!} \right) d\tau, & (21) \\ \frac{p}{p-\alpha} 1 &= e^{\alpha t}, & \frac{1}{p-\alpha} f(t) &= \int_0^t f(\tau) e^{\alpha(t-\tau)} d\tau, \\ \frac{p}{(p-\alpha)^n} 1 &= \frac{t^{n-1}}{(n-1)!} e^{\alpha t}, & \frac{1}{(p-\alpha)^n} f(t) &= \int_0^t f(\tau) \frac{(t-\tau)^{n-1}}{(n-1)!} e^{\alpha(t-\tau)} d\tau. \end{aligned}$$

Преимущество этого обозначения в том, что в последних двух уравнениях, операторы которых могут быть выражены единичным членом, в числителе имеются  $p$  или  $1$  вместо  $Q^{n-1}$  или  $Q^n$ .

Мы получаем также немедленно с помощью прямого разложения

$$\frac{p^2}{p^2 + n^2} 1 = \cos nt, \quad \frac{np}{p^2 + n^2} 1 = \sin nt, \quad (22)$$

$$\frac{p^2}{p^2 - n^2} 1 = \operatorname{ch} nt, \quad \frac{np}{p^2 - n^2} 1 = \operatorname{sh} nt. \quad (23)$$

Возвращаясь к (6), мы видим, что поскольку решение является чисто алгебраическим процессом, то, если мы напомним  $p^{-1}$  вместо  $Q$  в каждом из уравнений (6), затем формально умножим их на  $p$  и завершим решение с помощью

алгебры, мы придем к тому же самому решению при условии, что будем придерживаться основного правила — операторы должны быть разложены по нулевой и отрицательным степеням  $p$  перед их интерпретацией с помощью приведенных выше формул. Но по этому правилу мы получаем вместо (6)

$$(a_{rs}p + b_{rs})y_s = pa_{rs}u_s + S_r. \quad (24)$$

Эти уравнения называются *вспомогательными уравнениями*. Они образуются, как можно убедиться проверкой, из дифференциальных уравнений следующим образом.

Напишем  $p$  вместо  $d/dt$  в левой части уравнения; к правой части каждого уравнения прибавим результат, получаемый при отбрасывании  $b_{rs}$  в левой части и замене  $y_s$  на их начальные значения. Получаемые вспомогательные уравнения должны быть решены по правилам линейной алгебры, как если бы  $p$  было числом, и результат нужно интерпретировать с помощью разложения по убывающим степеням  $p$ , понимая  $p^{-1}$  как операцию интегрирования от 0 до  $t$ .

**7.052. Правило элементарных дробей.** Так как операторное решение (11) разложимо по степеням  $Q$  или  $p^{-1}$ , начиная с постоянного члена, то оператор должен иметь вид  $F(p)/G(p)$ , где  $F(p)$  — полином относительно  $p$  той же самой степени, что и  $G(p)$ , или ниже. Если  $p$  заменить числом  $z$ , то  $F(z)/G(z)$  является рациональной функцией от  $z$  и поэтому может быть разложена на элементарные дроби. Каждую такую дробь можно разложить по убывающим степеням  $z$ , начиная, возможно, с константы, и сумма разложений есть разложение функции  $F(z)/G(z)$ . Следовательно, если мы формально разобьем оператор на элементарные дроби и применим их по отдельности к данной функции, то сумма результатов даст результат применения разложения оператора  $F(p)/G(p)$  к той же самой функции.

Решение особенно просто, когда  $G(z)$  имеет только простые нули вида  $z = \alpha$  и не равно нулю при  $z = 0$ . Мы имеем тогда алгебраическое тождество

$$\frac{F(p)}{pG(p)} = \frac{F(0)}{pG(0)} + \sum_{\alpha} \frac{F(\alpha)}{\alpha G'(\alpha)} \frac{1}{p - \alpha},$$

откуда

$$\begin{aligned} \frac{F(p)}{G(p)} &= \frac{F(0)}{G(0)} + \sum_{\alpha} \frac{F(\alpha)}{\alpha G'(\alpha)} \frac{p}{p - \alpha}, \\ \frac{F(p)}{G(p)} 1 &= \frac{F(0)}{G(0)} + \sum_{\alpha} \frac{F(\alpha)}{\alpha G'(\alpha)} e^{\alpha t}. \end{aligned} \quad (25)$$

Поэтому та часть решения, которая зависит от начальных условий, выражается непосредственно с помощью конечного числа членов. Несколько иная форма является более удобной, когда функция, над которой производится операция, отлична от константы; мы можем написать

$$\frac{F(p)}{G(p)} = \lim_{z \rightarrow \infty} \frac{F(z)}{G(z)} + \sum_{\alpha} \frac{F(\alpha)}{G'(\alpha)} \frac{1}{p - \alpha},$$

$$\frac{F(p)}{G(p)} S(t) = \lim_{z \rightarrow \infty} \frac{F(z)}{G(z)} S(t) + \sum_{\alpha} \frac{F(\alpha)}{G'(\alpha)} \int_0^t S(\tau) e^{\alpha(t-\tau)} d\tau; \quad (26)$$

представление (25) часто называется теоремой разложения Хевисайда. Но его методы содержат также две другие теоремы разложения, а именно разложение по степеням  $p^{-1}$  и по степеням  $e^{-ph}$ , где  $h$  — константа, и в данном случае первое разложение играет основную роль. Поэтому в настоящей работе мы будем называть разложение (25) *правилом элементарных дробей*. Оно читается следующим образом: *разделим на  $p$ , разложим на элементарные дроби, умножим на  $p$  и запишем результат применения оператора к константе*.

Если  $G(p)$  имеет кратные корни или  $p$  сомножителем, выражение  $F(p)/pG(p)$  через элементарные дроби все же можно еще получить, но оно будет содержать члены вида  $p^{-s}$  или  $(p - \alpha)^{-s}$  и  $F(p)/G(p)$  будет содержать слагаемые вида  $p^{-(s-1)}$  или  $p/(p - \alpha)^s$ . Эти члены можно применять к функциям с помощью (21). Следовательно, если только функции  $S$ , интегрируемы и начальные значения неизвестных независимы, решение может быть получено операционными методами и результат может быть выражен в худшем случае с помощью конечного числа однократных интегралов.

Удобный способ нахождения членов вида  $(p - \alpha)^{-s}$  можно проиллюстрировать на следующем примере:

$$F(p) = \frac{p}{(p+1)^2(p+2)}.$$

Когда  $z \rightarrow -1$ , дробь  $1/(z+2)$  стремится к 1; тогда

$$\begin{aligned} F(p) - \frac{p}{(p+1)^2} &= \frac{p}{(p+1)^2} \left( \frac{1}{p+2} - 1 \right) = \\ &= -\frac{p}{(p+1)(p+2)} = -\frac{p}{p+1} + \frac{p}{p+2}, \\ F(p) 1 &= \left( \frac{p}{(p+1)^2} - \frac{p}{p+1} + \frac{p}{p+2} \right) 1 = te^{-t} - e^{-t} + e^{-2t}. \end{aligned}$$

С помощью вычитания мы можем таким способом снижать наивысший показатель в знаменателе на 1 на каждом шаге. Каждый шаг проверяет алгебраические выкладки предыдущего шага.

**7.053. Принцип суперпозиции.** Не всегда, однако, бывает удобно использовать два различных разложения на элементарные дроби в зависимости от того, является ли функция, к которой применяется оператор, константой или нет. Это неудобство можно обойти с помощью принципа суперпозиции. Пусть мы имеем в  $p$ -представлении

$$F(p) = a_0 + a_1 p^{-1} + a_2 p^{-2} + \dots,$$

$$F(p) 1 = a_0 + a_1 t + a_2 \frac{t^2}{2} + \dots + a_n \frac{t^n}{n!} + \dots = f(t);$$

тогда

$$f'(t) = a_1 + a_2 t + \dots + \frac{a_n t^{n-1}}{(n-1)!} + \dots$$

и из третьей строки (21)

$$F(p) \varphi(t) = a_0 \varphi(t) + \int_0^t \varphi(\tau) f'(t-\tau) d\tau. \quad (27)$$

Интегрируя по частям, имеем

$$F(p) \varphi(t) = \varphi(0) f(t) + \int_{\tau=0}^t f(t-\tau) d\varphi(\tau), \quad (28)$$

так как  $f(0) = a_0$ . Отсюда если мы знаем  $F(p) 1$ , то вычисление  $F(p) \varphi(t)$  сводится к однократному интегрированию. Используя этот результат, мы можем вывести 7.052 (26) из (25), и для этого нам нужно лишь одно разложение на элементарные дроби.

Этой теореме можно дать следующую физическую интерпретацию. Мы можем рассматривать систему, описываемую дифференциальными уравнениями и подверженную возмущениям, представляемым нашими  $S_r(t)$ . Но если система была в состоянии  $y_s = 0$  до момента времени  $t = 0$  и затем  $y_s$  скачком изменились до  $u_s$ , мы могли бы представить, что это произошло благодаря некоторому множеству импульсных возмущений, и, таким образом, фактически включить начальные значения в  $QS$ ,

ценой того, что мы считаем  $\int_0^\delta S_r(\tau) d\tau = a_{rs} u_s$  в пределе, когда  $\delta$  делается произвольно малым. Тогда член  $\varphi(0) f(t)$  можно

рассматривать как остаточный эффект в момент времени  $t$  от импульсного возмущения  $\varphi(0)$  в момент времени  $t=0$ . Второй член возмущения появился благодаря  $S_t$  или  $\varphi(t)$ , и его можно рассматривать как результирующий эффект многочисленных малых возмущений  $S_t d\tau$  или  $d\varphi(\tau)$  в момент времени  $d\tau$ . Каждое из них дает свой остаточный эффект в момент времени  $t$ , но интервал действия возмущения теперь отстоит от рассматриваемого момента на  $t-\tau$  вместо  $t$ . Следовательно, полный вклад возмущений равен сумме элементов вида  $f(t-\tau)d\varphi(\tau)$ , что и дает интеграл. Не обязательно требовать, чтобы  $\varphi(t)$  была дифференцируемой, но если она не дифференцируема, то интеграл является не обычным интегралом Римана, а интегралом обобщенного вида — интегралом Стильтьеса (ср. 1.10, 1.102).

**7.054.** Третий метод часто является наиболее удобным, когда функция, к которой применяется оператор, сама может быть выражена в виде  $G(p)l = g(t)$ . Тогда

$$F(p)g(t) = F(p)G(p)l \quad (29)$$

и мы можем непосредственно приступить к выписыванию результата применения оператора в правой части с помощью правила элементарных дробей.

**7.06. Уравнения более высокого порядка.** Описанный метод очень легко распространяется на уравнения более высокого порядка сведением их к уравнениям первого порядка. Таким образом, если мы имеем уравнение второго порядка, такое, как

$$\frac{d^2x}{dt^2} + a \frac{dx}{dt} + bx = 1, \quad (1)$$

причем  $x = x_0$ ,  $dx/dt = x_1$  при  $t = 0$ , то мы вводим новую переменную  $y$  с помощью уравнения

$$\frac{dx}{dt} - y = 0 \quad (2)$$

и первоначальное уравнение заменяется на

$$\frac{dy}{dt} + bx + ay = 1. \quad (3)$$

Теперь (2) и (3) — уравнения первого порядка и вспомогательными уравнениями будут

$$px - y = px_0, \quad (4)$$

$$py + bx + ay = px_1 + 1. \quad (5)$$

Исключая  $y$  с помощью алгебраических выкладок, мы получаем

$$x = \frac{(p^2 + ap)x_0 + px_1 + 1}{p^2 + ap + b} 1, \quad (6)$$

что можно расписать далее с помощью правила элементарных дробей, полагая

$$p^2 + ap + b = (p - \alpha)(p - \beta). \quad (7)$$

**7.061.** Если мы имеем  $n$  дифференциальных уравнений второго порядка, возможно, с различными функциями в правых частях, мы поступим точно так же. Если типичное уравнение записывается в виде

$$a_{rs} \frac{d^2 y_s}{dt^2} + b_{rs} \frac{dy_s}{dt} + c_{rs} = S_r(t), \quad (1)$$

то мы возьмем

$$z_s = \frac{dy_s}{dt}. \quad (2)$$

Определяя, таким образом, набор из  $n$  новых переменных  $z_s$ , можно записать (1) в виде

$$a_{rs} \frac{dz_s}{dt} + b_{rs} z_s + c_{rs} y_s = S_r(t), \quad (3)$$

и вместо  $n$  уравнений второго порядка мы имеем теперь  $2n$  уравнений первого порядка. Операционные методы решения будут пригодны при условии, что начальные значения всех  $y_s$  и  $z_s$  независимы. Если они равны  $u_s$  и  $v_s$ , мы можем записать вспомогательные уравнения

$$p y_s - z_s = p u_s, \quad (4)$$

$$(a_{rs} p + b_{rs}) z_s + c_{rs} y_s = S_r(t) + a_{rs} p v_s. \quad (5)$$

Первый шаг в процессе решения состоит в исключении  $z_s$  из этих двух систем уравнений; тогда

$$(a_{rs} p + b_{rs}) p (y_s - u_s) + c_{rs} y_s = S_r(t) + a_{rs} p v_s. \quad (6)$$

т. е.

$$(a_{rs} p^2 + b_{rs} p + c_{rs}) y_s = S_r(t) + (a_{rs} p^2 + b_{rs} p) u_s + a_{rs} p v_s. \quad (7)$$

Эти уравнения могут быть решены относительно  $y_s$  как система уравнений первого порядка и затем применены те же самые правила интерпретации. Начальные значения  $y_s$  и  $dy_s/dt$  в этой схеме входят через  $u_s$  и  $v_s$  в правой части равенства.

