Р .Курант

КУРС ДИФФЕРЕНЦИАЛЬНОГО И ИНТЕГРАЛЬНОГО ИСЧИСЛЕНИЯ

Том 2

Книга представляет собой мастерски написанный крупным математиком курс математического анализа, адресуемый автором «будущим учителям и научным работникам в области математики, физики и других естественных наук, а также инженерам». Первый том был впервые издан на русском языке в 1931 г. Последнее, 4-е издание первого тома, переработанное и значительно дополненное, вышло в конце 1967 г.

Второй том посвящен главным образом дифференциальному и интегральному исчислению функций многих переменных. По сравнению с первым русским изданием, вышедшим в 1931 г., настоящий перевод содержит многочисленные добавления автора , появившиеся в последних изданиях на немецком и английском языках.

Книга может служить полезным учебным пособием для студентов и преподавателей университетов, педагогических институтов и втузов с повышенным курсом математики.

Содержание

Предметный указатель

ПРЕДИСЛОВИЕ КО ВТОРОМУ РУССКОМУ ИЗДАНИЮ

Второй том немецкого оригинала этого курса вышел первым изданием в 1928 г. С него был сделан русский перевод, опубликованный в 1931 г. Дополнения, внесенные автором во второе немецкое издание (1931 г.), не успели включить в русское издание.

В 1936 г. был напечатан при участии автора английский перевод второго тома с многочисленными изменениями и дополнениями, а в 1955 г. вышло третье исправленное и дополненное немецкое издание. Что интерес к этой книге не ослабевал, доказывают многочисленные последующие допечатки (уже без изменений) обоих вариантов второго тома (английского до 1945 г. и немецкого до 1963 г.).

В настоящем русском издании соединен весь материал, содержащийся в английском варианте второго тома и в его третьем немецком издании. Как и в первом томе наибольшие по объему дополнения взяты из английского издания — это главы VII и VIII, посвященные элементам вариационного исчисления и теории функций комплексной переменной, а затем многочисленные задачи и упражнения ко всему содержанию книги, а также ответы и указания к ним.

В работе над переводом мы руководствовались главным образом интересами самого широкого круга советских читателей и прежде всего нуждами наших студентов. Для того чтобы возможно лучше приспособить книгу к их потребностям, мы позволили себе сделать кое-какие изменения и перестановки. Кроме того, книга снабжена добавлениями, вставками, пояснительными примечаниями. Для облегчения чтения все это помещено в тексте в квадратных скобках. Замеченные недосмотры исправлены без специальных оговорок.

Нет сомнения, что и второй том будет полезным пособием для широкого круга преподавателей математики, а также для аспирантов и научных работников в различных областях физики и техники; его с интересом и пользой будут штудировать студенты инженерно-физических вузов.

С удовольствием благодарю И. С. Аршона и А. М. Олевского за весьма существенные замечания и советы, а также А. Ф. Лапко за внимательное редактирование книги.

К несчастью, под этим предисловием уже не может подписаться Ю. Л. Рабинович. 18 марта 1968 г. Юлий Лазаревич скончался.

Москва,
6 марта 1969 г.

З. Г. Либин

ИЗ ПРЕДИСЛОВИЯ К ПЕРВОМУ НЕМЕЦКОМУ ИЗДАНИЮ

Настоящий заключительный том моих лекций по дифференциальному и интегральному исчислению посвящен главным образом учению о функциях многих переменных. Я старался следовать тому же дидактическому принципу, что и в первом томе: мотивировать построение понятий и методов их естественным происхождением из наглядных истоков и везде по возможности облегчать доступ к приложениям. Эти стремления вполне совместимы с требованиями строгости.

Для начинающих математиков я хочу подчеркнуть, что было бы вредным педантизмом требовать штудирования такой книги страницу за страницей в предложенном порядке. Тот, кто хочет быстро усвоить самые необходимые сведения, может сначала прочитать вторую, затем четвертую главу и лишь потом заполнить пробелы чтением третьей главы и дополнений к прочитанным главам. А некоторым читателям вообще нет необходимости предварительно систематически изучать первую главу.

Геттинген, ноябрь 1928 г. — *Р. Курант*

ИЗ ПРЕДИСЛОВИЯ КО ВТОРОМУ НЕМЕЦКОМУ ИЗДАНИЮ

Во втором издании этого тома, помимо улучшений и исправлений помещено некоторое число значительных по объему добавлений. Укажу, в частности, на расширение главы о дифференциальных уравнениях и на включение нового раздела о гамма-функции.

Геттинген, сентябрь 1931 г. — *Р. Курант*

ИЗ ПРЕДИСЛОВИЯ К АНГЛИЙСКОМУ ИЗДАНИЮ

... Английское издание во многом отличается от немецкого и содержит много дополнительного материала. В частности, весьма расширена глава о дифференциальных уравнениях; добавлены главы о вариационном исчислении и о функциях комплексной переменной, а также приложение о действительных числах [1]).

... Я должен выразить благодарность друзьям и коллегам, которые помогли мне в подготовке рукописи к печати, чтении корректур и составлении упражнений, в первую очередь д-ру Фриц Джону, мисс Маргарет Кеннеди, а также д-ру Шенбергу.

Нью-Рошель, Нью-Йорк,
март 1936 г.

Р. Курант

ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ НЕМЕЦКОМУ ИЗДАНИЮ

Третье издание отличается от второго главным образом большим числом дополнений, взятых из вышедшего тем временем в свет английского издания.

Нью-Рошель,
январь 1955 г.

Р. Курант

---

[1]) В русском издании это приложение напечатано в конце первого тома *(Прим. перев.)*

ГЛАВА I

КРАТКИЙ ОБЗОР ОСНОВНЫХ ПОНЯТИЙ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ И ВЕКТОРНОГО ИСЧИСЛЕНИЯ

Математические факты, составляющие главное содержание этого второго тома, часто находят наглядную иллюстрацию и применение с помощью простых основных понятий аналитической геометрии и векторного исчисления. Поэтому, хотя мы и вправе предполагать, что читатель уже знаком с этими дисциплинами, представляется целесообразным дать сводку их элементов в краткой вводной главе. Однако необязательно изучить эту главу перед чтением остальной части книги; читателю рекомендуется обращаться к изложенному здесь материалу всякий раз, когда в этом встретится надобность при изучении дальнейших разделов.

§ 1. Прямоугольные координаты и векторы

1. Системы координат. Для указания положения точки на плоскости или в пространстве, как известно, обычно пользуются прямоугольной системой координат. На плоскости выбирают две взаимно перпендикулярные прямые, ось x и ось y, в пространстве — три взаимно перпендикулярные прямые, оси x, y и z. Берут общую единицу длины для всех осей и всякой точке плоскости обычным способом приводят в соответствие координаты x и y, а каждой точке пространства — координаты x, y, z (рис. 1). Обратно, каждой системе значений (x, y) или (x, y, z) соответствует одна и только одна точка плоскости или пространства. Всякая точка вполне определяется своими координатами. Запись $P(x, y, z)$ и $Q(x, y)$ будет обозначать: точка P пространства имеет координаты x, y, z и точка Q на плоскости имеет координаты x, y.

Расстояние d между двумя точками плоскости с координатами (x_1, y_1) и (x_2, y_2) выражается, на основании теоремы Пифагора, формулой

$$d=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2},$$

а *расстояние между двумя точками* (x_1, y_1, z_1) и (x_2, y_2, z_2) *в пространстве* — формулой

$$d=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}.$$

В частности, расстояние точки $P(x, y, z)$ от начала координат есть

$$r=\sqrt{x^2+y^2+z^2}.$$

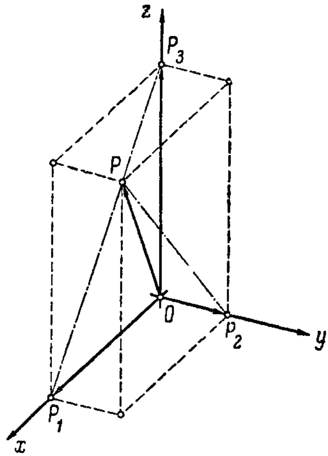


Рис. 1.

При построении прямоугольной системы координат следует уделить особое внимание вопросу об *ориентировке* или *ориентации* этой системы.

В каждой системе координат (x, y) на плоскости положительным считается то направление вращения, при котором поворот положительной оси

x на 90° вокруг начала координат совмещает ее кратчайшим путем с положительной осью y. Если это положительное вращение происходит против часовой стрелки, то система координат называется *правой* (рис. 2), если же по часовой стрелке, то — *левой системой* (рис. 3).

Не выходя из плоскости, невозможно перевести жестко связанную правую систему координат в левую систему никаким движением в этой плоскости.

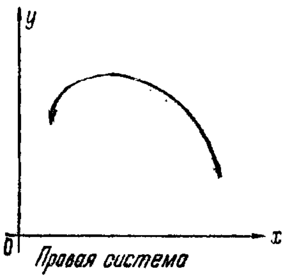


Рис. 2.

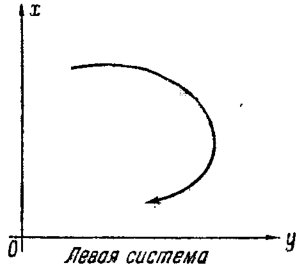


Рис. 3.

Совершенно такую же классификацию можно установить для систем координат в пространстве. Вообразим себе наблюдателя, стоящего на плоскости xy так, что голова его направлена в сторону положительной оси z; тогда возможны две пространственные системы, различающиеся по той ориентации, которую имеет система координат в плоскости xy с точки зрения упомянутого наблюдателя. Если эта система — правая, то и пространственная

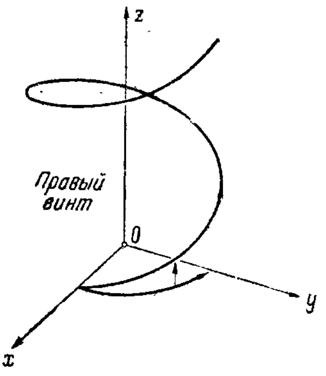


Рис. 4.

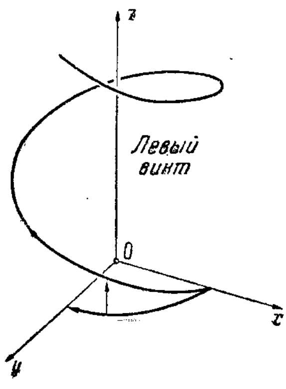


Рис. 5.

система x, y, z называется *правой* (рис. 4), если же система x, y представляется наблюдателю левой, то и система x, y, z называется *левой* (рис. 5).

Две различные ориентации пространственной системы можно наглядно иллюстрировать с помощью правого и левого винта. Если привести плоскость xy во вращательное движение в себе самой вокруг оси z в направлении, указанном ее ориентировкой, и сообщить ей одновременно поступательное движение в сторону положительной оси z, то результирующее движение будет для правой системы движением обычного правого винта (рис. 4), а для левой системы — движением левого винта (рис. 5). Никаким движением в пространстве жестко связанной системы осей x, y, z невозможно левую систему сделать правой.

В дальнейшем мы всегда будем пользоваться только правой системой координат.

Разумеется, и любой системе трех произвольных осей, проходящих через одну точку и не лежащих в одной плоскости, можно приписать определенную ориентацию таким же путем, как это было сделано для системы прямоугольных координат x, y, z.

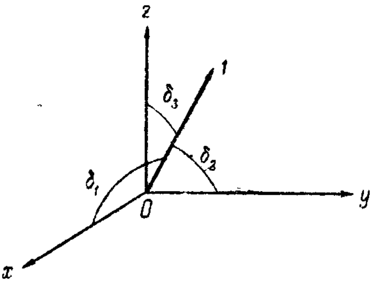


Рис. 6.

2. Направления и векторы. Всякую ориентированную, т. е. направленную в определенную сторону прямую l в пространстве или на плоскости мы будем называть осью. Любая такая ось указывает некоторое *направление*. Всякая другая ориентированная прямая, которую можно совместить с осью l параллельным переносом, сохраняя при этом ориентацию, считается представительницей того же направления. Для того чтобы задать какое-либо направление l относительно данной системы координат, проводят из начала координат ориентированную полупрямую того же направления и на ней берут точку на расстоянии 1 от начала; координаты (α, β, γ) этой точки равны косинусам углов δ_1, δ_2 и δ_3, образуемых ориентированной прямой l с положительными осями x, y, z[1]) и называются *направляющими косинусами* данного направления l (рис. 6). Согласно формуле расстояния точки от начала, эти направляющие косинусы удовлетворяют соотношению

$$\alpha^2+\beta^2+\gamma^2=1.$$

Если ограничиваться плоскостью xy, то всякое направление в этой плоскости определяется углами δ_1 и δ_2 $\left(\delta_1+\delta_2=\frac{\pi}{2}\right)$, которые ориентированная полупрямая l, идущая из начала координат в указанном направлении, образует с положительными осями x и y, либо направляющими косинусами $\alpha=\cos\delta_1$ и $\beta=\cos\delta_2$, которые теперь удовлетворяют соотношению (рис. 7):

$$\alpha^2+\beta^2=1.$$

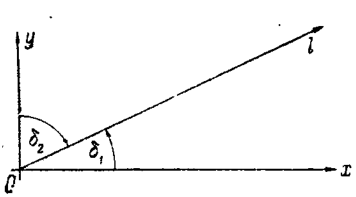


Рис. 7.

Отрезок данной длины и данного направления называется *вектором*; точнее, *связанным* вектором, если его начальная точка фиксирована в пространстве, и *свободным вектором*, если положение начальной точки безразлично. На последующих страницах, а также на протяжении почти всей книги, мы будем часто опускать прилагательное «свободный» и (если не сделана специальная оговорка) будем всегда считать векторы *свободными*. Векторы мы будем обозначать полужирным латинским шрифтом: $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{x}$, $\boldsymbol{A}$, а на письме — черточкой над буквой, обозначающей вектор. Вектор с начальной точкой A и конечной точкой B будем обозначать символом $\overline{AB}$.

---

[1]) Угол, образуемый одной ориентированной прямой в пространстве с другой, всегда можно брать между 0 и π, так как в дальнейшем придется иметь дело только с косинусами таких углов.

Два (свободных) вектора называются *равными*, если один из них можно совместить с другим путем параллельного перенесения. Длину вектора $\boldsymbol{a}$ называют также его *модулем* и обозначают символом $|\boldsymbol{a}|$ или той же буквой a обычного шрифта.

Если из начальной и конечной точки вектора $\boldsymbol{a}$ опустить перпендикуляры на ось l, то получим на ней соответствующий вектору $\boldsymbol{a}$ направленный отрезок. Если направление этого отрезка совпадает с ориентировкой оси l, то его длина называется *проекцией вектора* $\boldsymbol{a}$ *на ось* l; если же оба направления противоположны, то *проекцией вектора* $\boldsymbol{a}$ *на ось* l называется *длина* этого отрезка, взятая *со знаком минус*. Проекцию вектора $\boldsymbol{a}$ на ось l обозначают символом a_l. Пусть δ — угол между вектором $\boldsymbol{a}$ и осью l (рис. 8). [На рис. 8 вектор $\boldsymbol{a}$ перенесен так, что его начало лежит в одной из точек оси проекций. В пространственном случае плоскость чертежа есть плоскость, проходящая через ось проекций и новое положение вектора $\boldsymbol{a}$]. Тогда в обоих случаях

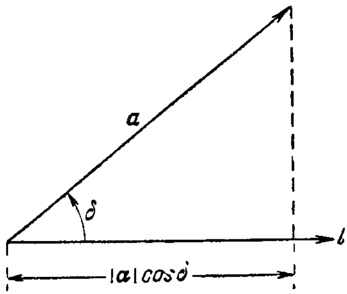


Рис. 8.

$$a_l = |\boldsymbol{a}| \cos\delta.$$

Проекции вектора $\boldsymbol{a}$ на три оси какой-либо системы координат обозначают через a_1, a_2, a_3, а углы этого вектора с осями x, y, z через $\delta_1, \delta_2, \delta_3$. Согласно последней формуле,

$$a_1 = |\boldsymbol{a}| \cos\delta_1, \qquad a_2 = |\boldsymbol{a}| \cos\delta_2, \qquad a_3 = |\boldsymbol{a}| \cos\delta_3. \tag{*}$$

Возвышая эти равенства в квадрат и складывая, получим

$$a_1^2 + a_2^2 + a_3^2 = |\boldsymbol{a}|^2 (\cos^2\delta_1 + \cos^2\delta_2 + \cos^2\delta_3) = a^2,$$

откуда

$$a = |\boldsymbol{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}. \text{ (На плоскости } a = \sqrt{a_1^2 + a_2^2}.)$$

Проекции a_1, a_2, a_3 вектора $\boldsymbol{a}$ на оси x, y, z называются *координатами* этого вектора. Вектор вполне определяется своими координатами. Поэтому равенство двух векторов

$$\boldsymbol{a} = \boldsymbol{b}$$

равносильно системе трех численных равенств

$$a_1 = b_1, \qquad a_2 = b_2, \qquad a_3 = b_3.$$

Мы будем пользоваться записью $\boldsymbol{a} = \{a_1, a_2, a_3\}$ в следующем смысле: $\boldsymbol{a}$ есть вектор с координатами a_1, a_2, a_3.

На плоскости xy вектор $\boldsymbol{a} = \{a_1, a_2\}$, причем для указания направления достаточно одного угла δ_1 (угол поворота от положительной полуоси x до вектора $\boldsymbol{a}$), ибо $\delta_2 = \frac{\pi}{2} - \delta_1$. Вектор $\boldsymbol{a}$ на плоскости вполне определяется также своим модулем a и углом $\delta_1 = \theta$ (азимутом или полярным углом вектора), если принять начало 0 за полюс, а положительную полуось x — за полярную ось. Формулы перехода от полярных координат вектора к его прямоугольным координатам: $a_1 = a\cos\theta$, $a_2 = a\sin\theta$.

Вектор, длина которого равна 1, называется единичным вектором. Его проекция на любую ось равна косинусу угла между этим вектором и осью. Единичный вектор, имеющий направление вектора $\boldsymbol{a}$, мы часто будем обозначать $\boldsymbol{a}^\circ$. Ясно, что $\boldsymbol{a}^\circ = \{\cos\delta_1, \cos\delta_2, \cos\delta_3\}$, где $\delta_1, \delta_2, \delta_3$ — углы, составляемые вектором $\boldsymbol{a}^\circ$, а значит и $\boldsymbol{a}$, с осями координат.

Пользование векторными обозначениями естественно и выгодно по двум различным причинам. Во-первых, многочисленные геометрические понятия и еще большее количество физических понятий, как, например, сила, скорость, ускорение, непосредственно возникают как векторы, независимо от выбранной системы координат. Во-вторых, для вычислений с векторами можно установить простые правила, аналогичные правилам действий над числами, и при помощи этих векторных операций многие рассуждения и выкладки значительно упрощаются и могут проводиться в форме, не зависящей от случайно выбранной системы координат.

3. Сложение векторов. Начнем с определения *суммы двух векторов* $\boldsymbol{a}$ *и* $\boldsymbol{b}$. Для этой цели переместим вектор $\boldsymbol{b}$ параллельно самому себе так, чтобы его начальная точка совпала с конечной точкой вектора $\boldsymbol{a}$. Тогда начальная точка вектора $\boldsymbol{a}$ и конечная точка вектора $\boldsymbol{b}$ определяют новый

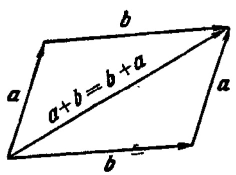


Рис. 9.

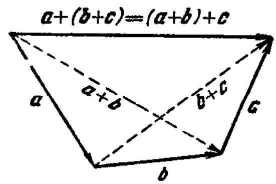


Рис. 10.

вектор $\boldsymbol{c}$, начальная точка которого совпадает с начальной точкой вектора $\boldsymbol{a}$, а конечная точка — с конечной точкой $\boldsymbol{b}$. Вектор $\boldsymbol{c}$ называют *суммой векторов* $\boldsymbol{a}$ и $\boldsymbol{b}$ и записывают это так: $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{c}$.

Рассмотрение рис. 9 и 10 показывает, что сумма векторов подчиняется *переместительному закону*:

$$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a},$$

и *сочетательному закону*:

$$\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})=(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}.$$

Прямым следствием определения сложения векторов является следующее векторное тождество:

$$\overline{A_1A_2}+\overline{A_2A_3}+\ldots+\overline{A_{n-1}A_n}=\overline{A_1A_n},$$

справедливое при любом расположении точек $A_1, A_2, \ldots, A_n$.

Из определения сложения векторов непосредственно вытекает теорема о проекциях: *проекция суммы двух или нескольких векторов на ось l равна сумме проекций слагаемых векторов на ту же ось:*

$$(\boldsymbol{a}+\boldsymbol{b})_l=a_l+b_l.$$

В частности, проекции суммы $\boldsymbol{a}+\boldsymbol{b}$ на координатные оси равны соответственно

$$a_1+b_1, \quad a_2+b_2, \quad a_3+b_3,$$

так что

$$\boldsymbol{a}+\boldsymbol{b}=\{a_1+b_1,\ a_2+b_2,\ a_3+b_3\}.$$

Это значит, что для нахождения суммы двух векторов имеется следующее простое правило: *любая координата вектора, являющегося суммой нескольких векторов, равна сумме соответствующих координат слагаемых векторов.*

Всякая точка $P(x, y, z)$ может быть определена своим *радиус-вектором* ($r = \overline{OP}$), т. е. вектором, идущим от начала координат к определяемой точке; координатами радиус-вектора являются как раз координаты точки P.

Возьмем три единичных вектора, каждый из которых имеет положительное направление одной из осей координат: e_1 по направлению оси x, e_2 по направлению оси y и e_3 по направлению оси z. Единичные векторы e_1, e_2, e_3 называются *ортами* координатных осей. Если вектор a имеет координаты a_1, a_2, a_3, то вектор

$$a = a_1 e_1 + a_2 e_2 + a_3 e_3.$$

Векторы $a_1 = a_1 e_1$, $a_2 = a_2 e_2$, $a_3 = a_3 e_3$ называются *составляющими* или *компонентами* вектора a по осям координат (см. ниже п° 5).

4. Преобразование координат. С помощью формулированной выше теоремы о проекциях легко получить *формулы преобразования координат*, выражающие координаты x', y', z' данной точки P относительно системы осей Ox', Oy', Oz' через координаты x, y, z этой точки относительно другой системы осей Ox, Oy, Oz, предполагая, что обе системы координат — прямоугольные и имеют общее начало O. (Поворот системы координат вокруг начала O является частным случаем такого преобразования.)

Три новые оси образуют с тремя старыми осями углы, *косинусы* которых даны в следующей таблице, не нуждающейся в пояснении:

	x	y	z
x'	α_1	β_1	γ_1
y'	α_2	β_2	γ_2
z'	α_3	β_3	γ_3

Из нее видно, например, что косинус угла между осью x' и осью z равен γ_1, и т. д.

