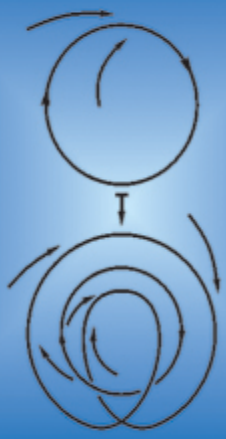
В. И. Арнольд

ГЕОМЕТРИЧЕСКИЕ МЕТОДЫ В ТЕОРИИ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ



R&C Dynamics

В. И. Арнольд

ГЕОМЕТРИЧЕСКИЕ МЕТОДЫ В ТЕОРИИ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Издание второе, исправленное и дополненное

Редакция журнала
«Регулярная и хаотическая динамика»

МЦНМО, ВКМ НМУ

1999

УДК 517.2
ББК 22.161.6
А 84

А 84 **Арнольд В. И.**
Геометрические методы в теории обыкновенных дифференциальных уравнений. — Ижевск: Ижевская республиканская типография. 2000. — 400 с.

В книге изложен ряд основных идей и методов, применяемых для исследования обыкновенных дифференциальных уравнений. Элементарные методы интегрирования рассматриваются с точки зрения общематематических понятий (разрешение особенностей, группы Ли симметрий, диаграммы Ньютона и т. д.). Теория уравнений с частными производными первого порядка изложена на основе геометрии контактной структуры.

В книгу включены классические и современные результаты теории динамических систем: структурная устойчивость, У-системы, аналитические методы локальной теории в окрестности особой точки или периодического решения (нормальные формы Пуанкаре), теория бифуркации фазовых портретов при изменении параметров (мягкое и жесткое возбуждение автоколебаний при потере устойчивости), удвоение периода Фейгенбаума, теорема Дюлака и др.

Книга рассчитана на широкий круг математиков и физиков — от студентов до преподавателей и научных работников.

ISBN 5-89806-028-4 **ББК 22.161.6**

Содержание

Предисловие 5

Некоторые используемые обозначения 9

Глава 1. **Специальные уравнения** 11
§ 1. Дифференциальные уравнения, инвариантные относительно групп симметрий 11
§ 2. Разрешение особенностей дифференциальных уравнений 19
§ 3. Уравнения, не разрешенные относительно производных . 25
§ 4. Нормальная форма уравнения, не разрешенного относительно производной, в окрестности регулярной особой точки 37
§ 5. Стационарное уравнение Шредингера 44
§ 6. Геометрия дифференциального уравнения второго порядка и геометрия пары полей направлений в трехмерном пространстве 57

Глава 2. **Уравнения с частными производными первого порядка** 75
§ 7. Линейные и квазилинейные уравнения с частными производными первого порядка 75
§ 8. Нелинейное уравнение с частными производными первого порядка 85
§ 9. Теорема Фробениуса 104

Глава 3. **Структурная устойчивость** 108
§ 10. Понятие структурной устойчивости 109
§ 11. Дифференциальные уравнения на торе 117
§ 12. Аналитическое приведение к повороту аналитических диффеоморфизмов окружности 136
§ 13. Введение в гиперболическую теорию 144
§ 14. У-системы 151
§ 15. Структурно устойчивые системы не всюду плотны . . . 166

Глава 4. **Теория возмущений** 169
§ 16. Метод усреднения 170
§ 17. Усреднение в одночастотных системах 174
§ 18. Усреднение в многочастотных системах 179
§ 19. Усреднение в гамильтоновых системах 192
§ 20. Адиабатические инварианты 196
§ 21. Усреднение в слоении Зейферта 202

Глава 5. **Нормальные формы** 209
§ 22. Формальное приведение к линейной нормальной форме . 209
§ 23. Резонансный случай 213
§ 24. Области Пуанкаре и Зигеля 217
§ 25. Нормальная форма отображения в окрестности неподвижной точки 223
§ 26. Нормальная форма уравнения с периодическими коэффициентами 226
§ 27. Нормальная форма окрестности эллиптической кривой . 235
§ 28. Доказательство теоремы Зигеля 250

Глава 6. **Локальная теория бифуркаций** 258
§ 29. Семейства и деформации 258
§ 30. Матрицы, зависящие от параметров, и особенности декремент-диаграмм 276
§ 31. Бифуркации особых точек векторного поля 301
§ 32. Версальные деформации фазовых портретов 307
§ 33. Потеря устойчивости положения равновесия 312
§ 34. Потеря устойчивости автоколебаний 330
§ 35. Версальные деформации эквивариантных векторных полей на плоскости 349
§ 36. Перестройки топологии при резонансах 372
§ 37. Классификация особых точек 388

Образцы экзаменационных задач 394

Предисловие

Основное открытие Ньютона, то, которое он счел нужным засекретить и опубликовал лишь в виде анаграммы, состоит в следующем:

«Data aequatione quotcunque fluentes quantitae involvente fluxiones invenire et vice versa». В переводе на современный математический язык это означает: «Полезно решать дифференциальные уравнения».

В настоящее время теория дифференциальных уравнений представляет собой трудно обозримый конгломерат большого количества разнообразных идей и методов, в высшей степени полезный для всевозможных приложений и постоянно стимулирующий теоретические исследования во всех отделах математики.

Большая часть путей, связывающих абстрактные математические теории с естественнонаучными приложениями, проходит через дифференциальные уравнения. Многие разделы теории дифференциальных уравнений настолько разрослись, что стали самостоятельными науками; проблемы теории дифференциальных уравнений имели большое значение для возникновения таких наук, как линейная алгебра, теория групп Ли, функциональный анализ, квантовая механика и т. д. Таким образом, дифференциальные уравнения лежат в основе естественнонаучного математического мировоззрения.

При отборе материала для настоящей книги автор старался изложить основные идеи и методы, применяемые для изучения дифференциальных уравнений. Особые усилия были приложены к тому, чтобы основные идеи, как правило простые и наглядные, не загромождались техническими деталями. С наибольшей подробностью рассматриваются наиболее фундаментальные и простые вопросы, в то время как изложение более специальных и трудных частей теории носит характер обзора.

Книга начинается с исследования некоторых специальных дифференциальных уравнений, интегрируемых в квадратурах. При этом основное внимание уделяется не формально-рецептурной стороне элементарной теории интегрирования, а ее связям с общематематическими идеями, методами и понятиями (разрешение особенностей, группы Ли,

диаграммы Ньютона), с одной стороны, и естественнонаучным приложениям — с другой.

Теория уравнении с частными производными первого порядка рассматривается при помощи естественной контактной структуры в многообразии 1-струй функций. Попутно излагаются необходимые элементы геометрии контактных структур, делающие всю теорию независимой от других источников.

Значительную часть книги занимают методы, обычно называемые качественными. Современное развитие основанной А. Пуанкаре качественной теории дифференциальных уравнений привело к пониманию того, что, подобно тому, как явное интегрирование дифференциальных уравнений, вообще говоря, невозможно, невозможным оказывается и качественное исследование сколько-нибудь общих дифференциальных уравнений с многомерным фазовым пространством. В книге обсуждается анализ дифференциальных уравнений с точки зрения структурной устойчивости, то есть устойчивости качественной картины по отношению к малым изменениям дифференциальных уравнений. Изложены основные результаты, полученные после первых работ А. А. Андронова и Л. С. Понтрягина в этой области: начала теории структурно устойчивых У-систем Аносова, все траектории которых экспоненциально неустойчивы, и теорема Смейла о неплотности множества структурно устойчивых систем. Обсуждается также вопрос о значении этих математических открытий для приложений (речь идет об описании устойчивых хаотических режимов движения, вроде турбулентных).

К наиболее мощным и часто применяемым методам исследования дифференциальных уравнений относятся различные асимптотические методы. В книге изложены основные идеи метода усреднения, восходящего к работам основоположников небесной механики и широко используемого во всех областях приложений, где нужно отделить медленную эволюцию от быстрых осцилляций (Н. Н. Боголюбов, Ю. А. Митропольский и др.).

Несмотря на обилие исследований по усреднению, в вопросе об эволюции даже для простейших многочастотных систем далеко не все ясно. В книге дается обзор работ о прохождении резонансов и о захвате в резонанс, направленных к выяснению этого вопроса.

Основой метода усреднения является идея уничтожения возмущений посредством подходящего выбора системы координат. Эта же идея

лежит в основе теории нормальных форм Пуанкаре. Метод нормальных форм является основным методом локальной теории дифференциальных уравнений, описывающей поведение фазовых кривых в окрестности особой точки или замкнутой фазовой кривой. В книге изложены основы метода нормальных форм Пуанкаре, включая доказательство фундаментальной теоремы Зигеля о линеаризации голоморфного отображения.

Важные применения метод нормальных форм Пуанкаре находит не только при исследовании отдельного дифференциального уравнения, но и в теории бифуркаций, когда предметом изучения является семейство уравнений, зависящих от параметров.

Теория бифуркаций изучает изменения качественной картины при изменении параметров, от которых зависит система. При общих значениях параметров обычно приходится иметь дело с системами общего положения (все особые точки простые и т. д.). Однако, если система зависит от параметров, то при некоторых значениях параметров неизбежно встречаются вырождения (например, слияние двух особых точек векторного поля).

В однопараметрическом семействе общего положения встречаются лишь простейшие вырождения (те, от которых нельзя избавиться малым шевелением семейства). Таким образом возникает иерархия вырождений по коразмерностям соответствующих поверхностей в функциональном пространстве всех изучаемых систем: в однопараметрических семействах общего положения встречаются лишь вырождения, соответствующие поверхностям коразмерности один, и т. д.

В последние годы в теории бифуркаций наблюдается значительный прогресс, связанный с применением идей и методов общей теории особенностей дифференцируемых отображений Х. Уитни.

Книга заканчивается главой о теории бифуркаций, в которой применяются развитые в предыдущих главах методы и описаны результаты, полученные в этой области, начиная с основополагающих работ А. Пуанкаре и А. А. Андронова.

При изложении всех вопросов автор стремился избежать аксиоматически-дедуктивного стиля, характерным признаком которого являются немотивированные определения, скрывающие фундаментальные идеи и методы; подобно притчам, их разъясняют лишь ученикам наедине.

Продолжающаяся, как утверждают, уже более 50 лет аксиоматизация и алгебраизация математики привела к неудобочитаемости столь большого числа математических текстов, что стала реальностью всегда угрожающая математике угроза полной утраты контакта с физикой и естественными науками. Автор старался вести изложение таким образом, чтобы книгой могли пользоваться не только математики, но все потребители теории дифференциальных уравнений.

У читателя настоящей книги предполагаются лишь очень небольшие общематематические представления в объеме примерно первых двух курсов университетской программы; достаточно (но не необходимо), например, знакомство с учебником В. И. Арнольда «Обыкновенные дифференциальные уравнения», М., 1974.[1]

Изложение построено таким образом, чтобы читатель мог пропускать места, оказавшиеся для него трудными, без большого ущерба для понимания дальнейшего: были приняты меры к тому, чтобы по возможности избегать ссылок из главы в главу и даже из параграфа в параграф.

Содержание настоящей книги составил материал ряда обязательных и специальных курсов, читавшихся автором на механико-математическом факультете МГУ в 1970–1976 годах для студентов-математиков II–III курсов, для слушателей факультета повышения квалификации и на экспериментальном потоке математиков естественнонаучного профиля.

Автор выражает благодарность студентам О. Е. Хадину, А. К. Ковальджи, Е. М. Кагановой и доц. Ю. С. Ильяшенко, чьи конспекты были очень полезными при подготовке этой книги. Составленный Ю. С. Ильяшенко конспект специального курса, а также конспекты лекций на экспериментальном потоке в течение ряда лет находились в библиотеке факультета. Автор благодарен многочисленным читателям и слушателям этих курсов за ряд ценных замечаний, использованных при подготовке книги. Автор благодарен рецензентам Д. В. Аносову и В. А. Плиссу за тщательное рецензирование рукописи, способствовавшее ее улучшению.

Июнь 1977 г. *В. Арнольд*

[1] При изложении нескольких отдельных вопросов используются или упоминаются также самые первоначальные сведения о дифференциальных формах, группах Ли и функциях комплексного переменного. Для понимания большей части книги знакомство с этими сведениями не обязательно.

Некоторые используемые обозначения

$\mathbb{R}$ — множество всех вещественных чисел.
$\mathbb{C}$ — множество всех комплексных чисел.
$\mathbb{Z}$ — множество всех целых чисел.
$\mathbb{R}^n$ — n-мерное вещественное линейное пространство.
$\exists$ — существует.
$\forall$ — для всякого.
$a \in A$ — элемент a множества A.
$A \subset B$ — подмножество A множества B.
$A \cap B$ — пересечение (общая часть) множеств A и B.
$A \cup B$ — объединение множеств A и B.
$A \setminus B$ — разность множеств A и B (часть A вне B).
$A \times B$ — прямое произведение множеств A и B (множество пар (a, b), $a \in A$, $b \in B$).
$A \oplus B$ — прямая сумма линейных пространств.
$f\colon A \to B$ — отображение f из A в B.
$x \mapsto y$ или $y = f(x)$ — отображение f переводит элемент x в элемент y.
$\operatorname{Im} f$ или $f(A)$ — образ отображения f (но $\operatorname{Im} z$ — мнимая часть z).
$f^{-1}(y)$ — полный прообраз точки y при отображении f (множество всех x, для которых $f(x) = y$).
$\operatorname{Ker} f$ — ядро линейного оператора f (полный прообраз нуля).
$\dot{f}$ — скорость изменения функции f (производная по времени t).
f', f_*, $\dfrac{df}{dx}$, $\dfrac{Df}{Dx}$ — производная отображения f.
T_xM — касательное пространство к многообразию M в точке x.
$A \Rightarrow B$ — из утверждения A следует B.
$A \Leftrightarrow B$ — утверждения A и B равносильны.
$\omega_1 \wedge \omega_2$ — внешнее произведение дифференциальных форм ω_1 и ω_2.
$f \circ g$ — суперпозиция отображений ($(f \circ g)(x) = f(g(x))$).
$L_v f$ — производная функции f по направлению векторного поля v.