Определитель коэффициентов при  $dy_s/dt$  и  $dz_s/dt$  в (2) и (3) равен

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & \dots \\ a_{21} & \dots & \dots & \dots & 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & a_{nn} & 1 & 0 & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & 1 \end{vmatrix} = \|a_{rs}\|,$$

так что решение возможно с произвольными начальными значениями всех  $y_s$  и  $dy_s/dt$  снова при том же условии  $\|a_{rs}\| \neq 0$ .

Физический смысл условия  $A = \|a_{rs}\| \neq 0$  ясен в этом случае. Переменные  $y_s$  могут быть координатами некоторой динамической системы и будут удовлетворять дифференциальным уравнениям второго порядка по времени. Далее, условие, что начальные значения  $y_s$  и  $dy_s/dt$  могут быть выбраны независимо, равносильно тому, что выбранные координаты и их скорости изменения могут иметь произвольные начальные значения.

**7.062.** Мы уже знаем правила

$$\frac{p^2}{p^2 + n^2} 1 = \cos nt, \quad \frac{np}{p^2 + n^2} 1 = \sin nt. \quad (1)$$

Их можно проверить и с помощью правила элементарных дробей. Если мы продифференцируем по  $n$  разложение в ряд этих операторов по степеням  $p^{-1}$ , то получим сходящийся ряд; поэтому

$$\begin{aligned} \frac{2n^2 p^2}{(p^2 + n^2)^2} 1 &= nt \sin nt, \\ \left[ \frac{p}{p^2 + n^2} - \frac{2n^2 p}{(p^2 + n^2)^2} \right] 1 &= t \cos nt, \end{aligned} \quad (2)$$

откуда

$$\frac{2n^4 p}{(p^2 + n^2)^2} 1 = \sin nt - nt \cos nt. \quad (3)$$

Равенство **7.04** (10.б) можно записать в виде

$$e^{-\alpha t} \frac{1}{p - \alpha} f(t) = \frac{1}{p} [e^{-\alpha t} f(t)]. \quad (4)$$

Откуда

$$e^{-\alpha t} \frac{1}{(p - \alpha)^n} f(t) = \frac{1}{p^n} [e^{-\alpha t} f(t)], \quad (5)$$

$$e^{-\alpha t} F(p - \alpha) f(t) = F(p) [e^{-\alpha t} f(t)]. \quad (6)$$

В частности,

$$e^{-\alpha t} \frac{p(p-\alpha)}{(p-\alpha)^2 + \beta^2} 1 = \frac{p(p+\alpha)}{p^2 + \beta^2} e^{-\alpha t} = \\ = \frac{p(p+\alpha)}{p^2 + \beta^2} \frac{p}{p+\alpha} 1 = \frac{p^2}{(p^2 + \beta^2)} 1 = \cos \beta t, \quad (7)$$

и поэтому

$$\frac{p(p-\alpha)}{(p-\alpha)^2 + \beta^2} 1 = e^{\alpha t} \cos \beta t. \quad (8)$$

Кроме того,

$$e^{-\alpha t} \frac{\beta p}{(p-\alpha)^2 + \beta^2} 1 = \frac{\beta(p+\alpha)}{p^2 + \beta^2} e^{-\alpha t} = \frac{\beta p}{p^2 + \beta^2} 1 = \sin \beta t$$

и поэтому

$$\frac{\beta p}{(p-\alpha)^2 + \beta^2} 1 = e^{\alpha t} \sin \beta t. \quad (9)$$

Эти правила приведены здесь как иллюстрация 7.04 и 7.054. Другой способ состоит в том, что мы можем применить правило элементарных дробей непосредственно:

$$\frac{p}{p-\alpha-i\beta} 1 = e^{(\alpha+i\beta)t} = e^{\alpha t} (\cos \beta t + i \sin \beta t),$$

где мы разделяем действительную и мнимые части.

**7.07.** Иногда бывает нужно определить предел функции  $[F(p)/G(p)] 1$  или ее интеграла при  $t \rightarrow \infty$ . Задача индукционного равновесия из следующей главы является примером этого. Эти пределы можно просто найти с помощью правила элементарных дробей. Необязательно рассматривать кратные множители, так как мы можем разделить их, немного изменив константы. Все корни  $\alpha$  должны иметь отрицательные действительные части, в противном случае решение будет содержать неограниченно возрастающие экспоненциальные члены с положительными показателями или тригонометрические члены, осциллирующие в конечных пределах.

Таким образом, мы имеем

$$\frac{F(p)}{G(p)} 1 = \frac{F(0)}{G(0)} + \sum \frac{F(\alpha)}{\alpha G'(\alpha)} e^{\alpha t} \quad (10)$$

и предел при  $t \rightarrow \infty$  равен  $F(0)/G(0)$ . Далее, если интеграл от  $[F(p)/G(p)] 1$  должен иметь конечный предел, то  $F(0)/G(0)$  должно быть нулем; тогда

$$\int_0^\infty \sum \frac{F(\alpha)}{\alpha G'(\alpha)} e^{\alpha t} dt = - \sum \frac{F(\alpha)}{\alpha^2 G'(\alpha)} = \\ = \lim_{\lambda \rightarrow 0} \sum \frac{F(\alpha)}{\alpha G'(\alpha)} \frac{1}{\lambda - \alpha} = \lim_{\lambda \rightarrow 0} \frac{F(\lambda)}{\lambda G(\lambda)}. \quad (11)$$



Отсюда

$$\lim_{t \rightarrow \infty} \frac{F(p)}{G(p)} 1 = \frac{F(0)}{G(0)}, \quad (12)$$

$$\lim_{t \rightarrow \infty} p^{-1} \frac{F(p)}{G(p)} 1 = \lim_{\lambda \rightarrow 0} \frac{F(\lambda)}{\lambda G(\lambda)} \quad (13)$$

при условии, что пределы справа существуют и что все нули  $G(p)$  имеют отрицательные действительные части.

**7.08.** В динамических приложениях  $G(p)$  часто является четной функцией  $p$  с простыми нулями в  $\pm in$ . Мы можем разделить  $F(p)$  на четную и нечетную части, так что

$$F(p) = \frac{1}{2} [F(p) + F(-p)] + \frac{1}{2} [F(p) - F(-p)] = S(p) + pT(p), \quad (14)$$

где  $S(p)$  и  $T(p)$  — четные функции; и далее

$$\frac{F(p)}{G(p)} 1 = \frac{S(p) + pT(p)}{G(p)} 1 = \frac{S(0)}{G(0)} + \sum \frac{S(in) + inT(in)}{inG'(in)} e^{int}. \quad (15)$$

Объединяя члены с  $e^{+int}$  и  $e^{-int}$  вместе, мы получим

$$\frac{S(0)}{G(0)} + \sum \frac{2S(in)}{inG'(in)} \cos nt - \sum \frac{2nT(in)}{inG'(in)} \sin nt, \quad (16)$$

но

$$\frac{d}{dp^2} G(p) = \frac{1}{2p} G'(p), \quad inG'(in) = -2n^2 \left[ \frac{d}{dp^2} G(p^2) \right]_{p^2 = -n^2}, \quad (17)$$

$$\frac{F(p)}{G(p)} 1 = \frac{S(0)}{G(0)} - \sum \frac{S(in) \cos nt - nT(in) \sin nt}{n^2 \left[ \frac{d}{dp^2} G(p^2) \right]_{p^2 = -n^2}}. \quad (18)$$

Так как  $S$  и  $T$  — полиномы от  $p$ , содержащие только члены четных степеней, то полученное выражение дает результат применения оператора  $F(p)/G(p)$  к 1 прямо в действительном виде через тригонометрические функции.

**7.09. Единичная функция Хевисайда.** Эта функция определяется соотношениями

$$H(t) = 0 \quad (t < 0); \quad H(t) = 1 \quad (t > 0). \quad (1)$$

Очевидно,

$$p^{-n} H(t) = p^{-n} 1 \quad (t > 0); \quad p^{-n} H(t) = 0 \quad (t < 0). \quad (2)$$

Поэтому если

$$F(p) 1 = f(t),$$

то

$$F(p) H(t) = f(t) \quad (t > 0), \quad F(p) H(t) = 0 \quad (t < 0), \quad (3)$$

и вообще

$$F(p)H(t) = f(t)H(t). \quad (4)$$

Нас обычно интересуют лишь положительные значения  $t$ , а в этом случае неважно, применяется ли оператор  $F(p)$  к 1 или к  $H(t)$ . Далее, вместо  $F(p)1$  или  $F(p)H(t)$  обычно пишут просто  $F(p)$  и опускают тот подразумеваемый факт, что оператор  $F(p)$  применяется к 1 или к  $H(t)$ .

### ПРИМЕРЫ

Решить следующие дифференциальные уравнения с заданными начальными условиями:

$$1. \frac{d^2x}{dt^2} + 4 \frac{dx}{dt} + 3x = 1, \quad x_0 = 3, \quad x_1 = -2;$$

$$2. \frac{d^2x}{dt^2} + 5 \frac{dx}{dt} + 6x = 12, \quad x_0 = 2, \quad x_1 = 0;$$

$$3. \frac{dx}{dt} + 3x = e^{-2t}, \quad x_0 = 0;$$

$$4. \frac{d^2x}{dt^2} + 4 \frac{dx}{dt} + 4x = t^2 e^{-2t}, \quad x_0 = 0, \quad x_1 = 0;$$

$$5. \frac{d^4y}{dx^4} + 6 \frac{d^3y}{dx^3} + 11 \frac{d^2y}{dx^2} + 6 \frac{dy}{dx} = 20e^{-2x} \sin x;$$

если дано, что

$$y = 6, \quad \frac{dy}{dx} = 0, \quad \frac{d^2y}{dx^2} = 4, \quad \frac{d^3y}{dx^3} = 0 \quad \text{при} \quad x = 0. \quad (\text{M/c, 1930.})$$

6. Решить уравнения:

$$\frac{dx}{dt} + 5x + 2y = e^{-t},$$

$$\frac{dy}{dt} + 2x + 2y = 0,$$

если дано, что  $x = 1$ ,  $y = 0$ , когда  $t = 0$ .

(Prelim., 1945.)

7. Если

$$F(p)1 = f(t),$$

то доказать, что

$$\left[ \frac{1}{p} F(p) - F'(p) \right] 1 = tf(t).$$

### ЛИТЕРАТУРА

1. Caque J., J. Math., (2) 9, 185—222 (1864).
2. Jeffreys H., Operational Methods in Mathematical Physics, 1927.

## ФИЗИЧЕСКИЕ ПРИЛОЖЕНИЯ ОПЕРАЦИОННОГО МЕТОДА

Кончай разговоры, пошли к лошадям.

*Английская поговорка*

**8.01. Зарядка конденсатора.** Электрический контур содержит источник тока, конденсатор, катушку с самоиндукцией и сопротивлением. Вначале контур разомкнут. Необходимо определить, как будет изменяться во времени заряд на пластинах конденсатора.