Из точки P опустим перпендикуляры на оси Ox, Oy, Oz; пусть основаниями этих перпендикуляров будут точки P_1, P_2, P_3 (см. рис. 1 на стр. 15). Векторы $\overline{OP_1}$, $\overline{OP_2}$, $\overline{OP_3}$ являются составляющими вектора $\overline{OP}$ по направлениям осей x, y, z. Поэтому

$$\overline{OP} = \overline{OP_1} + \overline{OP_2} + \overline{OP_3}.$$

Согласно теореме о проекциях, абсцисса x' точки P, как проекция вектора $\overline{OP}$ на ось x', равна сумме проекций на ось x' векторов $\overline{OP_1}$, $\overline{OP_2}$, $\overline{OP_3}$; так как направляющие косинусы оси x' в системе x, y, z суть α_1, β_1, γ_1, то проекции этих векторов на ось x' равны соответственно $\alpha_1 x$, $\beta_1 y$, $\gamma_1 z$, а стало быть,

$$x' = \alpha_1 x + \beta_1 y + \gamma_1 z.$$

Проделав аналогичные рассуждения для y' и z' (в системе x, y, z направляющие косинусы оси y' равны α_2, β_2, γ_2, а направляющие косинусы оси z' — α_3, β_3, γ_3), получим следующие *формулы преобразования координат:*

$$\begin{aligned} x' &= \alpha_1 x + \beta_1 y + \gamma_1 z, \\ y' &= \alpha_2 x + \beta_2 y + \gamma_2 z, \\ z' &= \alpha_3 x + \beta_3 y + \gamma_3 z \end{aligned}$$

и обратную систему формул:

$$\begin{aligned} x &= \alpha_1 x' + \alpha_2 y' + \alpha_3 z', \\ y &= \beta_1 x' + \beta_2 y' + \beta_3 z', \\ z &= \gamma_1 x' + \gamma_2 y' + \gamma_3 z'. \end{aligned}$$

Так как координаты связанного вектора $\boldsymbol{v}$ выражаются формулами

$$v_1 = x_2 - x_1, \qquad v_2 = y_2 - y_1, \qquad v_3 = z_2 - z_1,$$

где (x_1, y_1, z_1) — координаты начальной точки вектора $\boldsymbol{v}$, а (x_2, y_2, z_2) — координаты его конечной точки, то для *координат вектора* получаются *те же формулы преобразования*, что и для координат точки:

$$\begin{aligned} v_1' &= \alpha_1 v_1 + \beta_1 v_2 + \gamma_1 v_3, \\ v_2' &= \alpha_2 v_1 + \beta_2 v_2 + \gamma_2 v_3, \\ v_3' &= \alpha_3 v_1 + \beta_3 v_2 + \gamma_3 v_3. \end{aligned}$$

5. Умножение вектора на число. Для суммы *равных* векторов естественно ввести следующее обозначение: $\boldsymbol{v} + \boldsymbol{v} = 2\boldsymbol{v}$, $\boldsymbol{v} + \boldsymbol{v} + \boldsymbol{v} = 3\boldsymbol{v}$ и т. д. В соответствии с определением сложения векторов, в порядке обобщения этого обозначения, вводим следующее определение умножения вектора на число. *Произведением $c\boldsymbol{v}$ или $\boldsymbol{v}c$ вектора $\boldsymbol{v}$ на число c* называется новый вектор, длина которого равна $|c| \cdot |\boldsymbol{v}|$ и направление которого при $c > 0$ совпадает с направлением $\boldsymbol{v}$, а при $c < 0$ противоположно направлению вектора $\boldsymbol{v}$; если $c = 0$, то произведение $c\boldsymbol{v} = \boldsymbol{v}c$ есть нуль-вектор с координатами $\{0, 0, 0\}$. Нуль-вектор обозначают просто символом 0 (нуль).

Если вектор $\boldsymbol{v} = \{v_1, v_2, v_3\}$, то вектор-произведение $c\boldsymbol{v} = \{cv_1, cv_2, cv_3\}$.

Умножение вектора на число подчиняется, по определению, *переместительному закону* $c\boldsymbol{v} = \boldsymbol{v}c$. Нетрудно доказать, что оно подчиняется и *распределительному* закону в следующих двух видах:

$$(\alpha + \beta)\,\boldsymbol{u} = \alpha\boldsymbol{u} + \beta\boldsymbol{u}, \qquad \alpha\,(\boldsymbol{u} + \boldsymbol{v}) = \alpha\boldsymbol{u} + \alpha\boldsymbol{v}.$$

Легко также показать, что $\alpha\,(\beta\boldsymbol{u}) = (\alpha\beta)\,\boldsymbol{u}$.

Произведение $(-1)\,\boldsymbol{v}$ есть вектор, *равнопротивоположный* вектору $\boldsymbol{v}$, т. е. вектор, имеющий тот же модуль, что и $\boldsymbol{v}$, и направленный в противоположную сторону. Его обозначают короче так:

$$(-1)\,\boldsymbol{v} = -\boldsymbol{v}.$$

Вектор $\boldsymbol{v}$ и единичный вектор $\boldsymbol{v}^\circ$ того же направления связаны равенствами $\boldsymbol{v} = |\boldsymbol{v}|\,\boldsymbol{v}^\circ = v\boldsymbol{v}^\circ$ и $\boldsymbol{v}^\circ = \dfrac{1}{v}\boldsymbol{v}$.

6. Скалярное произведение двух векторов. Действие *умножения двух векторов* можно ввести таким образом, что это «умножение» будет подчиняться законам, отчасти аналогичным законам умножения чисел. Существуют два различных вида умножения векторов. Сначала мы введем более простое и более важное для нас *скалярное произведение*.

Скалярным произведением $\boldsymbol{ab}$ векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ называется число, равное произведению их модулей и косинуса угла δ между их направлениями:

$$\boldsymbol{ab} = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cos\delta.$$

Стало быть, скалярное произведение двух векторов равно проекции одного из этих векторов на направление другого, помноженной на модуль этого другого вектора.

Из определения скалярного произведения сразу вытекает *переместительный или коммутативный закон:*

$$\boldsymbol{ab} = \boldsymbol{ba},$$

а с помощью теоремы о проекциях нетрудно вывести и *распределительный закон* скалярного умножения:

$$(\boldsymbol{a}+\boldsymbol{b})\,\boldsymbol{c}=\boldsymbol{a}\boldsymbol{c}+\boldsymbol{b}\boldsymbol{c}.$$

Вместе с тем между формальными свойствами скалярного умножения векторов и обыкновенного умножения чисел имеется существенное различие. Дело в том, что *скалярное произведение может обратиться в нуль и в том случае, когда ни один из сомножителей не равен нулю.*

Действительно, *скалярное произведение* $\boldsymbol{ab}=0$ *в том и только в том случае, если векторы* $\boldsymbol{a}$ *и* $\boldsymbol{b}$ *взаимно перпендикулярны либо по крайней мере один из них есть нуль-вектор.*

Для скалярного произведения *ортов* координатных осей имеем, очевидно:

$$\boldsymbol{e}_1\boldsymbol{e}_1=\boldsymbol{e}_2\boldsymbol{e}_2=\boldsymbol{e}_3\boldsymbol{e}_3=1, \qquad \boldsymbol{e}_1\boldsymbol{e}_2=\boldsymbol{e}_2\boldsymbol{e}_3=\boldsymbol{e}_1\boldsymbol{e}_3=0.$$

7. Выражение скалярного произведения через координаты перемножаемых векторов. Обозначим координаты перемножаемых векторов через a_1, a_2, a_3 и b_1, b_2, b_3, а их составляющие по направлениям осей координат через $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ и $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$. Тогда

$$\boldsymbol{a}=\boldsymbol{a}_1+\boldsymbol{a}_2+\boldsymbol{a}_3, \qquad \boldsymbol{b}=\boldsymbol{b}_1+\boldsymbol{b}_2+\boldsymbol{b}_3$$

и

$$\boldsymbol{ab}=(\boldsymbol{a}_1+\boldsymbol{a}_2+\boldsymbol{a}_3)(\boldsymbol{b}_1+\boldsymbol{b}_2+\boldsymbol{b}_3).$$

На основании распределительного закона можно выполнить умножение почленно; при этом произведения $\boldsymbol{a}_1\boldsymbol{b}_2$, $\boldsymbol{a}_1\boldsymbol{b}_3$, $\boldsymbol{a}_2\boldsymbol{b}_1$, $\boldsymbol{a}_2\boldsymbol{b}_3$, $\boldsymbol{a}_3\boldsymbol{b}_1$ и $\boldsymbol{a}_3\boldsymbol{b}_2$ равны нулю, так как их сомножители взаимно перпендикулярны, а стало быть,

$$\boldsymbol{ab}=\boldsymbol{a}_1\boldsymbol{b}_1+\boldsymbol{a}_2\boldsymbol{b}_2+\boldsymbol{a}_3\boldsymbol{b}_3.$$

Но в этом выражении $\boldsymbol{a}_1\boldsymbol{b}_1=(a_1\boldsymbol{e}_1)(b_1\boldsymbol{e}_1)=a_1b_1\boldsymbol{e}_1\boldsymbol{e}_1=a_1b_1$, $\boldsymbol{a}_2\boldsymbol{b}_2=(a_2\boldsymbol{e}_2)(b_2\boldsymbol{e}_2)=a_2b_2$ и аналогично $\boldsymbol{a}_3\boldsymbol{b}_3=a_3b_3$. Следовательно,

$$\boldsymbol{ab}=a_1b_1+a_2b_2+a_3b_3.$$

Это — важное и удобное правило для вычисления скалярного произведения двух векторов по их координатам. Но это равенство можно было бы принять и за исходный пункт в качестве определения скалярного произведения.

Если, в частности, подставить вместо $\boldsymbol{a}$ и $\boldsymbol{b}$ их единичные векторы $\boldsymbol{a}^\circ=\{\alpha_1, \alpha_2, \alpha_3\}$ и $\boldsymbol{b}^\circ=\{\beta_1, \beta_2, \beta_3\}$, то скалярное произведение $\boldsymbol{a}^\circ\boldsymbol{b}^\circ$ будет равно косинусу угла δ между векторами $\boldsymbol{a}$ и $\boldsymbol{b}$, и мы получим для косинуса угла между двумя векторами формулу

$$\cos\delta=\boldsymbol{a}^\circ\boldsymbol{b}^\circ=\alpha_1\beta_1+\alpha_2\beta_2+\alpha_3\beta_3.$$

Скалярное произведение вектора $\boldsymbol{a}$ на самого себя обозначают символом $\boldsymbol{a}^2$ и называют *скалярным квадратом* или просто *квадратом* вектора $\boldsymbol{a}$. Очевидно, $\boldsymbol{a}^2=aa\cos 0=a^2$, т. е. скалярный квадрат вектора равен квадрату его модуля.

Физическое значение скалярного произведения видно, например, из того факта, известного из элементарной механики, что сила $\boldsymbol{F}$, действующая на материальную точку на вектор-перемещении $\boldsymbol{s}$, совершает при этом работу $\boldsymbol{Fs}$.

8. Уравнение прямой на плоскости и уравнение плоскости в пространстве. Пусть дана прямая на плоскости xy или плоскость в пространстве xyz. Для того чтобы вывести их уравнения, восставим к прямой (или плоскости) перпендикуляр (нормаль), и на нем выберем определенное *«положительное» направление нормали* к прямой или плоскости; при этом безразлично, какое из двух возможных направлений принять за положительное (см. рис. 11). Единичный вектор, имеющий направление положительной нормали, обозначим через $\boldsymbol{n}^\circ$. Точки прямой (или плоскости) характеризуются тем, что идущий к ним из начала координат радиус-вектор имеет постоянную проекцию p на направление нормального вектора $\boldsymbol{n}^\circ$; другими словами, скалярное произведение радиус-вектора $\boldsymbol{r}$ на единичный нормальный вектор $\boldsymbol{n}^\circ$ равно постоянному числу p:

$$\boldsymbol{r}\boldsymbol{n}^\circ = p.$$

Это и есть уравнение прямой на плоскости (или плоскости в пространстве) в векторной записи. Геометрический смысл числа p следующий: абсолютная величина $|p|$ постоянной p есть расстояние прямой (или плоскости) от начала координат; $p=0$, если прямая (или плоскость) проходит через начало координат; если же прямая (или плоскость) не проходит через начало координат, то $p<0$, когда нормальный вектор $\boldsymbol{n}^\circ$ направлен от прямой (плоскости) в ту сторону от нее, где находится начало координат, и $p>0$, если вектор $\boldsymbol{n}^\circ$ направлен в сторону, противоположную от начала координат.

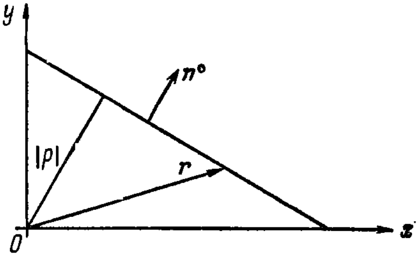


Рис. 11.

Пусть $\boldsymbol{n}^\circ = \{\alpha, \beta\}$ для прямой на плоскости и $\boldsymbol{n}^\circ = \{\alpha, \beta, \gamma\}$ для плоскости в пространстве, а x, y (x, y, z) — текущие координаты на прямой (на плоскости), так что $\boldsymbol{r} = \{x, y\}$ $(\boldsymbol{r} = \{x, y, z\})$. Тогда уравнение $\boldsymbol{r}\boldsymbol{n}^\circ = p$ запишется в координатном виде так:

$$\alpha x + \beta y - p = 0 \text{ (уравнение прямой на плоскости)}, \tag{1}$$

$$\alpha x + \beta y + \gamma z - p = 0 \text{ (уравнение плоскости в пространстве)}. \tag{2}$$

Обратно, если заданные числа α, β (α, β, γ) являются направляющими косинусами некоторого направления, то уравнение (1) определяет прямую на плоскости xy, а уравнение (2) — плоскость в пространстве, имеющие нормальный вектор $\boldsymbol{n}^\circ = \{\alpha, \beta\}$ $(\boldsymbol{n}^\circ = \{\alpha, \beta, \gamma\})$ и находящиеся на расстоянии $|p|$ от начала координат.

Выведенные здесь уравнения (1) и (2) называются *уравнениями в нормальном виде* или *нормальными уравнениями* прямой на плоскости и плоскости в пространстве. Левая часть нормального уравнения имеет определенный геометрический смысл и для точек P, не лежащих на прямой (на плоскости). Действительно, из того факта, что выражение $\alpha x + \beta y$ или $\alpha x + \beta y + \gamma z$ равно проекции радиус-вектора $\boldsymbol{r} = \overline{OP}$ на положительное направление нормали, легко вывести, что выражение $\alpha x + \beta y - p$ $(\alpha x + \beta y + \gamma z - p)$ есть расстояние точки $P(x, y, z)$ от прямой (плоскости), взятое со знаком плюс, если точка P лежит на той стороне от прямой (плоскости), куда указывает нормальный вектор $\boldsymbol{n}^\circ$, и со знаком минус, если точка P лежит на другой стороне.

Из нормального уравнения, умножая его на произвольный не равный нулю множитель, можно получить другие виды уравнения прямой или плоскости. Обратно, любое линейное уравнение

$$Ax + By + D = 0 \qquad (\text{или } Ax + By + Cz + D = 0)$$

представляет прямую линию на плоскости xy (или плоскость) при условии, что не все коэффициенты A, B (или A, B, C) равны нулю [1]). Действительно,

[1]) Если $A = B = 0$ (или $A = B = C = 0$), то и D должно равняться нулю и уравнению удовлетворяют все точки плоскости (или пространства).

разделим, например, второе из этих уравнений на $\sqrt{A^2+B^2+C^2}$ и положим

$$\alpha=\frac{A}{\sqrt{A^2+B^2+C^2}},\quad \beta=\frac{B}{\sqrt{A^2+B^2+C^2}},\quad \gamma=\frac{C}{\sqrt{A^2+B^2+C^2}};$$
$$p=-\frac{D}{\sqrt{A^2+B^2+C^2}}.$$

Тогда получится уравнение, которое, согласно изложенному выше, представляет плоскость, находящуюся на расстоянии $|p|$ от начала координат и имеющую единичный нормальный вектор $n^\circ=\{\alpha, \beta, \gamma\}$. Аналогичным путем можно привести к нормальному виду и первое уравнение.

Из того, что коэффициенты A, B (A, B, C) пропорциональны координатам единичного нормального вектора n°, вытекает, что и вектор $N=\{A, B\}$ $(N=\{A, B, C\})$ является нормальным вектором прямой на плоскости (плоскости в пространстве), но уже не единичным.

В дальнейшем нам понадобится *формула для угла δ между двумя плоскостями*, заданными нормальными уравнениями

$$\alpha x+\beta y+\gamma z-p=0,$$
$$\alpha' x+\beta' y+\gamma' z-p'=0.$$

Так как угол δ между плоскостями равен углу между их единичными нормальными векторами, то $\cos\delta$ равен скалярному произведению этих единичных векторов

$$\cos\delta=n^\circ n^{\circ\prime}=\alpha\alpha'+\beta\beta'+\gamma\gamma'.$$

Аналогично для *угла δ между прямыми на плоскости xy*

$$\alpha x+\beta y-p=0 \quad\text{и}\quad \alpha' x+\beta' y-p'=0$$

имеем

$$\cos\delta=\alpha\alpha'+\beta\beta'.$$

9. Уравнение прямой в пространстве. Прямую в пространстве можно задать с помощью любых двух различных плоскостей, проходящих через эту прямую. Таким образом, для пространственной прямой получаются аналитически два линейных уравнения

$$A_1x+B_1y+C_1z+D_1=0,$$
$$A_2x+B_2y+C_2z+D_2=0,$$

которым удовлетворяют координаты x, y, z любой точки прямой. Так как через данную прямую проходит бесчисленное множество плоскостей, то этот способ задания пространственной прямой не является однозначным.

Для аналитического представления прямой в пространстве часто удобнее пользоваться *параметрическим заданием* с помощью параметра t. Рассмотрим три целые линейные функции от t:

$$\left.\begin{aligned} x&=a_1+b_1t,\\ y&=a_2+b_2t,\\ z&=a_3+b_3t,\end{aligned}\right\}\qquad(1)$$

где b_1, b_2, b_3 не равны одновременно нулю. Когда t пробегает числовую ось, точка (x, y, z) описывает прямую линию. Это видно из того, что, исключая t из двух пар уравнений, мы получаем два линейных уравнения для x, y, z.

Направляющие косинусы этой прямой можно найти следующим путем. Возьмем две точки прямой P_1 и P_2 с координатами

$$P_1:\quad x_1=a_1+b_1t_1,\quad y_1=a_2+b_2t_1,\quad z_1=a_3+b_3t_1;$$
$$P_2:\quad x_2=a_1+b_1t_2,\quad y_2=a_2+b_2t_2,\quad z_2=a_3+b_3t_2.$$

Проектируя вектор $\overline{P_1P_2}$ на оси координат, имеем

$$|\overline{P_1P_2}|\cos\delta_1 = x_2 - x_1 = b_1(t_2 - t_1),$$
$$|\overline{P_1P_2}|\cos\delta_2 = y_2 - y_1 = b_2(t_2 - t_1),$$
$$|\overline{P_1P_2}|\cos\delta_3 = z_2 - z_1 = b_3(t_2 - t_1),$$

где δ_1, δ_2 и δ_3 — углы, образуемые вектором $\overline{P_1P_2}$, а следовательно и прямой, с осями координат. Стало быть,

$$\cos\delta_1 = \rho b_1, \qquad \cos\delta_2 = \rho b_2, \qquad \cos\delta_3 = \rho b_3, \qquad \text{где } \rho = \frac{t_2 - t_1}{|\overline{P_1P_2}|},$$

т. е. направляющие косинусы прямой пропорциональны коэффициентам b_1, b_2, b_3 при t в ее параметрических уравнениях. Так как $\cos^2\delta_1 + \cos^2\delta_2 + \cos^2\delta_3 = 1$, то отсюда вытекает

$$\cos\delta_1 = \frac{b_1}{\sqrt{b_1^2 + b_2^2 + b_3^2}}, \qquad \cos\delta_2 = \frac{b_2}{\sqrt{b_1^2 + b_2^2 + b_3^2}}, \qquad \cos\delta_3 = \frac{b_3}{\sqrt{b_1^2 + b_2^2 + b_3^2}},$$

причем двузначность квадратного корня выражает тот факт, что на прямой можно выбрать любое из двух возможных направлений. Любой вектор, имеющий направление прямой, называется *направляющим вектором прямой*. Так, направляющими векторами прямой являются вектор $\overline{P_1P_2}$ и вектор $S = \{b_1, b_2, b_3\}$. Вектор $s = \{\cos\delta_1, \cos\delta_2, \cos\delta_3\}$ является *единичным* направляющим вектором прямой. С помощью направляющих косинусов прямой $\alpha = \cos\delta_1$, $\beta = \cos\delta_2$, $\gamma = \cos\delta_3$ можно привести параметрические уравнения прямой к следующему виду:

$$\left.\begin{aligned} x &= x_0 + \alpha\tau, \\ y &= y_0 + \beta\tau, \\ z &= z_0 + \gamma\tau, \end{aligned}\right\} \tag{2}$$

где (x_0, y_0, z_0) — фиксированная точка на прямой. Действительно, пусть $P_0(x_0, y_0, z_0)$ — произвольная, но фиксированная точка прямой, а $P(x, y, z)$ — переменная ее точка. Тогда вектор

$$\overline{P_0P} = s\tau,$$

где $s = \{\alpha, \beta, \gamma\}$ — единичный направляющий вектор прямой, а τ — переменный параметр, равный расстоянию от точки P_0 до P, взятому со знаком плюс, если вектор $\overline{P_0P}$ направлен в ту же сторону, что и вектор s, и со знаком минус — в противном случае. Проектируя обе части последнего векторного уравнения на оси координат, получим

$$x - x_0 = \alpha\tau, \qquad y - y_0 = \beta\tau, \qquad z - z_0 = \gamma\tau,$$

т. е. искомые параметрические уравнения (2).

Связь между прежним параметром t и новым параметром τ можно получить, сравнивая, например, выражения для x из обеих систем уравнений (1) и (2):

$$a_1 + b_1 t = x_0 + \alpha\tau.$$

Из уравнений (2) можно получить представление координат переменной точки $P(x, y, z)$ нашей прямой, отнесенное к двум точкам этой прямой: $P_0(x_0, y_0, z_0)$ и $P_1(x_1, y_1, z_1)$. Пусть переменной точке P прямой соответствует значение τ параметра, а точке P_1 — значение τ_1. Тогда $x_1 = x_0 + \alpha\tau_1$, откуда $\alpha = \dfrac{x_1 - x_0}{\tau_1}$. Подставив это выражение для α в уравнение $x = x_0 + \alpha\tau$, получим $x=\left(1-\frac{\tau}{\tau_1}\right)x_0+\frac{\tau}{\tau_1}x_1=\lambda_0x_0+\lambda_1x_1$, где положено $1-\frac{\tau}{\tau_1}=\lambda_0$ и $\frac{\tau}{\tau_1}=\lambda_1$. Выполнив то же самое для y и z, получим следующую систему уравнений:

$$x=\lambda_0x_0+\lambda_1x_1,\qquad y=\lambda_0y_0+\lambda_1y_1,\qquad z=\lambda_0z_0+\lambda_1z_1.$$

Положение любой точки прямой характеризуется парой чисел λ_0 и λ_1, причем $\lambda_0+\lambda_1=1$. Нетрудно убедиться из равенств $1-\frac{\tau}{\tau_1}=\lambda_0$ и $\frac{\tau}{\tau_1}=\lambda_1$, учитывая сказанное выше о знаке параметра τ, что для внутренних точек отрезка P_0P_1 числа λ_0 и λ_1 положительны. [Для точек прямой, лежащей вне этого отрезка, знаки λ_0 и λ_1 различны, причем за точкой P_0 будет $\lambda_0>0$ и $\lambda_1<0$, а за точкой P_1 будет $\lambda_0<0$, $\lambda_1>0$.]