Пусть $(x_1, \ldots, x_n)$ — координатные функции. Вектор v задается тогда своими компонентами $v_1, \ldots, v_n$. Производная по направлению поля v задается формулой

$$L_v f = v_1 \frac{\partial f}{\partial x_1} + \ldots + v_n \frac{\partial f}{\partial x_n}.$$

При фиксированной системе координат $(x_1, \ldots, x_n)$ используются следующие обозначения:

dx_k — функция от вектора, равная его k-й компоненте.

$\dfrac{\partial}{\partial x_k}$ — векторное поле, k-я компонента которого равна 1, а остальные компоненты равны нулю.

Для дифференциального уравнения $\dot{x} = v(x)$ область определения правой части называется *фазовым пространством*, точка x называется *фазовой точкой*, вектор $v(x)$ называется *вектором фазовой скорости*, v называется *векторным полем фазовой скорости*. Если $x = \varphi(t)$ — решение уравнения, то образ отображения φ называется *фазовой кривой*, а график отображения φ — *интегральной кривой*.

Для дифференциального уравнения $\dot{x} = v(x, t)$ область определения правой части называется *расширенным фазовым пространством*; v задает в расширенном фазовом пространстве *поле направлений*; если $x = \varphi(t)$ — решение, то график отображения φ называется *интегральной кривой*.

Глава 1
Специальные уравнения

При исследовании дифференциальных уравнений применяются методы всех отделов математики. В настоящей главе обсуждаются отдельные специальные уравнения и типы уравнений. Особое внимание обращается, с одной стороны, на значение рассматриваемых уравнений для приложений и с другой — на связи методов исследования с различными общематематическими вопросами (разрешение особенностей, диаграммы Ньютона, группы Ли симметрий и т. д.). Глава заканчивается элементарной теорией стационарного одномерного уравнения Шредингера и геометрической теорией нелинейного уравнения второго порядка.

§ 1. Дифференциальные уравнения, инвариантные относительно групп симметрий

В этом параграфе изложены общие соображения, на которых основаны методы интегрирования дифференциальных уравнений в явном виде. В качестве примера обсуждается теория подобия, т. е. теория однородных и квазиоднородных уравнений.

А. Группы симметрий дифференциальных уравнений. Рассмотрим векторное поле v в фазовом пространстве U.

Определение. Диффеоморфизм $g\colon U \to U$ называется *симметрией поля* v, если он переводит поле v в себя:

$$v(gx) = g_{*x}v(x).$$

Поле v называется тогда *инвариантным* относительно диффеоморфизма g.

Пример 1. Векторное поле с независящими от x компонентами на плоскости с координатами (x, y) инвариантно относительно сдвигов вдоль оси x (рис. 1).

Пример 2. Векторное поле $x\partial_x + y\partial_y$ на евклидовой плоскости (x, y) инвариантно относительно растяжений $g(x, y) = (\lambda x, \lambda y)$ и относительно поворотов.

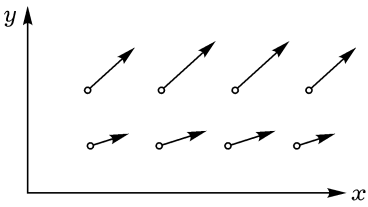


Рис. 1

Все симметрии данного поля образуют группу.

Задача. Найти группу симметрии поля $x\partial_x + y\partial_y$ на плоскости с координатами (x, y).

Рассмотрим поле направлений в расширенном фазовом пространстве.

Определение. Диффеоморфизм расширенного фазового пространства называется *симметрией поля направлений,* если он переводит поле в себя. Поле направлений называется тогда *инвариантным* относительно этого диффеоморфизма.

Пример 1. Поле направлений уравнения $\dot{x} = v(x)$ инвариантно относительно сдвигов вдоль оси t (рис. 2 а).

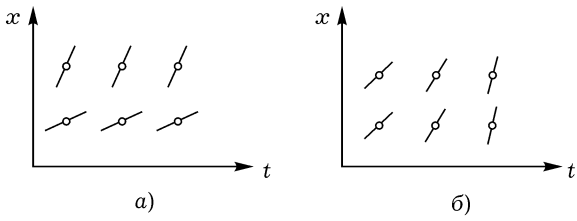


Рис. 2

Пример 2. Поле направлений уравнения $\dot{x} = v(t)$ инвариантно относительно сдвигов вдоль оси x (рис. 2 б).

Определение. Дифференциальное уравнение $\dot{x} = v(x)$ (соответственно, $\dot{x} = v(x, t)$) называется *инвариантным* относительно диффеоморфизма g фазового пространства (соответственно, расширенного фазового пространства), если векторное поле v (соответственно, поле направлений v) инвариантно относительно этого диффеоморфизма g; диффеоморфизм g называется тогда *симметрией* данного уравнения.

Теорема. *Симметрия уравнения переводит фазовые (интегральные) кривые уравнения в фазовые (интегральные) кривые того же уравнения.*

Доказательство.

Пусть $x = \varphi(t)$ — решение уравнения $\dot{x} = v(x)$ и g — симметрия. Тогда $x = g(\varphi(t))$ — тоже решение, поэтому симметрия переводит фазовую кривую в фазовую кривую. Для интегральных кривых доказательство аналогично. ■

Пример. Семейство интегральных кривых уравнения $\dot{x} = v(t)$ переходит в себя при сдвигах вдоль оси x, а уравнения $\dot{x} = v(x)$ — при сдвигах вдоль оси t.

Следующие примеры часто встречаются в приложениях под названием «теории подобия», «теории размерностей» или «соображений автомодельности».

Б. Однородные уравнения.

Определение. Поле направлений на плоскости без точки O называется *однородным,* если оно инвариантно относительно всех растяжений

$$g^\lambda(x, y) = (e^\lambda x, e^\lambda y), \qquad \lambda \in \mathbb{R}.$$

Дифференциальное уравнение $\dfrac{dy}{dx} = v(x, y)$ называется *однородным*, если его поле направлений однородно (рис. 3).

Иными словами, направления поля во всех точках каждого луча, выходящего из начала координат, должны быть параллельны:

$$v(e^\lambda x, e^\lambda y) \equiv v(x, y).$$

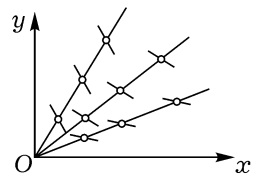


Рис. 3

Пример. Функция f называется однородной степени d, если $f(e^\lambda x, e^\lambda y) \equiv e^{\lambda d} f(x, y)$. Примером является любая форма (однородный многочлен) степени d. Пусть P и Q — две формы степени d от x и y. Дифференциальное уравнение

$$\dot{x} = P, \qquad \dot{y} = Q$$

задается векторным полем на плоскости. Соответствующее поле направлений в области $P \neq 0$ является полем направлений однородного уравнения

$$\frac{dy}{dx} = \frac{Q}{P}\left(\text{например, } \frac{dy}{dx} = \frac{ax+by}{cx+dy},\ \frac{dy}{dx} = \frac{x^2-y^2}{x^2+y^2} \text{ и т. п.}\right)$$

Замечание. Областью определения однородного поля не обязательно должна быть вся плоскость без точки O — однородные поля можно рассматривать в любой однородной (= инвариантной относительно растяжений) области, например, в угле с вершиной O и т. п.

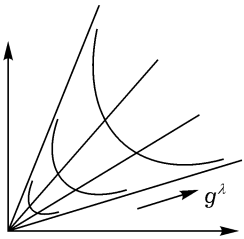


Рис. 4

Теорема. *Интегральная кривая однородного уравнения под действием растяжений g^λ переходит в интегральную кривую того же уравнения.*

Таким образом, для исследования однородного уравнения достаточно исследовать по одной интегральной кривой в каждом секторе плоскости.

Доказательство получается непосредственным применением теоремы п. А.

Задача. Докажите, что фазовые кривые системы $\dot{x} = P$, $\dot{y} = Q$, где P и Q формы степени d, получаются друг из друга при помощи гомотетий (рис. 4).

Если какая-либо из этих кривых замкнута и проходится за время T, то при растяжении g^λ из нее получится замкнутая фазовая кривая с периодом обращения $\dfrac{T}{e^{\lambda(d-1)}}$.

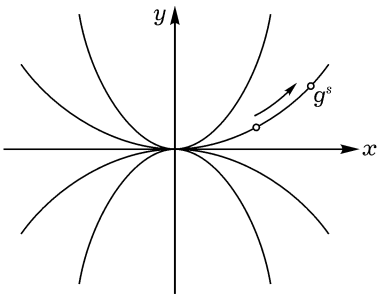


Рис. 5

В. Квазиоднородные уравнения и «соображения размерностей».

Зафиксируем вещественные числа α и β и рассмотрим семейство растяжений в разное число раз по разным направлениям на плоскости

$$g^s(x, y) = (e^{\alpha s}x,\ e^{\beta s}y). \qquad (1)$$

Заметим, что формула (1) задает однопараметрическую группу линейных преобразований плоскости (рис. 5).

Определение. Функция f называется *квазиоднородной степени* d, если

$$f(g^s(x, y)) \equiv e^{ds} f(x, y).$$

Пример. Если $\alpha = \beta = 1$, то получаются обычные однородные функции степени d.

Квазиоднородные степени складываются при умножении функций. Они называются также *весами*. Таким образом, x имеет вес α, y — вес β, x^2y — вес $2\alpha + \beta$ и т. д. Все квазиоднородные одночлены фиксированной степени легко увидеть на следующей *диаграмме Ньютона* (рис. 6). Будем изображать одночлен $x^p y^q$ точкой (p, q) на квадранте целочисленной решетки. Тогда показатели всех одночленов степени d — это целые точки на отрезке с уравнением $d = \alpha p + \beta q$ на плоскости показателей (p, q).

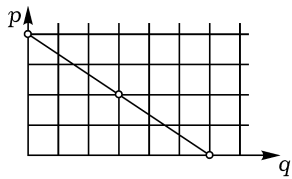


Рис. 6

Задача. Подобрать веса так, чтобы функция $x^2 + xy^3$ была квазиоднородной.

Определение. Дифференциальное уравнение $\dfrac{dy}{dx} = v(x, y)$ называется *квазиоднородным* (*с весами* α, β), если поле направлений v инвариантно относительно растяжений (1).

Из общей теоремы п. А о симметриях вытекает

Теорема. *Интегральные кривые квазиоднородного уравнения получаются друг из друга под действием растяжений* (1).

Задача. Докажите, что функция $v(x, y)$ задает квазиоднородное дифференциальное уравнение (с весами (α, β)), если и только если она квазиоднородная степени $d = \beta - \alpha$.

Замечание. Приведенные определения и теоремы легко переносятся на случай большего двух числа переменных и на дифференциальные уравнения порядка выше 1. В частности, легко доказывается

Теорема. *Пусть на плоскости* (x, y) *дана кривая* $\gamma\colon y = y(x)$, *и в точке* (x_0, y_0) *имеем* $\dfrac{d^k y}{dx^k} = F$. *Тогда для кривой* $g^s\gamma$ *в соответствующей точке будет*

$$\frac{d^k y}{dx^k} = e^{(\beta - k\alpha)s} F.$$

Иными словами, $\dfrac{d^k y}{dx^k}$ *преобразуется при преобразовании* (1) *как* $\dfrac{y}{x^k}$, чем и объясняется удобство обозначения $\dfrac{d^k y}{dx^k}$.

Задача. Докажите, что если частица в силовом поле с однородной степени d силой проходит траекторию Γ за время T, то эта же частица пройдет гомотетичную траекторию $\lambda\Gamma$ за время

$$T' = \lambda^{(1-d)/2} T.$$

Решение. Уравнение Ньютона $\dfrac{d^2x}{dt^2} = F(x)$, в котором F однородна степени d, переходит в себя при подходящих преобразованиях (1). Именно для этого нужно взять веса α (для x) и β (для t) так, чтобы $\alpha - 2\beta = \alpha d$. Итак, $\beta = \dfrac{1-d}{2}\alpha$. Поэтому растяжению $x' = \lambda x$ соответствует $T' = \lambda^{(1-d)/2} T$.

Задача. Докажите третий закон Кеплера: квадраты времен прохождения подобных траекторий в поле тяготения относятся как кубы линейных размеров.

Решение. Из решения предыдущей задачи при $d = -2$ (закон всемирного тяготения) получаем $T' = \lambda^{3/2} T$.

Задача. Выясните, как зависит от амплитуды период колебаний в случае возвращающей силы, пропорциональной отклонению (линейный осциллятор) и кубу отклонения (мягкая сила).

Ответ. Для линейного маятника период не зависит от амплитуды, а для мягкого обратно пропорционален амплитуде.