Пусть  $y$  — заряд конденсатора,  $t$  — время,  $C$  — емкость конденсатора,  $L$  — самоиндукция,  $R$  — сопротивление контура,  $E$  — электродвижущая сила элемента. Ток равен  $\dot{y}$ , и зарядка конденсатора создаст разность потенциалов  $y/C$ , стремящуюся противодействовать первоначальной ЭДС. Тогда  $y$  удовлетворяет дифференциальному уравнению

$$E - \frac{y}{C} = L\ddot{y} + R\dot{y}. \quad (1)$$

В начальный момент  $y$  и  $\dot{y}$  равны нулю. Отсюда вспомогательное уравнение будет просто

$$\left(Lp^2 + Rp + \frac{1}{C}\right)y = E, \quad (2)$$

и операторное решение равно

$$y = \frac{E}{Lp^2 + Rp + \frac{1}{C}} = \frac{E}{L(p + \alpha)(p + \beta)}. \quad (3)$$

Разложением на элементарные дроби приходим к

$$y = \frac{E}{L\alpha\beta} + \frac{Ee^{-\alpha t}}{L(-\alpha)(-\alpha + \beta)} + \frac{Ee^{-\beta t}}{L(-\beta)(-\beta + \alpha)} = \\ = EC + \frac{E}{L(\alpha - \beta)} \left( \frac{1}{\alpha} e^{-\alpha t} - \frac{1}{\beta} e^{-\beta t} \right). \quad (4)$$

Так как и  $\alpha + \beta$  и  $\alpha\beta$  положительны,  $\alpha$  и  $\beta$  либо оба действительны и положительны, либо являются комплексно сопряженными числами с положительными действительными частями.

В любом случае  $y$  стремится к пределу  $CE$ , как и следовало ожидать.

Заметим, что если контур не содержит емкости или самоиндукции, дифференциальное уравнение упростится:

$$R\dot{y} = E. \quad (5)$$

Следовательно, если найдено решение для обычного сопротивления, самоиндукция и емкость могут быть учтены заменой  $R$  на  $Lp + R + 1/Cp$ . Это выражение иногда называют оператором сопротивления и операционный метод — методом операторов сопротивления. Экспоненциальные члены в решении быстро становятся пренебрежимо малыми, хотя они важны в экспериментах, где мы хотим узнать, как быстро достигается стационарное состояние. Они часто называются *переходными процессами*.

**8.02. Переменная ЭДС, приложенная к катушке с самоиндукцией.** Пусть  $x$  — генерируемый ток. ЭДС равна  $v \cos nt$ , т. е. действительной части от  $ve^{int}$ . Тогда мы должны решить уравнение

$$L\dot{z} + Rz = ve^{int} = \frac{vp}{p - in},$$

и если  $z$  вначале равно нулю, начальные условия ничего не вносят во вспомогательное уравнение. Тогда операторное решение будет

$$z = \frac{vp}{(Lp + R)(p - in)} = \frac{v}{L} \left( \frac{e^{int}}{in + R/L} + \frac{e^{-Rt/L}}{-R/L - in} \right) = \frac{v(R - Lin)}{L^2n^2 + R^2} (e^{int} - e^{-Rt/L})$$

и его действительная часть равна

$$x = \frac{v}{L^2n^2 + R^2} (R \cos nt + Ln \sin nt - Re^{-Rt/L}).$$

Два первых члена дают периодические колебания тока, сдвинутые по фазе относительно ЭДС. Последний член описывает переходный процесс, который затухает спустя время порядка  $L/R$ . Гармоническая часть решения имеет амплитуду  $v/(L^2n^2 + R^2)^{1/2}$  и может быть записана как действительная часть

$$\frac{ve^{int}}{R + Lin}.$$

Это основа так называемой „векторной диаграммы“ в электротехнике, не имеющей ничего общего с векторами, а являю-

щейся особым случаем геометрического представления комплексных величин. Этот случай обычно связывают с именем Агрона, хотя он был рассмотрен Уоллисом и Весселом.

**8.03. Разряд конденсатора в одном из двух взаимно связанных контуров.** Положим, что мы имеем два одинаковых контура, каждый с самоиндукцией  $L$ , конденсатором емкостью  $C$  и пренебрежимо малым сопротивлением. Один конденсатор вначале заряжен до  $x_0$ , другой не заряжен. Коэффициент взаимной индукции равен  $M$ . Первый контур замыкается; требуется найти результирующие изменения зарядов.

Пусть  $CL = 1/\alpha^2$ ,  $M = L\beta$ ; если  $x$ ,  $y$  обозначают заряды конденсаторов в двух контурах, то

$$L(\ddot{x} + \beta\ddot{y} + \alpha^2 x) = 0, \quad (1)$$

$$L(\beta\ddot{x} + \ddot{y} + \alpha^2 y) = 0. \quad (2)$$

Вначале

$$x = x_0, \quad \dot{x} = 0, \quad y = 0, \quad \dot{y} = 0, \quad (3)$$

и вспомогательные уравнения будут

$$(p^2 + \alpha^2)x + \beta p^2 y = p^2 x_0, \quad (4)$$

$$\beta p^2 x + (p^2 + \alpha^2)y = \beta p^2 x_0. \quad (5)$$

Выполнив алгебраические действия, получим

$$\frac{x}{p^2(p^2 + \alpha^2) - \beta^2 p^4} = \frac{y}{\beta p^2(p^2 + \alpha^2) - \beta p^4} = \frac{x_0}{(p^2 + \alpha^2)^2 - \beta^2 p^4} \quad (6)$$

и

$$\frac{x}{(1 - \beta^2)p^4 + p^2\alpha^2} = \frac{y}{\beta\alpha^2 p^2} = \frac{x_0}{[(1 + \beta)p^2 + \alpha^2][(1 - \beta)p^2 + \alpha^2]}. \quad (7)$$

Тогда

$$\begin{aligned} x &= \frac{p^2(1 - \beta^2) + \alpha^2}{[(1 + \beta)p^2 + \alpha^2][(1 - \beta)p^2 + \alpha^2]} p^2 x_0 = \\ &= \left[ \frac{(1 + \beta)p^2}{(1 + \beta)p^2 + \alpha^2} + \frac{(1 - \beta)p^2}{(1 - \beta)p^2 + \alpha^2} \right] \frac{1}{2} x_0 = \\ &= \left[ \cos \frac{\alpha}{\sqrt{1 + \beta}} t + \cos \frac{\alpha}{\sqrt{1 - \beta}} t \right] \frac{1}{2} x_0, \end{aligned} \quad (8)$$

$$\begin{aligned} y &= \frac{\beta\alpha^2}{[(1 + \beta)p^2 + \alpha^2][(1 - \beta)p^2 + \alpha^2]} p^2 x_0 = \\ &= \left[ \frac{(1 + \beta)p^2}{(1 + \beta)p^2 + \alpha^2} - \frac{(1 - \beta)p^2}{(1 - \beta)p^2 + \alpha^2} \right] \frac{1}{2} x_0 = \\ &= \left[ \cos \frac{\alpha}{\sqrt{1 + \beta}} t - \cos \frac{\alpha}{\sqrt{1 - \beta}} t \right] \frac{1}{2} x_0. \end{aligned} \quad (9)$$

Если мы примем

$$\frac{\alpha}{\sqrt{1-\beta}} = \gamma + \delta, \quad \frac{\alpha}{\sqrt{1+\beta}} = \gamma - \delta, \quad (10)$$

решения будут иметь вид

$$x = x_0 \cos \gamma t \cos \delta t, \quad y = x_0 \sin \gamma t \sin \delta t. \quad (11)$$

Если  $M$  мало,  $\delta$  тоже мало, и возмущение состоит из быстрых колебаний с периодом  $2\pi/\gamma$  и амплитудами, варьирующими так, что колебания передаются из одного контура в другой за время  $\pi/2\delta$ . Таким образом, это случай биений из-за слабой связи. Хорошо известны подобные явления у двух маятников, висящих на одной и той же не абсолютно жесткой опоре; один маятник влияет на другой, смещая опору. Происходит то же явление передачи колебаний от одного маятника к другому через регулярные интервалы.

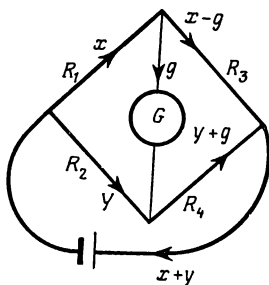


Рис. 31.

Если связь сильная, так что  $\beta \approx 1$ , два периода  $2\pi \sqrt{(1 \pm \beta)}/\alpha$  очень сильно различаются, и изменение заряда состоит из быстрого колебания, наложенного на

медленное колебание равной амплитуды. Медленное колебание имеет одну и ту же фазу в обоих контурах; фазы быстрых колебаний противоположны.

#### 8.04. Определение самоиндукции по методу Римингтона [1, 2].

В этом методе неизвестная индуктивность помещается в первое плечо мостика Уитстона; четвертое плечо шунтировано емкостью известной величины.

Рассмотрим сначала обычный мостик Уитстона с сопротивлениями в плечах  $R_1, R_2, R_3, R_4$ ; сопротивление гальванометра  $G$ ; сопротивление батареи и проводников  $b$ ;  $x$  — ток в  $R_1$ ,  $y$  — ток в  $R_2$ ,  $g$  — ток через гальванометр. Тогда

$$R_1 x - R_2 y + G g = 0, \quad (1)$$

$$R_3 x - R_4 y - (R_3 + R_4 + G) g = 0, \quad (2)$$

$$b(x + y) + R_2 y + R_4(y + g) = E, \quad (3)$$

и, решая (1) и (2), мы находим

$$\frac{g}{R_2 R_3 - R_1 R_4} = \frac{x + y}{G(R_1 + R_2 + R_3 + R_4) + (R_1 + R_2)(R_3 + R_4)}. \quad (4)$$

Если  $g$  мало по сравнению с  $x$  и  $y$ , мы имеем приближенно

$$x + y = \frac{E}{b + \frac{(R_1 + R_3)(R_2 + R_4)}{R_1 + R_2 + R_3 + R_4}}, \quad (5)$$

$$g = \frac{(x + y)(R_2 R_3 - R_1 R_4)}{G(R_1 + R_2 + R_3 + R_4)}. \quad (6)$$

Важная особенность схемы в том, что  $g = 0$ , если  $R_2 R_3 = R_1 R_4$ , независимо от точности аппроксимации (5).

Согласно нашему первому результату, мы можем учесть самоиндукцию в первом плече, заменив  $R_1$  на  $Lp + R_1$ . Пусть схема четвертого плеча такая, как на рисунке. Сопротивление главного проводника равно  $R_4$ , шунтированной части  $r$ . Шунт имеет сопротивление  $S$ . Тогда эффективное сопротивление целого плеча

$$R_4 - r + \frac{rS}{r + S} = R_4 - \frac{r^2}{r + S}.$$

Если шунт содержит емкость  $C$ , мы учтем это, заменив  $S$  на  $S + 1/Cp$ . Отсюда в формуле (6) для  $g$  мы должны заменить  $R_1$  на  $Lp + R_1$  и  $R_4$  на

$$R_4 - \frac{r^2}{r + S + 1/Cp} = R_4 - \frac{r^2 Cp}{(r + S)Cp + 1}. \quad (7)$$

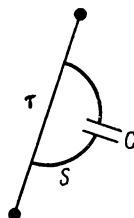


Рис. 32.

Результат описывает ток через гальванометр, когда контур с источником внезапно замыкается.

Можно показать, что в реальных условиях  $g$  не может быть равным нулю во все моменты времени. Достаточным условием для этого было бы тождественное равенство нулю модифицированного оператора  $R_2 R_3 - R_1 R_4$ ; тогда  $g$  равно нулю при любых значениях остающегося множителя. Нетрудно показать, что это условие является также и необходимым. Тогда этот множитель заменяется на

$$R_2 R_3 - (Lp + R_1) \left( R_4 - \frac{r^2 Cp}{(r + S)(Cp + 1)} \right). \quad (8)$$

Перемножив и приравняв коэффициенты при степенях  $p$  нулю, мы найдем

$$R_4(r + S) = r^2, \quad (9)$$

$$-LR_4 + (R_2 R_3 - R_1 R_4)(r + S)C + R_1 r^2 C = 0, \quad (10)$$

$$R_2 R_3 - R_1 R_4 = 0. \quad (11)$$

Согласно схеме,  $r \leq R_4$ ,  $S \geq 0$ . Отсюда соотношение (9) может выполняться только, если  $r = R_4$  и  $S = 0$ . Шунтирующая проволока

должна быть прикреплена к концам  $R_4$  и иметь нулевое сопротивление. Выражение (11) — обычное условие для балансировки; подставляя его в (10), мы имеем

$$L = R_1 R_4 C. \quad (12)$$

Эти условия не могут быть удовлетворены полностью. Но изменения тока при замыкании батарейного контура настолько быстрые, что обычный гальванометр не успевает на них реагировать. Если стрелка гальванометра устанавливается на нулевом значении, мы имеем обычные условия для балансировки, но если результирующий ток меняет направление, он будет действовать на гальванометр как импульс и произойдет баллистический отброс стрелки гальванометра. Условие, что  $g$  стремится к нулю и не произойдет баллистический отброс стрелки, следующее:

$$g \rightarrow 0, \quad \int_0^{\infty} g dt = 0. \quad (13)$$

Но оно выполнено, согласно 7.07, если постоянный член и член с  $p$  в операторной форме для  $g$  исчезают; и опять, независимо от приближения (5), мы можем использовать (8). Формальный предел при  $p \rightarrow 0$  будет

$$R_2 R_3 - R_1 R_4 = 0, \quad (14)$$

тот же, что и ранее. Коэффициент при  $p$  равен

$$L R_4 - R_1 r^2 C; \quad (15)$$

исчезновение этого коэффициента и есть условие отсутствия баллистического отброса стрелки. Условие (9), которое связано с членами с  $p^2$ , больше не возникает. Метод, следовательно, заключается в следующем: вначале сбалансируем мостик обычным образом, удовлетворив этим (14); затем будем присоединять шунт с емкостью к различным точкам в плече  $R_1$  так, чтобы варьировать  $r$ . Когда найдено положение, при котором не происходит баллистического отброса стрелки,  $r$  найдено и из (15) определяется  $L$ .