Задача. Составим уравнение плоскости, проходящей через точку $M_0(X_0, Y_0, Z_0)$ и перпендикулярной к прямой, заданной параметрическими уравнениями

$$x=a_1+b_1t,\qquad y=a_2+b_2t,\qquad z=a_3+b_3t.$$

За нормальный вектор плоскости можно принять направляющий вектор данной прямой, имеющий координаты b_1, b_2, b_3. Поэтому уравнение искомой плоскости можно записать в следующем виде:

$$b_1x+b_2y+b_3z+D=0.$$

Так как точка (X_0, Y_0, Z_0) лежит на этой плоскости, то имеем тождество

$$b_1X_0+b_2Y_0+b_3Z_0+D=0.$$

Выразив из этого тождества неизвестный член D и подставив в только что найденное уравнение, получим уравнение плоскости, проходящей через точку $M_0(X_0, Y_0, Z_0)$ перпендикулярно к прямой с направляющим вектором $\{b_1, b_2, b_3\}$:

$$b_1(x-X_0)+b_2(y-Y_0)+b_3(z-Z_0)=0.$$

Аналогично, уравнение прямой на плоскости xy, проходящей через точку $M_0(X_0, Y_0)$ перпендикулярно прямой с направляющим вектором $\{b_1, b_2\}$, есть

$$b_1(x-X_0)+b_2(y-Y_0)=0.$$

Упражнения

1. Показать, что числа $\alpha_1, \beta_1, \gamma_1$; $\alpha_2, \beta_2, \gamma_2$; $\alpha_3, \beta_3, \gamma_3$ (n° 4), определяющие преобразование одной прямоугольной системы координат в другую, удовлетворяют соотношениям:

$$\begin{aligned}
&\alpha_1\alpha_2+\beta_1\beta_2+\gamma_1\gamma_2=0, &\quad &\alpha_1^2+\beta_1^2+\gamma_1^2=1,\\
&\alpha_2\alpha_3+\beta_2\beta_3+\gamma_2\gamma_3=0, &\quad &\alpha_2^2+\beta_2^2+\gamma_2^2=1,\\
&\alpha_3\alpha_1+\beta_3\beta_1+\gamma_3\gamma_1=0, &\quad &\alpha_3^2+\beta_3^2+\gamma_3^2=1.
\end{aligned}$$

2. Даны два вектора $\boldsymbol{a}$ и $\boldsymbol{b}$ с общим началом O и конечными точками A и B. Доказать, что вектор, идущий из точки O к точке M, делящей отрезок AB в отношении $\frac{AM}{MB}=\lambda$, дается формулой

$$\overline{OM}=\frac{\boldsymbol{a}+\lambda\boldsymbol{b}}{1+\lambda}.$$

3. Центр массы двух точек A и B (в которых сосредоточены равные массы) находится в середине отрезка AB. Центр массы вершин треугольника ABC (в которых сосредоточены равные массы) может быть определен как точка M, делящая отрезок CD в отношении $CM : MD = 2 : 1$, где D — середина стороны AB. Показать, что это определение не зависит от того порядка, в котором взяты вершины треугольника.

4. Центр массы вершин тетраэдра $PQRS$ (в которых сосредоточены равные массы) может быть определен как точка N, делящая отрезок SM в отношении $SN : NM = 3 : 1$, где M — центр массы вершин P, Q, R. Показать, что это определение не зависит от того, в каком порядке брать вершины тетраэдра, и что оно находится в согласии с общим определением центра массы системы материальных точек (т. I, стр. 327, n° 9).

5. В тетраэдре $PQRS$ середины ребер PQ, RS, PR, QS, PS, QR помечены соответственно буквами A, A', B, B', C, C'. Доказать, что все прямолинейные отрезки AA', BB' и CC' проходят через центр массы вершин тетраэдра и делятся в нем пополам.

6. В n точках $P_1, P_2, \ldots, P_n$ в пространстве сосредоточены соответственно массы m_1, $m_2, \ldots, m_n$. Пусть точка G — их центр массы. Обозначим через $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$ векторы, имеющие общее начало G и концы в точках $P_1, P_2, \ldots, P_n$. Доказать, что $m_1\boldsymbol{p}_1 + m_2\boldsymbol{p}_2 + \ldots + m_n\boldsymbol{p}_n = 0$.

§ 2. Площадь треугольника. Векторное умножение. Объем тетраэдра

1. Площадь треугольника, построенного на векторах $\boldsymbol{a}$ и $\boldsymbol{b}$ в плоскости xy. Для удобства вычисления приложим векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ к началу координат. Тогда наш треугольник займет положение OP_1P_2 (рис. 12 или 13), где стороны $OP_1 = |\boldsymbol{a}| = a$, $OP_2 = |\boldsymbol{b}| = b$ и угол между ними $\delta = \theta_2 - \theta_1$. Если ввести полярные координаты (r, θ) с полюсом O и полярной осью Ox,

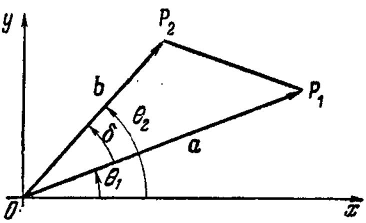


Рис. 12.

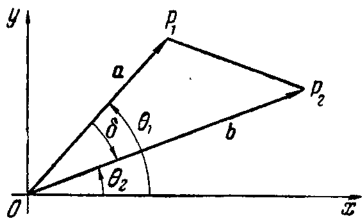


Рис. 13.

то (a, θ_1) — будут полярные координаты вершины P_1, а (b, θ_2) — полярные координаты вершины P_2. Угол $\delta > 0$ при расположении рис. 12 и $\delta < 0$ на рис. 13. Площадь нашего треугольника равна абсолютной величине выражения

$$\frac{1}{2} ab \sin \delta = \frac{1}{2} ab \sin (\theta_2 - \theta_1) = \frac{1}{2} ab (\sin \theta_2 \cos \theta_1 - \cos \theta_2 \sin \theta_1) =$$

$$= \frac{1}{2} (a \cos \theta_1 \cdot b \sin \theta_2 - b \cos \theta_2 \cdot a \sin \theta_1) = \frac{1}{2} (a_1 b_2 - a_2 b_1).$$

Это выражение может оказаться как положительным, так и отрицательным; оно меняет знак при перестановке вершин P_1 и P_2 (ср. рис. 12 с рис. 13). Ясно, что выражение

$$F = \frac{1}{2} (a_1 b_2 - a_2 b_1)$$

дает площадь треугольника, построенного на векторах $\boldsymbol{a}$ и $\boldsymbol{b}$, снабженное знаком плюс, если кратчайший поворот от вектора $\boldsymbol{a}$ до вектора $\boldsymbol{b}$ совпадает с принятым в нашей системе координат положительным направлением вращения, и знаком минус, если — с отрицательным. Но это направление кратчайшего поворота совпадает с направлением обхода вершин OP_1P_2. Поэтому можно сказать и так: площадь F положительна, если обход вершин треугольника происходит в положительном направлении, и отрицательна, если обход вершин совершается в отрицательном направлении. Вместе с тем видно, что знак площади треугольника соответствует общему соглашению о знаке площади, установленному в т. I, гл. V, § 2, n° 1 (стр. 311).

Выражение $a_1b_2-a_2b_1$, дающее удвоенную, снабженную знаком площадь ориентированного треугольника, принято записывать символически в следующем виде:

$$a_1b_2-a_2b_1=\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

и называть *определителем* (или *детерминантом*) *второго порядка*.

Теперь легко получить формулу для площади треугольника с вершинами в точках $A(x_0, y_0)$, $B(x_1, y_1)$, $C(x_2, y_2)$. Этот треугольник построен на векторах $\overline{AB}=\{x_1-x_0, y_1-y_0\}$ и $\overline{AC}=\{x_2-x_0, y_2-y_0\}$. Следовательно, его площадь

$$F=\frac{1}{2}\begin{vmatrix} x_1-x_0 & y_1-y_0 \\ x_2-x_0 & y_2-y_0 \end{vmatrix},$$

причем $F>0$, если обход вершин ABC совершается в положительном направлении, и $F<0$ в противном случае.

2. Векторное умножение двух векторов. Наряду со скалярным умножением важную роль играет *векторное умножение* двух векторов. Векторное произведение $[\boldsymbol{ab}]$ вектора $\boldsymbol{a}$ на вектор $\boldsymbol{b}$ определяется следующим образом [1]) (рис. 14).

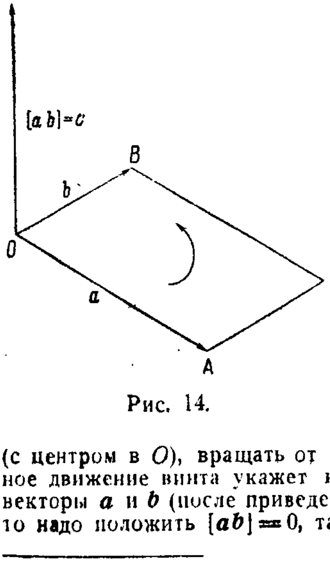


Рис. 14.

Приводим векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ к общему началу O и строим на них в пространстве параллелограмм. Векторным произведением $[\boldsymbol{ab}]=\boldsymbol{c}$ называется новый вектор $\boldsymbol{c}$, который перпендикулярен к плоскости параллелограмма, построенного на $\boldsymbol{a}$ и $\boldsymbol{b}$, имеет модуль, численно равный площади этого параллелограмма, и направлен в такую сторону, что тройка векторов $\boldsymbol{a}$, $\boldsymbol{b}$ и $\boldsymbol{c}=[\boldsymbol{ab}]$, взятая в этом порядке, имеет такую же ориентацию (правую или левую), как принятая в основу система координат x, y, z. Так как в этой книге используется только правая система координат, то у нас тройка векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $[\boldsymbol{ab}]$ будет всегда правой системой, т. е. если шляпку правого винта, лежащую в плоскости векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ (с центром в O), вращать от $\boldsymbol{a}$ к $\boldsymbol{b}$ по кратчайшему пути, то поступательное движение винта укажет направление вектора $[\boldsymbol{ab}]$ (см. рис. 14). Если векторы $\boldsymbol{a}$ и $\boldsymbol{b}$ (после приведения к общему началу) лежат на одной прямой, то надо положить $[\boldsymbol{ab}]=0$, так как площадь параллелограмма равна нулю.

---

[1]) В литературе встречаются различные обозначения векторного произведения, как $\boldsymbol{a}\times\boldsymbol{b}$, $\boldsymbol{a}\wedge\boldsymbol{b}$ и другие.

Свойства векторного произведения.

1) Если $a \neq 0$ и $b \neq 0$, то $[ab] = 0$ в том и только в том случае, когда векторы a и b имеют либо одинаковые, либо противоположные направления.

Действительно, при этом и только при этом условии равна нулю площадь параллелограмма, построенного на векторах a и b.

2) Вместо переместительного закона имеет место закон, который можно назвать *антипереместительным*:

$$[ba] = -[ab].$$

Он вытекает непосредственно из определения направления векторного произведения.

3) Если α и β — действительные числа, то

$$[\alpha a\, \beta b] = \alpha\beta\, [ab].$$

Если хоть одно из чисел α и β равно нулю, то в обеих частях этого равенства стоит нуль-вектор. Если $\alpha > 0$ и $\beta > 0$, то это равенство вытекает из того факта, что при этом вектор αa имеет одинаковое направление с a, вектор βa — то же самое направление, что и b, а параллелограмм, построенный на векторах αa и βb, имеет площадь, бо́льшую в $\alpha\beta$ раз площади параллелограмма, построенного на a и b, и лежит в той же плоскости, что и последний. Нетрудный по существу, но кропотливый разбор всех возможных комбинаций знаков чисел α и β обнаруживает справедливость свойства 3 при всех действительных значениях α и β.

4) Векторное произведение подчиняется *распределительному* закону:

$$[a\,(b + c)] = [ab] + [ac], \qquad [(b + c)\,a] = [ba] + [ca].$$

Достаточно доказать первую из этих формул, так как вторая вытекает из первой на основании свойства 2 — антипереместительного закона.

Мы дадим для векторного произведения $[ab]$ такое геометрическое построение, которое сразу обнаруживает справедливость распределительного закона.

Проведем через общее начало O векторов a и b плоскость E, перпендикулярную к вектору a. На рис. 15 эта плоскость перпендикулярна к плоскости чертежа. Спроектируем вектор b на плоскость E и получим вектор $b' = \overline{OB'}$ — компоненту вектора b в плоскости E. Тогда $[ab'] = [ab]$, так как, во-первых, параллелограмм, построенный на a и b', равновелик параллелограмму, построенному на a и b, а во-вторых, векторы $[ab']$ и $[ab]$ имеют одинаковое направление, ибо a, b, b' лежат в одной плоскости и направление вращения (по кратчайшему пути) от a к b' то же самое, что от a до b.

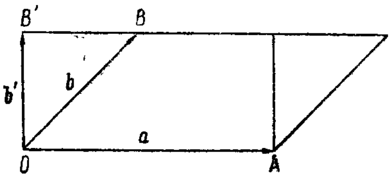


Рис. 15.

Так как a и b' взаимно перпендикулярны, то $|[ab']| = |[ab]| = |a| \cdot |b'|$. Поэтому, удлинив вектор b' в $|a|$ раз, получим вектор b'', имеющий тот же модуль, что и $[ab'] = [ab]$. Но $[ab]$ перпендикулярен как к вектору a, так и к b; поэтому вектор $[ab]$ получится из b'' поворотом на 90° вокруг вектора a как оси. Притом из конца вектора a это вращение должно выглядеть как положительное; такое вращение мы будем называть *положительным вращением вокруг вектора a как оси*. [Это значит следующее: если в основу положена, как в этой книге, правая система координат, то положительное вращение вокруг a соответствует направлению вектора a

по правилу правого винта, а если в основу положена левая система координат, то — по правилу левого винта.]

Итак, векторное произведение $[\boldsymbol{ab}]$ можно построить следующим образом: сперва ортогонально проектировать вектор $\boldsymbol{b}$ на плоскость E, перпендикулярную к $\boldsymbol{a}$, полученную компоненту удлинить в $|\boldsymbol{a}|$ раз и затем повернуть вокруг вектора $\boldsymbol{a}$ в положительном направлении на 90°.

Теперь приступаем к доказательству распределительного закона: $[\boldsymbol{a}(\boldsymbol{b}+\boldsymbol{c})]=[\boldsymbol{ab}]+[\boldsymbol{ac}]$. Векторы $\boldsymbol{b}=\overline{OB}$ и $\boldsymbol{c}=\overline{OC}$ являются сторонами параллелограмма $OBDC$, диагональ которого $\overline{OD}=\boldsymbol{b}+\boldsymbol{c}$. Три операции: проектирование, удлинение и поворот мы теперь выполним не над отдельными векторами $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{b}+\boldsymbol{c}$, а сразу над всем параллелограммом $OBDC$.

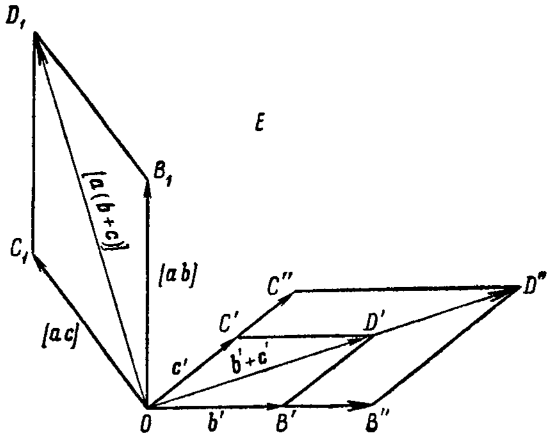


Рис. 16.

На рис. 16 изображена проекция $OB'D'C'$ этого параллелограмма на плокость E, перпендикулярную к вектору $\boldsymbol{a}$; эта плоскость E представлена теперь плоскостью чертежа, а вектор $\boldsymbol{a}$, не изображенный на рисунке, направлен вертикально вверх. Векторы $\overline{OB'}=\boldsymbol{b}'$, $\overline{OC'}=\boldsymbol{c}'$ и $\overline{OD'}=\boldsymbol{b}'+\boldsymbol{c}'$ являются компонентами векторов $\overline{OB}$, $\overline{OC}$ и $\overline{OD}$ на плоскость E. Над параллелограммом $OB'D'C'$ производим преобразование подобия с коэффициентом подобия, равным $|\boldsymbol{a}|$. Получился параллелограмм $OB''D''C''$, который мы затем вращаем в плоскости E вокруг вектора $\boldsymbol{a}$ как оси на 90° так, что направление вращения образует с вектором $\boldsymbol{a}$ правый винт. В итоге получится параллелограмм $OB_1D_1C_1$, сторонами которого являются векторы $\overline{OB_1}=[\boldsymbol{ab}]$ и $\overline{OC_1}=[\boldsymbol{ac}]$, а диагональный вектор $\overline{OD_1}=[\boldsymbol{a}(\boldsymbol{b}+\boldsymbol{c})]$. Тем самым подтверждается равенство $[\boldsymbol{ab}]+[\boldsymbol{ac}]=[\boldsymbol{a}(\boldsymbol{b}+\boldsymbol{c})]$.

В физике векторное произведение двух векторов используется для определения момента силы относительно точки. Сила $\boldsymbol{F}$, приложенная в точке с радиус-вектором $\boldsymbol{r}$, имеет относительно начала координат момент $[\boldsymbol{rF}]$.

3. Вычисление координат векторного произведения по координатам перемножаемых векторов. Найдем сначала попарные векторные произведения ортов координатных осей. На основании свойства 1

$$[\boldsymbol{e}_1\boldsymbol{e}_1]=[\boldsymbol{e}_2\boldsymbol{e}_2]=[\boldsymbol{e}_3\boldsymbol{e}_3]=0,$$

а из определения векторного произведения вытекает, что

$$[e_1e_2]=e_3, \qquad [e_2e_3]=e_1, \qquad [e_3e_1]=e_2,$$
$$[e_2e_1]=-e_3, \qquad [e_3e_2]=-e_1, \qquad [e_1e_3]=-e_2.$$

Пусть теперь векторы a и b заданы своими координатами: $a=\{a_1, a_2, a_3\}$, $b=\{b_1, b_2, b_3\}$. Выразим эти векторы через орты осей координат

$$a=a_1e_1+a_2e_2+a_3e_3, \qquad b=b_1e_1+b_2e_2+b_3e_3.$$

На основании распределительного закона и свойства 3 имеем

$$[ab]=[(a_1e_1+a_2e_2+a_3e_3)(b_1e_1+b_2e_2+b_3e_3)]=$$
$$=a_1b_1[e_1e_1]+a_1b_2[e_1e_2]+a_1b_3[e_1e_3]+a_2b_1[e_2e_1]+a_2b_2[e_2e_2]+a_2b_3[e_2e_3]+$$
$$+a_3b_1[e_3e_1]+a_3b_2[e_3e_2]+a_3b_3[e_3e_3].$$

Подставив сюда только что найденные парные векторные произведения ортов координатных осей, получаем

$$[ab]=(a_2b_3-a_3b_2)e_1+(a_3b_1-a_1b_3)e_2+(a_1b_2-a_2b_1)e_3.$$

Отсюда вытекает, что координаты векторного произведения $c=[ab]$ выражаются так:

$$c_1=\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \qquad c_2=\begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}=-\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \qquad c_3=\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

4. Объем тетраэдра. Рассмотрим тетраэдр (рис. 17), одна вершина которого есть общее начало M векторов $a=\{a_1, a_2, a_3\}$, $b=\{b_1, b_2, b_3\}$ и $c=\{c_1, c_2, c_3\}$, а остальными вершинами являются концы A, B, C этих векторов (на рис. 17 Mx, My, Mz — прямые, параллельные осям координат.) Для того чтобы выразить объем этого тетраэдра через координаты определяющих его векторов, поступим так. Примем за основание тетраэдра треугольник, построенный на векторах $a=\overline{MA}$ и $b=\overline{MB}$, площадь которого S численно равна $\frac{1}{2}|[ab]|$. Это векторное произведение направлено по перпендикуляру, опущенному из вершины C на плоскость треугольника MAB. Длина h этого перпендикуляра (высота тетраэдра) равна, с точностью до знака, проекции вектора $c=\overline{MC}$ на направление вектора $[ab]$, т. е. скалярному произведению $n^\circ c$, где n° — единичный вектор, указывающий направление $[ab]$. Само векторное произведение $[ab]=|[ab]|n^\circ$. Так как объем V тетраэдра равен $\frac{1}{3}Sh$, а $S=\frac{1}{2}|[ab]|$, то

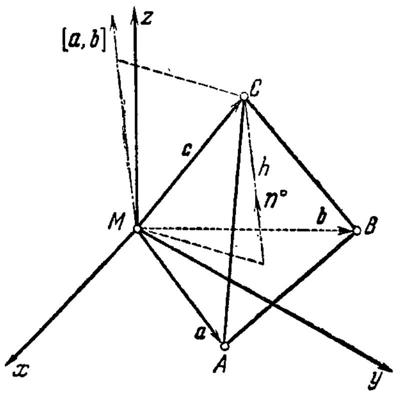


Рис. 17.

$$V=\frac{1}{6}|[ab]|n^\circ c=\frac{1}{6}[ab]c,$$

причем это выражение может оказаться как положительным, так и отрицательным, ибо (см. выше) h может отличаться знаком от скалярного произведения $n^\circ c$. Так как координаты вектора $[ab]$ равны соответственно

$$\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \quad -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix},$$

а координаты вектора c суть c_1, c_2, c_3, то, подставив в последнюю формулу координатное выражение скалярного произведения $[ab]$ на c, получим

$$V=\frac{1}{6}\left(c_1\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}-c_2\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}+c_3\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\right).$$

Полученные формулы верны и в том случае, если точки M, A, B (а значит, и векторы a и b) лежат на одной прямой. Правда, наше рассуждение уже непригодно, однако в этом случае $V=0$, а правая часть формулы тоже равна нулю, ибо $[ab]=0$, а стало быть, равны нулю и все координаты этого вектора.

Как и площадь треугольника по формуле, выведенной в n° 1, объем тетраэдра здесь также получается с определенным знаком. Покажем, что этот знак — положительный, если векторы a, b, c (взятые в этом порядке) образуют систему той же ориентации (правой или левой), что и применяемая система координат, и отрицательный, если обе системы имеют различную ориентацию. В самом деле, в первом случае угол δ между векторами $[ab]$ и c — острый, а во втором случае — тупой (что сразу видно из определения векторного произведения), а объем

$$V=\frac{1}{6}|[ab]|\cdot|c|\cos\delta.$$

Выражение, фигурирующее в скобках в координатной формуле для V, принято записывать короче с помощью символа

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = c_1\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}-c_2\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}+c_3\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix};$$

это выражение и обозначающий его символ называют *определителем или детерминантом третьего порядка*. Раскрывая содержащиеся в правой части определители второго порядка, получаем

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_2b_3c_1 - a_3b_2c_1 + a_3b_1c_2 - a_1b_3c_2 + a_1b_2c_3 - a_2b_1c_3.$$

Выведем теперь формулу для объема тетраэдра с вершинами в точках $A_0(x_0, y_0, z_0)$, $A_1(x_1, y_1, z_1)$, $A_2(x_2, y_2, z_2)$, $A_3(x_3, y_3, z_3)$. Этот тетраэдр можно считать построенным на векторах

$$\overline{A_0A_1}=\{x_1-x_0,\ y_1-y_0,\ z_1-z_0\},$$
$$\overline{A_0A_2}=\{x_2-x_0,\ y_2-y_0,\ z_2-z_0\},$$
$$\overline{A_0A_3}=\{x_3-x_0,\ y_3-y_0,\ z_3-z_0\}.$$

Следовательно, его объем (с надлежащим знаком)

$$V=\frac{1}{6}\begin{vmatrix} x_1-x_0 & y_1-y_0 & z_1-z_0 \\ x_2-x_0 & y_2-y_0 & z_2-z_0 \\ x_3-x_0 & y_3-y_0 & z_3-z_0 \end{vmatrix}.$$

Упражнения[1])

1. Вычислить расстояние от точки $P(x_0, y_0, z_0)$ до прямой l, заданной параметрическими уравнениями

$$x = at + b, \quad y = ct + d, \quad z = et + f.$$

2. Найти условие, при котором три вектора $\boldsymbol{a} = \{a_1, a_2, a_3\}$, $\boldsymbol{b} = \{b_1, b_2, b_3\}$, $\boldsymbol{c} = \{c_1, c_2, c_3\}$ параллельны одной плоскости.