Задача. Известно, что волчок с вертикальной осью имеет критическую угловую скорость: если угловая скорость больше критической, волчок стоит вертикально устойчиво, а если меньше — то падает.

Как изменится критическая угловая скорость, если перенести волчок на Луну, где ускорение силы тяжести в шесть раз меньше земного?

Ответ. Уменьшится в $\sqrt{6}$ раз.

Г. Применения однопараметрических групп симметрий к понижению порядка.

Теорема. *Если известна однопараметрическая группа симметрий поля направлений в* $\mathbb{R}^n$, *то задача интегрирования соответствующего дифференциального уравнения сводится к задаче интегрирования уравнения в* $\mathbb{R}^{n-1}$.

В частности, если известна однопараметрическая группа симметрий поля направлений на плоскости, то соответствующее уравнение $\dfrac{dy}{dx} = f(x, y)$ *интегрируется явно.*

Доказательство.

Пусть $\{g^s\}$ — данная группа симметрий. Рассмотрим орбиты $\{g^s x\}$ потока $\{g^s\}$. Можно (по меньшей мере локально) определить $(n-1)$-мерное *пространство орбит* (фактор-пространство по действию g^s) и отображение p исходного пространства на фактор-пространство (орбиты потока $\{g^s\}$ p переводит в точки). Оказывается, исходное поле направлений переходит при отображении p в некоторое новое поле направлений в $(n-1)$-мерном пространстве орбит; его только и остается проинтегрировать. ■

Точнее говоря, рассмотрим какую-нибудь точку $x_0 \in \mathbb{R}^n$; мы предполагаем, что проходящая через точку x_0 орбита группы симметрий $\{g^s\}$ является кривой σ. Проведем через точку x_0 какую-либо $(n-1)$-мерную локальную трансверсаль Σ к кривой σ. В окрестности точки x_0 введем локальную систему координат (s, u), где паре $s \in \mathbb{R}$, $u \in \Sigma$ отвечает точка $g^s u$ исходного пространства. Тогда отображение p проектирования на пространство орбит и действие группы симметрий g^s задаются в окрестности точки x_0 формулами

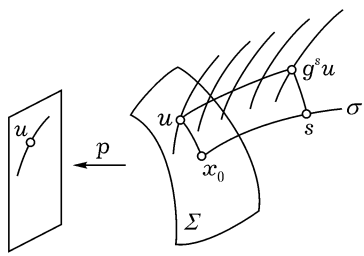


Рис. 7

$$p(s, u) = u, \qquad g^{s_1}(s_2, u) = (s_1 + s_2, u)$$

(точки поверхности Σ параметризуют локальные орбиты).

Заметим, что если группа g^s явно известна, то координаты (s, u) можно явно найти. Запишем в этих координатах наше исходное дифференциальное уравнение. Если наше поле направлений в точке x_0 не касается поверхности Σ (чего всегда можно добиться выбором Σ), то в окрестности этой точки наше уравнение принимает вид

$$\frac{du}{ds} = v(s, u).$$

При этом группа симметрий $\{g^s\}$ имеет вид сдвигов вдоль оси s, поэтому *функция v от s не зависит.* Векторное поле $v(u)$ на Σ определяет на этой $(n-1)$-мерной поверхности поле направлений; зная его интегральные кривые, мы найдем (квадратурой) решения уравнения $\dfrac{du}{ds} = v(u)$, а значит — интегральные кривые исходного уравнения.

В частном случае $n = 2$ выбор координат (s, u) сразу приводит к интегрируемому уравнению $\dfrac{du}{ds} = v(u)$.

Замечание. Практически часто бывает удобно вместо координаты s использовать подходящую функцию z переменной s. В такой системе координат уравнение, допускающее группу симметрий $\{g^s\}$, будет записываться в виде уравнения

$$\frac{du}{dz} = v(u)\, f(z)$$

(в случае $n = 2$ — в виде уравнения с разделяющимися переменными).

Пример. Однородное уравнение приводится к уравнению с разделяющимися переменными в полярной системе координат, а также в системе координат $u = \dfrac{y}{x}$, $z = x$ (рис. 8 а).

Здесь $\{g^s\}$ — однопараметрическая группа растяжений в e^s раз; Σ для полярной системы — окружность $x^2 + y^2 = 1$, а для второй системы координат — прямая $x = 1$; $z = e^s$.

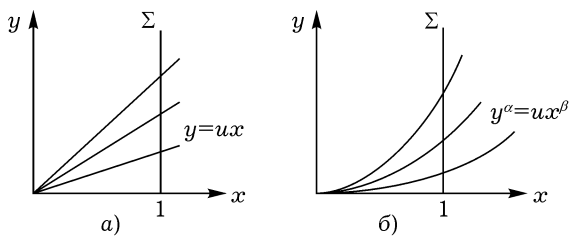


Рис. 8

Задача. В каких координатах интегрируется явно квазиоднородное уравнение

$$\frac{dy}{dx} = v(x, y),$$

где вес x равен α, вес y равен β (так что v — квазиоднородная функция степени $\beta - \alpha$).

Решение. Можно взять $u = \dfrac{y^\alpha}{x^\beta}$, $z = x$ (в области, где $x \neq 0$). См. рис. 8 б.

Задача. Выписать явно уравнение с разделяющимися переменными, к которому приводится уравнение предыдущей задачи в координатах (u, z).

Решение. $y^\alpha = ux^\beta$, поэтому $\alpha y^{\alpha-1}dy = x^\beta du + \beta u x^{\beta-1}dx$. Если $dy = v\,dx$, то $\alpha y^{\alpha-1}v\,dx = x^\beta du + \beta u x^{\beta-1}\,dx$, т. е.

$$\frac{du}{dx} = \frac{\alpha y^{\alpha-1}v - \beta u x^{\beta-1}}{x^\beta}.$$

Но $v(x, y) = x^{(\beta/\alpha)-1}w(u)$, поэтому

$$\frac{du}{dx} = \frac{\alpha w(u) - \beta u}{x}.$$

§ 2. Разрешение особенностей дифференциальных уравнений

Здесь коротко описан один важный общематематический прием, называемый разрешением особенностей, раздутием или σ-процессом.

А. σ-процесс.

Вблизи неособой точки все векторные поля устроены просто и одинаково.

Для исследования мелких деталей всевозможных математических объектов вблизи особых точек разработан специальный аппарат, имеющий, подобно микроскопу, большую разрешающую силу — т. н. *разрешение особенностей.* С аналитической точки зрения речь идет о выборе таких систем координат вблизи особой точки, в которых малым перемещениям вблизи особенности соответствуют большие изменения координат.

Этим свойством обладает уже полярная система координат, но переход к полярным координатам требует трансцендентных (тригонометрических) функций, поэтому алгебраически часто удобнее другая процедура — так называемый σ-процесс, или раздутие особенности.

Мы начнем с одной вспомогательной конструкции. Пусть $p\colon \mathbb{R}^2 \setminus O \to \mathbb{R}\mathrm{P}^1$ — стандартное расслоение, определяющее проективную прямую. (Проективная прямая — это многообразие, точками которого являются прямые на плоскости, проходящие через начало координат. Отображение p сопоставляет точке плоскости прямую, соединяющую ее с началом координат.)

Рассмотрим *график* Γ отображения p. Этот график представляет собой гладкую поверхность в прямом произведении $(\mathbb{R}^2 \setminus O) \times \mathbb{R}\mathrm{P}^1$ (рис. 9). Вкладывая плоскость без точки в плоскость, мы можем рассматривать график как гладкую поверхность Γ в прямом произведении $\mathbb{R}^2 \times \mathbb{R}\mathrm{P}^1$.

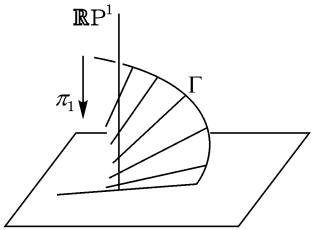


Рис. 9

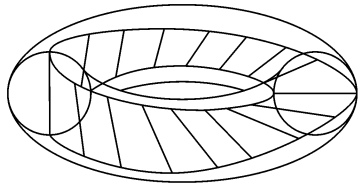


Рис. 10

Естественная проекция $\pi_1\colon \mathbb{R}^2 \times \mathbb{R}\mathrm{P}^1 \to \mathbb{R}^2$ диффеоморфно отображает график Γ на проколотую плоскость $\mathbb{R}^2 \setminus O$.

(Чтобы яснее представить себе все это, полезно заметить, что Γ локально имеет вид винтовой лестницы; в целом проективная прямая диффеоморфна окружности S^1, а произведение $\mathbb{R}^2 \times \mathbb{R}\mathrm{P}^1$ диффеоморфно внутренности баранки.)

Теорема. *Замыкание графика* Γ *отображения* p *в* $\mathbb{R}^2 \times \mathbb{R}\mathrm{P}^1$ *является гладкой поверхностью* $\Gamma_1 = \Gamma \cup (O \times \mathbb{R}\mathrm{P}^1)$. *Поверхность* Γ_1 *диффеоморфна листу Мёбиуса* (рис. 10).

Доказательство.

Пусть (x, y) — координаты на плоскости, $u = \dfrac{y}{x}$ — аффинная локальная координата в $\mathbb{R}\mathrm{P}^1$. Тогда (x, y, u) — локальная система координат в $\mathbb{R}^2 \times \mathbb{R}\mathrm{P}^1$. В этой системе координат Γ задается уравнением $y = ux$, $x \neq 0$, поэтому Γ_1 в этой системе координат задается локальным уравнением $y = ux$. Эта поверхность гладкая; она получается добавлением к части Γ, покрытой нашей системой координат, попавшей туда части проективной прямой $O \times \mathbb{R}\mathrm{P}^1$.

Доказательство гладкости Γ_1 завершается рассмотрением второй локальной системы координат (x, y, v), где $x = vy$.

Проекция $\pi_2\colon \mathbb{R}^2 \times \mathbb{R}\mathrm{P}^1 \to \mathbb{R}\mathrm{P}^1$ расслаивает Γ_1 на прямые. При полном обходе окружности $\mathbb{R}\mathrm{P}^1$ соответствующая прямая на $\mathbb{R}^2$ поворачивается на угол π; отсюда следует, что Γ_1 — лист Мёбиуса. ■

Определение. Переход от $\mathbb{R}^2$ к Γ_1 называется *σ-процессом с центром* O, или *раздутием* точки O в прямую $O \times \mathbb{R}\mathrm{P}^1$. Отображение $\pi_1\colon \Gamma_1 \to \mathbb{R}^2$ называется *антисигма-процессом* или *сжатием* окружности $O \times \mathbb{R}\mathrm{P}^1$ в точку O.

Отображение $\pi_1 : \Gamma_1 \to \mathbb{R}^2$, суженное на Γ, является диффеоморфизмом на проколотую плоскость. Поэтому всевозможные геометрические объекты на плоскости, имеющие особенность в точке O, переносятся на Γ_1. При этом особенности могут упрощаться или «разрешаться».

Пример. Рассмотрим три прямые, проходящие через точку O. На Γ_1 им соответствуют три прямые, пересекающие $\mathbb{R}\mathrm{P}^1$ уже в разных точках (рис. 11).

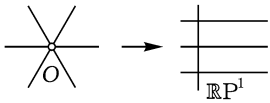


Рис. 11

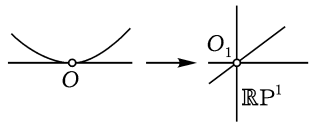


Рис. 12

Задача. Рассмотрим две кривые, имеющие в точке O касание порядка n (например, $y = 0$ и $y = x^2$, $n = 2$). Доказать, что на Γ_1 им соответствуют две кривые, имеющие в соответствующей точке O_1 касание порядка $n - 1$ (рис. 12).

Если после σ-процесса особенности не сводятся к трансверсальным пересечениям, то можно сделать еще σ-процесс в полученных особых точках, и т. д., пока все не сведется к трансверсальным пересечениям. Можно доказать, что особенности всякой алгебраической кривой могут быть таким образом разрешены (сведены к трансверсальным пересечениям) за конечное число шагов.

Задача. Разрешить особенность кривой $x^2 = y^3$.

Ответ. См. рис. 13.

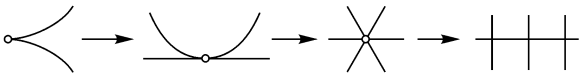


Рис. 13

Б. Формулы разрешения.

Практически σ-процесс означает переход от координат (x, y) к координатам $\left(x, u = \frac{y}{x}\right)$ там, где $x \neq 0$, и к координатам $\left(v = \frac{x}{y}, y\right)$ там, где $y \neq 0$ (рис. 14). Посмотрим, что происходит при этом с дифференциальным уравнением, заданным векторным полем на плоскости (x, y). Мы будем предполагать, что точка O — особая точка нашего векторного поля.

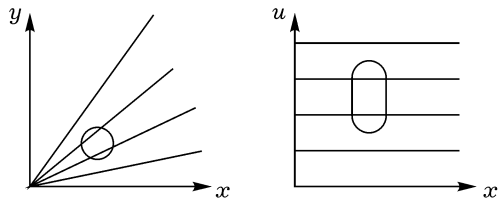


Рис. 14

Теорема. *Гладкое векторное поле w с особой точкой O превращается после σ-процесса в векторное поле на Γ, продолжающееся до гладкого поля на Γ_1.*

Доказательство.