**8.05. Сейсмограф.** В принципе большинство сейсмографов — маятники Эйлера, т. е. такие маятники, у которых опора жестко связана с землей; при движении грунта точка опоры смещается по горизонтали и маятник возмущается. Сейсмограф отличается от маятника Эйлера, описываемого в учебниках динамики, двумя особенностями. Вместо свободы совершать колебания в вертикальной плоскости он вынужден качаться, как



ворота, вокруг почти вертикальной оси, так что период существенно удлинен. Для получения сил трения, пропорциональных относительной скорости, вводится демпфирование жидкостью или электромагнитное затухание. Смещение массы по отношению к почве удовлетворяет уравнению

$$\ddot{x} + 2\kappa\dot{x} + n^2x = \lambda\dot{\xi}, \quad (1)$$

где  $\xi$  — смещение грунта и  $\kappa$ ,  $n$ ,  $\lambda$  — константы инструмента. Некоторые приборы, такие, как сейсмографы Вихерта и Вуда — Андерсона, не используют принцип маятника Эйлера, но описываются тем же уравнением. Другие приборы предназначены для записи вертикальных движений грунта; для этого требуется тяжелая масса с упругим подвесом, что удобно для короткопериодных колебаний грунта, подобных регистрируемым в сейсмической разведке. Труднее сконструировать инструмент для регистрации более длинных периодов, где  $x$  удовлетворяло бы линейному дифференциальному уравнению. Эти трудности преодолены несколькими различными способами, и соответствующие дифференциальные уравнения имеют форму (1).

Первая цель прибора: регистрировать с максимально достижимой точностью время любого внезапного изменения скорости грунта. Вторая цель: когда такое изменение записано, прибор должен как можно быстрее вернуться к первоначальной позиции, чтобы быть готовым записать любое более позднее возмущение.

Положим сначала, что грунт внезапно приобретает конечную скорость, равную 1. Тогда  $\dot{\xi}$  меняется скачком от 0 до 1 и, следовательно,  $\dot{x}$  от 0 до  $\lambda$ . Начальные условия имеют вид

$$x = 0, \quad \dot{x} = \lambda, \quad (2)$$

и вспомогательное уравнение

$$(p^2 + 2\kappa p + n^2)x = \lambda p \quad (t > 0), \quad (3)$$

$$p^2 + 2\kappa p + n^2 = (p + \alpha)(p + \beta). \quad (4)$$

Тогда

$$x = \frac{\lambda p}{(p + \alpha)(p + \beta)} \quad (t > 0), \quad (5)$$

$$x = \frac{\lambda}{\alpha - \beta} (e^{-\beta t} - e^{-\alpha t}). \quad (6)$$

Зарегистрированное смещение  $x$  вначале возрастает с конечной скоростью  $\lambda$ , достигает максимума  $\lambda(\beta/\alpha)^{1/(\alpha-\beta)}$  спустя время  $\frac{1}{\alpha - \beta} \ln \frac{\alpha}{\beta}$  и затем асимптотически стремится к нулю.

Если  $\alpha$  и  $\beta$  действительны и  $\beta < \alpha$ , главное влияние на запись через значительный интервал времени имеет член  $e^{-\beta t}$ : чтобы ограничить эффект возмущения возможно более коротким интервалом времени, мы должны максимально увеличить  $\beta$ . Но

$$\beta = \kappa - \sqrt{\kappa^2 - n^2} = \frac{n^2}{\kappa + \sqrt{\kappa^2 - n^2}}, \quad (7)$$

и для данного  $n$  максимальное  $\beta$  (при условии, что оно действительно и, следовательно,  $\kappa \geq n$ ) достигается, когда  $\kappa = n$ .

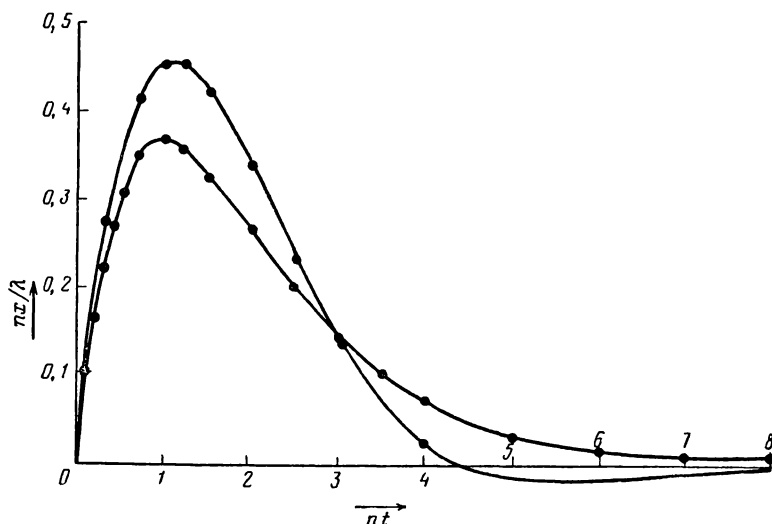


Рис. 33.

Это условие аперiodичности. При этом решение сводится к

$$x = \frac{\lambda p}{(p+n)^2} = \lambda t e^{-nt} \quad (t > 0). \quad (8)$$

Максимальное смещение наступает при этом в момент времени  $1/n$  и равно  $\lambda/en$ .

Если  $\kappa < n$ , мы можем положить

$$n^2 - \kappa^2 = \gamma^2. \quad (9)$$

Тогда (6) принимает вид

$$x = \frac{\lambda}{\gamma} e^{-\kappa t} \sin \gamma t, \quad (10)$$

и движение затухает тем быстрее, чем больше  $\kappa$  в пределах рассмотренного диапазона. Следовательно, аperiodическое состояние  $\kappa = n$  дает наименьшее движение на больших временах для данного  $n$ .

Однако на практике  $\kappa$  обычно делают несколько меньше  $n$ . Так, например, в приборе Милна — Шоу  $\kappa \approx 0,7n$ . Тогда движение носит колебательный характер, но отношение первого максимального отклонения ко второму равно  $e^{\kappa\pi/\gamma} \approx 20$ . Но  $x$  проходит через нуль спустя  $\pi/\gamma$ , или  $\sim 4/n$ , после начала и после этого составляет малую долю первого максимума. Уменьшение затухания на больших временах менее важно, чем быстрое возвращение к нулю после первого максимума. Время первого максимума, отсчитанное от начала движения, равно 1,  $1/n$ ; для аperiodического прибора оно равно  $1/n$  и для недемпфированного сейсмографа составляет  $1,57/n$ .

Подобным образом устроен и сейсмограф Голицына, но движение маятника не записывается непосредственно, а генерирует за счет электромагнитной индукции электрический ток, проходящий через гальванометр. Если  $x$  — смещение маятника и  $y$  — смещение зеркальца гальванометра, дифференциальные уравнения имеют вид

$$\ddot{x} + 2\kappa_1\dot{x} + n_1^2x = \lambda\ddot{\xi}, \quad (11)$$

$$\ddot{y} + 2\kappa_2\dot{y} + n_2^2y = \mu\dot{x}, \quad (12)$$

где мы пренебрегаем реакцией индуцированного тока на маятник. Положив опять, что движение грунта начинается с единичной скоростью, мы имеем

$$y = \frac{\lambda_1 p^2}{(p^2 + 2\kappa_1 p + n_1^2)(p^2 + 2\kappa_2 p + n_2^2)} \quad (t > 0). \quad (13)$$

В первых приборах  $\kappa$  и  $n$  были сделаны одинаковыми в обеих взаимодействующих системах, каждая из которых являлась аperiodической, так что

$$\kappa_1 = \kappa_2 = n_1 = n_2 = n. \quad (14)$$

Тогда

$$y = \frac{\lambda \mu p^2}{(p + n)^4} \quad (t > 0),$$

$$y = \lambda \mu \frac{d}{dt} \left( \frac{t^3}{3!} e^{-nt} \right) = \frac{1}{2} \lambda \mu \left( t^2 - \frac{1}{3} n t^3 \right) e^{-nt}. \quad (15)$$

Следовательно, индикатор начинает двигаться не с конечной скоростью, как маятник, а с конечным ускорением. Максимальное смещение наступает через  $(3 - \sqrt{3})/n = 1,27/n$  от начала, зеркальце проходит через положение равновесия в момент  $3/n$ .

и максимальное смещение в противоположном направлении происходит в момент  $4,73\pi$ . Затем зеркальце асимптотически возвращается к положению равновесия. Отношение двух экстремумов смещения равно  $e^{2\sqrt{3}}/(2 + \sqrt{3})^2 = 2,3$ . По сравнению с частично затухающим прибором с прямой записью такого типа, как в приборе Милна — Шоу, у прибора Голицына первый максимум достигается немного позднее, первый нуль — несколько раньше и отношение последующего экстремума смещения к первому больше. Из рис. 34 видно, что, несмотря на

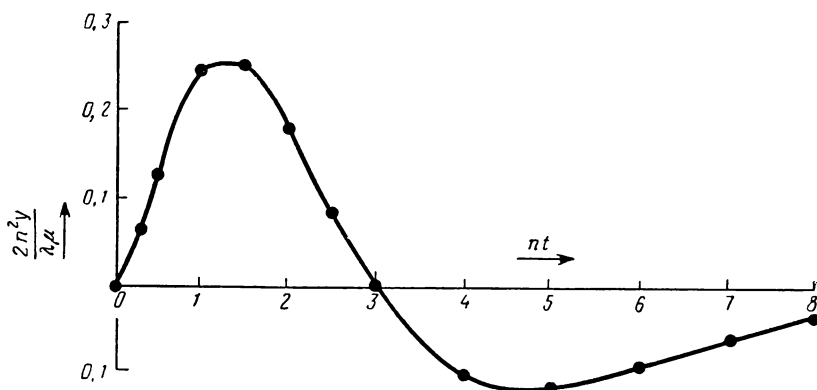


Рис. 34.

то что  $y$  растет пропорционально  $t^2$ , а не  $t$  при малых  $t$ , быстрое увеличение  $y$  наступает весьма скоро, так что можно очень точно определить начало движения.

В более поздних модификациях прибора Голицына соотношение (14) было нарушено уменьшением затухания и выбором гальванометра с меньшим собственным периодом, чем маятник. Для гармонических движений грунта это уменьшает зависимость увеличения от периода вынуждающей силы. Однако, по-видимому, не удастся уменьшить лишние качания после импульсивного изменения скорости грунта. В некоторых последних конструкциях нельзя уже пренебрегать реакцией маятника на индуцированный ток \*).

**8.06. Резонанс.** Простой маятник, находящийся первоначально в состоянии равновесия, возмущается в течение конечного интервала времени гармонической силой с периодом, рав-

---

\*) Более полная теория дана в [3].

ным собственному периоду маятника. Нужно найти движение маятника после прекращения действия силы.

Дифференциальное уравнение будет

$$\ddot{x} + n^2 x = f \sin nt = f \frac{np}{p^2 + n^2} \quad (0 < t < T)$$

с  $x = 0$ ,  $\dot{x} = 0$  при  $t = 0$ . Тогда

$$x = f \frac{np}{(p^2 + n^2)^2} = \frac{f}{2n^2} (\sin nt - nt \cos nt).$$

Движение можно рассматривать как гармоническое с непрерывно увеличивающейся амплитудой. Пусть возмущение действует в течение времени  $T = r\pi/n$ , где  $r$  — целое. В конце этого времени

$$x = -\frac{f}{2n^2} r\pi(-1)^r, \quad \dot{x} = 0.$$

Последующее движение описывается тогда формулой

$$x = -(-1)^r \frac{r\pi f}{2n^2} \cos(nt - r\pi) = -\frac{r\pi f}{2n^2} \cos nt$$

и представляет собой гармоническое движение с амплитудой, пропорциональной длительности возмущения.