3. Вывести условие того, что две прямые линии

$$\begin{aligned} x &= a_1 t + x_1, \\ y &= a_2 t + y_1, \\ z &= a_3 t + z_1 \end{aligned} \qquad \text{и} \qquad \begin{aligned} x &= b_1 t + x_2, \\ y &= b_2 t + y_2, \\ z &= b_3 t + z_2 \end{aligned}$$

либо пересекаются, либо параллельны.

4*. Найти кратчайшее расстояние между двумя прямыми l и l', заданными уравнениями

$$\begin{aligned} x &= at + b, \\ y &= ct + d, \\ z &= et + f \end{aligned} \qquad \text{и} \qquad \begin{aligned} x &= a't + b', \\ y &= c't + d', \\ z &= e't + f'. \end{aligned}$$

5. Показать, что плоскость, проходящая через три точки $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$ и $P_3(x_3, y_3, z_3)$, имеет уравнение

$$\begin{vmatrix} x_1 - x & y_1 - y & z_1 - z \\ x_2 - x & y_2 - y & z_2 - z \\ x_3 - x & y_3 - y & z_3 - z \end{vmatrix} = 0.$$

6. Тело вращается равномерно с угловой скоростью ω вокруг оси, проходящей через начало координат и имеющей направляющие косинусы (α, β, γ). Найти скорость точки $P(x, y, z)$.

7. Доказать тождество Лагранжа

$$[\boldsymbol{ab}]^2 = |\boldsymbol{a}|^2 \cdot |\boldsymbol{b}|^2 - (\boldsymbol{ab})^2.$$

8. Площадь выпуклого многоугольника на плоскости xy с вершинами $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, ..., $P_n(x_n, y_n)$ равна половине абсолютной величины выражения

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} + \ldots + \begin{vmatrix} x_{n-1} & y_{n-1} \\ x_n & y_n \end{vmatrix} + \begin{vmatrix} x_n & y_n \\ x_1 & y_1 \end{vmatrix}.$$

9*. а) Доказать неравенство

$$D = \begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} \leqslant \sqrt{(a^2 + b^2 + c^2)(a_1^2 + b_1^2 + c_1^2)(a_2^2 + b_2^2 + c_2^2)}.$$

б) При каком условии имеет место знак равенства?

§ 3. Элементарные сведения об определителях второго и третьего порядка

1. Законы составления и основные свойства. Определители второго и третьего порядка, участвующие в формулах для площади треугольника и объема тетраэдра, и их обобщение — *определитель n-го порядка*, играют важную роль во всех областях математики, в качестве аппарата для произ-

[1]) Более трудные упражнения отмечены звездочкой.

водства различных формальных вычислений в сжатом и легко обозримом виде. Свойства определителей мы здесь выведем на примере определителей второго и третьего порядка; определители более высокого порядка нам редко понадобятся. Следует, однако, подчеркнуть, что все основные теоремы, принадлежащей формулировке, справедливы и для определителей любого порядка. Для ознакомления с их теорией читателю следует обратиться к курсам алгебры и теории определителей [1]).

Определители

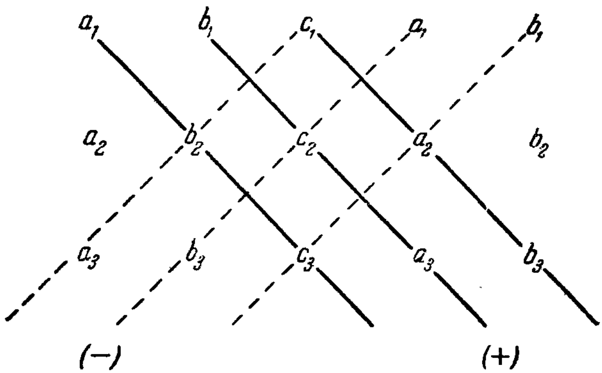
$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

представляют собой выражения, составленные по установленному правилу из своих элементов a_1, b_1, a_2, b_2 и a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3. Горизонтальные ряды (например, a_1, b_1, c_1) называются *строками*, а вертикальные ряды (например, c_1, c_2, c_3) — *столбцами* определителя.

О вычислении определителя второго порядка все сказано в определяющей его формуле

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1, \tag{1}$$

а для определителя третьего порядка приведем *правило Сарруса*, которое хорошо показывает симметричный характер закона составления этого определителя:



После третьего столбца вновь переписывают первые два столбца; затем вычисляют произведение каждой тройки чисел, стоящей на одной из шести прямых, составляющих угол 45° с горизонталью; три произведения — тех троек, которые идут сверху и слева направо и вниз, снабжают знаком плюс, а остальные три произведения — знаком минус, и все шесть членов складывают. В итоге получится значение определителя:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - a_1c_2b_3 - b_1a_2c_3. \tag{2}$$

[1]) См., например, Каган В. Ф., Основания теории определителей, Одесса, 1922; Курош А. Г., Курс высшей алгебры, изд. 1963; Сушкевич А. К., Основы высшей алгебры; Шрейер О. и Шпернер Е., Теория матриц, М.—Л., 1936.

Докажем теперь несколько теорем, выражающих основные свойства определителей.

1) *Если заменить в определителе строки столбцами, а столбцы строками, то значение его не изменится, т. е.*

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}, \qquad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Это легко проверить с помощью данных выше для определителей развернутых выражений (1) и (2).

2) *Если в определителе поменять местами две строки или два столбца, то его значение изменит только знак, т. е. помножится на* (—1).

На основании свойства 1 эту теорему достаточно доказать для столбцов, и ее легко проверить с помощью тех же выражений (1) и (2).

3) На стр. 32 мы ввели определители третьего порядка с помощью равенства

$$\Delta = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = c_1 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Принято говорить, что это равенство дает *разложение определителя третьего порядка по элементам третьей строки.*

Поменяем теперь местами вторую и третью строку определителя Δ и для полученного определителя выпишем аналогичное разложение по элементам третьей строки. Так как этот новый определитель равен $-\Delta$ (на основании свойства 2), то получается равенство

$$\Delta = -b_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}.$$

Таким же путем получим

$$\Delta = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

Последние два тождества называются *разложениями определителя третьего порядка по элементам* соответственно *второй и первой строки.*

Заменяя столбцы строками (на основании свойства 1 значение определителя при этом не изменяется), мы получим следующие три *разложения по элементам* первого, второго и третьего *столбцов*:

$$\Delta = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix},$$

$$\Delta = -a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix},$$

$$\Delta = a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Из этих разложений сразу вытекает следующая теорема.

4) *Если все элементы одной строки или одного столбца помножить на число k, то значение определителя помножится на k.*

Из свойств 2 и 4 можно вывести следующую теорему.

5) *Если элементы двух строк (или двух столбцов) пропорциональны,* т. е. если каждый элемент одной строки (или столбца) равен произведению соответствующего элемента другой строки (или столбца) на один и тот же множитель k, *то определитель равен нулю.*

Действительно, согласно теореме 4, множитель k можно вынести за знак определителя; если теперь поменять местами те строки (или столбцы), которые оказались одинаковыми, то определитель не изменится, тогда как, по теореме 2, он должен изменить свой знак. Следовательно, определитель равен нулю.

В частности, определитель, у которого какая-нибудь строка или столбец состоит из одних нулей, равен нулю; впрочем, это сразу видно из первоначального выражения определителя.

6) *Сумма двух определителей одинакового порядка, отличающихся только элементами одной строки (или одного столбца) равна определителю, в котором сохранены все общие строки (или столбцы) слагаемых определителей, а элементы оставшейся строки (или столбца) равны суммам соответствующих элементов слагаемых определителей.*

Пример.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b'_1 & c_1 \\ a_2 & b'_2 & c_2 \\ a_3 & b'_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1+b'_1 & c_1 \\ a_2 & b_2+b'_2 & c_2 \\ a_3 & b_3+b'_3 & c_3 \end{vmatrix}.$$

В самом деле, разложим оба слагаемых определителя по элементам той строки (или того столбца), которыми они отличаются (в нашем примере — по элементам второго столбца), а затем сложим результаты; получится выражение

$$-(b_1+b'_1)\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + (b_2+b'_2)\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - (b_3+b'_3)\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix},$$

которое, очевидно, является разложением определителя

$$\begin{vmatrix} a_1 & b_1+b'_1 & c_1 \\ a_2 & b_2+b'_2 & c_2 \\ a_3 & b_3+b'_3 & c_3 \end{vmatrix}$$

по элементам второго столбца. Это и доказывает теорему 6.

Такое же рассуждение можно провести и для определителя второго порядка.

7) *Если к элементам одного ряда (строки или столбца) определителя прибавить соответствующие элементы параллельного ряда, умноженные на один и тот же множитель k, то значение определителя не изменится.*

Действительно, по теореме 6, видоизмененный определитель равен сумме первоначального определителя и определителя, в котором элементы двух параллельных рядов пропорциональны; но, согласно теореме 5, этот второй определитель равен нулю.

Покажем на примерах приложение этих теорем к вычислению определителей.

Пример 1.

$$\begin{vmatrix} a_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & c_3 \end{vmatrix} = a_1 b_2 c_3,$$

что можно получить, разлагая определитель по элементам, например, первой строки. Определитель, у которого отличны от нуля только элементы «главной диагонали», равен произведению этих элементов.

Пример 2.

$$\begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{vmatrix} = 2\begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} = -4.$$

Сначала к элементам первой строки прибавлены элементы второй, а затем новый определитель разложен по элементам первой строки.

Пример 3.

$$D = \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = \begin{vmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 1 & z & z^2 \end{vmatrix} = \begin{vmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 0 & z-x & z^2-x^2 \end{vmatrix}.$$

Разлагая последний определитель по элементам первого столбца, получим

$$D = \begin{vmatrix} y-x & y^2-x^2 \\ z-x & z^2-x^2 \end{vmatrix} = (y-x)(z-x)\begin{vmatrix} 1 & y+x \\ 1 & z+x \end{vmatrix} = (y-x)(z-x)(z-y).$$

2. Понятие об определителе четвертого и вообще любого порядка. Правилом разложения по элементам строки или столбца можно воспользоваться для введения, в порядке обобщения, понятия определителя четвертого и вообще какого угодно порядка. Если дана квадратная таблица из 16 величин

$$\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix},$$

то этим символом обозначают выражение

$$a_1\begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix} - d_1\begin{vmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{vmatrix}$$

и называют это выражение *определителем четвертого порядка*. Таким же путем можно последовательно ввести определители пятого, шестого и вообще n-го порядка, так что определитель порядка n выражается рекуррентной формулой через определители порядка $n-1$. Оказывается, что все доказанные нами существенные свойства определителей второго или третьего порядка распространяются на определители любого порядка. (Исключением является только правило Сарруса, которое пригодно только для определителей третьего порядка.) Проведение такой программы выходит, однако, за пределы этого курса.

3. Приложение к системе линейных уравнений. Определители играют фундаментальную роль в теории систем линейных уравнений. Для того чтобы решить систему двух уравнений

$$a_1x + a_2y = A,$$
$$b_1x + b_2y = B$$

с двумя неизвестными x и y, помножим первое уравнение на b_2, второе — на a_2 и вычтем второе уравнение из первого; затем умножим первое уравнение на b_1, второе — на a_1 и вычтем первое уравнение из второго. В итоге получим

$$(a_1b_2-a_2b_1)\,x=Ab_2-Ba_2,$$
$$(a_1b_2-a_2b_1)\,y=Ba_1-Ab_1$$

или

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} x=\begin{vmatrix} A & a_2 \\ B & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} y=\begin{vmatrix} a_1 & A \\ b_1 & B \end{vmatrix}. \qquad (*)$$

Предположим, что определитель

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0.$$

Тогда из последних уравнений сразу получаются формулы решения

$$x=\frac{\begin{vmatrix} A & a_2 \\ B & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}, \qquad y=\frac{\begin{vmatrix} a_1 & A \\ b_1 & B \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}},$$

которые должно и можно проверить подстановкой. Определитель, являющийся общим знаменателем обеих дробей, называется *определителем системы.*

Если же, напротив, определитель $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}=0$ и при этом хотя бы один из определителей $\begin{vmatrix} A & a_2 \\ B & b_2 \end{vmatrix}$ и $\begin{vmatrix} a_1 & A \\ b_1 & B \end{vmatrix}$ отличен от нуля, то уравнения (*) приводят к противоречию.

Наконец, если все три определителя $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$, $\begin{vmatrix} A & a_2 \\ B & b_2 \end{vmatrix}$ и $\begin{vmatrix} a_1 & A \\ b_1 & B \end{vmatrix}$ равны нулю, то уравнения (*) ничего не говорят о решении системы.

Особенно важен для нас следующий полученный выше факт: система уравнений вышеуказанного вида, определитель которой отличен от нуля, всегда имеет одно и только одно решение.

Если определитель системы не равен нулю, а система *однородна*, т. е. $A=B=0$, то этим единственным решением является $x=0$, $y=0$.

Возьмем теперь систему трех линейных уравнений с тремя неизвестными

$$a_1x+a_2y+a_3z=A,$$
$$b_1x+b_2y+b_3z=B,$$
$$c_1x+c_2y+c_3z=C.$$

Аналогичное рассуждение приводит здесь к аналогичному результату.

Помножим первое уравнение на $\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}$, второе на $-\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}$, а третье на $\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}$ и сложим полученные три уравнения:

$$x\left\{a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-b_1\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}+c_1\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\right\}+$$
$$+y\left\{a_2\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-b_2\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}+c_2\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\right\}+$$
$$+z\left\{a_3\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-b_3\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}+c_3\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\right\}=$$
$$=A\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-B\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}+C\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}.$$

Правая часть, а также каждое из трех выражений, заключенных в фигурные скобки, есть разложение какого-то определителя по элементам одного из своих столбцов, и последнее уравнение можно записать в следующем виде:

$$x\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}+y\begin{vmatrix} a_2 & a_2 & a_3 \\ b_2 & b_2 & b_3 \\ c_2 & c_2 & c_3 \end{vmatrix}+z\begin{vmatrix} a_3 & a_2 & a_3 \\ b_3 & b_2 & b_3 \\ c_3 & c_2 & c_3 \end{vmatrix}=\begin{vmatrix} A & a_2 & a_3 \\ B & b_2 & b_3 \\ C & c_2 & c_3 \end{vmatrix}.$$

На основании теоремы 5 коэффициенты при y и z равны нулю, так что

$$x\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=\begin{vmatrix} A & a_2 & a_3 \\ B & b_2 & b_3 \\ C & c_2 & c_3 \end{vmatrix}.$$

Таким же способом выводим еще уравнения:

$$y\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=\begin{vmatrix} a_1 & A & a_3 \\ b_1 & B & b_3 \\ c_1 & C & c_3 \end{vmatrix},$$

$$z\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=\begin{vmatrix} a_1 & a_2 & A \\ b_1 & b_2 & B \\ c_1 & c_2 & C \end{vmatrix}.$$

Если определитель

$$\Delta=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix},$$

называемый *определителем системы*, отличен от нуля, то последние три уравнения дают искомые значения неизвестных. Полученные значения нуждаются еще в проверке, которую можно выполнить подстановкой. Итак, если $\Delta \neq 0$, то система трех линейных уравнений с тремя неизвестными имеет одно и только одно решение. В частности, если наша система уравнений однородна, т. е. $A=B=C=0$, а $\Delta \neq 0$, то это единственное решение есть $x=y=z=0$.

Если же, напротив, определитель системы $\Delta=0$, то из последних уравнений вытекает, что и их правые части должны равняться нулю, если только существует решение. Таким образом, при $\Delta=0$ заданная система имеет решение в том и только в том случае, если постоянные члены A, B, C удовлетворяют определенным условиям, выражающимся в том, что определители в правых частях последних трех уравнений должны обратиться в нуль.

Кроме рассмотренных нами систем, в которых число уравнений равно числу неизвестных, в дальнейшем встретятся порою системы двух однородных уравнений с тремя неизвестными

$$a_1x+a_2y+a_3z=0, \quad b_1x+b_2y+b_3z=0.$$

Если три определителя

$$D_1=\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \qquad D_2=\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \qquad D_3=\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

не равны одновременно нулю, то из наших уравнений можно выразить две неизвестные через третью; если, например, $D_3 \neq 0$, то

$$x=\frac{zD_1}{D_3}, \qquad y=-\frac{zD_2}{D_3}.$$

Введя обозначение $\frac{z}{D_3}=t$, получим

$$x=tD_1, \qquad y=-tD_2, \qquad z=tD_3.$$

Таким образом, наша система уравнений — неопределенная и имеет семейство решений, выражающихся через один параметр t, причем $x:y:z=D_1:(-D_2):D_3$.

Решенная только что задача имеет простой геометрический смысл. Рассмотрим два известных вектора $\boldsymbol{a}=\{a_1, a_2, a_3\}$ и $\boldsymbol{b}=\{b_1, b_2, b_3\}$ и неизвестный вектор $\boldsymbol{r}=\{x, y, z\}$. Наша система уравнений получает следующую векторную запись:

$$\boldsymbol{ar}=0, \qquad \boldsymbol{br}=0,$$

и геометрический смысл задачи таков: требуется найти вектор $\boldsymbol{r}$, перпендикулярный к двум данным векторам $\boldsymbol{a}$ и $\boldsymbol{b}$. Очевидно,

$$\boldsymbol{r}=t\,[\boldsymbol{ab}]$$

при любом значении числа t.

Упражнения

1. Показать, что определитель $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$ всегда можно привести к виду $\begin{vmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{vmatrix}$ путем повторного применения следующих операций: 1) перестановка двух строк или столбцов, 2) прибавление к элементам одной строки (или одного столбца) чисел, пропорциональных элементам другой строки (или другого столбца).

2. Если три определителя

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}, \qquad \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}, \qquad \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

не равны все нулю, то необходимым и достаточным условием существования решения системы трех уравнений

$$a_1x+a_2y=A,$$
$$b_1x+b_2y=B,$$
$$c_1x+c_2y=C,$$

с двумя неизвестными x и y является

$$D=\begin{vmatrix} a_1 & a_2 & A \\ b_1 & b_2 & B \\ c_1 & c_2 & C \end{vmatrix}=0.$$

3. Решить упражнение 3 стр. 33 чисто алгебраическим рассуждением.

4*. Доказать теоремы 1—7 (n° 1) для определителей четвертого порядка.

5. Доказать, что площадь треугольника на плоскости с вершинами (x_1, y_1), (x_2, y_2), (x_3, y_3) выражается формулой

$$F=\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

6. Доказать, что объем тетраэдра с вершинами (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4) выражается формулой

$$V=\frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}.$$

7. Вычислить следующие определители:

$$\text{а)}\ \begin{vmatrix} 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{vmatrix};\qquad \text{б)}\ \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{vmatrix};\qquad \text{в)}\ \begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & -1 & 7 \end{vmatrix};\qquad \text{г)}\ \begin{vmatrix} 1 & x & x^3 \\ 1 & y & y^3 \\ 1 & z & z^3 \end{vmatrix}.$$

8. Найти условие, которому должны удовлетворять числа a, b, c, чтобы система уравнений

$$\begin{aligned} 3x+4y+5z&=a,\\ 4x+5y+6z&=b,\\ 5x+6y+7z&=c \end{aligned}$$

могла иметь решение.

9. Решить систему уравнений

$$\begin{aligned} 2x-3y+4z&=4,\\ 4x-9y+16z&=10,\\ 8x-27y+64z&=34. \end{aligned}$$

§ 4. Аффинные преобразования и умножение определителей

В заключение этого краткого обзора мы исследуем простейшие свойства так называемых аффинных преобразований и попутно получим важную теорему об умножении определителей.

1. Аффинное преобразование плоскости и пространства. Под *отображением* или *преобразованием* какой-либо части пространства (или плоскости) разумеют закон, по которому каждой точке приводится в соответствие в качестве ее *изображения* другая точка пространства (или плоскости). Исходную точку называют *оригиналом* или *прототипом*. Понятию отображения можно дать наглядное физическое истолкование: представим себе, что рассматриваемая часть пространства (или плоскости) заполнена каким-нибудь деформируемым веществом, и что наше преобразование описывает процесс деформации, при котором каждая частица этого вещества перемещается из своего начального положения в некоторое конечное положение.

Пользуясь прямоугольной системой координат, обозначим через (x, y, z) координаты исходной точки, а через (x', y', z') — координаты ее изображения.

Самыми простыми и наглядными являются аффинные преобразования, которые, к тому же, имеют большое значение в общей теории преобразований.

Аффинным преобразованием называется такое преобразование, при котором координаты изображения x', y', z' (а на плоскости — x', y') выражаются линейно (точнее, в виде целых линейных функций) через координаты оригинала x, y, z (или x, y — на плоскости). Стало быть, аффинное преобразование в пространстве задается тремя уравнениями

$$\begin{aligned} x'&=a_1x+b_1y+c_1z+m_1,\\ y'&=a_2x+b_2y+c_2z+m_2,\\ z'&=a_3x+b_3y+c_3z+m_3, \end{aligned}$$

а в плоскости — двумя уравнениями

$$x' = a_1 x + b_1 y + m_1,$$
$$y' = b_1 x + b_2 y + m_2$$

с постоянными коэффициентами a_i, b_i, c_i, m_i. Каждой точке пространства (или плоскости) эти уравнения приводят в соответствие другую точку — ее изображение. И сразу возникает вопрос: возможно ли обратить это соответствие между оригиналом и его изображением? Другими словами, соответствует ли обратно каждой точке пространства (или плоскости) определенная точка в качестве ее оригинала. Для этого, очевидно, необходимо и достаточно, чтобы уравнения

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= x' - m_1, \\ a_2 x + b_2 y + c_2 z &= y' - m_2, \\ a_3 x + b_3 y + c_3 z &= z' - m_3, \end{aligned} \qquad \text{или} \qquad \begin{aligned} a_1 x + b_1 y &= x' - m_1, \\ a_2 x + b_2 y &= y' - m_2 \end{aligned}$$

были однозначно разрешимы относительно x, y, z (или x, y) при всех значениях x', y', z' (или x', y'). Поэтому, согласно § 3, n° 3, аффинное преобразование обратимо, и притом однозначно обратимо (т. е. каждое изображение имеет один и только один оригинал), если его определитель

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \qquad \text{или} \qquad \Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

отличен от нуля. Только такие преобразования мы и будем рассматривать и не станем касаться вопроса, что происходит, когда $\Delta = 0$.