Пусть w — поле, задающее систему $\dot{x} = P(x, y)$, $\dot{y} = Q(x, y)$. В координатах $\left(x, u = \frac{y}{x}\right)$ находим

$$\dot{x} = P(x, ux), \qquad \dot{u} = \frac{Q(x, ux) - uP(x, ux)}{x}.$$

Правые части гладкие, так как $P(0, 0) = Q(0, 0) = 0$. Во второй системе координат $\left(v = \frac{x}{y}, y\right)$ тоже получится гладкое поле. ■

Замечание. Может оказаться, что полученное векторное поле обращается в нуль на всей вклеенной при σ-процессе прямой. В таком случае поле можно поделить на x в области первой системы координат и на y в области второй. Деление не меняет направлений векторов поля. Поэтому на Γ_1 возникает поле направлений с особыми точками, лежащими на вклеенной прямой, но не заполняющими ее целиком. В окрестности каждой особой точки поле направлений задается гладким векторным полем.

Каждому «направлению входа» фазовых кривых исходного поля в O соответствует особая точка полученного поля, лежащая на вклеенной при σ-процессе прямой $\mathbb{R}\mathrm{P}^1$.

Если эти особые точки O_i устроены недостаточно просто, можно сделать в них σ-процессы. Продолжая таким же образом, можно в конце концов прийти к случаю, когда хотя бы одно из собственных чисел линеаризации поля в каждой особой точке отлично от нуля.

Во многих случаях уже первый σ-процесс позволяет разобраться в поведении фазовых или интегральных кривых вблизи особой точки.

Например, интегральные кривые однородного уравнения переходят при наших заменах координат $(x,\, y) \longmapsto \left(x,\, u = \dfrac{y}{x}\right)$ в интегральные кривые уравнений с разделяющимися переменными.

В. Пример. Исследование маятника с трением.

Проиллюстрируем метод на тривиальном примере линейного уравнения. Уравнение маятника с коэффициентом трения k имеет вид $\ddot{x} + k\dot{x} + x = 0$. Уравнение эквивалентно системе с фазовой плоскостью $(x,\, y)$:

$$\dot{x} = y \qquad \dot{y} = -ky - x.$$

Мы приходим к однородному уравнению

$$\frac{dy}{dx} = -k + \frac{y}{x}.$$

Согласно общей теории, после σ-процесса, т. е. в системе координат $\left(x, u{=}\dfrac{y}{x}\right)$, переменные должны разделяться. Действительно, $\dfrac{du}{dx} = -\dfrac{u^2 + ku + 1}{ux}$. Вводя еще $\ln|x| = z$, получаем

$$\frac{du}{dz} = -k - \left(u + \frac{1}{u}\right).$$

Исследуем интегральные кривые этого уравнения при различных значениях коэффициента $k > 0$. График функции $f = u + \dfrac{1}{u}$ — гипербола (рис. 15). Следовательно, график функции $-k{-}f(u)$ имеет вид, изображенный на рис. 16. Соответственно, интегральные кривые уравнения $\dfrac{du}{dz} = -k - f(u)$ будут иметь вид, изображенный на рис. 17. Возвращаясь на фазовую плоскость $(x,\, y)$, получаем рис. 18.

Итак, при малых значениях коэффициента трения $(0 < k < 2)$ маятник совершает бесконечное число колебаний, а при $k \geqslant 2$ направление движения маятника меняется не более одного раза.

Задача. Построить фазовые кривые уравнений $\dot{z} = az^n$ и $\dot{z} = a\overline{z}^n$, $z \in \mathbb{C}$.

Г. Пример. Период малых колебаний.

Теорема. *Предположим, что все фазовые кривые, проходящие через близкие к положению равновесия O точки, замкнуты. Тогда предел периода колебаний вблизи O при стремлении амплитуды колебаний к 0 равен периоду колебаний в линеаризованной системе.*

Доказательство.

После σ-процесса замкнутые фазовые кривые, обходящие один раз вокруг O, перейдут в кривые на листе Мёбиуса, замыкающиеся после двух

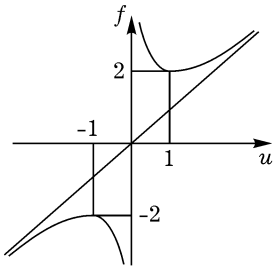


Рис. 15

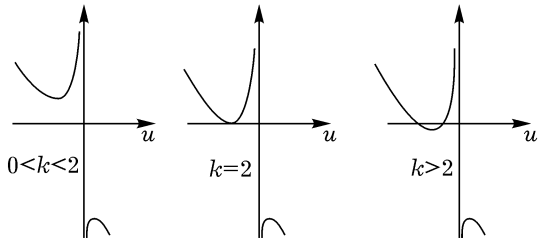


Рис. 16

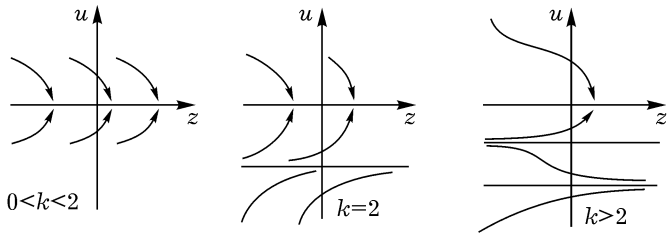


Рис. 17

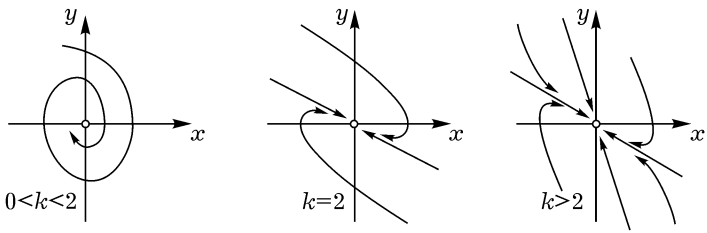


Рис. 18

оборотов; стремление амплитуды колебаний к 0 соответствует стремлению фазовой кривой на листе Мёбиуса к вклеиваемой при σ-процессе проективной прямой (к средней линии листа Мёбиуса).

По теореме о непрерывной зависимости решения от начального условия предел периода колебаний при стремлении амплитуды к 0 равен удвоенному периоду обращения по вклеенной прямой $\mathbb{R}\mathrm{P}^1$ в системе, полученной при σ-процессе. Но скорости движения по вклеенной прямой для данного по-

ля и для его линеаризации одинаковы (см. уравнение для $\dot{u}$ в п. Б). Легко проверить, что все фазовые кривые линеаризованного уравнения замкнуты. Эти замкнутые кривые в линейной системе проходятся за одинаковое время, так как линейное векторное поле переходит в себя при растяжениях фазовой плоскости. Следовательно, предел периода колебаний в исходной системе равен пределу периода колебаний в линеаризованной системе и, значит, равен просто периоду колебаний в линеаризованной системе. ■

Замечание. Предел, о котором шла речь, называется *периодом малых колебаний.*

Задача. Вычислить период малых колебаний маятника $\ddot{x} = -\sin x$ вблизи положения равновесия $x = 0$.

§ 3. Уравнения, не разрешенные относительно производных

В этом параграфе основные понятия теории дифференциальных уравнений, не разрешенных относительно производных, рассматриваются с точки зрения общей теории особенностей гладких отображений и геометрии пространства струй.

А. Основные определения.

Речь идет об уравнении

$$F(x, y, p) = 0, \tag{1}$$

где $p = \dfrac{dy}{dx}$.

Примеры. 1) $p^2 = x$; 2) $p^2 = y$; 3) $y = px + p^2$.

Трехмерное пространство с координатами (x, y, p) называется *пространством* 1-*струй функций* $y(x)$. (Две гладкие функции y_1, y_2 имеют в точке x_0 одинаковую k-струю, если $|y_1(x) - y_2(x)| = o(|x - x_0|^k)$; таким образом, 1-струя функции определяется выбором точки x, выбором значения y функции в этой точке и выбором значения p производной.)

Уравнение (1) задает в пространстве струй поверхность. Оказывается, на этой поверхности возникает поле направлений. Вот как оно строится. Рассмотрим какую-либо точку в пространстве струй. Компоненты вектора ξ, приложенного в этой точке, будем обозначать $dx(\xi)$, $dy(\xi)$, $dp(\xi)$. Таким образом, dx, dy и dp — это не какие-то мистические бесконечно малые величины, а вполне определенные линейные функции от вектора ξ.

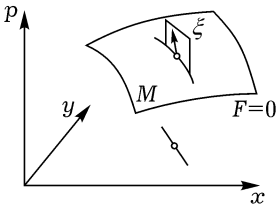


Рис. 19

В точке (x, y, p) пространства струй рассмотрим плоскость, составленную из векторов ξ, для которых $dy = p\,dx$. Иными словами, вектор ξ, приложенный в точке (x, y, p), попадает в указанную плоскость (рис. 19), если его проекция на евклидову плоскость (x, y) имеет направление с тангенсом угла наклона к оси x, равным p. Построенная плоскость называется *контактной плоскостью.* Таким образом, в каждой точке пространства 1-струй приложена контактная плоскость; все вместе они образуют контактное поле плоскостей (или, как еще говорят, контактную структуру) в пространстве 1-струй.

Задача*. Существуют ли поверхности в пространстве 1-струй, касающиеся в каждой своей точке приложенной в этой точке контактной плоскости?

Ответ. Нет.

Предположим, что поверхность в пространстве 1-струй, заданная уравнением (1), гладкая. [Это — не очень большое ограничение, так как для гладкой (= бесконечно дифференцируемой) функции F общего положения значение 0 не критическое и множество уровня 0 гладкое; если для данной функции это не так, то при почти всяком сколь угодно малом изменении функции F множество уровня 0 становится гладким: например, достаточно прибавить к F малую константу (см. теорему Сарда, § 10 п. E).]

Рассмотрим какую-либо точку на гладкой поверхности M, заданной уравнением (1), и предположим, что в этой точке касательная плоскость к поверхности не совпадает с контактной плоскостью. Тогда эти две плоскости пересекаются по прямой. Более того, касательные и контактные плоскости во всех близких точках поверхности пересекаются по прямым, так что в окрестности рассматриваемой точки возникает поле направлений на M.

Интегральными кривыми уравнения (1) называются интегральные кривые полученного поля направлений на поверхности M. Решить (или исследовать) уравнение (1) — значит найти (или исследовать) эти кривые. Связь интегральных кривых на M с графиками решений уравнения (1) на плоскости (x, y) обсуждается ниже; подчеркнем, что интегральные кривые на M определяются не в терминах решений уравнения (1), а в терминах контактных плоскостей.

Б. Регулярные точки и дискриминантная кривая.

Направление оси p в пространстве струй будем называть *вертикальным направлением.* Пусть M — гладкая поверхность в пространстве струй, заданная уравнением (1). Рассмотрим отображение проектирования вдоль вертикального направления

$$\pi\colon M \to \mathbb{R}^2, \qquad \pi(x, y, p) = (x, y).$$

Определение. Точка поверхности M называется *регулярной*, если она не является критической точкой отображения π.

Иными словами, точка поверхности M регулярна, если касательная плоскость в этой точке не вертикальна, или еще — если отображение проектирования на плоскость (x, y) в окрестности этой точки — диффеоморфизм.

Множество критических значений отображения π (т. е. проекция множества критических точек) называется *дискриминантной кривой* уравнения (1).

Пример. Для уравнения $p^2 = x$ дискриминантной кривой является ось y, а для уравнения $p^2 = y$ — ось x (рис. 20).

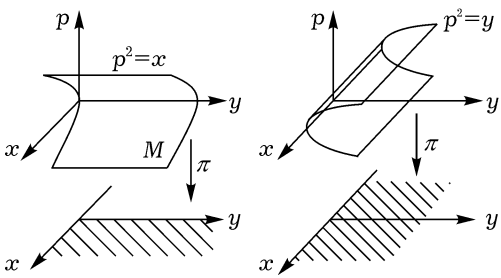


Рис. 20

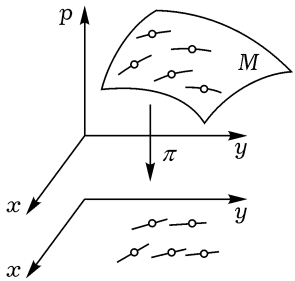


Рис. 21

Рассмотрим регулярную точку поверхности M. По теореме о неявной функции в окрестности этой точки M является графиком гладкой функции $p = v(x, y)$.

Теорема. *Проектирование на плоскость* (x, y) *переводит интегральные кривые уравнения* (1) *на* M *в окрестности регулярной точки*

в точности в интегральные кривые уравнения

$$\frac{dy}{dx} = v(x, y) \tag{2}$$

в окрестности проекции этой точки (рис. 21).

Доказательство.

По определению контактной плоскости, ее проекция на плоскость (x, y) есть прямая поля направлений (2). Поэтому поле направлений уравнения (1) в окрестности рассматриваемой регулярной точки на M переходит при локальном диффеоморфизме π в поле направлений уравнения (2); значит переходят друг в друга и интегральные кривые. ■

Замечание. В целом проекции интегральных кривых уравнения (1) на плоскость (x, y) не являются, вообще говоря, интегральными кривыми никакого поля направлений. Проекции интегральных кривых уравнения (1) на плоскость (x, y) имеют на дискриминантной кривой в общем случае точки возврата, но для некоторых уравнений (1) эти проекции остаются гладкими и в точках дискриминантной кривой.

В. Примеры.