Линейные дифференциальные уравнения в динамике получаются, если опустить в разложении квадраты и более высокие степени смещения. В результате при совпадении периода вынуждающей силы и собственного периода амплитуда будет расти до такого уровня, при котором уже необходимо учесть отброшенные члены.

**8.07.** Три частицы с массами  $m$ ,  $\frac{21}{20}m$  и  $m$  прикреплены последовательно к легкой растянутой струне длиной  $4l$  так, что делят ее на равные интервалы. Одна из частиц массы  $m$  возбуждается поперечным (к струне) импульсом  $I$ . Найти последующее движение средней частицы. (Переходные экзамены в колледже, 1923.)

Если  $x_1$ ,  $x_2$ ,  $x_3$  — смещения трех частиц и  $P$  — натяжение, мы найдем обычным путем уравнения движения

$$\begin{aligned} \ddot{x}_1 &= -\lambda(2x_1 - x_2), \\ \frac{21}{20}\ddot{x}_2 &= -\lambda(-x_1 + 2x_2 - x_3), \\ \ddot{x}_3 &= -\lambda(-x_2 + 2x_3), \end{aligned} \quad (1)$$

где  $\lambda = P/ml$ . Вначале, когда все смещения нулевые,  $\dot{x}_2 = \dot{x}_3 = 0$ ;  $\dot{x}_1 = l/m$ . Тогда вспомогательные уравнения будут

$$\begin{aligned}(p^2 + 2\lambda)x_1 - \lambda x_2 &= \frac{pl}{m}, \\ -\lambda x_1 + \left(\frac{21}{20}p^2 + 2\lambda\right)x_2 - \lambda x_3 &= 0, \\ -\lambda x_2 + (p^2 + 2\lambda)x_3 &= 0.\end{aligned}\quad (2)$$

Так как нам нужно найти только изменения  $x_2$ , то мы исключим  $x_1$  и  $x_3$ . Мы имеем

$$x_1 = \frac{\lambda}{p^2 + 2\lambda}x_2 + \frac{p}{p^2 + 2\lambda}\frac{l}{m}, \quad x_3 = \frac{\lambda}{p^2 + 2\lambda}x_2 \quad (3)$$

и тогда

$$\left(\frac{21}{20}p^2 + 2\lambda - \frac{2\lambda^2}{p^2 + 2\lambda}\right)x_2 = \frac{\lambda p}{p^2 + 2\lambda}\frac{l}{m}. \quad (4)$$

Упрощая, получим

$$(7p^2 + 4\lambda)(3p^2 + 10\lambda)x_2 = \frac{20\lambda pl}{m}. \quad (5)$$

Тогда

$$x_2 = \frac{20}{58}\frac{l}{m}\left(\frac{7p}{7p^2 + 4\lambda} - \frac{3p}{3p^2 + 10\lambda}\right), \quad (6)$$

$$x_2 = \frac{10}{29}\frac{l}{m}\left(\frac{\sin \alpha t}{\alpha} - \frac{\sin \beta t}{\beta}\right), \quad (7)$$

где

$$\alpha^2 = \frac{4}{7}\lambda, \quad \beta^2 = \frac{10}{3}\lambda. \quad (8)$$

Движения других частиц могут быть найдены при желании с помощью (3) и разложения на элементарные дроби. Они будут содержать члены с теми же периодами, как в (7), но, кроме того, члены с периодом  $2\pi/\sqrt{2\lambda}$ . Они соответствуют третьей нормальной моде, в которой средняя частица не движется. Это иллюстрирует одно большое преимущество операционного метода. Мы ищем только  $x_2$ , и метод прямо дает его. Пользуясь обычным методом, мы должны были бы определить амплитуды всех трех нормальных мод отдельно, даже хотя одна из них не имеет отношения к поставленному вопросу.

**8.08. Малые колебания в динамике.** Рассмотрим динамическую систему с лагранжевой функцией вида

$$2L = a_{rs}\dot{x}_r\dot{x}_s - c_{rs}x_rx_s, \quad (1)$$

так что уравнения движения будут

$$a_{rs}\ddot{x}_s + c_{rs}x_s = S_r, \quad (2)$$

где  $S_r$  — компонента любой обобщенной силы, приложенной к  $x_r$  и не учитываемой в потенциальной энергии. Если система первоначально находится в покое и только одно из  $S_r$  отлично от нуля, мы можем написать

$$e_{ms} = a_{ms}p^2 + c_{ms}, \quad (3)$$

и вспомогательные уравнения имеют вид

$$e_{ms}x_s = 0 \quad (m \neq r), \quad e_{ms}x_s = S_r \quad (m = r). \quad (4)$$

Обозначая через  $\Delta$  определитель  $e_{ms}$  и через  $E_{rs}$  алгебраическое дополнение  $e_{rs}$  в этом определителе, мы получаем операторное решение

$$x_s = \frac{E_{rs}}{\Delta} S_r. \quad (5)$$

Отметим, что  $r$  — частный индекс, а не индекс, по которому проведено суммирование. Далее, определитель  $\Delta$  симметрический, так что  $E_{rs} = E_{sr}$ . Таким образом, сила, действующая по координате  $x_r$ , будет производить точно такие же изменения в  $x_s$ , какие та же сила, действующая по  $x_s$ , производит в  $x_r$ . Мы получили теорему взаимности, применимую ко всем негирскопическим и свободным от трения системам.

Легко видеть, что трение не повлияет на результаты, если оно может быть выражено диссипативной функцией  $F = \frac{1}{2} b_{rs} \dot{x}_r \dot{x}_s$ .

В частности, результат справедлив для электрических цепей.

Теперь предположим, что сила сводится к импульсу  $J_r$  в момент  $t=0$  и что  $\Delta$  не имеет кратных множителей; мы можем написать

$$\Delta = A \prod (p^2 + \alpha^2) \quad (6)$$

и заменить  $S_r$  на  $pJ_r$ . Тогда

$$x_s = \frac{E_{rs}p}{A \prod (p^2 + \alpha^2)} J_r = \sum_a \frac{E_{rs}(-\alpha^2)pJ_r}{\prod'(-\alpha^2)p^2 + \alpha^2}, \quad (7)$$

где  $E_{rs}(-\alpha^2)$  и  $\prod'(-\alpha^2)$  обозначают результаты подстановки  $-\alpha^2$  вместо  $p^2$  в  $E_{rs}$  и  $d\Delta/dp^2$ . Отсюда

$$x_s = \sum_a \frac{E_{rs}(-\alpha^2)}{\alpha \prod'(-\alpha^2)} J_r \sin \alpha t \quad (8)$$

Отдельные члены имеют разные периоды, а члены с одним и тем же периодом в различных координатах образуют нормальную моду системы.

Непосредственным следствием отсюда является то, что для некоторых  $s$  и  $\alpha$ , скажем  $\alpha_1$ ,  $E_{rs}(-\alpha^2) = 0$  для всех  $r$ , а  $x_s$  не содержит членов с  $\sin \alpha_1 t$ , каков бы ни был прикладываемый импульс. Другими словами, если  $x_s$  — смещение частицы системы, в данной моде мы имеем узел. Но тогда, если мы рассматриваем импульс  $J_s$ , приложенный к  $x_s$ , мы будем иметь

$$x_r = \sum_{\alpha} \frac{E_{sr}(-\alpha)^2}{\alpha \prod' (-\alpha^2)} J_s \sin \alpha t, \quad (9)$$

и опять, так как  $E_{sr} = E_{rs}$ , член с  $\sin \alpha_1 t$  будет иметь нулевой коэффициент в каждой координате. Отсюда следует другая общая теорема взаимности: нельзя возбудить моду импульсами силы в любом узле этой моды. Аналогичным образом можно показать, что, если начальные значения координат определены начальными условиями, но скорости равны нулю, последующие значения координат содержат члены с множителями  $E_{rs}(-\alpha^2) \cos \alpha t$  и начальное смещение в узле любой моды не оказывает влияния на такие члены этой моды в последующем движении.

Этот принцип, приложенный к непрерывной системе, позволяет провести критическую проверку вопроса о существовании глубокофокусных землетрясений. Большинство землетрясений происходит на глубинах не свыше 50 км и возбуждает, кроме волн, проходящих сквозь землю, два типа поверхностных волн, теоретически объясненных Рэлеем и Лявом. Они похожи на волны в глубокой воде тем, что смещения быстро затухают с глубиной и незначительны на глубинах больше одной длины волны, которая в этом случае равна примерно 50—100 км. Согласно этому принципу, они не должны возбуждаться возмущениями на больших глубинах. Тэрнер определил по времени пробега объемных волн, что некоторые землетрясения происходят на глубинах до 400 км, однако имеются возможности и для другой интерпретации. Изучение сейсмограмм этих землетрясений, проведенное Стоунли, показало, что поверхностные волны на них отсутствуют. Этот факт необъясним при других предположениях, но вполне вытекает из принципа взаимности, если эти землетрясения произошли на больших глубинах.

**8.09. Случай кратных корней.** При исследовании колебаний динамических систем относительно положения равновесия обычный метод поиска решений в форме  $x_s = \lambda_s e^{vt}$  встречает труд-



ности, когда детерминантное уравнение для  $\gamma^2$  имеет кратные корни. Обычно, если мы имеем систему дифференциальных уравнений с  $n$  переменными и исключаем их последовательно, мы получаем уравнение для одного выбранного переменного. Если мы подставим в него  $e^{\gamma t}$ , то получим уравнение для  $\gamma$ ; в случае кратного корня мы будем иметь второе решение  $te^{\gamma t}$ . Если это происходит в теории малых колебаний, кратное значение  $\gamma$  ведет к членам в форме  $t \cos \kappa t$ ,  $t \sin \kappa t$  ( $\kappa = i\gamma$ ) и, за исключением случаев специальных начальных условий, малые колебания будут неограниченно расти. В действительности это никогда не происходит, а если бы произошло, то противоречило бы основному принципу, что потенциальная энергия минимальна в положении равновесия и что, если начальные смещения и скорости достаточно малы, но не равны нулю, существует предел, ограничивающий допустимые смещения. Это озадачило Лапласа; объяснение было дано Раусом [4] и Хевисайдом [5]. Если система не диссипативна и кратных корней нет, то, согласно 4.082 и 4.09, нули минора любого элемента на главной диагонали находятся между нулями исходного определителя, и если определитель  $\Delta$  содержит множитель  $(p^2 + \alpha^2)^k$ , то каждый первый минор содержит множитель  $(p^2 + \alpha^2)^{k-1}$ . Когда мы оцениваем вклад начальных условий в операторное решение, из 7.061 (7) имеем

$$x_m = \frac{E_{rm}}{\Delta} a_{rs} (p^2 u_s + p v_s);$$

множитель  $(p^2 + \alpha^2)^{k-1}$  сократится, и в знаменателе останется только множитель  $(p^2 + \alpha^2)$ . То же будет с каждым кратным корнем, и результат будет содержать только члены в форме  $\cos \alpha t$  и  $\sin \alpha t$ . Изменения  $u_s$  и  $v_s$  будут влиять на отношения коэффициентов при этих тригонометрических множителях для различных координат; при этом не появляются члены типа  $t \cos \alpha t$  и  $t \sin \alpha t$ .

**8.10. Диссипативные и гироскопические системы.** В таких системах теорема о разделении корней может не выполняться. Тогда операторное решение может иметь кратные множители в знаменателе и в решениях появятся члены вида  $te^{-\alpha t}$ ,  $t \cos \alpha t$ ,  $t \sin \alpha t$ . Мы рассмотрели в качестве простого примера такой системы аperiодический сейсмограф. Эти члены не повлияют на устойчивость, если недемпфированная система устойчива и не гироскопическая, так как решения экспоненциально уменьшаются и стремятся к 0 с ростом  $t$ . Но если система поддерживается в устойчивом состоянии только гироскопическими

силами, совпадение корней может нарушить устойчивость. Положим, что уравнения для двух координат  $x_1$ ,  $x_2$  имеют вид

$$\ddot{x}_1 - b\dot{x}_2 + c_1x_1 = 0, \quad \ddot{x}_2 + b\dot{x}_1 + c_2x_2 = 0. \quad (1)$$

Положим

$$x_1 = \lambda_1 e^{\gamma t}, \quad x_2 = \lambda_2 e^{\gamma t}. \quad (2)$$

Мы видим, что  $\gamma$  должно удовлетворять следующему детерминантному уравнению:

$$\begin{vmatrix} \gamma^2 + c_1 & -b\gamma \\ b\gamma & \gamma^2 + c_2 \end{vmatrix} = 0, \quad (3)$$

т. е.

$$\gamma^4 + (c_1 + c_2 + b^2)\gamma^2 + c_1c_2 = 0. \quad (4)$$

Необходимое условие устойчивости заключается в том, что оба значения  $\gamma^2$  должны быть действительными и меньше 0. Отсюда  $c_1c_2 > 0$  и имеем два случая в зависимости от того, положительны или отрицательны  $c_1$  и  $c_2$ .