Путем введения промежуточной точки x'', y'', z'' (или x'', y'') можно общее аффинное преобразование разбить на два преобразования

$$\begin{aligned} x'' &= a_1 x + b_1 y + c_1 z, \\ y'' &= a_2 x + b_2 y + c_2 z, \\ z'' &= a_3 x + b_3 y + c_3 z \end{aligned} \qquad \text{и} \qquad \begin{aligned} x' &= x'' + m_1, \\ y' &= y'' + m_2, \\ z' &= z'' + m_3, \end{aligned}$$

или

$$\begin{aligned} x'' &= a_1 x + b_1 y, \\ y'' &= a_2 x + b_2 y \end{aligned} \qquad \text{и} \qquad \begin{aligned} x' &= x'' + m_1, \\ y' &= y'' + m_2. \end{aligned}$$

Сначала точка (x, y, z) отображается на (x'', y'', z''), а затем точка (x'', y'', z'') отображается на (x', y', z'). Так как второе преобразование есть просто параллельное перемещение пространства (или плоскости) как целого и, стало быть, не вызывает никаких вопросов, то можно ограничиться исследованием одного лишь первого преобразования. Поэтому мы будем рассматривать только аффинные преобразования вида

$$\begin{aligned} x' &= a_1 x + b_1 y + c_1 z, \\ y' &= a_2 x + b_2 y + c_2 z, \\ z' &= a_3 x + b_3 y + c_3 z \end{aligned} \qquad \text{или} \qquad \begin{aligned} x' &= a_1 x + b_1 y, \\ y' &= a_2 x + b_2 y \end{aligned}$$

с определителем, отличным от нуля.

Теоремы о линейных уравнениях, доказанные в § 3, n° 3, позволяют выразить *обратное преобразование* в виде следующих формул:

$$\begin{aligned} x &= a_1' x' + b_1' y' + c_1' z', \\ y &= a_2' x' + b_2' y' + c_2' z', \\ z &= a_3' x' + b_3' y' + c_3' z' \end{aligned} \qquad \text{или} \qquad \begin{aligned} x &= a_1' x' + b_1' y', \\ y &= a_2' x' + b_2' y', \end{aligned}$$

в которых постоянные a'_i, b'_i, ... представляют собой определенные выражения, составленные из коэффициентов a_i, b_i, ... Из этих уравнений вытекают обратно первоначальные формулы. В частности, из $x=y=z=0$ вытекает $x'=y'=z'=0$, и обратно.

Геометрические свойства аффинного преобразования характеризуются следующими теоремами:

1) *В пространстве: изображение плоскости есть плоскость; на плоскости: изображение прямой есть прямая.*

Действительно, уравнение плоскости можно записать в виде

$$Ax+By+Cz+D=0,$$

а уравнение прямой на плоскости — в виде

$$Ax+By+D=0,$$

причем коэффициенты A, B, C (или A, B) не равны одновременно нулю. Координаты изображений точек данной плоскости удовлетворяют уравнению

$$A(a'_1x'+b'_1y'+c'_1z')+B(a'_2x'+b'_2y+c'_2z')+C(a'_3x'+b'_3y'+c'_3z')+D=0,$$

а координаты изображений точек данной прямой на плоскости удовлетворяют уравнению

$$A(a'_1x'+b'_1y')+B(a'_2x'+b'_2y')+D=0.$$

Следовательно, эти изображающие точки x', y', z' (или x', y') сами заполняют плоскость (или прямую на плоскости), так как коэффициенты при текущих координатах x', y', z' (или x', y')

$$\begin{aligned} A' &= Aa'_1+Ba'_2+Ca'_3,\\ B' &= Ab'_1+Bb'_2+Cb'_3,\\ C' &= Ac'_1+Bc'_2+Cc'_3 \end{aligned} \qquad \left(\text{или}\quad \begin{aligned} A' &= Aa'_1+Ba'_2,\\ B' &= Ab'_1+Bb'_2 \end{aligned}\right)$$

не могут все одновременно обратиться в нуль. В противном случае мы имели бы

$$\begin{aligned} Aa'_1+Ba'_2+Ca'_3 &= 0,\\ Ab'_1+Bb'_2+Cb'_3 &= 0,\\ Ac'_1+Bc'_2+Cc'_3 &= 0 \end{aligned} \qquad \left(\text{или}\quad \begin{aligned} Aa'_1+Ba'_2 &= 0,\\ Ab'_1+Bb'_2 &= 0 \end{aligned}\right),$$

и эти уравнения можно было бы рассматривать как однородную систему уравнений с неизвестными A, B, C (или A, B), а из этой системы вытекало бы, что $A=B=C=0$ (или $A=B=0$), что противоречит условию.

2) *Изображение прямой в пространстве есть прямая.*

В самом деле, прямую в пространстве можно рассматривать как линию пересечения двух плоскостей; поэтому, согласно теореме 1, изображение прямой тоже является линией пересечения двух плоскостей, т. е. прямой.

3) *Две параллельные плоскости в пространстве (или две параллельные прямые на плоскости) имеют параллельные изображения.*

Ибо если бы эти изображения пересекались, то и исходные две плоскости (или прямые на плоскости) пересекались бы в точках, являющихся оригиналами точек пересечения изображений.

4) *Две параллельные прямые в пространстве имеют своими изображениями две параллельные же прямые.*

Действительно, так как данные две прямые лежат в одной плоскости и не пересекаются, то, по теоремам 1 и 2, тем же свойством обладают и прямые, их изображающие; стало быть, прямые эти тоже параллельны.

5) *Изображение вектора $\mathfrak{v}$ есть другой вектор $\mathfrak{v}'$, началом которого является изображение начальной точки, а концом — изображение конечной точки вектора $\mathfrak{v}$.*

Так как координаты вектора равны разностям одноименных координат конечной и начальной точки, то при общем аффинном преобразовании они преобразуются по той же схеме, что и координаты точки:

$$v_1' = a_1v_1 + b_1v_2 + c_1v_3,$$
$$v_2' = a_2v_1 + b_2v_2 + c_2v_3,$$
$$v_3' = a_3v_1 + b_3v_2 + c_3v_3.$$

2. Умножение аффинных преобразований и разложение общего аффинного преобразования на примитивные преобразования. Если изображение (x', y', z') точки (x, y, z), полученное преобразованием

$$x' = a_1x + b_1y + c_1z,$$
$$y' = a_2x + b_2y + c_2z,$$
$$z' = a_3x + b_3y + c_3z,$$

подвергнем второму преобразованию

$$x'' = a_1'x' + b_1'y' + c_1'z',$$
$$y'' = a_2'x' + b_2'y' + c_2'z',$$
$$z'' = a_3'x' + b_3'y' + c_3'z',$$

то нетрудно убедиться, что новая изображающая точка (x'', y'', z'') может быть получена непосредственно по точке (x, y, z) с помощью «результирующего», тоже аффинного, преобразования

$$x'' = \alpha_1x + \beta_1y + \gamma_1z,$$
$$y'' = \alpha_2x + \beta_2y + \gamma_2z,$$
$$z'' = \alpha_3x + \beta_3y + \gamma_3z,$$

коэффициенты которого выражаются через коэффициенты составляющих преобразований следующими формулами:

$$\alpha_1 = a_1'a_1 + b_1'a_2 + c_1'a_3, \quad \beta_1 = a_1'b_1 + b_1'b_2 + c_1'b_3, \quad \gamma_1 = a_1'c_1 + b_1'c_2 + c_1'c_3,$$
$$\alpha_2 = a_2'a_1 + b_2'a_2 + c_2'a_3, \quad \beta_2 = a_2'b_1 + b_2'b_2 + c_2'b_3, \quad \gamma_2 = a_2'c_1 + b_2'c_2 + c_2'c_3,$$
$$\alpha_3 = a_3'a_1 + b_3'a_2 + c_3'a_3, \quad \beta_3 = a_3'b_1 + b_3'b_2 + c_3'b_3, \quad \gamma_3 = a_3'c_1 + b_3'c_2 + c_3'c_3.$$

Процесс последовательного проведения двух преобразований называется *умножением* преобразований, а «результирующее» преобразование называется *произведением* данных двух преобразований. Если определители первых двух преобразований отличны от нуля, то эти преобразования обратимы; поэтому произведение этих преобразований тоже обратимо. Последние формулы показывают, что коэффициенты результирующего преобразования получаются из коэффициентов составляющих преобразований по следующему правилу:

Коэффициент, стоящий в i-м столбце и k-й строке результирующего преобразования, равен сумме произведений коэффициентов i-го столбца первого преобразования на соответствующие коэффициенты k-й строки второго преобразования.

Совершенно таким же путем умножение преобразований

$$x' = a_1x + b_1y, \qquad x'' = a_1'x' + b_1'y',$$
$$y' = a_2x + b_2y \quad \text{и} \quad y'' = a_2'x' + b_2'y'$$

дает результирующее преобразование (произведение данных двух преобразований)

$$x'' = (a_1 a'_1 + a_2 b'_1)\, x + (b_1 a'_1 + b_2 b'_1)\, y,$$
$$y'' = (a_1 a'_2 + a_2 b'_2)\, x + (b_1 a'_2 + b_2 b'_2)\, y.$$

Преобразование называется *примитивным*, если оно оставляет неизменными две из трех (на плоскости — одну из двух) координат оригинала. В наглядной форме это можно выразить так: примитивным называется такое преобразование, при котором пространство (или плоскость) подвергается растяжению только вдоль одного направления, так что все точки перемещаются вдоль семейства параллельных прямых. (Однако величина растяжения может изменяться от точки к точке.) Примитивное аффинное преобразование, при котором перемещение происходит параллельно оси x, выражается аналитически формулами следующего вида:

$$\begin{aligned} x' &= ax + by + cz, \\ y' &= y, \\ z' &= z \end{aligned} \qquad \text{или} \qquad \begin{aligned} x' &= ax + by, \\ y' &= y. \end{aligned}$$

Общее аффинное преобразование на плоскости

$$x' = a_1 x + b_1 y,$$
$$y' = a_2 x + b_2 y,$$

с отличным от нуля определителем, можно представить в виде произведения примитивных преобразований.

Приступая к доказательству, можно предположить, что $a_1 \neq 0$ [1]). Введем промежуточное примитивное преобразование

$$\xi = a_1 x + b_1 y, \qquad \eta = y,$$

определитель которого равен a_1 и, стало быть, отличен от нуля. Из ξ, η получаем x', y' с помощью второго примитивного преобразования

$$x' = \xi, \qquad y' = \frac{a_2}{a_1}\xi + \frac{a_1 b_2 - a_2 b_1}{a_1}\eta$$

с определителем $\dfrac{a_1 b_2 - a_2 b_1}{a_1} = \dfrac{1}{a_1}\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$. Таким образом, данное преобразование представлено как произведение примитивных аффинных преобразований.

Совершенно аналогичным путем *можно и общее аффинное преобразование в пространстве*

$$x' = a_1 x + b_1 y + c_1 z,$$
$$y' = a_2 x + b_2 y + c_2 z,$$
$$z' = a_3 x + b_3 y + c_3 z$$

с отличным от нуля определителем представить в виде произведения примитивных аффинных преобразований.

---

[1]) Если $a_1 = 0$, то $b_1 \neq 0$, и к случаю $a_1 \neq 0$ можно вернуться, изменив обозначения: x на y и y на x. Это изменение обозначений, выражающееся преобразованием $X = y$, $Y = x$, можно в свою очередь осуществить с помощью последовательного проведения примитивных преобразований

$$\begin{aligned} x_1 &= x - y, \\ y_1 &= y, \end{aligned} \qquad \begin{aligned} x_2 &= x_1, \\ y_2 &= x_1 + y_1 = x, \end{aligned} \qquad \begin{aligned} X &= -x_2 + y_2 = y, \\ Y &= y_2 = x. \end{aligned}$$

Сначала заметим, что из трех определителей второго порядка

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}, \quad \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

по крайней мере один отличен от нуля, так как в противном случае определитель преобразования

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

был бы равен нулю, как показывает его разложение по элементам третьей строки. Как и в плоском случае, можно допустить, не жертвуя общностью, что 1) $a_1 \neq 0$ и 2) $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$.

Вводим первое преобразование, дающее промежуточное изображение (ξ, η, ζ):

$$\begin{aligned} \xi &= a_1x + b_1y + c_1z, \\ \eta &= \quad\quad\;\; y \quad\quad\;\;, \\ \zeta &= \quad\quad\quad\quad\quad\;\; z. \end{aligned}$$

Определитель этого примитивного преобразования есть $a_1 \neq 0$.

Для второго преобразования, которое приведет к точке (ξ', η', ζ'), положим $\xi' = \xi$, $\zeta' = \zeta$, а выражение для η' выберем так, чтобы было $\eta' = y'$. Для этого подставим в $\eta' = y' = a_2x + b_2y + c_2z$ вместо x, y, z их выражения через ξ, η, ζ и в итоге получится второе примитивное преобразование

$$\begin{aligned} \xi' &= \xi, \\ \eta' &= \frac{a_2}{a_1}\xi + \frac{1}{a_1}\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}\eta + \frac{1}{a_1}\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}\zeta, \\ \zeta' &= \zeta. \end{aligned}$$

Определитель преобразования есть $\frac{1}{a_1}\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$.

Третье преобразование предрешено двумя первыми:

$$\begin{aligned} x' &= \xi', \\ y' &= \eta', \\ z' &= -\frac{\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}\xi' + \frac{\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}\eta' + \frac{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}\zeta'. \end{aligned}$$

3. Геометрический смысл определителя преобразования и теорема умножения определителей. Полученные результаты позволяют выяснить геометрический смысл определителя аффинного преобразования и доказать алгебраическую теорему об умножении определителей.

Рассмотрим треугольник на плоскости xy с вершинами в точках $(0, 0)$, (x_1, y_1), (x_2, y_2). Согласно § 2, n° 1, его площадь

$$F = \frac{1}{2}\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.$$

Подвергнем плоскость xy примитивному аффинному преобразованию

$$x' = ax + by, \qquad y' = y.$$

Тогда нашему треугольнику будет соответствовать на плоскости изображений треугольник с вершинами $(0, 0)$, $(ax_1 + by_1, y_1)$ и $(ax_2 + by_2, y_2)$. Каково отношение площади F' этого преобразованного треугольника к площади F данного треугольника? По той же формуле из § 2, n° 1

$$F' = \frac{1}{2}\begin{vmatrix} x_1' & x_2' \\ y_1' & y_2' \end{vmatrix} = \frac{1}{2}\begin{vmatrix} ax_1 + by_1 & ax_2 + by_2 \\ y_1 & y_2 \end{vmatrix} =$$

$$= \frac{1}{2}\begin{vmatrix} ax_1 & ax_2 \\ y_1 & y_2 \end{vmatrix} + \frac{1}{2}\begin{vmatrix} by_1 & by_2 \\ y_1 & y_2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} ax_1 & ax_2 \\ y_1 & y_2 \end{vmatrix}.$$

Следовательно,

$$F' = \frac{1}{2}\, a \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = aF.$$

Если же взять примитивное преобразование

$$x' = x, \qquad y' = \alpha x + \beta y,$$

то таким же путем получим

$$F' = \beta F.$$

Оказывается, стало быть, что примитивное аффинное преобразование вызывает умножение площади треугольника на постоянный множитель, не зависящий от выбора треугольника. То обстоятельство, что мы взяли треугольник, одна из вершин которого находится в начале координат, не существенно: тот же результат легко получить для любого треугольника.

Так как общее аффинное преобразование может быть представлено в виде произведения примитивных преобразований, то этим же свойством обладает любое аффинное преобразование. *При аффинном преобразовании плоскости отношение площади преобразованного треугольника к площади исходного треугольника есть постоянная величина, не зависящая от треугольника и полностью определяемая коэффициентами преобразования.*

Для того чтобы вычислить это постоянное отношение, возьмем треугольник с вершинами $(0, 0)$, $(1, 0)$ и $(0, 1)$, площадь которого $F = \frac{1}{2}$. Преобразование

$$x' = a_1 x + b_1 y,$$
$$y' = a_2 x + b_2 y$$

переводит его в треугольник с вершинами $(0, 0)$, (a_1, a_2) и (b_1, b_2) с площадью (по формуле § 2, n° 1)

$$F' = \frac{1}{2}\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = F\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Стало быть, *постоянное отношение площадей $F' : F$ при любом аффинном преобразовании равно определителю преобразования.*

Аналогичное исследование можно провести для пространственных аффинных преобразований. Рассмотрим тетраэдр с вершинами $(0, 0, 0)$, (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3) и совершим примитивное преобразование

$$\begin{aligned} x' &= ax + by + cz, \\ y' &= \phantom{ax + {}} y \phantom{{} + cz}, \\ z' &= \phantom{ax + by + {}} z. \end{aligned}$$

Преобразованный тетраэдр имеет вершины $(0, 0, 0)$, $(ax_1+by_1+cz_1, y_1, z_1)$, $(ax_2+by_2+cz_2, y_2, z_2)$, $(ax_3+by_3+cz_3, y_3, z_3)$ и объем

$$V' = \frac{1}{6}\begin{vmatrix} ax_1+by_1+cz_1 & ax_2+by_2+cz_2 & ax_3+by_3+cz_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = \frac{1}{6}a\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

Следовательно,

$$V' = aV,$$

где V — объем исходного тетраэдра. Если взять примитивное преобразование

$$x' = x, \qquad y' = \alpha x + \beta y + \gamma z, \qquad z' = z,$$

то объем V' преобразованного тетраэдра получится тем же путем:

$$V' = \beta V,$$

а примитивное преобразование

$$x' = x, \qquad y' = y, \qquad z' = Ax + By + Cz$$

дает

$$V' = CV.$$

Отсюда вытекает, что и любое аффинное преобразование вызывает умножение объема тетраэдра на постоянный коэффициент. (Мы взяли тетраэдр, одна из вершин которого лежит в начале координат, но нетрудно показать, что это справедливо для любого тетраэдра.)

Для того чтобы вычислить постоянное отношение $V' : V$ для общего аффинного преобразования

$$x' = a_1x + b_1y + c_1z,$$
$$y' = a_2x + b_2y + c_2z,$$
$$z' = a_3x + b_3y + c_3z,$$

возьмем конкретный тетраэдр с вершинами $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ и $(0, 0, 1)$. Его объем $V=\frac{1}{6}$. Преобразованный тетраэдр имеет вершины $(0, 0, 0)$, (a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3) и его объем

$$V' = \frac{1}{6}\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \frac{1}{6}\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Стало быть, искомое постоянное отношение $V' : V$ равно определителю преобразования.

Имеет геометрический смысл и знак определителя преобразования. В § 2, n° 1 и n° 4 мы установили связь между направлением обхода периметра треугольника и знаком его площади и между ориентацией тройки векторов, на которых построен параллелепипед, и знаком его объема. Из этой связи вытекает, что *аффинное преобразование плоскости с положительным определителем сохраняет направление обхода, а преобразование с отрицательным определителем изменяет направление обхода на противоположное*; *пространственное аффинное преобразование сохраняет или изменяет ориентацию тетраэдра смотря по тому, какой знак имеет определитель преобразования — плюс или минус.*

Рассмотрим теперь два последовательных аффинных преобразования

$$\begin{aligned} x' &= a_1x + b_1y + c_1z, \\ y' &= a_2x + b_2y + c_2z, \\ z' &= a_3x + b_3y + c_3z \end{aligned} \quad \text{и} \quad \begin{aligned} x'' &= \alpha_1x' + \beta_1y' + \gamma_1z', \\ y'' &= \alpha_2x' + \beta_2y' + \gamma_2z', \\ z'' &= \alpha_3x' + \beta_3y' + \gamma_3z'. \end{aligned}$$

Их произведением будет преобразование

$$\begin{aligned} x'' &= (\alpha_1a_1 + \beta_1a_2 + \gamma_1a_3)\,x + (\alpha_1b_1 + \beta_1b_2 + \gamma_1b_3)\,y + (\alpha_1c_1 + \beta_1c_2 + \gamma_1c_3)\,z, \\ y'' &= (\alpha_2a_1 + \beta_2a_2 + \gamma_2a_3)\,x + (\alpha_2b_1 + \beta_2b_2 + \gamma_2b_3)\,y + (\alpha_2c_1 + \beta_2c_2 + \gamma_2c_3)\,z, \\ z'' &= (\alpha_3a_1 + \beta_3a_2 + \gamma_3a_3)\,x + (\alpha_3b_1 + \beta_3b_2 + \gamma_3b_3)\,y + (\alpha_3c_1 + \beta_3c_2 + \gamma_3c_3)\,z. \end{aligned}$$

При переходе от x, y, z к x', y', z' объем тетраэдра умножается на

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix},$$

а при последующем переходе от x', y', z' к x'', y'', z'' он умножается на

$$\begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix}.$$

Если же совершить прямой переход от x, y, z к x'', y'', z'', то объем тетраэдра помножится на определитель

$$\begin{vmatrix} \alpha_1a_1 + \beta_1a_2 + \gamma_1a_3 & \alpha_1b_1 + \beta_1b_2 + \gamma_1b_3 & \alpha_1c_1 + \beta_1c_2 + \gamma_1c_3 \\ \alpha_2a_1 + \beta_2a_2 + \gamma_2a_3 & \alpha_2b_1 + \beta_2b_2 + \gamma_2b_3 & \alpha_2c_1 + \beta_2c_2 + \gamma_2c_3 \\ \alpha_3a_1 + \beta_3a_2 + \gamma_3a_3 & \alpha_3b_1 + \beta_3b_2 + \gamma_3b_3 & \alpha_3c_1 + \beta_3c_2 + \gamma_3c_3 \end{vmatrix}.$$

Отсюда вытекает, что

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix} =$$

$$= \begin{vmatrix} a_1\alpha_1 + a_2\beta_1 + a_3\gamma_1 & b_1\alpha_1 + b_2\beta_1 + b_3\gamma_1 & c_1\alpha_1 + c_2\beta_1 + c_3\gamma_1 \\ a_1\alpha_2 + a_2\beta_2 + a_3\gamma_2 & b_1\alpha_2 + b_2\beta_2 + b_3\gamma_2 & c_1\alpha_2 + c_2\beta_2 + c_3\gamma_2 \\ a_1\alpha_3 + a_2\beta_3 + a_3\gamma_3 & b_1\alpha_3 + b_2\beta_3 + b_3\gamma_3 & c_1\alpha_3 + c_2\beta_3 + c_3\gamma_3 \end{vmatrix}.$$

Это тождество выражает *теорему умножения определителей.* Произведение двух определителей третьего порядка равно определителю того же порядка, у которого на пересечении l-го столбца и k-й строки стоит сумма произведений элементов i-го столбца множимого на соответствующие элементы k-й строки множителя. Для краткости говорят, что определитель-произведение получается «умножением» столбцов первого определителя на строки второго. Так как строки определителя можно заменить столбцами, а столбцы — строками, то произведение двух определителей можно также составить, умножая строки на столбцы, столбцы на столбцы или строки на строки. Аналогичное правило справедливо для произведения двух определителей второго порядка [и вообще двух определителей n-го порядка]. Например,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} = \begin{vmatrix} a_1\alpha_1 + b_1\beta_1 & a_1\alpha_2 + b_1\beta_2 \\ a_2\alpha_1 + b_2\beta_1 & a_2\alpha_2 + b_2\beta_2 \end{vmatrix}$$

(умножение строк на строки).

Упражнения

1. Какие условия должны быть выполнены, чтобы аффинное преобразование

$$x' = ax + by, \qquad y' = a_1x + b_1y$$

оставляло неизменным расстояние между любыми двумя точками?