Пример 1. $p^2 = x$ (рис. 22).

Поверхность M — параболический цилиндр. Дискриминантная кривая — ось y. Чтобы найти интегральные кривые, удобно взять за координаты на M не x и y, а p и y (тем более что последняя система координат — глобальная).

Запишем условия, наложенные на компоненты dx, dy, dp вектора ξ, приложенного в точке (x, y, p) поверхности M и принадлежащего нашему полю направлений:

$$\begin{cases} p^2 = x & \text{(условие принадлежности } M\text{)}; \\ 2p\,dp = dx & \text{(условие касания } M\text{)}; \\ dy = p\,dx & \text{(условие принадлежности контактной плоскости)}. \end{cases}$$

Следовательно, в координатах (p, y) интегральные кривые определяются из уравнения $dy = 2p^2\,dp$.

Итак, интегральные кривые на M даются соотношениями $y + C = \frac{2}{3}p^3$, $x = p^2$. Их проекции на плоскость (x, y) — полукубические параболы.

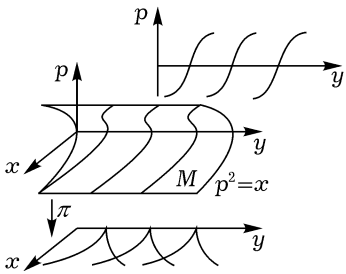


Рис. 22

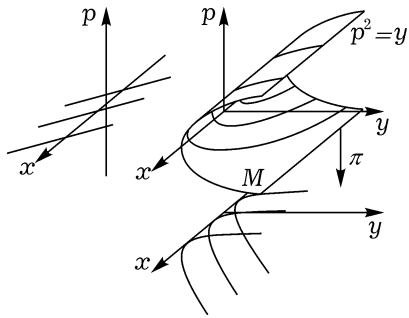


Рис. 23

Пример 2. $p^2 = y$ (рис. 23).

Поступая как в предыдущем примере, получаем

$$\begin{cases} p^2 = y, \\ 2p\,dp = dy, \\ dy = p\,dx. \end{cases}$$

Выбирая в качестве координат на поверхности M на этот раз x и p, получаем $p(dx - 2dp) = 0$, откуда либо $p = 0$, $y = 0$, либо

$$x = 2p + C, \quad y = p^2.$$

Проекции этих кривых на плоскость (x, y) — параболы, касающиеся дискриминантной кривой $y = 0$.

Пример 3 (Уравнение Клеро). $y = px + f(p)$ (рис. 24).

Поверхность M — линейчатая (ее пересечения с плоскостями $p = \mathrm{const}$ — прямые). За координаты на M удобно взять x и p. Интегральные кривые ищем из соотношений

$$\begin{cases} y = px + f(p), \\ dy = p\,dx + x\,dp + f'\,dp, \\ dy = p\,dx. \end{cases}$$

Находим $(x + f')\,dp = 0$. Точки, где $x + f' = 0$ — критические, а остальные регулярные. На плоскости с координатами (x, p) интегральные кривые — прямые $p = \mathrm{const} = C$; эти прямые, вообще говоря, пересекают линию критических точек $(x + f' = 0)$.

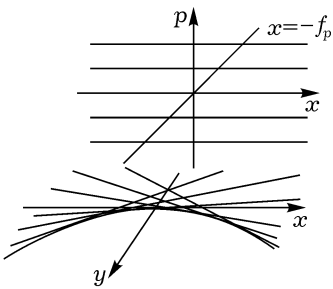


Рис. 24

Проекции интегральных кривых на плоскость (x, y) — прямые $y = Cx + f(C)$, касающиеся дискриминантной кривой. [Строго говоря, точки пересечения с критической линией не входят в интегральные кривые на M, так как для данного уравнения в этих точках поле направлений не определено: контактная плоскость касается M.]

Дискриминантная кривая находится из условий

$$y = px + f(p), \qquad x + f' = 0.$$

Например, если $f(p) = -\dfrac{p^2}{2}$, то дискриминантная кривая — парабола $y = \dfrac{x^2}{2}$, а проекции интегральных кривых — ее касательные.

Теория уравнения Клеро связана с важными общематематическими понятиями: преобразованием Лежандра и проективной двойственностью.

Г. Преобразование Лежандра.

Пусть дана функция f переменной x. *Преобразованием Лежандра*[1] этой функции называется новая функция g новой переменной p, которая определяется следующим образом. Рассмотрим график f на плоскости (x, y). Проведем прямую $y = px$ с тангенсом угла наклона к оси x, равным p. Найдем точку, где график всего дальше от прямой по направлению оси ординат. Рассмотрим разность ординат точек прямой и графика. Эта разность и есть значение функции g в точке p (рис. 25):

$$g(p) = \sup_x (px - f(x)).$$

Пример 1. Пусть $f(x) = \dfrac{x^2}{2}$. Вычисляя преобразование Лежандра, получаем $g(p) = \dfrac{p^2}{2}$.

[1] В литературе преобразованиями Лежандра называют несколько разных объектов, связываемых также с именами Минковского и Юнга, но мы не будем стремиться к полной педантичности терминологии.

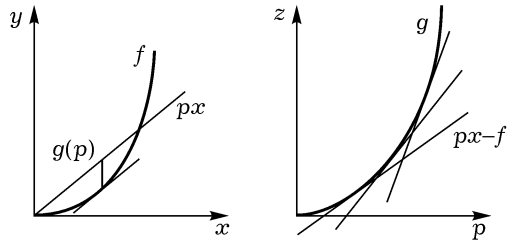


Рис. 25

Пример 2. Пусть $f(x) = \frac{x^\alpha}{\alpha}$. Тогда $g(p) = \frac{p^\beta}{\beta}$, где $\alpha^{-1} + \beta^{-1} = 1$. (Здесь x, p, α и β неотрицательны).

Если функция f строго выпуклая ($f'' > 0$), и ее производная задает диффеоморфизм прямой на прямую, то функция g также строго выпуклая; при этом верхняя грань достигается в единственной точке x, где $f'(x) = p$. Такие точки x и p называются *соответствующими* друг другу при преобразовании Лежандра.

Теорема. *Имеет место неравенство*

$$f(x) + g(p) \geqslant px.$$

Если f строго выпукла и f' — диффеоморфизм на, то равенство достигается если и только если точки x и p соответственные.

Доказательство.

Функция $px - f(x)$ не превосходит своей верхней грани $g(p)$. ■

Пример. При любых неотрицательных x, p имеет место неравенство $px \leqslant \frac{x^\alpha}{\alpha} + \frac{p^\beta}{\beta}$.

В формулируемых ниже следствиях предполагается, что рассматриваемые функции f и g строго выпуклы и их производные задают диффеоморфизмы прямой на прямую.

Следствие. *Преобразование Лежандра инволютивно: преобразованием Лежандра функции $g(p)$ является (при подходящем обозначении координаты) функция $f(x)$.*

Доказательство.

Действительно, неравенство предыдущей теоремы симметрично относительно f и g. ■

Следствие. *Переход от строго выпуклой функции g, задающей уравнение Клеро $y = px - g(p)$, к функции f, задающей огибающую решений по формуле $y = f(x)$, есть преобразование Лежандра.*

Доказательство.

График функции f является огибающей своих касательных. ■

Замечание. Преобразование Лежандра для функций n переменных определяется совершенно аналогично и обладает теми же свойствами. Если x — точка $\mathbb{R}^n$, то p — точка двойственного линейного пространства (пространства $\mathbb{R}^{n*}$ линейных функций на $\mathbb{R}^n$).

Д. Проективная двойственность.

Преобразование Лежандра является частным случаем одной общей конструкции проективной геометрии. Рассмотрим проективное пространство размерности n. Это пространство обозначается через $\mathbb{R}\mathrm{P}^n$.

Точка проективного пространства задается ненулевым вектором x из аффинного пространства $\mathbb{R}^{n+1}$, определенным с точностью до умножения на число, отличное от нуля. Это определение записывается кратко так:

$$\mathbb{R}\mathrm{P}^n = (\mathbb{R}^{n+1} \setminus O)/(\mathbb{R} \setminus O).$$

Гиперплоскость в проективном пространстве состоит из всех точек проективного пространства, для которых соответствующие точки аффинного пространства принадлежат одной гиперплоскости, проходящей через нуль.

Рассмотрим множество всех гиперплоскостей в n-мерном проективном пространстве. Это множество само естественно является n-мерным проективным пространством.

Действительно, гиперплоскость в проективном пространстве задается однородным уравнением

$$(a, x) = 0, \qquad x \in \mathbb{R}^{n+1}, \qquad a \in \mathbb{R}^{n+1*} \setminus O,$$

где $\mathbb{R}^{n+1*}$ — пространство линейных функций в $\mathbb{R}^{n+1}$ (это пространство само линейно, имеет размерность $n+1$ и называется двойственным или дуальным к исходному линейному пространству $\mathbb{R}^{n+1}$).

Таким образом, гиперплоскости в проективном пространстве соответствует ненулевой вектор в $\mathbb{R}^{n+1*}$, определенный с точностью до умножения на отличное от нуля число. Следовательно, *множество всех гиперплоскостей в* $\mathbb{R}\mathrm{P}^n$ *имеет естественную структуру проективного пространства размерности* n:

$$\mathbb{R}\mathrm{P}^{n*} = (\mathbb{R}^{n+1*} \setminus O)/(\mathbb{R} \setminus O).$$

Проективное пространство гиперплоскостей в данном проективном пространстве $\mathbb{R}\mathrm{P}^n$ называется *двойственным к* $\mathbb{R}\mathrm{P}^n$ *пространством* и обозначается через $\mathbb{R}\mathrm{P}^{n*}$. Например, пространство всех прямых на проективной плоскости само является проективной плоскостью, двойственной исходной.

Заметим, что двойственность является взаимным понятием, т. е. $\mathbb{R}\mathrm{P}^{n**} = \mathbb{R}\mathrm{P}^n$. Это следует из симметрии a и x в уравнении гиперплоскости $(a, x) = 0$.

Примеры. Все прямые, проходящие через одну точку проективной плоскости, образуют, как нетрудно сообразить, прямую на двойственной плоскости. Все прямые, проходящие через одну точку проективной плоскости внутри угла с вершиной в этой точке, образуют, как нетрудно видеть, отрезок на двойственной плоскости.

Все касательные к невырожденной кривой второго порядка на проективной плоскости образуют невырожденную кривую второго порядка на двойственной плоскости. Вообще, все касательные к любой гладкой кривой образуют (не обязательно гладкую) кривую на двойственной плоскости. Эта кривая называется двойственной к исходной.

Теорема. *Графики строго выпуклой функции и ее преобразования Лежандра проективно двойственны друг другу.*

Доказательство.

Рассмотрим всевозможные прямые на аффинной плоскости с координатами (x, y), не параллельные оси y. Эти прямые сами образуют плоскость: можно задать прямую уравнением $y = px - z$ и рассматривать (p, z) как аффинные координаты на новой плоскости. В таком случае преобразование Лежандра сводится к переходу от графика функции f к семейству касательных к этому графику: когда точка на плоскости (x, y) пробегает график функции f, касательная к графику функции f пробегает на плоскости (p, z) кривую, являющуюся графиком преобразования Лежандра, $z = g(p)$. ■

Таким образом, *преобразование Лежандра есть не что иное, как переход от кривой к проективно двойственной кривой, записанный в аффинных координатах.*

Пример. Пусть график f — выпуклая (вниз) ломаная. *Опорной прямой* называется прямая, от которой график лежит вверх, но с которой он имеет общую точку[1]. Рассмотрим все опорные прямые к выпуклой ломаной.

Легко проверить, что они сами образуют выпуклую ломаную на двойственной плоскости. Действительно, опорные прямые в каждой вершине исходной ломаной заполняют угол и, следовательно, образуют отрезок на двойственной плоскости. Точно так же вершины на двойственной плоскости получаются из отрезков исходной ломаной.

Проективная двойственность позволяет рассматривать более общие случаи, чем преобразование Лежандра.

Задача. Построить кривую, проективно двойственную кривой рис. 26.

Указание. Двойным касательным исходной кривой отвечают точки самопересечения двойственной кривой — следовательно, двойственная кривая имеет 4 точки самопересечения.

Точкам перегиба исходной кривой соответствуют точки возврата двойственной кривой. Действительно, если $f = x^3$, то касательная к графику в точке $x = t$ задается координатами $p = 3t^2$, $z = 2t^3$. Эти соотношения определяют на плоскости с координатами (p, z) кривую с точкой возврата. Итак, двойственная кривая имеет 8 точек возврата, по две между каждыми последовательными самопересечениями.

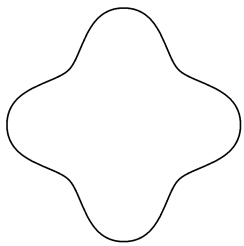


Рис. 26

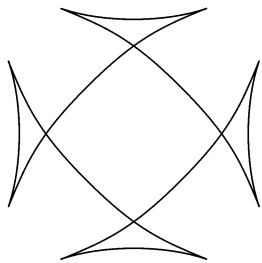


Рис. 27

[1]Вообще, *опорная гиперплоскость* к выпуклому телу есть плоскость, имеющая с телом общую точку и такая, что тело лежит в одном из полупространств, на которые плоскость делит пространство.

Далее, рассмотрим исходную кривую как пару пересекающихся эллипсов, слегка сглаженных вблизи точек пересечения (части каждого эллипса внутри другого выкинуты).