Рассмотрим сначала случай, когда  $c_1$  и  $c_2$  положительны. Тогда система устойчива даже при  $b = 0$ . Условие для кратных корней имеет вид

$$(c_1 + c_2 + b^2)^2 = 4c_1c_2, \quad (5)$$

т. е.

$$(c_1 - c_2)^2 + 2b^2(c_1 + c_2) + b^4 = 0. \quad (6)$$

При  $c_1, c_2 > 0$  это может быть выполнено только, если  $b = 0$  и, следовательно,  $c_1 = c_2$ . Отсюда, если (3) имеет кратные корни, равные  $\gamma^2$ , все элементы определителя исчезают. Этого следовало ожидать, так как  $c_1x_1^2 + c_2x_2^2$  в этом случае положительная форма и теорема о разделении корней справедлива для гироскопической системы, когда члены  $c_{rs}x_r x_s$  являются положительной формой.

Если, однако,  $c_1$  и  $c_2$  оба отрицательны, то (6) может быть выполнено при условии, что

$$\begin{aligned} b^2 &= -(c_1 + c_2) \pm \sqrt{(c_1 + c_2)^2 - (c_1 - c_2)^2} = \\ &= -(c_1 + c_2) \pm 2\sqrt{c_1c_2} = [\sqrt{-c_1} \pm \sqrt{-c_2}]^2. \end{aligned} \quad (7)$$

Таким образом, детерминантное уравнение может в этом случае иметь равные корни при  $b$ , не равном 0. Оно сведется теперь к

$$\gamma^4 \pm 2\sqrt{c_1c_2}\gamma^2 + c_1c_2 = 0$$

и

$$\gamma^2 = \mp \sqrt{c_1c_2}. \quad (8)$$

В этом случае отдельные элементы (3) не исчезают, хотя  $\gamma^2$  все еще действительно и отрицательно. При нижних знаках в (7) и (8)  $\gamma^2$  будет положительным, а система с очевидностью неустойчивой. Поэтому мы возьмем минус в (8) и плюс в (7). Чтобы оценить, что произойдет с системой в этом случае, если  $x_1 = u_1$ ,  $\dot{x}_1 = 0$ ,  $x_2 = 0$ ,  $\dot{x}_2 = 0$  при  $t = 0$ , мы запишем

$$c_1 = -\alpha^2, \quad c_2 = -\beta^2, \quad b = \alpha + \beta; \quad (9)$$

$$\begin{aligned} (p^2 - \alpha^2)x_1 - (\alpha + \beta)px_2 &= p^2u_1, \\ (\alpha + \beta)px_1 + (p^2 - \beta^2)x_2 &= (\alpha + \beta)pu_1. \end{aligned} \quad (10)$$

Операторное решение будет

$$x_1 = \frac{p^4 + (\alpha^2 + 2\alpha\beta)p^2}{(p^2 + \alpha\beta)^2} u_1, \quad x_2 = -\frac{(\alpha + \beta)\alpha^2 p}{(p^2 + \alpha\beta)^2} u_1, \quad (11)$$

и

$$\frac{x_1}{u_1} = \frac{p^2}{p^2 + \alpha\beta} + \frac{\alpha(\alpha + \beta)p^2}{(p^2 + \alpha\beta)^2} = \cos \sqrt{\alpha\beta} t + \frac{\alpha(\alpha + \beta)}{2\sqrt{\alpha\beta}} t \sin \sqrt{\alpha\beta} t, \quad (12)$$

$$\frac{x_2}{u_1} = -\frac{(\alpha + \beta)\alpha^2}{2(\alpha\beta)^{3/2}} [\sin \sqrt{\alpha\beta} t - \sqrt{\alpha\beta} t \cos \sqrt{\alpha\beta} t]. \quad (13)$$

Следовательно, система сохраняет устойчивость только благодаря гироскопическим членам; и если их коэффициенты таковы, что периоды равны, устойчивость может быть нарушена, так как амплитуда возмущения будет линейно расти во времени. Это соответствует волчку с  $C^2 n^2 = 4AMgh$ .

**8.101. Гироскопическая система с малым трением.** Мы не будем здесь рассматривать произвольное начальное возмущение, а ограничимся анализом уравнения частот. Если уравнения движения имеют вид

$$\begin{aligned} \ddot{x}_1 + f\dot{x}_1 + c_1x_1 - b\dot{x}_2 &= 0, \\ \ddot{x}_2 + f\dot{x}_2 + c_2x_2 + b\dot{x}_1 &= 0, \end{aligned} \quad (1)$$

где  $c_1 < 0$ ,  $c_2 < 0$ ,  $f \geq 0$  и  $b$  достаточно велико, чтобы обеспечить устойчивость при  $f = 0$ , мы положим, что решения пропорциональны  $e^{\gamma t}$ , тогда  $\gamma$  удовлетворяет уравнению

$$\gamma^4 + 2f\gamma^3 + \gamma^2(c_1 + c_2 + b^2 + f^2) + f\gamma(c_1 + c_2) + c_1c_2 = 0. \quad (2)$$

Отсюда

$$\sum \gamma = -2f \leq 0, \quad \sum \frac{1}{\gamma} = -f \left( \frac{1}{c_1} + \frac{1}{c_2} \right) \geq 0. \quad (3)$$

Обозначим корни для  $f = 0$  через  $\pm in_1$ ,  $\pm in_2$ , где  $n_1 > n_2$ ; пусть для малого положительного  $f$  они равны  $\pm in_1 - \alpha_1$ ,  $\pm in_2 - \alpha_2$

с точностью до членов порядка  $f$ . С той же точностью

$$\alpha_1 + \alpha_2 = f > 0,$$

$$\frac{1}{in_1 - \alpha_1} + \frac{1}{-in_1 - \alpha_1} + \frac{1}{in_2 - \alpha_2} + \frac{1}{-in_2 - \alpha_2} \approx -\frac{2\alpha_1}{n_1^2} - \frac{2\alpha_2}{n_2^2} > 0.$$

Эти неравенства согласуются только, если  $\alpha_1$  и  $\alpha_2$  имеют противоположные знаки, и при  $n_1 > n_2$ ,  $\alpha_1 > 0$ ,  $\alpha_2 < 0$ . Следовательно, если система удерживается в устойчивом состоянии только гироскопическими силами, малое трение всегда приводит к потере устойчивости. Более быстрые собственные колебания будут затухать, но более медленные увеличивать амплитуду со временем.

Эта особенность гироскопического движения имеет теоретическое и практическое значение. Если мы используем обычный метод анализа малых колебаний относительно стационарного движения, пренебрегая трением, мы часто находим, что все корни  $\gamma$  чисто мнимые, и убеждаемся, что система устойчива. Если выражение  $c_{rs}x_r x_s$  существенно  $\geq 0$  и имеется слабое трение,  $a_{rs}\dot{x}_r \dot{x}_s + c_{rs}x_r x_s$  будет уменьшаться и колебания будут постепенно затухать. Такие системы называют системами с *вековой устойчивостью*, но если рассматриваемая форма не является существенно положительной и система удерживается в устойчивом состоянии только гироскопическими членами, то медленные колебания относительно стационарного движения будут постепенно расти по амплитуде. Настанет момент, когда их уже нельзя будет считать малыми, и это приведет к полному изменению характера движения. Такие системы называют системами с *ординарной устойчивостью*, но с *вековой неустойчивостью*. Инженеры стараются избегать таких систем. Они, возможно, имеют большое значение в развитии звездных систем, и в частности солнечной системы.

**8.11. Радиоактивный распад.** Урановое семейство элементов таково, что атом любого из них, за исключением последнего, способен распадаться на атом следующего элемента и атом гелия ( $\alpha$ -частица) или свободный электрон ( $\beta$ -частица). Излучаемые частицы ( $\alpha$  или  $\beta$ ) оставляют атом и не влияют на последующие стадии распада. Число атомов любого элемента, распадающихся в короткий отрезок времени, пропорционально временному интервалу и имеющемуся числу атомов этого элемента \*). Если  $u$ ,  $x_1$ ,  $x_2$ , ...,  $x_n$  — математические ожидания

\*) Строго говоря, так как радиоактивность — случайный процесс, это правило справедливо для математического ожидания (М. О.) числа распадающихся атомов. Истинное число будет отклоняться от М. О., но если М. О. велико, отклонениями можно пренебречь.

числа атомов различных элементов, имеющих в момент времени  $t$ , они удовлетворяют дифференциальным уравнениям

$$\begin{aligned}\frac{du}{dt} &= -\kappa u, \\ \frac{dx_1}{dt} &= \kappa u - \kappa_1 x_1, \\ \frac{dx_2}{dt} &= \kappa_1 x_1 - \kappa_2 x_2, \\ &\dots \dots \dots \\ \frac{dx_n}{dt} &= \kappa_{n-1} x_{n-1}.\end{aligned}\tag{1}$$

Положим, что при  $t = 0$  имеются только атомы урана, тогда  $u = u_0$  и все другие зависимые переменные равны нулю. Вспомогательные уравнения примут вид

$$\begin{aligned}(p + \kappa) u &= p u_0, \\ (p + \kappa_1) x_1 &= \kappa u, \\ (p + \kappa_2) x_2 &= \kappa_1 x_1, \\ &\dots \dots \dots \\ (p + \kappa_{n-1}) x_{n-1} &= \kappa_{n-2} x_{n-2}, \\ p x_n &= \kappa_{n-1} x_{n-1},\end{aligned}\tag{2}$$

и отсюда операторные решения будут

$$\begin{aligned}u &= \frac{p u_0}{p + \kappa}, \quad x_1 = \frac{\kappa p u_0}{(p + \kappa)(p + \kappa_1)}, \\ x_2 &= \frac{\kappa \kappa_1 p u_0}{(p + \kappa)(p + \kappa_1)(p + \kappa_2)}, \dots, \\ x_n &= \frac{\kappa \kappa_1 \dots \kappa_{n-1} u_0}{(p + \kappa)(p + \kappa_1) \dots (p + \kappa_{n-1})}.\end{aligned}\tag{3}$$

Используя правило разложения на элементарные дроби, получаем искомую временную зависимость. Действительно,

$$\begin{aligned}u &= u_0 e^{-\kappa t}, \quad x_1 = \frac{\kappa u_0}{\kappa_1 - \kappa} (e^{-\kappa t} - e^{-\kappa_1 t}), \\ x_2 &= \kappa \kappa_1 u_0 \left[ \frac{1}{(\kappa_1 - \kappa)(\kappa_2 - \kappa)} e^{-\kappa t} + \frac{1}{(\kappa - \kappa_1)(\kappa_2 - \kappa_1)} e^{-\kappa_1 t} + \right. \\ &\quad \left. + \frac{1}{(\kappa - \kappa_2)(\kappa_1 - \kappa_2)} e^{-\kappa_2 t} \right],\end{aligned}\tag{4}$$

$$x_n = u_0 - \frac{\kappa_1 \dots \kappa_{n-1} u_0}{(\kappa_1 - \kappa) \dots (\kappa_{n-1} - \kappa)} e^{-\kappa t} - \dots \dots \tag{5}$$

Из всех констант распада  $\kappa$  — наименьшая. Если прошло достаточно времени для того, чтобы все экспоненциальные множители, за исключением  $e^{-\kappa t}$ , стали пренебрежимо малыми, результат будет приближенно следующим:

$$u = u_0 e^{-\kappa t}, \quad x_1 = \frac{\kappa}{\kappa_1} u_0 e^{-\kappa t}, \quad x_2 = \frac{\kappa}{\kappa_2} u_0 e^{-\kappa t}, \dots, \quad (6)$$

$$x_n = u_0 (1 - e^{-\kappa t}). \quad (7)$$

Числа атомов различных элементов, за исключением последнего, уменьшаются, причем их отношения сохраняются примерно постоянными и равными обратным отношениям их констант распада.

С другой стороны, если прошло так мало времени, что единица еще хорошо в первом приближении аппроксимирует все экспоненты, мы можем разложить операторы по нисходящим степеням  $p$  и интерпретировать их поочередно. Следовательно, сначала  $x_1$  будет увеличиваться пропорционально  $t$ ,  $x_2$  — пропорционально  $t^2$ ,  $x_n$  — пропорционально  $t^n$ .