2. Доказать, что при аффинном преобразовании поверхность второго порядка

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Kz + L = 0$$

имеет своим изображением тоже поверхность второго порядка.

3*. Доказать, что аффинное преобразование

$$x' = a_1x + b_1y + c_1z,$$
$$y' = a_2x + b_2y + c_2z,$$
$$z' = a_3x + b_3y + c_3z$$

оставляет неизменным по крайней мере одно направление.

4. Вывести формулы преобразования для такого поворота на угол φ вокруг оси $\frac{x}{1} = \frac{y}{0} = \frac{z}{-1}$, что из точки $(-1, 0, 1)$ вращение в плоскости $x = z$ кажется происходящим в положительном направлении (т. е. в правой системе координат — против часовой стрелки).

5. Доказать, что при аффинном преобразовании центр массы системы материальных точек преобразуется в центр массы изображающих точек.

6. Пусть $\alpha_1, \ldots, \gamma_3$ обозначают девять направляющих косинусов (см. § 1, п° 4), определяющих преобразование одной прямоугольной системы координат в другую. Показать, что определитель

$$\Delta = \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix} = \pm 1,$$

причем $\Delta = +1$, если обе системы координат имеют одинаковую ориентацию, и $\Delta = -1$, если их ориентации различны.

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ I

1. Два вектора $\boldsymbol{a}$, $\boldsymbol{b}$ (или три вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$) называются *линейно независимыми*, если линейное соотношение

$$\alpha\boldsymbol{a} + \beta\boldsymbol{b} = 0 \qquad (\text{или} \quad \alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c} = 0)$$

возможно только при $\alpha = \beta = 0$ (или $\alpha = \beta = \gamma = 0$). Эти векторы называются *линейно зависимыми*, если существует линейное соотношение такого вида, в котором по крайней мере один из коэффициентов α, β (или α, β, γ) не равен нулю. Доказать следующие предложения:

а) Три не равных нулю вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, из которых каждые два взаимно перпендикулярны, являются линейно независимыми.

б) Два вектора $\boldsymbol{a}$, $\boldsymbol{b}$ линейно независимы в том и только в том случае, если $[\boldsymbol{ab}] \neq 0$. Три вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ линейно независимы в том и только в том случае, если

$$\boldsymbol{a}\,[\boldsymbol{bc}] = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \neq 0.$$

в) Если два вектора $\boldsymbol{a}$, $\boldsymbol{b}$ плоскости линейно независимы, то любой вектор $\boldsymbol{v}$ этой плоскости можно представить в виде $\boldsymbol{v} = \alpha\boldsymbol{a} + \beta\boldsymbol{b}$. Аналогично, если три вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ линейно независимы, то любой вектор $\boldsymbol{v}$ в пространстве можно представить в виде $\boldsymbol{v} = \alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c}$.

2. Мы уже знаем, что для любых трех векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ их векторно-скалярное произведение

$$\boldsymbol{abc} = [\boldsymbol{ab}]\,\boldsymbol{c} = \boldsymbol{a}\,[\boldsymbol{bc}] = \boldsymbol{b}\,[\boldsymbol{ca}] = \boldsymbol{c}\,[\boldsymbol{ab}] = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

(Общее скалярное значение этих выражений удобно обозначать $\boldsymbol{abc}$ и называть *смешанным произведением* векторов $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$.)

Доказать следующие векторные тождества:

а) $$(\boldsymbol{abc})\,(\boldsymbol{xyz}) = \begin{vmatrix} \boldsymbol{ax} & \boldsymbol{ay} & \boldsymbol{az} \\ \boldsymbol{bx} & \boldsymbol{by} & \boldsymbol{bz} \\ \boldsymbol{cx} & \boldsymbol{cy} & \boldsymbol{cz} \end{vmatrix}.$$

б) $$[\boldsymbol{a}\,[\boldsymbol{bc}]] = \boldsymbol{b}\,(\boldsymbol{ac}) - \boldsymbol{c}\,(\boldsymbol{ab}); \qquad [[\boldsymbol{ab}]\,\boldsymbol{c}] = \boldsymbol{b}\,(\boldsymbol{ac}) - \boldsymbol{a}\,(\boldsymbol{bc}).$$

[Словесно: двойное векторное произведение трех векторов равно среднему по записи вектору, помноженному на скалярное произведение двух крайних, минус другой вектор внутренней скобки, помноженный на скалярное произведение оставшихся двух.]

в) Скалярное произведение двух векторных произведений

$$[\boldsymbol{ab}]\,[\boldsymbol{xy}] = (\boldsymbol{ax})\,(\boldsymbol{by}) - (\boldsymbol{ay})\,(\boldsymbol{bx}) = \begin{vmatrix} \boldsymbol{ax} & \boldsymbol{ay} \\ \boldsymbol{bx} & \boldsymbol{by} \end{vmatrix}.$$

Тождество Лагранжа (упр. 7, стр. 33) получается отсюда как частный случай.

г) Смешанное произведение трех двойных векторных произведений

$$[\boldsymbol{a}\,[\boldsymbol{bc}]]\,[\boldsymbol{b}\,[\boldsymbol{ca}]]\,[\boldsymbol{c}\,[\boldsymbol{ab}]] = 0.$$

Из последнего тождества вывести такое следствие: если через каждую из трех прямых, имеющих общую точку, провести плоскость, перпендикулярную к плоскости двух других прямых, то полученные три плоскости проходят через одну прямую.

3. На плоскости даны две системы прямоугольных осей координат (одинаковой ориентации) Ox, Oy и Ox', Oy'; при этом угол $\widehat{xOx'} = \alpha$. Вывести средствами векторной алгебры формулы перехода от одной системы координат к другой:

$$x = x'\cos\alpha - y'\sin\alpha, \qquad x' = x\cos\alpha + y\sin\alpha,$$
$$y = x'\sin\alpha + y'\cos\alpha, \qquad y' = -x\sin\alpha + y\cos\alpha.$$

4. Из формул поворота осей координат, полученных в упр. 3, вывести формулы сложения:

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta, \qquad \sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta.$$

5. Даны две прямоугольные координатные системы Ox, Oy, Oz и Ox', Oy', Oz'; косинусы углов между осями обеих систем даны в следующей таблице:

	x	y	z
x'	α_1	β_1	γ_1
y'	α_2	β_2	γ_2
z'	α_3	β_3	γ_3.

В упр. 1, стр. 26, были доказаны следующие соотношения:

$$\alpha_1^2+\beta_1^2+\gamma_1^2=1, \qquad \alpha_1\alpha_2+\beta_1\beta_2+\gamma_1\gamma_2=0,$$
$$\alpha_2^2+\beta_2^2+\gamma_2^2=1, \qquad \alpha_2\alpha_3+\beta_2\beta_3+\gamma_2\gamma_3=0,$$
$$\alpha_3^2+\beta_3^2+\gamma_3^2=1, \qquad \alpha_3\alpha_1+\beta_3\beta_1+\gamma_3\gamma_1=0.$$

Определитель третьего порядка

$$\Delta=\begin{vmatrix}\alpha_1 & \beta_1 & \gamma_1\\ \alpha_2 & \beta_2 & \gamma_2\\ \alpha_3 & \beta_3 & \gamma_3\end{vmatrix},$$

элементы которого удовлетворяют этим соотношениям, называется *ортогональным*, а соответствующее преобразование называется *ортогональным преобразованием*. В упр. 6, стр. 50, было доказано, что ортогональный определитель равен ± 1 (плюс — если обе системы координат имеют одинаковую ориентацию, минус — если их ориентация различна).

Доказать: а) что всякому ортогональному определителю Δ, $\Delta=+1$, соответствуют две системы координат Ox, Oy, Oz и Ox', Oy', Oz' одинаковой ориентации, для которых косинусы углов между осями координат обеих систем равны элементам определителя Δ;

б) что для всякого ортогонального определителя Δ выполняются также и следующие соотношения:

$$\alpha_1^2+\alpha_2^2+\alpha_3^2=1, \qquad \alpha_1\beta_1+\alpha_2\beta_2+\alpha_3\beta_3=0,$$
$$\beta_1^2+\beta_2^2+\beta_3^2=1, \qquad \beta_1\gamma_1+\beta_2\gamma_2+\beta_3\gamma_3=0,$$
$$\gamma_1^2+\gamma_2^2+\gamma_3^2=1, \qquad \gamma_1\alpha_1+\gamma_2\alpha_2+\gamma_3\alpha_3=0.$$

6*. Рассмотрим две системы координат Ox, Oy, Oz и Ox', Oy', Oz' одинаковой ориентации. Предполагаем, что оси Oz и Oz' не совпадают; пусть угол $\widehat{zOz'}=\theta$ $(0<\theta<\pi)$. Проведем полупрямую Ox_1 перпендикулярно как к оси Oz, так и к оси Oz' в такую сторону (из двух возможных), чтобы система Ox_1, Oz, Oz' имела ту же ориентацию, что и система Ox, Oy, Oz. Тогда Ox_1 идет по линии пересечения плоскостей Oxy и $Ox'y'$. Обозначим угол xOx_1 через φ, а угол x_1Ox' через ψ, отсчитывая каждый из этих углов в принятом в его плоскости, Oxy или $Ox'y'$, положительном направлении, так что $0\leqslant\varphi<2\pi$ и $0\leqslant\psi<2\pi$.

Доказать, что коэффициенты формул перехода от системы Ox, Oy, Oz к системе Ox', Oy', Oz' даются следующей схемой:

	x'	y'	z'
x	$\cos\varphi\cos\psi-$ $-\sin\varphi\sin\psi\cos\theta$	$-\cos\varphi\sin\psi-$ $-\sin\varphi\cos\psi\cos\theta$	$\sin\varphi\sin\theta$
y	$\sin\varphi\cos\psi+$ $+\cos\varphi\sin\psi\cos\theta$	$-\sin\varphi\sin\psi+$ $+\cos\varphi\cos\psi\cos\theta$	$-\cos\varphi\sin\theta$
z	$\sin\psi\sin\theta$	$\cos\psi\sin\theta$	$\cos\theta$

(Заметим, что этот результат остается в силе и при $\theta=0$, а также при $\theta=\pi$, когда φ и ψ становятся неопределенными, причем соответственно $\varphi+\psi=\angle xOx'$ или $\varphi-\psi=\angle xOx'$. Углы φ, ψ, θ — это так называемые *эйлеровы углы*, и этот результат вместе с упр. 5 показывает, что самое общее ортогональное преобразование с определителем $\Delta=+1$ можно выразить «параметрически» через *три* независимых переменных параметра φ, ψ, θ, удовлетворяющих неравенствам

$$0 \leqslant \theta \leqslant \pi, \qquad 0 \leqslant \varphi < 2\pi, \qquad 0 \leqslant \psi < 2\pi.)$$

7. Дан сферический треугольник ABC со сторонами a, b, c и углами A, B, C на «единичной сфере» (т. е. на шаровой поверхности радиуса 1). Из результата упр. 6 вывести «теорему косинусов» сферической тригонометрии

$$\cos a = \cos b \cos c + \sin b \sin c \cos A.$$

8. Найти угол φ между плоскостью $Ax + By + Cz + D = 0$ и прямой

$$x = x_0 + b_1 t, \qquad y = y_0 + b_2 t, \qquad z = z_0 + b_3 t.$$

9. Доказать тождество

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (bc - ad)^2$$

путем вычисления произведения определителей

$$\begin{vmatrix} a & b \\ -b & a \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} c & d \\ -d & c \end{vmatrix}.$$

10*. Доказать, что значение определителя

$$D = \begin{vmatrix} \cos(\theta + \alpha) & \cos(\theta + \beta) & \cos(\theta + \gamma) \\ \sin(\theta + \alpha) & \sin(\theta + \beta) & \sin(\theta + \gamma) \\ \sin(\beta - \gamma) & \sin(\gamma - \alpha) & \sin(\alpha - \beta) \end{vmatrix}$$

не зависит от θ.

11. Дано: $A = x^2 + y^2 + z^2$, $B = xy + yz + zx$. Показать, что

$$D = \begin{vmatrix} B & A & B \\ B & B & A \\ A & B & B \end{vmatrix} = (x^3 + y^3 + z^3 - 3xyz)^2.$$

12. Показать, что

$$\Delta = \begin{vmatrix} t_1 + x & a + x & a + x & a + x \\ b + x & t_2 + x & a + x & a + x \\ b + x & b + x & t_3 + x & a + x \\ b + x & b + x & b + x & t_4 + x \end{vmatrix}$$

есть выражение вида $A + Bx$, где A и B не зависят от x. Подставляя вместо x частные значения, доказать затем, что

$$A = \frac{af(b) - bf(a)}{a - b}, \qquad B = \frac{f(b) - f(a)}{a - b},$$

где $f(t) = (t_1 - t)(t_2 - t)(t_3 - t)(t_4 - t)$.

13*. Даны функции $u(x)$ и $v(x)$, причем $u = \frac{1}{v}$. Доказать, что $v''' = \frac{D}{u^4}$, где D есть определитель

$$D = \begin{vmatrix} u''' & 3u'' & 3u' \\ u'' & 2u' & u \\ u' & u & 0 \end{vmatrix}.$$

ГЛАВА II

ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ И ИХ ПРОИЗВОДНЫЕ

С функциями многих переменных мы уже познакомились в главе X первого тома и узнали там достаточно, чтобы оценить их важность и пользу. Теперь мы займемся более тщательным изучением этих функций, исследуем такие их свойства, которые не были затронуты в предыдущем томе, и докажем те теоремы, которые там получили только наглядное оправдание. Но в доказательствах этого тома мы не будем опираться на предварительное знание какого-либо доказательства, развитого в главе X первого тома. Тем не менее читателю рекомендуется прочитать ту главу, так как проведенное там наглядное рассмотрение поможет ему составить себе ясное представление о вещах, пожалуй, несколько абстрактных.

Как правило, теорема, доказанная для функций двух независимых переменных, может быть распространена на функции трех и большего числа переменных без существенного изменения хода рассуждения. Поэтому мы в дальнейшем ограничимся рассмотрением функций двух переменных, а функции большего числа переменных мы будем рассматривать лишь в тех случаях, когда существо вопроса потребует для них особого исследования.

§ 1. Понятие функции многих переменных

1. Функция и область ее задания. Записывая уравнения вида

$$u=x+y, \qquad u=x^2y^3 \text{ или } u=\ln(1-x^2-y^2),$$

мы приводим в соответствие *паре значений* (x, y) значение *функции* u. В первых двух из этих примеров значение функции u присваивается *любой* паре значений (x, y), а в третьем примере закон соответствия имеет смысл лишь для таких систем значений (x, y), которые удовлетворяют неравенству $x^2+y^2<1$.

Во всех этих случаях говорят, что u есть *функция независимых переменных* или *аргументов* x *и* y. Это выражение всегда применяют, когда какой-либо закон соответствия относит каждой паре значений (x, y) из некоторого заданного множества пар чисел определенное значение u в качестве *зависимой* переменной. Аналогично, если каждому набору значений $(x_1, x_2, \ldots, x_n)$ из некоторого заданного множества соответствует определенное значение величины u, то говорят, что u есть функция от n переменных $x_1, x_2, \ldots, x_n$.

Так, например, объем $u = xyz$ прямоугольного параллелепипеда есть функция длин трех его ребер x, y, z; магнитное склонение есть функция долготы, широты и времени; сумма $u = x_1 + x_2 + \ldots \ldots + x_n$ является функцией своих n слагаемых $x_1, x_2, \ldots, x_n$.

В случае функции двух аргументов $u = f(x, y)$ пару значений (x, y) изображают *точкой P* плоскости, имеющей координаты x и y относительно некоторой прямоугольной системы координат; мы будем иногда называть эту величину u *функцией точки P*, считая точку P аргументом функции. У функций $u = x + y$ и $u = x^2y^3$ эта точка-аргумент может пробегать всю плоскость xy; поэтому говорят, что эти функции определены на всей плоскости xy. У функции же

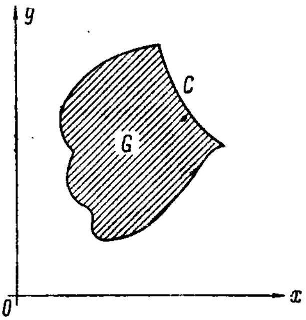


Рис. 18.

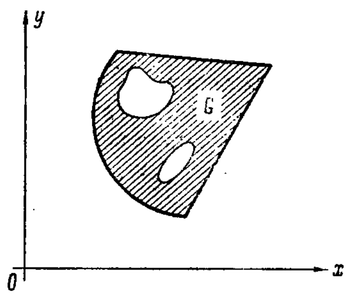


Рис. 19.

$u = \ln(1 - x^2 - y^2)$ точка-аргумент должна оставаться внутри окружности $x^2 + y^2 = 1$, где выполняется условие $x^2 + y^2 < 1$; в этом случае функция задана только во внутренней области круга.

Подобно тому, как это бывает у функции одной переменной, аргументы функции нескольких переменных могут изменяться либо *дискретно*, либо *непрерывно*. Так, например, средняя численность населения района города Москвы зависит от количества районов, т. е. от способа районирования и от общей численности населения города; оба эти аргумента — целые числа. С другой стороны, время, длины, веса и т. п. могут служить непрерывными аргументами. Мы будем иметь дело почти исключительно с непрерывно изменяющимися аргументами: например, точке (x, y) будет дозволено изменяться в определенной *области* плоскости xOy; здесь область играет ту же роль, что *интервал* у функции одного аргумента. Такой областью G двух независимых переменных может быть вся плоскость xy; область G может состоять из куска плоскости, ограниченного одной замкнутой кривой C, не пересекающей сама себя (*односвязная* область, рис. 18); наконец, область G может быть ограничена несколькими замкнутыми кривыми. В последнем случае она называется *многосвязной* областью, причем число граничных замкнутых кривых называется *степенью связности* или просто *связностью* области; на рис. 19 показана для примера *трехсвязная область*.

Граничные кривые, и вообще все рассматриваемые в этом томе кривые, мы будем предполагать *кусочно гладкими*. Это значит, что мы раз навсегда предполагаем, что всякая такая кривая состоит из конечного числа дуг, каждая из которых имеет в любой своей точке касательную, непрерывно изменяющуюся от одного конца дуги до другого ее конца. Стало быть, такие кривые могут иметь лишь конечное число угловых точек и точек заострения.

При изучении функций двух переменных чаще всего встречаются области следующих двух типов:

1) *Прямоугольная область* (рис. 20), характеризуемая двумя неравенствами вида

$$a \leqslant x \leqslant b, \qquad c \leqslant y \leqslant d;$$

в такой области каждая независимая переменная ограничена определенным интервалом, а точка-аргумент (x, y) пробегает прямоугольник со сторонами, параллельными осям координат.

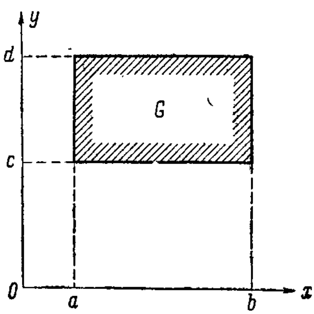


Рис. 20.

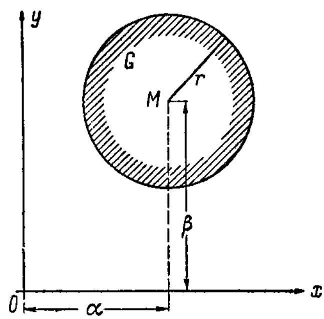


Рис. 21.

2) *Круговая область* (рис. 21), определяемая неравенством вида

$$(x-\alpha)^2+(y-\beta)^2 \leqslant r^2,$$

в которой точка-аргумент (x, y) пробегает круг радиуса r с центром (α, β).

Точка P, принадлежащая области G, называется *внутренней точкой* этой области, если существует круг с центром в P, целиком лежащий в G. Если P есть внутренняя точка области G, то говорят также, что G является *окрестностью* точки P. Таким образом, любая окрестность точки P содержит достаточно малый круг с центром в этой точке.

Теперь вкратце отметим, что у функций трех независимых переменных x, y, z положение дел точно такое же, только с тем отличием, что точка-аргумент непрерывно изменяется не в плоской, а в пространственной области. Эта область может быть, в частности, *прямоугольной областью*, определяемой тремя неравенствами вида

$$a_1 \leqslant x \leqslant a_2, \quad b_1 \leqslant y \leqslant b_2, \quad c_1 \leqslant z \leqslant c_2,$$

или *сферической* (*шаровой*) *областью*, характеризуемой одним неравенством вида

$$(x-\alpha)^2+(y-\beta)^2+(z-\gamma)^2 \leqslant r^2.$$

Отметим еще некоторую более тонкую детализацию понятия области, едва ли существенную для целей нашей книги, но имеющую важное значение при более глубоких исследованиях. Часто приходится рассматривать область, не содержащую своих *граничных точек* (или своей *границы*), т. е. точек тех кривых, которые эту область ограничивают. Такая область называется *открытой* или *незамкнутой* (ср. Дополнение к этой главе, § 1, n° 2). Так, например, область $x^2+y^2<1$ ограничена окружностью $x^2+y^2=1$, но не содержит этой окружности; стало быть, это — открытая область. Если же, напротив, граничные точки причисляют к области, как мы это чаще всего будем делать, то область называется *замкнутой*.

Когда число n независимых переменных больше трех, то наглядное представление уже не дает геометрического истолкования системы независимых переменных $x_1, x_2, \ldots, x_n$. Однако и в этом случае целесообразно пользоваться геометрической терминологией и называть набор n чисел *точкой n-мерного пространства*. В таком пространстве естественно называть *прямоугольной областью* множество точек, координаты которых удовлетворяют n неравенствам вида

$$a_1 \leqslant x_1 \leqslant b_1, \quad a_2 \leqslant x_2 \leqslant b_2, \ \ldots, \ a_n \leqslant x_n \leqslant b_n,$$

а шаровой областью — множество точек, удовлетворяющих неравенству вида

$$(x_1-\alpha_1)^2+(x_2-\alpha_2)^2+\ldots+(x_n-\alpha_n)^2 \leqslant r^2.$$

Теперь мы можем уточнить определение общего понятия функции следующим образом: *если задана область G, в которой независимые переменные $x_1, x_2, \ldots, x_n$ могут как угодно изменяться, и если каждой точке $(x_1, x_2, \ldots, x_n)$ этой области приводится в соответствие, с помощью какого-либо закона, определенное значение u, то величина u называется в этой области функцией непрерывных независимых переменных $x_1, x_2, \ldots, x_n$.*

Необходимо иметь в виду, что, как и в случае функции одной переменной, функциональное соответствие относит системе значений независимых переменных *единственное* значение функции u. Поэтому если функция определена с помощью *многозначного* аналитического выражения, например, $u=\operatorname{Arctg}\frac{y}{x}$, то функция таким выражением еще не вполне определена: необходимо еще дополнительно уточнить, какое из всех возможных значений этого выражения имеется в виду в определении функции; так, в приведенном примере требуется еще указать, надо ли брать то значение арктангенса, которое лежит между $-\frac{\pi}{2}$ и $\frac{\pi}{2}$, или значение между 0 и π или же дать какое-либо другое подобное уточнение. Всякое такое уточнение дает однозначную функцию: например, если выбрать значение арктангенса, лежащее между $-\frac{\pi}{2}$ и $\frac{\pi}{2}$, то этим выбором определяется однозначная функция, обозначаемая символом $u = \operatorname{arctg}\frac{y}{x}$ (с малой буквы).