Двойственная кривая также связана с парой эллипсов. Точкам пересечения исходных эллипсов соответствуют двойные касательные двойственных. Отсюда уже легко сообразить, какой перестройкой двойственная кривая получается из пары пересекающихся эллипсов с их двойными касательными (рис. 27).

E. Преобразование Лежандра и сопряженные нормы.

Определение. *Нормой* в $\mathbb{R}^n$ называется вещественная неотрицательная выпуклая положительно однородная четная функция первой степени, равная нулю лишь в начале координат:

$$f \geqslant 0, \quad f(x) = 0 \Leftrightarrow x = 0, \quad f(\lambda x) = |\lambda| f(x), \quad f(x+y) \leqslant f(x) + f(y).$$

Пусть f — норма в $\mathbb{R}^n$. Функция f задается множеством, где она равна единице. Это множество является выпуклой центрально-симметричной относительно нуля гиперповерхностью в $\mathbb{R}^n$. Обратно, каждое компактное выпуклое центрально-симметричное относительно нуля тело, содержащее нуль в $\mathbb{R}^n$, определяет единственную норму, равную 1 на его границе. Эта гиперповерхность $f = 1$ называется *единичной сферой* нормы f.

Задача 1. Найти единичные сферы следующих норм в $\mathbb{R}^3$:

$$\text{а) } f = \sqrt{(x, x)}, \quad \text{б) } f = \max |x_i|, \quad \text{в) } f = \sum |x_i|.$$

Рассмотрим пространство $\mathbb{R}^{n*}$, сопряженное к $\mathbb{R}^n$.

Определение. *Сопряженная норма* в $\mathbb{R}^{n*}$ определяется как

$$g(p) = \max_{f(x) \leqslant 1} |(p, x)|.$$

Легко проверить, что g действительно норма.

Отношение сопряженности взаимно, так как определяющее неравенство можно переписать в симметричном виде $|(p, x)| \leqslant f(x) g(p)$.

Сопоставим каждой точке p сопряженного пространства гиперплоскость $p = 1$ в исходном пространстве.

Теорема. *Единичная сфера сопряженной нормы есть множество опорных гиперплоскостей единичной сферы исходной нормы.*

Доказательство.

Условие $g(p) = 1$ означает, что (p, x) на исходном шаре имеет максимум, равный 1, т. е. что плоскость $p = 1$ опорная для исходной сферы. ■

Множество всех гиперплоскостей, опорных к данной выпуклой гиперповерхности, называется *двойственной выпуклой гиперповерхностью.* Таким образом, единичные сферы двойственных норм двойственны.

Задача 2. Найти поверхности в $\mathbb{R}^3$, двойственные а) сфере, б) тетраэдру, в) кубу, г) октаэдру.

Задача 3. Найти нормы, сопряженные нормам задачи 1.

Из всего сказанного выше видно, что переход от выпуклой гиперповерхности к двойственной ей локально задается преобразованием Лежандра.

Ж. Задача об огибающих семейства плоских кривых.

Две гладкие функции двух переменных

$$x = x(s, t), \qquad y = y(s, t)$$

задают семейство кривых на плоскости, параметризованных параметром t (указывающим точку на кривой) и занумерованных параметром s (указывающим номер кривой).

Задача. Нарисовать семейства кривых, заданных функциями
а) $x = (s+t)^2$, $y = t$,
б) $x = s + st + t^3$, $y = t^2$ (s и t малы),
в) $x = (s+t^2)^2$, $y = t$.

Ответ. См. рис. 28.

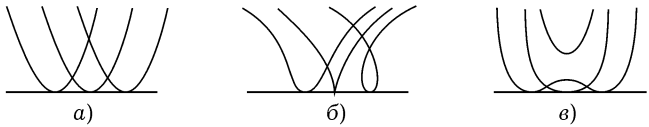


Рис. 28

Можно показать, что для семейства общего положения огибающая — кривая, единственные особенности которой точки возврата (как у полукубической параболы) и точки самопересечения; при этом в окрестности каждой точки гладкости огибающей семейство приводится к одной из нормальных форм а), б), в) гладкими заменами координат $X(x, y)$, $Y(x, y)$ и параметров $S(s)$, $T(s, t)$.

В то же время в окрестности общей точки дискриминантной кривой уравнение, не разрешенное относительно производной, приводится, вообще говоря, к нормальной форме

$$\left(\frac{dy}{dx}\right)^2 = x$$

гладкой заменой координат $X(x, y)$, $Y(x, y)$ (см. § 4).

В этих координатах проекциями интегральных кривых на плоскость (X, Y) являются полукубические параболы. Таким образом, дискриминантная кривая является огибающей проекций интегральных кривых лишь для исключительных уравнений (например, для уравнения Клеро). В частности, при небольшом общем изменении уравнения Клеро дискриминантная кривая из огибающей превратится в геометрическое место точек возврата проекций интегральных кривых.

§ 4. Нормальная форма уравнения, не разрешенного относительно производной, в окрестности регулярной особой точки

Здесь исследуются особенности семейства интегральных кривых для дифференциального уравнения общего положения, не разрешенного относительно производной.

А. Особые точки.

Рассмотрим уравнение

$$F(x, y, p) = 0, \qquad \text{где} \quad p = \frac{dy}{dx}, \tag{1}$$

заданное гладкой функцией F в некоторой области.

Мы предположим, что уравнение (1) задает гладкую поверхность в трехмерном пространстве струй с координатами (x, y, p). По теореме о неявной функции для этого достаточно, чтобы в точках, где $F = 0$, не обращался в нуль полный дифференциал функции F, что мы и будем предполагать.

Рассмотрим проекцию поверхности $F = 0$ на координатную плоскость (x, y) параллельно p-направлению.

Определение. Точка поверхности $F = 0$ называется *особой для уравнения* (1), если проектирование $(x, y, p) \longmapsto (x, y)$ поверхности на плоскость в окрестности этой точки не является локальным диффеоморфизмом поверхности на плоскость.

По теореме о неявной функции, особые точки — это те точки поверхности $F = 0$, в которых $\dfrac{\partial F}{\partial p} = 0$.

Б. Криминанта.

Рассмотрим множество всех особых точек уравнения (1). Это множество задается двумя уравнениями $F = 0$, $\dfrac{\partial F}{\partial p} = 0$ в трехмерном пространстве струй. Поэтому, вообще говоря, особые точки образуют кривую.

Определение. Множество особых точек уравнения $F = 0$ в трехмерном пространстве струй (x, y, p) называется *криминантой* уравнения.

По теореме о неявной функции, криминанта является гладкой кривой в трехмерном пространстве струй в окрестности каждой своей точки, в которой ранг производной отображения $(x, y, p) \longmapsto \left(F, \dfrac{\partial F}{\partial p}\right)$ трехмерного пространства на плоскость максимален (равен 2).

В. Дискриминантная кривая.

Определение. Проекция криминанты на плоскость (x, y) параллельно p-направлению называется *дискриминантной кривой*[1].

По теореме о неявной функции, окрестность точки криминанты диффеоморфно проектируется на плоскость (x, y) параллельно p-направлению, если криминанта в рассматриваемой точке не касается p-направления.

Замечание. Дискриминантная кривая может в этих условиях все же иметь особенности.

Они происходят от того, что в одну точку дискриминантной кривой могут проектироваться, вообще говоря, несколько точек криминанты. Эти особенности будут, вообще говоря, точками самопересечения дискриминантной кривой. Для «общего уравнения» в окрестности такой точки дискриминантная кривая состоит из двух ветвей, пересекающихся под ненулевым углом.

Точкам же, где криминанта касается p-направления, соответствуют «в общем случае» точки возврата на дискриминантной кривой.

Все более сложные особенности дискриминантной кривой, кроме точек самопересечения и возврата, устраняются малым шевелением уравнения. Особенности же этих двух типов сохраняются при малой деформации уравнения лишь немного смещаясь.

Г. Точки касания криминанты с контактной плоскостью.

В каждой точке (x, y, p) пространства струй имеется *контактная плоскость* $dy = p\,dx$. В частности, такая плоскость имеется в точках криминанты. Касательная к криминанте в данной точке может лежать в контактной плоскости или пересекать ее.

Определение. Точка криминанты называется *точкой касания с контактной плоскостью*, если касательная к криминанте в этой точке лежит в контактной плоскости.

Заметим, что точки касания криминанты с p-направлением являются точками касания с контактной плоскостью. Действительно, контактная плоскость в каждой точке содержит p-направление.

[1] Это определение — переформулировка определения дискриминантной кривой из § 3.

Д. Регулярные особые точки.

Определение. Особая точка уравнения (1) называется *регулярной*, если в этой точке выполнено условие гладкости криминанты[1]

$$\operatorname{rank}((x,\, y,\, p) \longmapsto (F,\, F_p)) = 2$$

и криминанта не касается контактной плоскости.

Пример. Рассмотрим уравнение $p^2 = x$. Криминанта задается уравнениями $p = 0$, $x = 0$. Это — ось y. Условие гладкости выполнено. Касательный к криминанте вектор $(0,\, 1,\, 0)$ не лежит в контактной плоскости $dy = 0\,dx$. Следовательно, каждая особая точка уравнения $p^2 = x$ регулярна.

Замечание. Для уравнения «общего положения» почти все особые точки регулярны: нерегулярные точки лежат на криминанте дискретно. Если для данного уравнения это не так, то во всяком случае этого можно добиться малым шевелением уравнения. (Обоснование этого и предыдущих «соображений общего положения» проводится с помощью теоремы Сарда, § 10.)

Е. Теорема о нормальной форме.

Теорема. *Пусть* (x_0, y_0, p_0) *— регулярная особая точка уравнения* $F(x, y, p) = 0$. *Тогда существует диффеоморфизм окрестности точки* $(x_0,\, y_0)$ *плоскости* $(x,\, y)$ *на окрестность точки* $(0,\, 0)$ *плоскости* $(X,\, Y)$, *приводящий уравнение* $F = 0$ *к виду* $P^2 = X$, *где* $P = \dfrac{dY}{dX}$.

Пояснение. Уравнение $F = 0$ задает поверхность в трехмерном пространстве линейных элементов на плоскости $(x,\, y)$. Диффеоморфизм плоскости переводит каждый линейный элемент в новый линейный элемент. Утверждается, что часть поверхности $F = 0$ вблизи регулярной особой точки можно перевести в часть поверхности $P^2 = X$ вблизи точки $(X = 0,\, Y = 0,\, P = 0)$.

Следствие. *Семейство интегральных кривых уравнения* (1) *в окрестности регулярной особой точки диффеоморфно как семейство кривых на плоскости* $(x,\, y)$ *семейству полукубических парабол* $y = x^{3/2} + C$.

Доказательство.

Указанный в теореме диффеоморфизм переводит интегральные кривые уравнения (1) на плоскости $(x,\, y)$ в интегральные кривые уравнения $P^2 = X$ на плоскости $(X,\, Y)$. Эти последние интегральные кривые являются полукубическими параболами с острием на дискриминантной кривой: $\dfrac{dY}{dX} = \sqrt{X}$, $Y = \dfrac{2}{3}X^{3/2} + C$. ■

[1]Рангом (rank) отображения называется ранг его производной.

Ж. Доказательство теоремы о нормальной форме.

Доказательство.

1°. *Редукция к случаю, когда криминантой является ось y.*

Пусть (x_0, y_0, p_0) — регулярная особая точка уравнения $F(x, y, p) = 0$. Тогда дискриминантная кривая в окрестности точки (x_0, y_0) гладкая. Рассмотрим проекции контактных плоскостей в точках криминанты на плоскость (x, y). Мы получим в окрестности точки (x_0, y_0) гладкое семейство прямых, не касающихся дискриминантной кривой.

Выберем теперь локальную систему координат на плоскости (x, y) вблизи точки (x_0, y_0) так, чтобы 1) дискриминантная кривая имела уравнение $x = 0$; 2) линии $y = \text{const}$ пересекали дискриминантную кривую по построенным только что направлениям.

Эти координаты мы будем по-прежнему обозначать через (x, y); производная $\dfrac{dy}{dx}$ по-прежнему будет обозначаться через p. Особая точка (x_0, y_0, p_0) получает теперь координаты $(0, 0, 0)$.

2°. *Анализ условий регулярности.*

Криминанта, согласно нашему выбору системы координат, является осью y: на ней $x = 0, p = 0$ ($y = \text{const}$). Из этого следует, что для нашего уравнения записанного в введенных координатах, $F(0, y, 0) = 0$, $F_p(0, y, 0) = 0$. Условие регулярности криминанты имеет теперь вид

$$\det \left| \frac{D(F, F_p)}{D(x, p)} \right| \neq 0, \quad \text{т. е.} \quad \det \begin{vmatrix} F_x & F_{xp} \\ F_p & F_{pp} \end{vmatrix} \neq 0,$$

(так как в точках криминанты $F_y = 0$, $F_{yp} = 0$). Далее, в точках криминанты $F_p = 0$. Следовательно, условие регулярности криминанты запишется в виде

$$F_x(0, y, 0) \neq 0, \qquad F_{pp}(0, y, 0) \neq 0.$$

Условие некасания с контактной плоскостью выполнено автоматически.

Разложим F в ряд Тейлора по p с остаточным членом степени 2:

$$F(x, y, p) = A(x, y) + pB(x, y) + p^2 C(x, y, p).$$

Из полученных выше соотношений следует, что $A(0, y) = 0$, $B(0, y) = 0$. Поэтому мы можем записать

$$A(x, y) = x\alpha(x, y), \qquad B(x, y) = x\beta(x, y),$$

где α и β — гладкие функции.