В эксперименте часто встречается промежуточная ситуация. Некоторые из экспонент могут стать малыми за время эксперимента, а другие все еще близки к 1. Мы имеем тогда

$$x_r = \frac{\kappa_{r-1} x_{r-1}}{p + \kappa_r} = \kappa_{r-1} (p^{-1} x_{r-1} - \kappa_r p^{-2} x_{r-1} + \dots), \quad (8)$$

и если  $\kappa_r t$  мало, мы можем пренебречь вторым и последующими членами по сравнению с первым. Отсюда в этом случае мы имеем

$$x_r \approx \kappa_{r-1} p^{-1} x_{r-1}. \quad (9)$$

Если  $x_{r-1}$  имеет вид  $t^s$ , то из (8) имеем

$$x_r = \frac{\kappa_{r-1} s!}{(p + \kappa_r) p^s} = \frac{\kappa_{r-1} s!}{p^{s+1}} \frac{p}{p + \kappa_r} = \frac{\kappa_{r-1} s!}{p^{s+1}} e^{-\kappa_r t}. \quad (10)$$

Если  $\kappa_r t$  мало, мы можем заменить экспоненту единицей и подтвердить (8). Но если  $\kappa_r t$  велико, то

$$p^{-1} e^{-\kappa_r t} = \int_0^t e^{-\kappa_r t} dt = \frac{1}{\kappa_r} + O(e^{-\kappa_r t}), \quad (11)$$

и, продолжая интегрирование, получим

$$p^{-s-1} e^{-\kappa_r t} \approx p^{-s} \frac{1}{\kappa_r} = \frac{1}{\kappa_r} \frac{t^s}{s!}. \quad (12)$$

Отсюда

$$x_r \approx \frac{\kappa_{r-1}}{\kappa_r} t^s = \frac{\kappa_{r-1}}{\kappa_r} x_{r-1}. \quad (13)$$

Разделяя элементы на долгоживущие и короткоживущие в зависимости от того, велико или мало  $\kappa_1 t$ , где  $t$  — длительность данного эксперимента, мы находим, что число продуктов распада первого из долгоживущих элементов увеличивается пропорционально  $t$ , второго —  $t^2$  и т. д. Число короткоживущих продуктов распада изменяется почти пропорционально числу предыдущего долгоживущего. Все продукты  $\beta$ -распада короткоживущие при не очень больших  $t$ .

Радий — третий продукт в ряде  $\alpha$ -распада урана. В образцах горных пород время, прошедшее с момента их образования, обычно таково, что справедливы соотношения (6). Наблюдения показывают, что число атомов радия и урана находится в постоянном отношении  $3,58 \cdot 10^{-7}$ . Это дает нам  $\kappa/\kappa_3$ . Скорость распада радия определяется непосредственно; а именно

$$\frac{1}{\kappa_3} = 2280 \text{ лет.}$$

Отсюда

$$\frac{1}{\kappa} = 6,37 \cdot 10^9 \text{ лет.}$$

Так мы находим скорость распада самого урана \*).

Содди аккуратно очистил ряд образцов соединений урана от радия и хранил их 10 лет. Было установлено, что образовалось новое количество радия; его количество росло пропорционально квадрату времени. Это говорит о том, что среди двух элементов, расположенных между ураном и радием, один долгоживущий (для срока в 10 лет), а другой короткоживущий. В действительности же независимым путем установлено, что оба они долгоживущие. Однако первый химически не отделим от обычного урана и, следовательно, присутствовал и в первоначальных образцах; значит, вначале вместо  $x_1 = 0$  мы имеем

$$x_1 \approx \frac{\kappa}{\kappa_1} u_0.$$

Для следующего элемента, иония,

$$x_2 \approx \kappa_1 p^{-1} x_1 \approx \kappa u_0 t,$$

а для радия

$$x_3 = \kappa_2 p^{-1} x_2 = \frac{1}{2} \kappa \kappa_2 u_0 t^2.$$

---

\*) Использованные здесь численные данные в более поздних экспериментальных работах пересмотрены; оказалось также, что последовательность разветвляется и затем частично соединяется. Поэтому учет современных результатов усложнил бы анализ без введения каких-либо новых принципов,

Содди [6] нашел, что 3 кг урана за 10,15 года дают  $202 \cdot 10^{-12}$  г радия. Отсюда, учитывая разницу в атомных весах

$$\frac{x_3}{u_0} = 7,1 \cdot 10^{-14}$$

и зная  $\kappa$ , имеем

$$\kappa_2 = 8,64 \cdot 10^{-6} \text{ год}^{-1}, \quad \frac{1}{\kappa_2} = 1,16 \cdot 10^5 \text{ лет.}$$

Это дает нам скорость исчезновения иония. Содди получил чуть меньшее значение  $1/\kappa_2$  по более многочисленным данным.

В другом случае, иногда встречающемся в эксперименте, используются естественные образцы, в которых различные элементы находились в соотношениях, определяемых (6). Затем уран и, возможно, некоторые более поздние продукты распада устраняются химическим путем и исследуется поведение оставшейся части. Решение для этого случая дано Седжвиком [7]. Операторный анализ, применить который предложил Кроссли, состоит в следующем. Пусть  $x_s$  при  $t=0$  равно 0 для  $s < r$  и  $u_s$  для  $s \geq r$ , где  $u_s = (\kappa_r/\kappa_s)u_r$ , за исключением  $r=n$ .

Вспомогательные уравнения теперь будут

$$(p + \kappa_r) x_r = p u_r, \quad (14)$$

$$(p + \kappa_s) x_s = p u_s + \kappa_{s-1} x_{s-1} \quad (r < s < n), \quad (15)$$

$$p x_n = p u_n + \kappa_{n-1} x_{n-1}. \quad (16)$$

Для  $r < s < n$

$$\begin{aligned} (p + \kappa_s) x_s &= (p + \kappa_s) u_s - \kappa_s u_s + \kappa_{s-1} x_{s-1} = \\ &= (p + \kappa_s) u_s - \kappa_{s-1} (u_{s-1} - x_{s-1}), \end{aligned} \quad (17)$$

т. е. если

$$u_s - x_s = y_s, \quad (18)$$

то

$$(p + \kappa_s) y_s = \kappa_{s-1} y_{s-1} \quad (19)$$

с

$$(p + \kappa_r) y_r = \kappa_r u_r, \quad p y_n = -\kappa_{n-1} u_{n-1} + \kappa_{n-1} y_{n-1}. \quad (20)$$

Операторные решения имеют вид

$$\begin{aligned} y_r &= \frac{\kappa_r u_r}{p + \kappa_r}, \quad y_{r+1} = \frac{\kappa_r^2 u_r}{(p + \kappa_r)(p + \kappa_{r+1})}, \\ y_{r+2} &= \frac{\kappa_r^2 \kappa_{r+1} u_r}{(p + \kappa_r)(p + \kappa_{r+1})(p + \kappa_{r+2})}, \\ y_n &= -\frac{1}{p} \kappa_r u_r + \frac{1}{p} \frac{\kappa_r^2 \kappa_{r+1} \dots \kappa_{n-2} u_r}{(p + \kappa_r) \dots (p + \kappa_{n-1})}. \end{aligned} \quad (21)$$



Следовательно,

$$y_r = u_r(1 - e^{-\kappa_r t}), \quad x_r = u_r e^{-\kappa_r t}; \quad (22)$$

и так как  $\kappa_r u_r = \kappa_s u_s$ , каждое  $y_s$ , за исключением  $y_n$ , стремится к  $u_s$ , когда  $t$  стремится к бесконечности. Выражение (22) получено с помощью правила разложения на элементарные дроби. Отсюда каждое  $x_s$ , кроме  $x_n$ , стремится к 0, как и ожидалось,

$$\begin{aligned} x_{r+1} = u_{r+1} \left( \frac{\kappa_{r+1}}{\kappa_r(\kappa_{r+1} - \kappa_r)} e^{-\kappa_r t} + \frac{\kappa_{r+1}}{\kappa_{r+1}(\kappa_r - \kappa_{r+1})} e^{-\kappa_{r+1} t} \right) = \\ = \frac{u_{r+1}}{\kappa_{r+1} - \kappa_r} \left( \frac{\kappa_{r+1}}{\kappa_r} e^{-\kappa_r t} - e^{-\kappa_{r+1} t} \right) \end{aligned} \quad (23)$$

и т. д.

Интересен случай, когда продолжительность эксперимента настолько велика, что некоторые из первых экспоненциальных членов становятся малыми, т. е. ранние короткоживущие элементы исчезают во время эксперимента. Если первый еще не ставший малым член — это  $e^{-\kappa_m t}$ , то мы получим приближенно

$$x_m = \frac{\kappa_r \kappa_{r+1} \dots \kappa_m u_m}{\kappa_m (\kappa_r - \kappa_m) (\kappa_{r+1} - \kappa_m) \dots (\kappa_{m-1} - \kappa_m)} e^{-\kappa_m t} \approx u_m e^{-\kappa_m t}, \quad (24)$$

так что этот элемент убывает почти так же, как если бы другие не присутствовали; убывание последующих элементов ряда следует тому же закону до тех пор, пока не появляется промежуточный элемент с  $\kappa_s \leq \kappa_m$ . Для долгоживущих элементов  $x_s$  содержит член  $e^{-\kappa_s t}$ , который может быть большим.

## ПРИМЕРЫ

1. Невесомая пружина длины  $3l$  находится под натяжением  $P$  между двумя закрепленными точками. Массы  $5m$  и  $8m$  делят пружину на три равные части. Частице с массой  $5m$  придается малое поперечное смещение со скоростью  $u$ . Доказать, что смещение другой частицы дается выражением

$$\frac{5}{14} \frac{u}{\alpha} \left( \sqrt{\frac{20}{3}} \sin \sqrt{\frac{3}{20}} \alpha t - \sqrt{2} \sin \frac{\alpha t}{\sqrt{2}} \right),$$

(М. Т., 1929.)

где  $\alpha^2 = P/ml$ .

2. Сейсмограф Голицына имеет параметры

$$\kappa_1 = \kappa_2 = \kappa, \quad n_1^2 = n_2^2 = 2\kappa^2.$$

Докажите, что отклик на единичный импульс скорости имеет вид

$$\frac{\lambda \mu}{2\kappa^2} (\kappa t \sin \kappa t - \sin \kappa t + \kappa t \cos \kappa t) e^{-\kappa t}.$$

3. Докажите, что если  $x$  — смещение на записи сейсмографа Голицына из-за импульсного изменения скорости почвы, то

$$\int_0^{\infty} x \, dt = 0$$

независимо от значений параметров прибора.

#### ЛИТЕРАТУРА

1. *Rimington E. C.*, Phil. Mag., (5), **24**, 54—60 (1887).
2. *Bronwich T. J.*, Phil. Mag., (6), **37**, 1919, 407—419.
3. *Rybner J.*, Gerlands Beitr., **31**, 259—281 (1931); **51**, 375—401 (1937); **55**, 303—313 (1939).
4. *Routh E. J.*, Stability of a given State of Motion, 1877.
5. *Heaviside O.*, Electrical Papers, **1**, 529.
6. *Soddy F.*, Phil. Mag., (6), **38**, 483—488 (1919).
7. *Sedgwick W. F.*, Proc. Cambr. Phil. Soc., **38**, 283 (1942).
8. *Carler G. W.*, The Simple Calculation of Electrical Transients, 1944.