В тех случаях, когда функция определена с помощью многозначного аналитического выражения, говорят, что это выражение определяет несколько различных однозначных *ветвей* функции (ср. т. I, стр. 32). Если желательно рассматривать одновременно все эти ветви, не отдавая предпочтения ни одной из них, то их можно рассматривать как ветви, образующие в совокупности *многозначную функцию*. В этой книге мы воспользуемся понятием многозначной функции только в главе VIII.

2. Простейшие типы функций. Как и у функций одной переменной, простейшими являются *целые рациональные* функции или *многочлены*. Самый общий многочлен первой степени, зависящий от двух переменных, имеет такой вид:

$$u = ax + by + c,$$

где a, b и c — постоянные (целая линейная функция). Общий вид многочлена второй степени

$$u = Ax^2 + Bxy + Cy^2 + Dx + Ey + F.$$

Вообще всякий многочлен любой степени есть сумма членов вида $a_{mn}x^m y^n$ с произвольными постоянными a_{mn}.

Частное двух многочленов называется *дробной рациональной* функцией. Простейшей из дробных рациональных функций является *дробно-линейная* функция

$$u = \frac{ax + by + c}{a'x + b'y + c'}.$$

Присоединяя к рациональным действиям извлечение корня, приходим от рациональных функций к известным *алгебраическим* функциям, например,

$$u = \sqrt{\frac{x-y}{x+y}} + \sqrt[3]{\frac{(x+y)^2}{x^3 + xy}}.$$

Точное определение термина «алгебраическая функция» будет дано на стр. 141.

Для построения более сложных функций от нескольких переменных почти всегда прибегают к хорошо известным функциям одной переменной, например,

$$u = \sin(x \arccos y) \quad \text{или} \quad u = \lg_x y.$$

(Ср. также § 5 о сложных функциях.)

3. Геометрическое изображение функций. В главе X первого тома мы рассмотрели два основных метода геометрического изображения функции двух независимых переменных: 1) поверхностью $u=f(x, y)$ в пространстве x, y, u, описываемой точкой (x, y, u), когда точка (x, y) пробегает область определения функции u, и 2) с помощью линий уровня в плоскости xOy, вдоль которых функция u сохраняет постоянное значение k. Мы не станем здесь повторять это рассмотрение. Если читатель еще не овладел полностью этими способами геометрического изображения функции, то ему рекомендуется прочитать изложение этой темы в главе X первого тома, § 1, n° 3.

§ 2. Непрерывность

1. Определение. Читатель, знакомый с функциями одного аргумента и видевший, какую важную роль играет у них понятие непрерывности, естественно ожидает, что соответствующее понятие будет занимать выдающееся место и в теории функций большего числа переменных. Более того, он наперед догадается, каков смысл утверждения, что функция $u=f(x, y)$ непрерывна в точке (ξ, η). Это выражение, грубо говоря, означает, что *для всех точек* (x, y), *близких к точке* (ξ, η), *значение* $f(x, y)$ *мало отличается от* $f(\xi, \eta)$, *и притом сколь угодно мало, если только точка* (x, y) *лежит достаточно близко от* (ξ, η). Выразим теперь эту мысль более точно:

Функция $f(x, y)$, определенная в области G, называется непрерывной в точке (ξ, η) этой области, если для всякого положительного числа ε возможно найти такое положительное число $\delta=\delta(\varepsilon)$ (вообще зависящее от ε и стремящееся к нулю вместе с ε), что для всех точек области G, расстояние которых от точки (ξ, η) не превышает δ, т. е. для которых

$$(x-\xi)^2+(y-\eta)^2 \leqslant \delta^2, \tag{1}$$

выполняется неравенство

$$|f(x, y)-f(\xi, \eta)| \leqslant \varepsilon. \tag{2}$$

Другими словами, неравенство

$$|f(\xi+h, \eta+k)-f(\xi, \eta)| \leqslant \varepsilon$$

выполняется для всякой пары значений (h, k), для которой $h^2+k^2 \leqslant \delta^2$, а точка $(\xi+h, \eta+k)$ принадлежит области G.

Примечание к условию (1). Если (ξ, η) — внутренняя точка области G, то точки (x, y) при достаточно малом δ заполняют весь круг $(x-\xi)^2+(y-\eta)^2 \leqslant \delta^2$, если же точка (ξ, η) — граничная, то точки (x, y) заполняют лишь ту часть этого круга, которая принадлежит области G. Рис. 22 иллюстрирует этот второй случай.

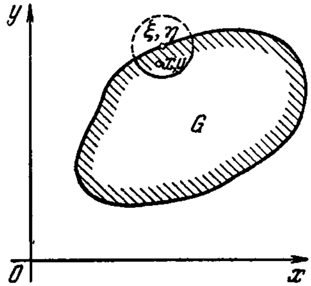


Рис. 22.

Если функция непрерывна в любой точке области G, то говорят, что она *непрерывна в области* G.

Условие $h^2+k^2 \leqslant \delta^2$, налагаемое в определении непрерывности на расстояние между точками (ξ, η) и (x, y), можно заменить следующим равносильным условием.

Для всякого сколь угодно малого $\varepsilon > 0$ должны существовать два таких положительных числа δ_1 и δ_2, что

$$|f(\xi+h, \eta+k)-f(\xi, \eta)| \leqslant \varepsilon,$$

коль скоро $|h| \leqslant \delta_1$ и $|k| \leqslant \delta_2$.

Оба условия действительно равносильны, ибо если первое условие выполнено, то выполняется и второе при $\delta_1=\delta_2=\frac{\delta}{\sqrt{2}}$, и, обратно, если выполнено второе условие, то выполняется и первое: надо только выбрать в качестве δ меньшее из двух чисел δ_1 и δ_2.

Читатель сам легко докажет следующие почти очевидные теоремы:

Сумма, разность и произведение непрерывных функций тоже непрерывны. Частное двух непрерывных функций непрерывно всюду, где знаменатель не обращается в нуль.

Менее очевидна следующая теорема; ее доказательство будет дано в § 5, n° 1 этой главы (стр. 85—86).

Непрерывная функция от непрерывной функции тоже непрерывна.

Как непосредственное следствие этих теорем имеем:

Все целые рациональные функции всюду непрерывны, а все дробные рациональные функции непрерывны всюду, где знаменатель не обращается в нуль.

Полезно отметить еще один почти очевидный факт. *Если функция $f(x, y)$ непрерывна в области G и отлична от нуля во внутренней точке P этой области, то можно окружить точку P такой окрестностью, принадлежащей полностью области G (например, достаточно малым кругом с центром P), в которой $f(x, y)$ нигде не обращается в нуль.* Действительно, если в точке P значение функции равно a, то, в силу непрерывности функции, возможно окружить P таким малым кругом, в котором значения функции отличаются от a меньше чем на $\frac{a}{2}$ и, стало быть, отличны от нуля.

Читателю необходимо познакомиться в этом месте с примерами точек разрыва, рассмотренными в т. I, гл. X, § 2, n° 2 (стр. 560). Особое внимание обратить на функцию $u=\frac{2xy}{x^2+y^2}$, имеющую своеобразный разрыв в точке (0, 0); этот пример понадобится на стр. 68.

Любопытно, что функция

$$u(x, y)=xy\frac{x^2-y^2}{x^2+y^2}, \qquad u(0, 0)=0,$$

напротив, непрерывна в начале координат. В самом деле,

$$|u(h, k)-u(0, 0)|=|hk|\frac{|h^2-k^2|}{h^2+k^2}\leqslant|hk|.$$

Стало быть, в определении непрерывности можно положить $\delta_1=\delta_2=\sqrt{\varepsilon}$. И с этой функцией мы еще встретимся ниже (стр. 72).

2. Понятие предела функции нескольких переменных. Понятие предела функции двух переменных тесно связано с понятием непрерывности. Предположим, что функция $f(x, y)$ определена в области G, и пусть точка (ξ, η) лежит либо внутри G, либо на ее границе. Тогда утверждение, что число l является пределом функции $f(x, y)$, когда x стремится к ξ и y стремится к η, имеет следующий смысл:

Для всякого $\varepsilon>0$ *существует такое число* $\delta>0$, *что*

$$|f(x, y)-l|<\varepsilon$$

для всех точек (x, y) *области* G, *для которых выполняется неравенство*

$$0<(x-\xi)^2+(y-\eta)^2\leqslant\delta^2.$$

Как и у функций одного аргумента, следует подчеркнуть, что переменная точка (x, y) должна быть отлична от точки (ξ, η).

Существование у функции предела l символически записывают так:

$$\lim_{\substack{x\to\xi\\ y\to\eta}} f(x, y)=l \quad \text{или} \quad f(x, y)\to l \quad \text{при} \quad (x, y)\to(\xi, \eta).$$

Для того чтобы подчеркнуть, что предел берется при одновременном стремлении $x\to\xi$ и $y\to\eta$, эту запись порой читают так: двойной предел функции $f(x, y)$, когда $x\to\xi$ и $y\to\eta$, равен l.

Наряду с выражением «функция имеет предел l» говорят также: «функция стремится к l» или «функция сходится к l».

Пользуясь языком пределов, можно дать другую формулировку понятию *непрерывности*:

Функция $f(x, y)$ *непрерывна в точке* (ξ, η) *в том и только в том случае, если*

$$\lim_{\substack{x\to\xi\\ y\to\eta}} f(x, y)=f(\xi, \eta).$$

Отсюда видно, что *определение непрерывности* содержит следующие два требования:

1) функция $f(x, y)$ должна стремиться к пределу l, когда переменная точка (x, y) стремится в области G к точке (ξ, η), и

2) этот предел l должен совпадать со значением функции в точке (ξ, η).

Понятие предела получает новое освещение, если ввести в рассмотрение *последовательности точек.* Говорят, что последовательность точек $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n), \ldots$ стремится или сходится к точке (ξ, η), если расстояние $\sqrt{(x_n-\xi)^2+(y_n-\eta)^2}$ стремится к нулю при безграничном возрастании n. Нетрудно показать (ср. т. I, стр. 69), что если $f(x, y) \to l$, когда $(x, y) \to (\xi, \eta)$, то $\lim\limits_{n\to\infty} f(x_n, y_n) = l$ для всякой последовательности точек области G, которая стремится к точке (ξ, η). Верна и обратная теорема: если $\lim\limits_{n\to\infty} f(x_n, y_n)$ существует и равен одному и тому же числу l для всякой последовательности точек (x_n, y_n) области G, стремящейся к точке (ξ, η), то двойной предел функции $f(x, y)$ при $x \to \xi$ и $y \to \eta$ существует и равен l. Доказательство опускаем.

В нашем определении предела мы давали точке (x, y) изменяться в области G. Однако при желании можно наложить ограничения на изменение точки (x, y). Можно, например, потребовать, чтобы она лежала в некоторой подобласти G' области G, или на кривой C, или на каком-либо множестве M точек области G. В таком случае говорят, что $f(x, y)$ стремится к l, когда точка (x, y) стремится к точке (ξ, η), изменяясь в подобласти G' (или на кривой C, или на множестве M). Для того чтобы это определение было состоятельно, подобласть G' (либо кривая C, либо множество M) должны содержать точки, сколь угодно близкие к точке (ξ, η).

Очевидно, и непрерывность функции можно определить тем же путем не только в области G, но также, например, вдоль кривой C.

3. Порядок малости функции. Пусть функция $f(x, y)$ непрерывна в точке (ξ, η). Это значит, что разность $f(x, y) - f(\xi, \eta)$ стремится к нулю, когда точка (x, y) стремится к точке (ξ, η). Введя в качестве новых переменных приращения координат $h = x - \xi$, $k = y - \eta$, можно ту же мысль выразить следующим образом: функция $\varphi(h, k) = f(\xi + h, \eta + k) - f(\xi, \eta)$, зависящая от переменных h и k, стремится к пределу 0, когда h и k стремятся одновременно к нулю. Такие функции $\varphi(h, k)$, которые стремятся к нулю вместе со своими аргументами h и k, часто встречаются [1]). Как и у функций одного

---

[1]) В литературе пользуются также выражением: $\varphi(h, k)$ становится *бесконечно малой* одновременно с h и k; этот оборот имеет совершенно точный смысл, а именно утверждение, что $\varphi(h, k)$ стремится вместе с h и k к нулю. Однако это выражение может привести к недоразумениям, и мы предпочитаем его избегать.

аргумента, здесь тоже полезно для многих целей описать более точно характер стремления к нулю функции $\varphi(h, k)$ при $h \to 0$ и $k \to 0$; для этого детализируют различные *порядки малости*, т. е. различные порядки стремления к нулю функции $\varphi(h, k)$ [говорят также: различные порядки обращения в нуль или различные порядки убывания]. Для этой цели пользуются в качестве масштаба для сравнения расстоянием

$$\rho = \sqrt{h^2 + k^2} = \sqrt{(x-\xi)^2 + (y-\eta)^2}$$

от точки $P(\xi, \eta)$ до точки Q с координатами $x = \xi + h$, $y = \eta + k$ (т. е. модулем вектора $\overline{PQ}$) и вводят следующее определение:

Функция $\varphi(h, k)$ имеет при $\rho \to 0$ тот же порядок малости или, точнее, по меньшей мере тот же порядок малости, что и $\rho = \sqrt{h^2 + k^2}$, если существует такая положительная постоянная C (не зависящая от h и k), что

$$\left| \frac{\varphi(h, k)}{\rho} \right| \leqslant C$$

для всех достаточно малых значений ρ; т. е. если можно указать такое число $\delta > 0$, что последнее неравенство выполняется при всех значениях h и k, для которых $0 < \sqrt{h^2 + k^2} < \delta$.

И далее: *функция $\varphi(h, k)$ имеет более высокий порядок малости, чем ρ, если частное $\dfrac{\varphi(h, k)}{\rho}$ стремится к нулю при $\rho \to 0$.*

Во избежание недоразумений подчеркнем, что более высокий порядок малости при $\rho \to 0$ означает более быстрое стремление к нулю, т. е. меньшую абсолютную величину значений функции в окрестности точки $\rho = 0$. Например, ρ^2 имеет более высокий порядок убывания, чем ρ, и вместе с тем $\rho^2 < \rho$ [не пишем символа абсолютной величины, так как $\rho^2 > 0$].

Утверждение, что $\varphi(h, k)$ имеет более высокий порядок малости, чем ρ, записывают символически так:

$$\varphi(h, k) = o(\rho).$$

Буква o выбрана потому, что это первая буква латинского слова ordo, означающего *порядок* (английского order, французского ordre, немецкого Ordnung). Для записи утверждения, что $\varphi(h, k)$ по меньшей мере того же порядка убывания, что и ρ (а не обязательно более высокого порядка), пользуются не малой буквой o, а большой буквой O:

$$\varphi(h, k) = O(\rho).$$

Впрочем, в этой книге мы последним символом не будем пользоваться.

Рассмотрим несколько примеров.

Координаты h и k вектора $\overline{PQ}$, идущего от точки $P(\xi, \eta)$ к точке $Q(\xi + h, \eta + k)$, имеют по меньшей мере тот же порядок малости, что его модуль, т. е. расстояние ρ, так как

$$\frac{|h|}{\sqrt{h^2 + k^2}} \leqslant 1 \quad \text{и} \quad \frac{|k|}{\sqrt{h^2 + k^2}} \leqslant 1.$$

То же утверждение справедливо для линейной однородной функции $ah+bk$ с постоянными коэффициентами a и b, а также для функции $\rho\sin\frac{1}{\rho}$.

Степень ρ^α расстояния ρ имеет при постоянном значении $\alpha>1$ более высокий порядок малости, чем ρ, т. е. $\rho^\alpha=o(\rho)$ при $\alpha>1$. Однородный многочлен второй степени $ah^2+bhk+ck^2$ от h и k имеет при $\rho\to 0$ порядок малости более высокий, чем ρ, так что

$$ah^2+bhk+ck^2=o(\rho).$$

Введем еще одно, более общее, определение: пусть $\omega(h, k)$ есть некоторая функция сравнения, определенная в достаточно малом круге с центром в точке $P(\xi, \eta)$ для всех значений h и k, не обращающихся одновременно в нуль, и отличная от нуля при всех этих значениях h и k. Тогда при $\rho\to 0$ *порядок малости функции* $\varphi(h, k)$ *не ниже порядка малости* $\omega(h, k)$, если существует такое постоянное C, что

$$\left|\frac{\varphi(h, k)}{\omega(h, k)}\right|\leqslant C;$$

этот факт символически записывают так: $\varphi(h, k)=O[\omega(h, k)]$. Аналогично $\varphi(h, k)$ имеет порядок малости более высокий, чем $\omega(h, k)$ или $\varphi(h, k)=o[\omega(h, k)]$, если $\frac{\varphi(h, k)}{\omega(h, k)}\to 0$ при $\rho\to 0$.

Например, порядок малости однородного трехчлена $ah^2+bhk+ck^2$ не ниже порядка (или по меньшей мере равен порядку) малости функции $\rho^2=h^2+k^2$, так как

$$|ah^2+bhk+ck^2|\leqslant\left(|a|+\frac{1}{2}|b|+|c|\right)(h^2+k^2),$$

т. е. $ah^2+bhk+ck^2=O(h^2+k^2)$.

Другой пример:

$$\rho=o\left(\frac{1}{\ln\rho}\right),\qquad \text{ибо}\qquad \lim_{\rho\to 0}\left(\rho:\frac{1}{\ln\rho}\right)=\lim_{\rho\to 0}(\rho\ln\rho)=0.$$

Упражнения

1. Исследовать поведение функций

$$\text{а) } x^3-3xy^2;\qquad \text{б) } x^4-6x^2y^2+y^4$$

в окрестности начала координат.

2. Сколько постоянных коэффициентов содержит общий вид многочлена $P(x, y)$ степени n?

3. Доказать, что функция

$$Ax^3+Bx^2y+Cxy^2+Dy^3$$

имеет в точке $x=y=0$ по меньшей мере тот же порядок малости, что и $\rho^3=(x^2+y^2)^{3/2}$.

4. Найти условие, при выполнении которого многочлен

$$P(x, y)=ax^2+2bxy+cy^2$$

имеет в окрестности точки $x=y=0$ в точности тот же порядок малости, что и ρ^2 $\left(\text{т. е. обе дроби } \frac{P(x, y)}{\rho^2} \text{ и } \frac{\rho^2}{P(x, y)} \text{ ограничены}\right)$.

5. Нижеследующие функции не определены в точке $x=y=0$. Какие из них можно доопределить так, чтобы они стали непрерывными в этой точке, и какие нет?

$$\text{а) } \frac{x^2-y^2}{x^2+y^2}; \quad \text{б) } \frac{x^2+2xy+y^2}{x^2+y^2}; \quad \text{в) } \frac{x^2+3xy+y^2}{x^2+4xy+y^2};$$

$$\text{г) } -\frac{|x-y|}{x^2-2xy+y^2}; \quad \text{д) } e^{-\frac{|x-y|}{x^2-2xy+y^2}};$$

$$\text{е) } |x|^y; \quad \text{ж) } |x|^{\left|\frac{1}{y}\right|}; \quad \text{з)* } |y|^{|x|} \frac{\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}+\left|\frac{y}{x}\right|}.$$

6. Найти соответствующие $\delta(\varepsilon)$ (стр. 59) для тех функций из упр. 5, которые, после дополнительного их определения в точке (0, 0), стали непрерывными.

§ 3. Частные производные от функции многих переменных

1. Частные производные и их геометрический смысл. Если в функции нескольких независимых переменных дать определенные численные значения всем аргументам, кроме одного, и позволить изменяться этому единственному аргументу, скажем x, то наша функция станет функцией от одной (оставшейся) независимой переменной.

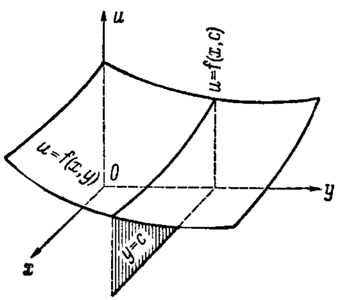


Рис. 23.

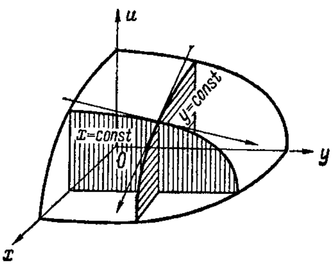


Рис. 24.

Рассмотрим, например, функцию $u=f(x, y)$ от двух аргументов x и y и припишем аргументу y определенное значение $y=y_0=c$. Полученную при этом функцию $u=f(x, y_0)$ от одной переменной x можно изобразить геометрически как сечение поверхности $u=f(x, y)$ плоскостью $y=y_0=c$ (рис. 23 и 24). Эта линия пересечения имеет в плоскости $y=y_0$ уравнение $u=f(x, y_0)$. Возьмем теперь обычным путем производную по x от функции $u=f(x, y_0)$ в точке $x=x_0$ (предполагая, что производная существует). Эта производная называется *частной производной от* $f(x, y)$ по x в точке (x_0, y_0). Согласно известному определению производной, частная производная по x есть следующий предел:

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}.$$

Если точка (x_0, y_0) лежит на границе области задания функции $u = f(x, y)$, то предельный переход по h к нулю должен совершаться таким образом, чтобы переменная точка $(x_0 + h, y_0)$ не выходила из этой области.

Геометрический смысл этой частной производной по x таков: она равна тангенсу угла между прямой, параллельной оси x, и касательной к кривой $y = f(x, y_0)$. Стало быть, частная производная по x дает *меру крутизны поверхности* $u = f(x, y)$ *по направлению* оси x.

Существует несколько различных обозначений для частных производных; укажем самые употребительные из них:

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = f_x(x_0, y_0) = u_x(x_0, y_0) = u'_x(x_0, y_0).$$

Если желательно, чтобы в самом обозначении оставался след того, что частная производная есть предел отношения приращений, то ее обозначают символом

$$\frac{\partial f}{\partial x} \quad \text{или} \quad \frac{\partial}{\partial x} f.$$

При этом пользуются круглой буквой ∂ вместо обыкновенной d, применяемой в обозначении производной от функции одного аргумента, желая этим показать, что речь идет о функции от многих переменных, которую дифференцируют по одной из них.

Иногда удобно пользоваться символом D, введенным Коши (о нем уже упоминалось в т. I, стр. 116), с добавлением индекса, указывающего, по какой переменной берут производную:

$$\frac{\partial f}{\partial x} = D_x f,$$

но мы будем лишь изредка пользоваться этим обозначением.

Точно таким же путем определяют частную производную от $f(x, y)$ по y в точке (x_0, y_0) соотношением

$$\lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} = f_y(x_0, y_0) = \frac{\partial f}{\partial y} = D_y f(x_0, y_0) =$$
$$= u_y(x_0, y_0) = u'_y(x_0, y_0).$$

Эта частная производная по y дает наклон касательной к линии пересечения поверхности $u = f(x, y)$ с плоскостью $x = x_0$ относительно прямой, параллельной оси y.

Будем теперь рассматривать точку (x_0, y_0), которая до сих пор считалась неподвижной, как переменную точку и в соответствии с этим опустим индексы 0. Другими словами, будем представлять себе, что дифференцирование выполнено в каждой точке области определения функции $f(x, y)$. Тогда обе частные производные сами окажутся функциями от x и y:

$$u_x(x, y) = f_x(x, y) = \frac{\partial f(x, y)}{\partial x} \quad \text{и} \quad u_y(x, y) = f_y(x, y) = \frac{\partial f(x, y)}{\partial y}.$$

Например, функция $u = x^2 + y^2$ имеет частные производные $u_x = 2x$ и $u_y = 2y$, потому что при дифференцировании по x член y^2 рассматривается как постоянный и имеет производную, равную нулю, а при дифференцировании по y член x^2 имеет производную, равную нулю. Аналогично, функция $u = x^3y$ имеет частные производные $u_x = 3x^2y$ и $u_y = x^3$.