Условия регулярности криминанты имеют вид $A_x(0, y) \neq 0$, $C(0, y, 0) \neq 0$. Мы можем даже предположить для дальнейшего, что $C > 0$, $A_x < 0$ (если это не так, сменим знаки F и/или x). Итак, $\alpha(0, 0) < 0$, $C(0, 0, 0) > 0$.

3°. *Исследование квадратного уравнения.*

Рассмотрим соотношение $F = 0$ как квадратное уравнение относительно p с коэффициентами C, B, A. Мы получаем

$$p = \frac{-B \pm \sqrt{B^2 - 4AC}}{2C} = \frac{-x\beta \pm \sqrt{x\gamma}}{2C},$$

где $\gamma = -4\alpha C + x\beta^2$ есть функция от (x, y, p); при этом $\gamma(0, 0, 0) = -4\alpha(0, 0, 0) \cdot C(0, 0, 0) > 0$.

Пусть, наконец, $x = \xi^2$. Тогда получаем, оставляя лишь знак «+» в «±»,

$$p = \frac{-\xi^2\beta(\xi^2, y) + \xi\sqrt{\gamma(\xi^2, y, p)}}{2C(\xi^2, y, p)}.$$

Применим к этому уравнению относительно $p(\xi, y)$ теорему о неявной функции. Получим решение $p = \xi\omega(\xi, y)$, где ω гладкая функция, $\omega(0, 0) \neq 0$.

4°. *Дифференциальное уравнение для $y(\xi)$.*

Заметим, что $p = \dfrac{dy}{dx} = \dfrac{1}{2\xi}\dfrac{dy}{d\xi}$. Поэтому мы получили дифференциальное уравнение для $y(\xi)$:

$$\frac{dy}{d\xi} = 2\xi^2\omega(\xi, y), \qquad \omega(0, 0) \neq 0. \tag{2}$$

Интегральные кривые на плоскости (ξ, y) пересекают ось $\xi = 0$, имея с линиями $y = \mathrm{const}$ касание второго порядка. Поэтому уравнение имеет первый интеграл вида $I = (\xi, y) = y - \xi^3 K(\xi, y)$, где K — гладкая функция, $K(0, 0) \neq 0$ (I — координата точки пересечения с осью $\xi = 0$; $K \neq 0$, так как $\omega \neq 0$).

5°. *Построение нормализующих координат.*

Разложим K на четную и нечетную по ξ части:

$$K(\xi, y) = L(\xi^2, y) + \xi M(\xi^2, y).$$

Здесь L и M — гладкие функции от x и y, $L(0, 0) \neq 0$. В этих обозначениях $I(\xi, y) = y - \xi^4 M(\xi^2, y) - \xi^3 L(\xi^2, y)$. Введем новые переменные Y, Ξ по формулам

$$\Xi = \xi\sqrt[3]{L(\xi^2, y)}, \qquad Y = y - \xi^4 M(\xi^2, y).$$

Тогда $I = Y - \Xi^3$.

Рассмотрим еще $X = \Xi^2$. Тогда

$$X = x\sqrt[3]{L^2(x, y)}, \qquad Y = y - x^2 M(x, y).$$

Эти формулы задают локальный диффеоморфизм плоскости в окрестности точки $(0, 0)$, так как $L(0, 0) \neq 0$. Первый интеграл принимает вид

$$I = Y - X^{3/2}.$$

Теперь $\left(\dfrac{dY}{dX}\right)^2 = \dfrac{9}{4}X$, и указанная в формулировке теоремы нормальная форма получается растяжением одной из координатных осей. ■

3. Замечания.

Основным моментом приведенного доказательства является подстановка $x = \xi^2$, т. е. переход к двулистному накрытию плоскости (x, y) с ветвлением вдоль дискриминантной кривой. Из топологических соображений (правда, в комплексной области) заранее ясно, что на этом двулистном накрытии двузначность $p(x, y)$ исчезает и уравнение распадается на два. Возня с квадратным уравнением нужна лишь для обоснования этого обстоятельства в вещественной области. Полученное на накрытии уравнение (2) остается привести к нормальной форме диффеоморфизмом, опускаемым с накрытия на исходную плоскость — что легко достигается разложением первого интеграла на четную и нечетную по ξ составляющие.

Первое доказательство теоремы нормальной форме было дано Ю. А. Бродским; оно основано на работе Р. Тома, который привел уравнение лишь к виду $p^2 = xE(x, y)$.

Замечание. Наше доказательство использовало представление четной функции в виде функции от квадрата аргумента. Для аналитических функций (или для формальных рядов) такое представление очевидно. В случае же гладких функций оно нуждается в обосновании.

Действительно, четную бесконечно дифференцируемую функцию можно рассматривать как функцию от квадрата аргумента, заданную на положительной полуоси. Она бесконечно дифференцируема во всех точках этой полуоси, включая нуль. Требуется же представить ее как сужение на положительную полуось функции, бесконечно дифференцируемой на всей оси.

Возможность такого представления означает возможность гладкого продолжения на отрицательную полуось. Она гарантируется теоремой (Э. Бореля) о существовании бесконечно дифференцируемых функций на прямой с любым рядом Тейлора в нуле. На доказательстве этой теоремы (впрочем, несложном) мы не останавливаемся.

Кроме регулярных особых точек, в отдельных точках гладкой дискриминантной кривой уравнения общего положения встречаются точки касания контактной плоскости с поверхностью уравнения. В окрестности такой точки уравнение общего положения диффеоморфизмом плоскости (x, y) приводится к нормальной форме $y = (p + kx)^2$ (*А. А. Давыдов.* Нормальная форма уравнения, не разрешенного относительно производной, в окрестности его особой точки. Функциональный анализ и его приложения. 1985, т. 19, вып. 2, 1–10; *А. В. Пхакадзе, А. А. Шестаков.* О классификации особых точек дифференциального уравнения первого порядка, не разрешенного относительно производной. Матем. сборник, 1959, т. 49, вып. 1, с. 3–12.)

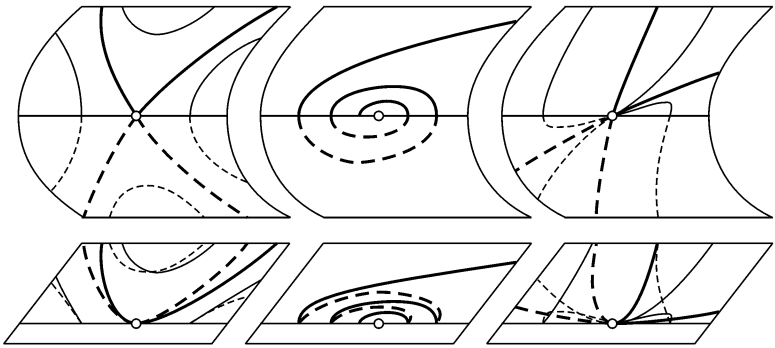


Рис. 29

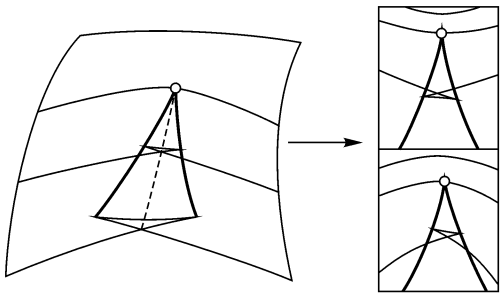


Рис. 30

Поле направлений на поверхности уравнения имеет в соответствующей точке криминанты такую же особенность, как поле направлений векторного поля на плоскости в окрестности обыкновенной особой точки типа седло, фокус или узел. Поэтому возникающие здесь особые точки неявных дифференциальных уравнений называются *сложенным седлом*, *фокусом* или *узлом*: они получаются из обычных при помощи отображения складывания. Интересно, что складывание не вносит новых модулей: параметр k в нормальной форме определяется отношением собственных чисел линеаризации векторного поля, при складывании фазового портрета которого получается сложенная особая точка (рис. 29).

Напротив, особые точки, в которых дискриминантная кривая имеет точку возврата (проекции точек сборки Уитни отображения поверхности уравнения на плоскость (x, y)) имеют функциональные модули относительно не

только диффеоморфизмов, но и даже относительно гомеоморфизмов плоскости (x, y) (см. цитированную работу Давыдова). Брюс заметил, что при проекции интегральных кривых на плоскость (x, y) в этом случае можно получить из семейства сечений ласточкиного хвоста $\{(a, b, c)\colon t^4 + at^2 + bt + c$ имеет кратный корень$\}$ плоскостями $a = \text{const}$ (рис. 30) при надлежащем гладком отображении ранга 2 пространства (a, b, c) на плоскость (x, y) (это следует из описанных перестроек волновых фронтов в *V. I. Arnold.* Wavefronts evolution and the equivariant Morse lemma. Comm. Pure and Appl. Math. 1976, **29**, 6, 557–582, *J. W. Bruce.* A note on first-order differential equations of degree greater than one and wavefront evolution. Bull. London Math. Soc. 1984, **16**, 139–144.

Нерегулярные особые точки общего положения исследованы подробно А. А. Давыдовым (*А. А. Давыдов.* Нормальная форма уравнения, не разрешенного относительно производной, в окрестности его особой точки.)

§ 5. Стационарное уравнение Шредингера

В этом параграфе изложены простейшие математические основы элементарной квантовой механики. Мы не останавливаемся на физических мотивировках вводимых определений, но используем физическую терминологию для описания свойств решений уравнения.

А. Определения и обозначения.

В физике уравнение

$$\frac{d^2\Psi}{dx^2} + (E - U(x))\Psi = 0 \tag{1}$$

называется *стационарным уравнением Шредингера.*

Независимая переменная x называется *декартовой координатой частицы.* Неизвестная функция Ψ, вообще говоря, комплексная, называется *волновой функцией частицы*; решения уравнения Шредингера называются *состояниями частицы.* Спектральный параметр E называется *энергией частицы.* Известная функция U называется *потенциалом* или *потенциальной энергией частицы. Квантовая механика* занимается в основном исследованием свойств уравнения (1) и обобщающих его уравнений и систем уравнений с частными производными.

Пример. Пусть $U = 0$. Тогда частица называется *свободной.* Уравнение Шредингера для свободной частицы с энергией $E = k^2$ имеет вид

$$\Psi_{xx} + k^2\Psi = 0. \tag{2}$$

Это уравнение имеет два линейно независимых решения

$$\Psi_+ = e^{ikx}, \qquad \Psi_- = e^{-ikx}.$$

Эти два решения называются *частицей, движущейся вправо* (с импульсом $k > 0$), и *частицей, движущейся влево*, соответственно. Таким образом, пространство состояний свободной частицы с энергией E — двумерное комплексное пространство.

Квадрат модуля волновой функции физики называют *плотностью вероятности* того, что частица находится в данном месте. Таким образом, свободная частица с импульсом k «с одинаковой вероятностью находится в любой точке» (этой терминологией можно пользоваться, не заботясь о том, что означают эти слова и как все это связано с теорией вероятностей).

Б. Потенциальные барьеры.

Предложим, что потенциал финитен (отличен от нуля лишь в некоторой области). Если $U \geqslant 0$, то говорят, что задан *потенциальный барьер*, а если $U \leqslant 0$ — *потенциальная яма*. Область, где потенциал отличен от нуля, называется *носителем* потенциала (рис. 31).

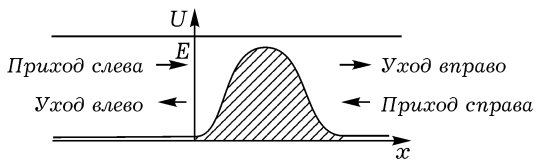


Рис. 31

Предположим, что энергия частицы $E = k^2$ *положительна*. Тогда левее носителя уравнение Шредингера (1) совпадает с уравнением свободной частицы (2). Следовательно, уравнение Шредингера имеет два решения, которые *левее носителя* совпадают с e^{ikx} и e^{-ikx} соответственно. Эти два решения называются *частицей, приходящей слева*, и *частицей, уходящей влево*, соответственно. Заметим, что решения эти определены при всех x, но совпадают с e^{ikx} и e^{-ikx} лишь левее носителя.

Точно так же существуют два решения, которые совпадают с e^{ikx} и e^{-ikx} *правее носителя*. Эти решения называются *частицей, уходящей вправо*, и *частицей, приходящей справа*, соответственно.

Задача. Может ли частица, пришедшая слева, целиком отразиться влево (т. е. может ли волновая функция справа от барьера быть нулем, а слева нет)? Целиком уйти вправо?

Ответ. Нет, да.

В. Оператор монодромии.

Определение. *Оператором монодромии* уравнения Шредингера (1) с финитным потенциалом называется линейный оператор, действующий из пространства состояний *свободной* частицы с энергией $E = k^2$ в себя, определенный следующим образом.

Решению уравнения свободной частицы (2) сопоставляется решение уравнения Шредингера, совпадающее с ним левее носителя, а этому решению — его значение правее носителя.