## УКАЗАТЕЛЬ

- Абель (Abel N. H.) 82  
 Абеля лемма 81, 82, 87, 88, 97, 98  
   — признак 82, 88, 103  
   — теорема 82  
 Абсолютная величина 27  
 Алгебраическое дополнение 224  
 Аргана диаграмма 269  
 Ассоциативность 125  
 Афинорное обозначение 156
- Баркилл (Burkill J. C.) 61  
 Безикович (Besicovitch A. S.) 100, 108, 313  
 Бейкер (Baker H. F.) 48, 373  
 Больцано (Bolzano N.) 48  
 Больцано — Вейерштрасса теорема 47  
 Больцман (Boltzmann L.) 272  
 Боннэ (Bonnet O.) 98  
 Борн (Born M.) 331  
 Брега условие 257  
 Бромвич (Bromwich T. J. I.A.) 91, 98  
 Буль (Boole G.) 373
- Вебер (Weber H.) 16  
 Вейерштрасса М-признак сходимости рядов 80, 84, 87  
   — теорема 45  
 Вектор единичный 117  
   — направляющий 117  
   — нулевой 121  
   — площади 128  
 Векторные функции 138  
 Векторы 111, 115  
   — безвихревые 320  
   — геометрическое представление 119  
   — деление 135  
   — компланарные 122  
   — направляющие 121  
   — скорость изменения 181  
   — соленоидальные 319
- Векторы, умножение 123, 126, 129 — 131  
 Величины физические 18, 20  
 Волчок 187  
 Вращение 277  
 Вязкость 178
- Гёдель (Gödel K.) 17  
 Гейне (Heine E.) 48  
 Гейне — Бореля теорема 46—49, 60, 107, 285, 288, 293  
 Гельдер (Hölder O.) 371  
 Гельдера условие 371  
 Гиббс (Gibbs J.) 272  
 Гирскопическая система 409, 411  
 Грани 35, 36  
 Грина задача 361  
   — лемма 314, 319, 329, 344, 345  
   — слой 357  
   — теорема 317, 326, 352, 355, 357  
 Гурса (Goursat E.) 48
- Давление 177  
 Даламбера принцип 148  
 Движение частицы 140  
 Дедекинд (Dedekind R.) 16, 22  
 Дедекиндовы сечения 23, 25  
 Деформация 208  
 Джеффрис (Jeffreys H.) 374  
 Диполь 337  
 Дирихле (Dirichlet P. G. L.) 82  
 Дирихле — Харди признак 82, 89  
 Диссипативная система 409  
 Дистрибутивность 125  
 Дифференцирование 63, 138, 373  
   — под знаком интеграла 323  
   — произведения 139  
   — функций 158  
 Дифференцируемость 45, 158, 290  
 Доказательство от противного 27  
 Достаточность 30  
 Дуга 282

- Дюбуа-Раймона условие 59, 60  
   — форма 98
- Евклид 21  
 Евклида аксиома 22, 23  
   — теорема 21, 22  
   — треугольник 23
- Жесткость 176  
 Жидкость 177  
   — движение 168
- Замена переменных 298
- Интеграл бесконечный 68  
   — вдоль спрямляемых кривых 306  
   — верхний предел 91  
   — двойной 293  
   — замена переменного 67  
   — кратный 278, 306  
   — несобственный 68, 71  
   — определенный 373  
   — поверхностный 307  
   — повторный 91, 294  
   — пределы 302  
   — сходящийся 71, 85  
   — Фурье 53, 98  
 Источник и сток 330
- Кантор (Cantor G.) 23, 34  
 Каратеодори (Carathéodory C.) 313  
 Карнап (Carnap R.) 17  
 Кватернионы 135  
 Кельвина теорема 354  
 Колебания малые 237, 239, 406  
 Коммутативность 125  
 Компоненты ковариантные 253  
   — контравариантные 253  
 Контур 232  
 Координаты декартовы 111  
   — инерциальная система 116  
   — косоугольные 251  
   — криволинейные 257  
   — нормальные 238  
   — ортогональные 259  
   — полярные 303  
   — преобразование 152  
 Корни кратные 408  
   — разделение 231  
 Коши — Буняковского неравенство 101  
 Коши теорема 48
- Кривые 278, 281  
   — спрямляемые 282  
 Кристаллы 254  
 Кронекера  $\delta$ -символ 115  
 Кроссли (Crossley A. F.) 416  
 Курант (Courant R.) 90  
 Кэмпбел (Campbell N. R.) 18
- Лагранжа метод 238  
   — тождество 131  
 Лапласа уравнение 327, 328, 333, 336, 338, 352, 361  
 Лармор (Larmor J.) 262  
 Лармора — Лоренца преобразование 262  
 Лебег (Lebesgue H.) 61, 313  
 Лебега интеграл 100  
 Ленард-Джонс (Lennard-Jones J. E.) 331  
 Липшица условие 99, 100, 371  
 Лоренц (Lorentz H. A.) 262  
 Лоренца сила 265  
 Ляв (Love A. E. H.) 408
- Магнитная проницаемость 180  
 Майкельсона — Морли опыты 261  
 Максвелла уравнения 263  
 Маллока машина 273  
 Маркова цепи 266  
 Математическая индукция 28  
 Матрица 193  
   — алгебраических дополнений 223  
   — блочная 215  
   — главные оси 162  
   — диагональная 198, 213  
   — единичная 196  
   — квазидиагональная 217  
   — комплексно сопряженная 197  
   — нулевая 196  
   — обратная 198  
   — ортогональная 199, 203, 204, 243  
   — присоединенная 198  
   — симметричная 197, 206  
   — собственные значения 212  
   — транспонированная 196  
   — унитарная 199, 243  
   — характеристическое уравнение 212  
   — эрмитова 198, 221, 227  
 Матрицы Дирака 249  
   — минор 209  
   — Паули 248  
   — подобные 213  
   — преобразования 201

- Матрицы ранг 208  
— форма 226, 227  
— Эддингтона 249  
Маятник Фуко 191  
— Эйлера 400  
Мёбиуса полоска 308, 309  
Мера нуль 61  
Миллера индексы 257  
Множество 28  
— замкнутое 29  
— минимальное 30  
— несчетное 30  
— счетное 29  
Момент импульса 165
- Направляющие косинусы 113  
Напряжение 171  
— гидростатическое 177  
— и деформация 174  
Необходимость 30  
Непрерывность 41, 42, 44, 48, 283, 288  
— равномерная 51  
Ньюмен (Newman M. H. A.) 49  
Ньютон 333  
Ньютона динамика 116  
— третий закон 148
- $O(x)$ ,  $o(x)$  52  
Области 278, 283, 284  
Оболочка сферическая 333  
Объемное сжатие 177  
Объемные силы 172  
Однозначность функции 41  
Окрестности 278  
Операторы, деление 379  
— интерпретация 373, 380, 385  
— композиция 377  
— ряды 231  
Операционные методы 372, 395  
Определители 200  
Ортогональность 234  
Остроградский М. В. 326
- $p$  (символ Хевисайда) 384  
Параллакс 189  
Пары чисел 24  
Пеано (Peano G.) 373  
Переходные процессы 398  
Пикар (Picard E.) 373, 374  
Плотность линейная 336  
Площадь поверхностей 307  
Поверхность 285
- Пограничный слой 328  
Поллард (Pollard S.) 63  
Порядок величины 52  
Последовательности 31  
— интервалов 23  
— неограниченные 31, 33  
— ограниченные 32, 33  
— осциллирующие 31—33  
— расходящиеся 33  
— сходящиеся 31, 32, 75  
Потенциал 327, 330  
— векторный 365  
— — диполя 368  
— — точечного заряда 368  
— и поле 340, 362  
— плоскости 361  
— поверхностная плотность 347  
— скоростей 328  
— сферы 360  
Поток 318  
Правило параллелограмма 113  
— суммирования 111, 114  
— частного 157  
— элементарных дробей 386  
Предельная точка 28  
Пределы функций 41  
Преобразование 114  
— обратное 115  
— ортогональное 276  
Принцип суперпозиции 388  
Произведение векторное 137  
— скалярное 124  
— смешанное 136  
Произведения бесконечные 99  
— сходимость 99  
— тройные 135  
Пуассона уравнение 335, 345, 371
- Радиоактивный распад 412  
Разрыв функции 42, 43  
Разряды конденсатора 397  
Распределения непрерывные 330  
Рассел (Russell B.) 17  
Растояние 278  
Расходимость 36, 37  
Резонанс 404  
Релей (Rayleigh) 408  
Релея принцип 236  
Релея — Ритца метод 356  
Римана  $\zeta$ -ряд 38  
— интеграл 56, 57, 60, 64, 293, 294  
Римингтона метод 398  
Ричардсон (Richardson L. F.) 357  
Ряды 37

- Ряды двойные 40, 91  
 — интегрируемость 76  
 — непрерывность 76  
 — операторов 375  
 — перестановки 40  
 — сравнение 80  
 — сумма 37  
 — сходимость 75, 79
- Самоиндукция 398
- Саутвелл (Southwell R. V.) 357
- Седжвик (Sedgwick W. F.) 416
- Сейсмографы 61, 400—403, 418
- Сжимаемость 177
- Сила 115
- Система гироскопическая 409, 411  
 — диссипативная 409  
 — уравнений 381
- Скаляр 111, 118, 124
- Скачок функции 49, 55
- Скорость 115  
 — угловая 142, 145, 168
- Слой двойной 338, 352
- Смитсис (Smithies F.) 108
- Содди (Soddy F.) 416
- Стационарность 236
- Стилтьеса интеграл 53, 56, 62, 63, 65, 307
- Струуд (Stroud W.) 21
- Сфера однородная 334
- Сходимость абсолютная 39  
 — интегралов 85  
 — неравномерная 78  
 — ограниченная 105  
 — признаки 79  
 — произведения 99  
 — равномерная 75, 79
- Твердое тело 147, 165, 167, 170, 174, 205
- Тензор антисимметричный 159, 161  
 — главные значения 164  
 — двухмерный 188  
 — диагональный вид 164  
 — изотропный 154  
 — инерции 165  
 — напряжений 180, 265  
 — первого порядка 152  
 — подстановки 115  
 — свертывание 152  
 — симметричный 159, 162  
 — скорости деформации 169
- Тензоры 152  
 — приложения к механике 185
- Теорема Гаусса — Остроградского 330, 344  
 — единственности 354  
 — о минимуме 353  
 — — ограниченной сходимости 84  
 — — покрытия 45  
 — — среднем 93, 95, 97  
 — Ролля 93  
 — Стокса 317, 319  
 — сходимости 105, 376  
 — Тейлора 95  
 — Тэннери 91  
 — эргодическая 271
- Титчмарш (Titchmarsh E. C.) 61
- Точка внешняя 283  
 — внутренняя 283, 360  
 — граничная 283  
 — разрыва 78
- Уайтхед (Whitehead A. N.) 17
- Углы Эйлера 185
- Угол пространственный 338
- Уитсона мостик 398
- Умова — Пойнтинга вектор 266
- Упругая деформация 171
- Уравнения вспомогательные 386  
 — движения 174  
 — детерминантные 212  
 — дифференциальные 379  
 — интегральные 272  
 — линейные 199, 208  
 — Эйлера 186  
 — ядро 273
- Ускорение 115, 178
- Устойчивость вековая 412  
 — ординарная 412
- Фукс (Fuchs L.) 373
- Функция векторная 158  
 — возрастающая 50  
 — двух переменных 72  
 — кусочно непрерывная 45  
 — непрерывная 50, 77, 86  
 — нескольких переменных 287  
 — обратная 50  
 — ограниченной вариации 52, 55  
 — однозначность 41  
 — скалярная 158  
 — убывающая 50  
 — Хевисайда 42, 43, 393
- Харди (Hardy G. H.) 82
- Хевисайд (Heaviside O.) 373
- Хобсон (Hobson E. W.) 49

Циркуляция 318

Числа действительные 21, 22

Шапка сферическая 335

Шварц (Schwarz H. A.) 309, 313

Шварца неравенство 101

Штольц (Stolz O.) 291

$\varepsilon$  27

$\varepsilon_{ikm}$ ,  $\varepsilon_{p\sigma m}$  134

Электрическая индукция 363

Электромагнитная теория 261

Юнга модули 156

Юнг В. (Young W. H.) 48, 100, 291

Юнг Г. (Young G. C.) 100

Якобиан 299

Якоби теорема 223, 232

# ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	5
ОТ АВТОРОВ	7
ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ	8
ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ	9
ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ	10
ГЛАВА 1. ДЕЙСТВИТЕЛЬНОЕ ПЕРЕМЕННОЕ	15
ГЛАВА 2. СКАЛЯРЫ И ВЕКТОРЫ	111
ГЛАВА 3. ТЕНЗОРЫ	152
ГЛАВА 4. МАТРИЦЫ	193
ГЛАВА 5. КРАТНЫЕ ИНТЕГРАЛЫ	278
ГЛАВА 6. ТЕОРИЯ ПОТЕНЦИАЛА	327
ГЛАВА 7. ОПЕРАЦИОННЫЕ МЕТОДЫ	372
ГЛАВА 8. ФИЗИЧЕСКИЕ ПРИЛОЖЕНИЯ ОПЕРАЦИОННОГО МЕТОДА	395

**Г. Джеффрис, Б. Свирлс**  
**МЕТОДЫ МАТЕМАТИЧЕСКОЙ ФИЗИКИ**

Редактор Э. А. Медушевская

Художник Е. М. Золотарев

Художественный редактор В. М. Варлашин

Технический редактор В. П. Сизова

Сдано в производство 1/VII 1969 г.

Подписано к печати 17/X 1969 г. 60×90<sup>1</sup>/<sub>16</sub> =

13,25 бум. л. 26,5 печ. л. Уч.-изд. л. 22,66 Изд. № 27/5051. Цена 1 р. 74 к. Зак. 231

---

ИЗДАТЕЛЬСТВО «МИР», Москва, 1-й Рижский пер., 2

---

Ленинградская типография № 2 имени Евгении Соколовой Главполиграфпрома  
Комитета по печати при Совете Министров СССР. Измайловский пр., 29.

Отпечатано в 11-й типографии Главполиграфпрома, зак. 793.