Оказывается, оператор монодромии обладает замечательными свойствами (1, 1)-унитарности. Чтобы их сформулировать, введем следующие

Обозначения. Обозначим через $\mathbb{R}^2$ *пространство вещественных решений уравнения Шредингера* (1). *Пространство состояний частицы* (т. е. пространство комплексных решений уравнения), является комплексификацией $\mathbb{R}^2$; мы его обозначим через $\mathbb{C}^2 = {}^{\mathbb{C}}\mathbb{R}^2$. Этому пространству принадлежат все четыре состояния приходящей и уходящей влево и вправо частицы.

Пространство вещественных решений уравнения свободной частицы (2) обозначим через $\mathbb{R}^2_0$ (так как для свободной частицы $U = 0$). В этом пространстве имеется естественный базис

$$e_1 = \cos kx, \qquad e_2 = \sin kx.$$

Пространство состояний свободной частицы мы обозначим через $\mathbb{C}^2_0$; это комплексификация пространства $\mathbb{R}^2_0$. Естественный базис образуют состояния частиц, движущихся вправо и влево соответственно. Мы обозначим их через

$$f_1 = e^{ikx}, \qquad f_2 = e^{-ikx}.$$

Заметим, что e_1, e_2 тоже определяют базис в пространстве состояний. Эти два базиса связаны соотношениями

$$f_1 = e_1 + ie_2, \qquad f_2 = e_1 - ie_2.$$

Определение. Группа SU(1, 1) (1, 1)-*унитарных унимодулярных матриц* состоит из всех комплексных матриц второго порядка с определителем 1, сохраняющих эрмитову форму $|z_1|^2 - |z_2|^2$. Иными словами, это матрицы $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, для которых

$$|a|^2 - |b|^2 = |c|^2 - |d|^2 = 1, \quad a\bar{c} - b\bar{d} = 0, \quad ad - bc = 1.$$

Теорема. *Матрица оператора монодромии в базисе* (f_1, f_2) *принадлежит группе* SU(1, 1).

Причина, по которой оператор монодромии принадлежит SU(1, 1), состоит в том, что фазовый поток уравнения (1) сохраняет площади. Для доказательства я напомню некоторые сведения о группе SU(1, 1).

Г. Алгебраическое отступление: группа SU(1, 1).

Рассмотрим вещественное линейное пространство $\mathbb{R}^2$ и его комплексификацию $\mathbb{C}^2$. Выберем в $\mathbb{R}^2$ элемент площади и будем обозначать через $[\xi, \eta]$ ориентированную площадь параллелограмма, натянутого на векторы ξ и η. Кососкалярное произведение $[\,,\,]$ называется *симплектической структурой.* Если в $\mathbb{R}^2$ фиксирован базис (e_1, e_2), на котором $[e_1, e_2] = 1$, то $[\xi, \eta]$ равно определителю, составленному из компонент векторов ξ, η в базисе e_1, e_2.

Комплексификация билинейной формы $[\,,]$ задает симплектическую структуру в $\mathbb{C}^2$; мы будем и ее обозначать теми же скобками.

Заметим, что форма $[\,,]$ невырождена: если $[\xi, \eta] = 0$ для всех ξ, то $\eta = 0$.

Рассмотрим в $\mathbb{C}^2$ *эрмитову форму* $\langle \xi, \eta \rangle = \frac{i}{2}[\xi, \overline{\eta}]$. Это действительно эрмитова форма: $\langle \lambda\xi, \eta \rangle = \lambda\langle \xi, \eta \rangle$, $\langle \xi, \eta \rangle = \overline{\langle \eta, \xi \rangle}$. Для дальнейшего полезно вычислить эрмитовы произведения векторов $f_1 = e_1 + ie_2$ и $f_2 = e_1 - ie_2$. Легко проверяется

Лемма. *Имеют место соотношения*

$$\langle f_1, f_1 \rangle = 1, \qquad \langle f_2, f_2 \rangle = -1, \qquad \langle f_1, f_2 \rangle = 0.$$

Доказательство.

Например,

$$\langle f_1, f_1 \rangle = \frac{i}{2}[f_1, \overline{f}_1] = \frac{i}{2}[f_1, f_2] = \frac{i}{2}\begin{vmatrix} 1 & i \\ 1 & -i \end{vmatrix} = 1. \qquad \blacksquare$$

Таким образом, эрмитова форма $\langle\,,\,\rangle$ имеет «тип $(1,1)$» (один положительный и один отрицательный квадрат в каноническом виде $\langle z,\, z\rangle = |z_1|^2 - |z_2|^2$).

Рассмотрим теперь линейные преобразования плоскости $\mathbb{C}^2$, сохраняющие эрмитову, симплектическую и вещественную структуры.

Определение. Группа линейных преобразований плоскости $\mathbb{C}^2$, сохраняющих эрмитову форму $\langle\ ,\ \rangle$, называется $(1,1)$-*унитарной группой* и обозначается через $\mathrm{U}(1,\,1)$.

Группа линейных преобразований плоскости $\mathbb{C}^2$, сохраняющих симплектическую структуру $[\ ,\ ]$, называется *специальной* (*или унимодулярной*) *линейной группой второго порядка* и обозначается через $\mathrm{SL}(2,\,\mathbb{C})$.

Группа всех вещественных линейных преобразований плоскости $\mathbb{C}^2$ (т. е. группа, элементы которой — комплексификации линейных преобразований в $\mathbb{R}^2$) называется *вещественной линейной группой второго порядка* и обозначается через $\mathrm{GL}(2,\,\mathbb{R})$.

Таким образом мы определили в группе $\mathrm{GL}(2,\,\mathbb{C})$ всех линейных преобразований в $\mathbb{C}^2$ три подгруппы: $(1,1)$-унитарную $\mathrm{U}(1,\,1)$, унимодулярную $\mathrm{SL}(2,\,\mathbb{C})$ и вещественную $\mathrm{GL}(2,\,\mathbb{R})$.

При этом эрмитова форма $\langle\ ,\ \rangle$, определяющая унитарную группу, симплектическая структура $[\,,]$, определяющая унимодулярную группу, и комплексное сопряжение, определяющее вещественную группу, связаны соотношением $\langle a,\, b\rangle = \frac{i}{2}[a,\, \overline{b}]$.

Теорема. *Пересечение любых двух из этих трех подгрупп совпадает с пересечением всех трех* (рис. 32).

Это пересечение называется *специальной* $(1,\,1)$-*унитарной группой*[1] и обозначается через $\mathrm{SU}(1,\,1)$. (Оно также называется вещественной унимодулярной группой и обозначается через $\mathrm{SL}(2,\,\mathbb{R})$. Оно называется также вещественной симплектической группой второго порядка и обозначается через $\mathrm{Sp}(1,\,\mathbb{R})$.)

Доказательство.

Если преобразование A вещественно и унимодулярно, то $[A\xi,\, A\eta] = [\xi, \eta]$ и $A\overline{\xi}=\overline{A\xi}$. Поэтому $\langle A\xi, A\eta\rangle=\frac{i}{2}[A\xi, \overline{A\eta}]=\frac{i}{2}[A\xi, A\overline{\eta}]=\frac{i}{2}[\xi, \eta]=\langle\xi, \eta\rangle$.

[1] Подчеркнем, что речь идет о группе операторов, а не матриц. Матрицы этих операторов принадлежат группе матриц $\mathrm{SU}(1,\,1)$ при специальном выборе базиса, указанном выше.

Если A вещественно и $(1, 1)$-унитарно, то $\langle A\xi, A\eta\rangle = \langle \xi, \eta\rangle$ и $A\overline{\xi} = \overline{A\xi}$. Поэтому $[A\xi, A\eta] = -2i\langle A\xi, \overline{A\eta}\rangle = -2i\langle A\xi, A\overline{\eta}\rangle = -2i\langle \xi, \overline{\eta}\rangle = [\xi, \eta]$.

Если A $(1,1)$-унитарно и унимодулярно, то $[A\xi, A\eta] = [\xi, \eta]$ и $[A\xi, \overline{A\eta}] = [\xi, \overline{\eta}]$. Поэтому $[A\xi, \overline{A\eta}] = [A\xi, A\overline{\eta}]$ для всех ξ и η. Следовательно, $[\xi, \overline{A\eta} - A\overline{\eta}] = 0$ для всех ξ и значит $\overline{A\eta} = A\overline{\eta}$ для всех η, т. е. A вещественно. ■

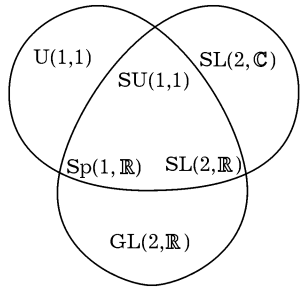


Рис. 32

Следствие. *Если матрица оператора в вещественном базисе* (e_1, e_2) *вещественная унимодулярная, то матрица этого оператора в комплексно сопряженном базисе* $(f_1 = e_1 + ie_2,\ f_2 = e_1 - ie_2)$ *специальная* $(1, 1)$*-унитарная, и обратно.*

Доказательство.

Эрмитов скалярный квадрат вектора $z_1 f_1 + z_2 f_2$ выражается через координаты (z_1, z_2) в базисе (f_1, f_2) по формуле $\langle z, z\rangle = |z_1|^2 - |z_2|^2$ (см. лемму). Поэтому (матрица A в базисе (e_1, e_2) вещественна и унимодулярна) $\iff$ $(A \in \mathrm{GL}(2, \mathbb{R}) \cap \mathrm{SL}(2, \mathbb{C}))$ $\iff$ $(A \in \mathrm{SU}(1, 1))$ $\iff$ $\iff$ (матрица A в базисе (f_1, f_2) $(1, 1)$-унитарна и унимодулярна). ■

Д. Геометрическое отступление: SU(1, 1) и геометрия Лобачевского.

Группы матриц $\mathrm{SL}(2, \mathbb{R})$ и $\mathrm{SU}(1, 1)$ следующим образом связаны с геометрией Лобачевского (рис. 33).

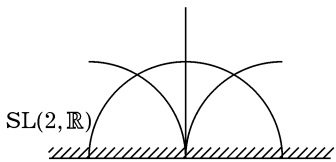


Рис. 33

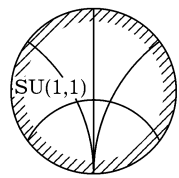


Рис. 34

Вещественная унимодулярная матрица второго порядка определяет дробно-линейное преобразование $z \longmapsto \dfrac{az+b}{cz+d}$, переводящее в себя верхнюю полуплоскость. Это преобразование является движением

плоскости Лобачевского, представленной в виде верхней полуплоскости. Все движения плоскости Лобачевского получаются таким способом. Группа движений плоскости Лобачевского изоморфна $\mathrm{SL}(2, \mathbb{R})/\pm E$.

Унимодулярная $(1,1)$-унитарная матрица второго порядка определяет дробно-линейное преобразование, переводящее в себя единичный круг. Действительно, конус $|z_1|^2 < |z_2|^2$ при $(1,1)$-унитарном унимодулярном преобразовании переходит в себя. При естественном отображении

$$\mathbb{C}^2 \setminus O \to \mathbb{C}\mathrm{P}^1, \qquad (z_1,\, z_2) \longmapsto w = \frac{z_1}{z_2}$$

этот конус переходит в единичный круг $|w| < 1$, а линейные преобразования в $\mathbb{C}^2$ переходят в дробно-линейные преобразования $\mathbb{C}\mathrm{P}^1$ (рис. 34).

Получающиеся из матриц из $\mathrm{SU}(1,\,1)$ дробно-линейные преобразования единичного круга в себя являются движениями плоскости Лобачевского, представленной как внутренность единичного круга. Все движения плоскости Лобачевского получаются таким способом. Группа движений плоскости Лобачевского изоморфна $\mathrm{SU}(1,\,1)/\pm E$.

Группы матриц $\mathrm{SL}(2,\,\mathbb{R})$ и $\mathrm{SU}(1,\,1)$ изоморфны: они получаются из одной и той же группы операторов. Матрицы этих операторов в вещественном базисе $(e_1,\, e_2)$ принадлежат $\mathrm{SL}(2,\,\mathbb{R})$, а в комплексно сопряженном базисе $(f_1,\, f_2)$ принадлежат $\mathrm{SU}(1,\,1)$. Переход от $\mathrm{SL}(2,\,\mathbb{R})$ к $\mathrm{SU}(1,\,1)$ соответствует переходу от вещественного базиса к комплексно сопряженному и от модели плоскости Лобачевского на верхней полуплоскости к модели в единичном круге.

Задача. Докажите, что группа $\mathrm{SL}(2,\,\mathbb{R})$ гомеоморфна полноторию $S^1 \times D^2$ (внутренности баранки).

Е. Свойства оператора вещественной монодромии.

Вернемся к оператору монодромии уравнения Шредингера (1). Кроме пространств $\mathbb{R}^2$ и $\mathbb{R}^2_0$ решений уравнения (1) и уравнения свободной частицы (2), рассмотрим еще *фазовую плоскость* $\mathbb{R}^2_\Phi$. Точки фазовой плоскости — это пары вещественных чисел $(\Psi,\, \Psi_x)$.

Зафиксируем $x \in \mathbb{R}$. Рассмотрим линейный оператор, сопоставляющий каждому (вещественному) решению уравнения (1) его начальное условие в точке x

$$B^x\colon \mathbb{R}^2 \to \mathbb{R}^2_\Phi, \qquad \Psi \longmapsto (\Psi(x),\, \Psi_x(x)).$$

Этот оператор — изоморфизм. Изоморфизм $g^{x_2}_{x_1} = B^{x_2}(B^{x_1})^{-1}$ называется *фазовым преобразованием* от x_1 к x_2.