Таким же образом дается определение частных производных при любом числе n независимых переменных

$$\frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_1} = \lim_{h \to 0} \frac{f(x_1 + h, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{h} =$$
$$= f_{x_1}(x_1, x_2, \ldots, x_n) = D_{x_1} f(x_1, x_2, \ldots, x_n),$$

конечно, в предположении, что этот предел существует.

Ясно, что можно также составить *частные производные высших порядков* от $f(x, y)$, вновь дифференцируя частные производные «первого порядка» $f_x(x, y)$ и $f_y(x, y)$ по каждой из независимых переменных и повторяя последовательно и дальше этот процесс. В какой последовательности дифференцируют по разным переменным отмечают порядком указателей (индексов) или же порядком символов ∂x и ∂y в «знаменателе» слева направо. Таким образом, для частных производных второго порядка установлены следующие символы:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f''_{xx} = f_{xx} = D^2_{xx} f,$$
$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x \partial y} = f''_{xy} = f_{xy} = D^2_{xy} f,$$
$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y \partial x} = f''_{yx} = f_{yx} = D^2_{yx} f,$$
$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f''_{yy} = f_{yy} = D^2_{yy} f.$$

По тому же правилу обозначают частные производные третьего порядка:

$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial^3 f}{\partial x^3} = f'''_{xxx} = f_{xxx},$$
$$\frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial^3 f}{\partial x^2 \partial y} = f'''_{xxy} = f_{xxy},$$
$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x} = f'''_{xyx} = f_{xyx} \text{ и т. д.}$$

и вообще производные порядка n:

$$\frac{\partial}{\partial x}\left(\frac{\partial^{n-1} f}{\partial x^{n-1}}\right)=\frac{\partial^n f}{\partial x^n}=f_{x^n}^{(n)}=f_{x^n},$$

$$\frac{\partial}{\partial y}\left(\frac{\partial^{n-1} f}{\partial x^{n-1}}\right)=\frac{\partial^n f}{\partial x^{n-1}\partial y}=f_{x^{n-1}y}^{(n)}=f_{x^{n-1}y} \quad \text{и т. д.}$$

Техника вычисления частных производных не содержит ничего такого, с чем бы читатель ранее не встречался. Действительно, согласно определению, надо считать постоянными все независимые переменные, кроме той, по которой берется производная. Поэтому остальные переменные следует рассматривать как постоянные и выполнять дифференцирование по тем же правилам, по которым дифференцируют функции одного аргумента. Тем не менее читателю полезно изучить примеры вычисления частных производных, данные в главе X первого тома (стр. 564 и след.).

Как и у функций одного аргумента, существование частных производных является специальным свойством функции, которым обладает даже не всякая непрерывная функция. Однако это свойство присуще всем практически важным функциям, за исключением, быть может, в отдельных особых точках.

Термин «дифференцируемая функция» будет объяснен в § 4, n° 1; у функций многих переменных он означает *больше*, чем одно только существование частных производных.

2. Существование частных производных по x и по y и непрерывность функции. Мы уже знаем, что у функции одного аргумента из существования производной в какой-либо точке вытекает непрерывность функции в этой точке (т. I, стр. 122). В противоположность этому уже у функции двух переменных из существования обеих частных производных еще не вытекает непрерывности функции. Возьмем, например, функцию $u(x, y)=\frac{2xy}{x^2+y^2}$ с дополнительным определением $u(0, 0)=0$. Эта функция имеет всюду обе частные производные. [Во всех точках, кроме начала координат, это вытекает из того факта, что в этих точках знаменатель не обращается в нуль, а в начале координат имеем

$$u_x(0, 0)=\lim_{h\to 0}\frac{u(h, 0)-u(0, 0)}{h}=0,$$

$$u_y(0, 0)=\lim_{k\to 0}\frac{u(0, k)-u(0, 0)}{k}=0.]$$

Между тем мы видели (т. I, стр. 560), что эта функция имеет разрыв в начале координат. Говоря геометрически, существование частных производных ограничивает поведение функции лишь по направлениям оси x и оси y, но не налагает ограничений на ее поведение во всех прочих направлениях.

Итак, существование у функции двух переменных обеих частных производных не позволяет делать вывод, что она непрерывна. Однако из существования в замкнутой области *ограниченных* частных производных *вытекает* непрерывность функции, как показывает следующая теорема:

Если функция $f(x, y)$ имеет всюду в области G частные производные $f_x(x, y)$ и $f_y(x, y)$, удовлетворяющие неравенствам

$$|f_x(x,y)|<M \quad \text{и} \quad |f_y(x, y)|<M,$$

где число M не зависит от x и y, то $f(x, y)$ непрерывна в области G.

Для доказательства рассмотрим две точки (x, y) и $(x+h, y+k)$, лежащие в области G. Предположим еще, что оба прямолинейных, взаимно перпендикулярных отрезка, соединяющих эти точки с точкой $(x+h, k)$, лежат целиком внутри области G. Это во всяком случае верно, если (x, y) есть внутренняя точка области G, а точка $(x+h, y+k)$ достаточно к ней близка. Тогда

$$f(x+h, y+k)-f(x, y)=$$
$$=\{f(x+h, y+k)-f(x+h, y)\}+\{f(x+h, y)-f(x, y)\}.$$

Оба члена в первой фигурной скобке правой части отличаются только значением аргумента y, а члены, заключенные во вторую фигурную скобку — только значением аргумента x. Поэтому каждую из этих фигурных скобок можно преобразовать по теореме дифференциального исчисления о среднем значении (т. I, стр. 130), рассматривая первую из этих скобок как приращение функции одного только y, а вторую — как приращение функции одного только x. В итоге получим

$$f(x+h, y+k)-f(x, y)=kf_y(x+h, y+\theta_1 k)+hf_x(x+\theta_2 h, y),$$

где θ_1 и θ_2 — два числа, лежащих между 0 и 1. Другими словами, производная по y берется в некоторой точке вертикального отрезка, соединяющего точки $(x+h, y)$ и $(x+h, y+k)$, а производная по x — в какой-то точке горизонтального отрезка от точки (x, y) до $(x+h, y)$.

Так как абсолютная величина каждой из этих производных, согласно условию теоремы, меньше числа M, то

$$|f(x+h, y+k)-f(x, y)|<M(|h|+|k|).$$

При достаточно малых значениях $|h|$ и $|k|$ делается сколь угодно малой и правая сторона этого неравенства, а это доказывает непрерывность функции $f(x, y)$.

3. Изменение порядка дифференцирования. Во всех примерах вычисления частных производных, приведенных в главе X первого тома (стр. 564—566), всегда $f_{yx}=f_{xy}$; другими словами, для всех рассмотренных там функций безразлично, производится ли дифференцирование сперва по x, а затем по y, или сначала по y, а затем по x. Это не случайное совпадение, а общее свойство обширного класса функций, выражающееся в следующей важной теореме:

Если «смешанные» вторые производные f_{xy} и f_{yx} функции $f(x, y)$ являются в области G непрерывными функциями от x и y, то внутри этой области всюду

$$f_{yx} = f_{xy};$$

другими словами, порядок дифференцирования по x и по y безразличен.

Доказательство этой теоремы, как и теоремы предыдущего номера, получается с помощью теоремы о среднем значении из дифференциального исчисления. Рассмотрим четыре точки: (x, y), $(x+h, y)$ $(x, y+k)$ и $(x+h, y+k)$, где $h \neq 0$ и $k \neq 0$. Если (x, y) является внутренней точкой области G, то при достаточно малых h и k все эти четыре точки, а также весь определяемый ими прямоугольник лежат внутри G. Составим теперь выражение

$$A = f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x, y).$$

Введя вспомогательную функцию

$$\varphi(x) = f(x, y+k) - f(x, y)$$

от одной переменной x, в которой рассматриваем y только как «параметр», приводим выражение A к следующему виду:

$$A = \varphi(x+h) - \varphi(x).$$

К правой части применяем теорему о среднем значении:

$$A = h\varphi'(x+\theta h),$$

где θ лежит между 0 и 1. Но

$$\varphi'(x) = f_x(x, y+k) - f_x(x, y),$$

и так как мы предположили, что смешанная вторая частная производная f_{xy} существует, то можно вновь применить теорему о среднем значении:

$$\varphi'(x) = kf_{xy}(x, y+\theta' k), \qquad 0 < \theta' < 1.$$

Теперь для A получается выражение

$$A = hkf_{xy}(x+\theta h, y+\theta' k),$$

где θ и θ' означают какие-то два числа между 0 и 1.

Однако с таким же успехом можно исходить из другой вспомогательной функции (на этот раз — от y)

$$\psi(y) = f(x+h, y) - f(x, y),$$

рассматривая x как параметр. Теперь выражение A запишется так:

$$A=\psi(y+k)-\psi(y).$$

Действуя таким же путем, как и раньше, после двукратного применения теоремы о среднем значении придем к равенству

$$A=khf_{yx}(x+\theta_1 h,\ y+\theta_1' k), \quad \text{где} \quad 0<\theta_1<1 \quad \text{и} \quad 0<\theta_1'<1.$$

Приравнивая оба выражения для A, получаем следующее соотношение:

$$f_{xy}(x+\theta h,\ y+\theta' k)=f_{yx}(x+\theta_1 h,\ y+\theta_1' k).$$

Заставим теперь h и k стремиться одновременно к нулю; принимая во внимание, что частные производные $f_{xy}(x, y)$ и $f_{yx}(x, y)$ непрерывны в точке (x, y), сразу получим то самое тождество

$$f_{xy}(x, y)=f_{yx}(x, y),$$

которое требовалось доказать.

Для более тонких исследований важно знать, что теорему о переместительности порядка дифференцирования можно доказать при более слабых ограничениях. Достаточно потребовать, чтобы существовала одна только из смешанных вторых производных (например, f_{xy}) в некоторой окрестности точки (x, y) и чтобы она была непрерывна в самой этой точке.

Для доказательства исходим из того же выражения, что и выше,

$$A=f(x+h,\ y+k)-f(x+h,\ y)-f(x,\ y+k)+f(x,\ y)= \\ =\psi(y+k)-\psi(y),$$

делим его на hk и заставляем стремиться к нулю одно только k. Тогда правая часть стремится к пределу $\frac{1}{h}\psi'(y)=\frac{f_y(x+h,\ y)-f_y(x,\ y)}{h}$. Следовательно,

$$\lim_{k\to 0}\frac{A}{hk}=\frac{f_y(x+h,\ y)-f_y(x,\ y)}{h}.$$

С другой стороны, выше было доказано, что из одного только предположения, что f_{xy} существует, вытекает равенство

$$\frac{A}{hk}=f_{xy}(x+\theta h,\ y+\theta' k), \quad \text{где} \quad 0<\theta, \theta'<1.$$

Так как мы предположили, что f_{xy} непрерывна, то при сколь угодно малом $\varepsilon>0$ и достаточно малых $|h|$ и $|k|$ имеем

$$f_{xy}(x, y)-\varepsilon<f_{xy}(x+\theta h,\ y+\theta' k)<f_{xy}(x, y)+\varepsilon,$$

откуда

$$f_{xy}(x, y)-\varepsilon\leqslant\frac{f_y(x+h,\ y)-f_y(x,\ y)}{h}\leqslant f_{xy}(x, y)+\varepsilon.$$

Следовательно,

$$\lim_{h\to 0}\frac{f_y(x+h,\ y)-f_y(x,\ y)}{h}=f_{xy}(x, y).$$

Так как, по определению производной, предел левой части равен f_{yx}, то имеем окончательно

$$f_{yx}(x, y) = f_{xy}(x, y).$$

Принципиально важно показать на примере, что если условие непрерывности второй производной f_{xy} или f_{yx} не выполнено, то наша теорема может и не быть справедливой, так что может случиться, что $f_{xy} \neq f_{yx}$. Такой пример дает функция, упомянутая уже на стр. 61, непрерывная в начале координат:

$$f(x, y) = xy\frac{x^2 - y^2}{x^2 + y^2} \quad \text{с добавочным условием} \quad f(0, 0) = 0.$$

У этой функции существуют все частные производные второго порядка, но они не являются непрерывными функциями.

Мы имеем

$$f_x(0, y) = \lim_{x \to 0} \frac{f(x, y) - f(0, y)}{x} = \lim_{x \to 0} y\frac{x^2 - y^2}{x^2 + y^2} = -y,$$

$$f_y(x, 0) = \lim_{y \to 0} \frac{f(x, y) - f(x, 0)}{y} = \lim_{y \to 0} x\frac{x^2 - y^2}{x^2 + y^2} = x,$$

откуда

$$f_{xy}(0, 0) = -1, \quad \text{а} \quad f_{yx}(0, 0) = 1.$$

Таким образом, эти два выражения не совпадают, что, согласно нашей теореме, может обусловливаться только разрывностью функции f_{xy} в начале координат.

Из теоремы о переместительности порядка дифференцирования вытекают далеко идущие следствия. Сразу замечаем, что число отличных друг от друга производных второго и высших порядков от функции многих переменных значительно меньше, чем можно было первоначально ожидать. Предположим, что все последовательно вычисляемые частные производные являются в рассматриваемой области непрерывными функциями от независимых переменных. Тогда, применяя нашу теорему, вместо функции $u = f(x, y)$, к функциям $f_x(x, y)$, $f_y(x, y)$, $f_{xy}(x, y)$ и т. д., придем к равенствам

$$f_{xxy} = f_{xyx} = f_{yxx},$$

$$f_{xyy} = f_{yxy} = f_{yyx},$$

$$f_{xxyy} = f_{xyxy} = f_{xyyx} = f_{yxxy} = f_{yxyx} = f_{yyxx} \text{ и т. д.}$$

Таким образом, получаем следующий совершенно общий результат:

При многократном дифференцировании функции двух переменных, можно как угодно изменять порядок дифференцирования, если только получающиеся производные являются непрерывными функциями.

Поэтому, если последнее условие выполнено, функция двух переменных имеет *три различных частных производных второго порядка*:

$$f_{xx}, \quad f_{xy}, \quad f_{yy},$$

четыре различных производных третьего порядка:

$$f_{xxx},\quad f_{xxy},\quad f_{xyy},\quad f_{yyy}$$

и вообще $(n+1)$ *различных производных n-го порядка:*

$$f_{x^n},\ f_{x^{n-1}y},\ f_{x^{n-2}y^2},\ \ldots,\ f_{xy^{n-1}},\ f_{y^n}.$$

Само собою разумеется, что аналогичная теорема справедлива и для функций от трех и большего числа независимых переменных. Действительно, наше рассуждение можно применить с тем же успехом к перемене порядка дифференцирования по x и z или по y и z и т. д., ибо каждая перестановка двух дифференцирований производится только по двум независимым переменным.

Упражнения

1. Сколько частных производных n-го порядка имеет функция трех переменных?

2. Доказать, что функция

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(x_1^2 + x_2^2 + \ldots + x_n^2)^{\frac{n-2}{2}}}$$

удовлетворяет уравнению Лапласа

$$f_{x_1x_1} + f_{x_2x_2} + \ldots + f_{x_nx_n} = 0.$$

3. Рассматривая определитель

$$\Delta = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

как функцию девяти независимых переменных $a_1, a_2, a_3, b_1, \ldots, c_3$, вывести формулы для частных производных этого определителя по своим элементам:

$$\frac{\partial\Delta}{\partial a_1} = \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix},\quad \frac{\partial\Delta}{\partial a_2} = -\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix},\ \ldots,\ \frac{\partial\Delta}{\partial c_3} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Вывести аналогичные формулы для частных производных определителя четвертого порядка по своим элементам (см. упр. 4, стр. 40).

4. Для любого определителя третьего и четвертого порядков доказать следующие теоремы:

а) Сумма произведений элементов какой-либо строки (какого-либо столбца) определителя на его соответствующие частные производные по элементам той же строки (того же столбца) равна значению определителя.

б) Сумма произведений элементов какой-либо строки (какого-либо столбца) определителя на его соответствующие частные производные по элементам другой строки (другого столбца) равна нулю.

в)* Для определителя Δ (упр. 3) доказать

$$\begin{vmatrix} \frac{\partial\Delta}{\partial a_1} & \frac{\partial\Delta}{\partial a_2} & \frac{\partial\Delta}{\partial a_3} \\ \frac{\partial\Delta}{\partial b_1} & \frac{\partial\Delta}{\partial b_2} & \frac{\partial\Delta}{\partial b_3} \\ \frac{\partial\Delta}{\partial c_1} & \frac{\partial\Delta}{\partial c_2} & \frac{\partial\Delta}{\partial c_3} \end{vmatrix} = \Delta^2.$$

§ 4. Полный дифференциал функции и его геометрический смысл

1. Понятие дифференцируемости. Для функций одной переменной существование производной теснейшим образом связано с возможностью приближенно выразить функцию $\eta = f(\xi)$, в окрестности точки x, с помощью некоторой линейной функции $\eta = \varphi(\xi)$. Эта линейная функция определяется формулой

$$\varphi(\xi) = f(x) + (\xi - x) f'(x).$$

Геометрически эта функция выражает, в текущих координатах ξ и η, касательную к кривой $\eta = f(\xi)$ в ее точке P с координатами $\xi = x$ и $\eta = f(x)$. Она характеризуется аналитически тем, что отличается от функции $f(\xi)$ в окрестности точки P на величину $o(h)$ более высокого порядка малости (стр. 63 и т. I, стр. 222), чем разность абсцисс $h = \xi - x$.

Поэтому

$$f(\xi) - \varphi(\xi) = f(\xi) - f(x) - (\xi - x) f'(x) = o(h),$$

или в другой записи

$$f(x+h) - f(x) - hf'(x) = o(h) = \varepsilon h,$$

где ε означает число, стремящееся к нулю вместе с h.

Величину $hf'(x)$, «линейную часть приращения» функции $f(x)$, соответствующую приращению h независимой переменной, мы назвали в первом томе (гл. II, § 3, n° 10) *дифференциалом* функции $f(x)$ и обозначили через

$$dy = df(x) = hy' = hf'(x)$$

(а также через $dy = y'dx$, так как для функции $y = x$ дифференциал $dy = dx = 1 \cdot h$).

Дифференциал является, как мы можем теперь сказать, функцией от двух независимых переменных x и h, причем относительно величины приращения h мы не делаем никаких предположений. Правда, этим понятием дифференциала пользуются только тогда, когда $|h|$ настолько мал, что дифференциал $hf'(x)$ представляет собою достаточно точное приближение приращения функции $f(x+h) - f(x)$ для изучаемого вопроса.

Обратно, вместо того чтобы исходить из понятия производной, можно положить в основу требование, чтобы для функции $\eta = f(\xi)$, в окрестности точки x, существовала такая приближенно ее выражающая линейная функция, чтобы разность между данной функцией и ее линейным приближением имела порядок малости более высокий, чем приращение h независимой переменной. Другими словами, мы требуем, чтобы для функции $f(\xi)$ при значении $\xi = x$ существовала такая величина A, зависящая от x, но не зависящая от h, что

$$f(x+h) - f(x) = Ah + o(h) = Ah + \varepsilon h,$$

где ε стремится к нулю вместе с h. Это требование равносильно требованию дифференцируемости $f(x)$ при значении x; в качестве величины A нужно тогда взять производную $f'(x)$ при значении x. В этом нетрудно убедиться, написав наше требование в форме

$$\frac{f(x+h) - f(x)}{h} = A + \varepsilon$$

и переходя к пределу при $h \to 0$. Таким образом, дифференцируемость функции одной переменной и возможность ее приближения с помощью линейной функции (в указанном смысле) являются равнозначащими свойствами.

Если учесть, что $A + \varepsilon = \alpha(x, h)$ как функция от h стремится к $A(x)$ при $h \to 0$, то придём к эквивалентному определению: функция $f(x)$ называется дифференцируемой в точке x, если $f(x+h) - f(x) = h\alpha(x, h)$, где $\alpha(x, h)$ является непрерывной функцией от h в точке $h = 0$.

Эти понятия можно совершенно естественным путем распространить на функции от двух или многих переменных.

Функция $u = f(x, y)$ называется дифференцируемой в точке (x, y), если в окрестности этой точки она может быть приближенно выражена с помощью линейной функции, т. е. может быть представлена следующим образом:

$$f(x+h, y+k) = f(x, y) + A(x, y)h + B(x, y)k + \varepsilon_1 h + \varepsilon_2 k,$$

где $A(x, y)$ и $B(x, y)$ не зависят от величин h и k, изменяющихся независимо друг от друга, а ε_1 и ε_2 стремятся к нулю вместе с h и k. Другими словами, разность между данной функцией $f(x+h, y+k)$ в точке $(x+h, y+k)$ и линейной относительно h и k функцией $f(x, y) + Ah + Bk$ [т. е. погрешность линейного приближения] должна быть при $\rho \to 0$ величиной порядка малости $o(\rho)$, т. е. порядка высшего, чем порядок малости расстояния $\rho = \sqrt{h^2 + k^2}$ между точками (x, y) и $(x+h, y+k)$, так что

$$f(x+h, y+k) = f(x, y) + Ah + Bk + \varepsilon\rho.$$

Эквивалентность обеих формулировок доказывается так. Всегда справедливо неравенство $|\varepsilon_1 h + \varepsilon_2 k| \leqslant \varepsilon\sqrt{h^2 + k^2}$, если через ε обозначим величину $\varepsilon = |\varepsilon_1| + |\varepsilon_2|$, стремящуюся к нулю вместе с ε_1 и ε_2. Тем самым из первой формулировки вытекает вторая. С другой стороны, так как $|\varepsilon\sqrt{h^2+k^2}| \leqslant |\varepsilon|(|h|+|k|)$, то при выполнении второго требования разность между функцией и ее линейным приближением имеет вид $\theta\varepsilon(|h|+|k|)$, где $-1 \leqslant \theta \leqslant +1$, а стало быть, выполняется и требование, содержащееся в первой формулировке.

Если такое приближенное представление существует, т. е. функция $f(x, y)$ дифференцируема в точке (x, y), то она имеет в этой точке обе частные производные, причем $f_x = A(x, y)$ и $f_y = B(x, y)$.

В самом деле, полагая $k = 0$ и разделив приращение функции на h, получим

$$\frac{f(x+h, y) - f(x, y)}{h} = A(x, y) + \varepsilon_1,$$

а так как ε_1 стремится к нулю вместе с h, то при предельном переходе $h \to 0$ левая часть действительно стремится к пределу $A(x, y)$, так что $f_x(x, y) = A(x, y)$. Таким же путем заключаем, что существует частная производная по y и что $f_y(x, y) = B(x, y)$.

Тем самым одновременно доказано, что линейное приближение дифференцируемой функции единственно.

Итак, функция, дифференцируемая в точке (x, y), может быть представлена в окрестности этой точки в следующем виде:

$$f(\xi, \eta) = f(x+h, y+k) = f(x, y) + f_x(x, y)h + f_y(x, y)k + \varepsilon\rho,$$

где $\varepsilon \to 0$ при $\rho \to 0$. Здесь $\xi = x + h$, $\eta = y + k$.

Из этой формулы можно вывести еще одно важное следствие. Для дифференцируемой функции нет нужды отдельно доказывать ее непрерывность: *если $f(x, y)$ дифференцируема в точке (x, y), то отсюда уже вытекает ее непрерывность в этой точке.*

Действительно, так как предельный переход $h \to 0$ и $k \to 0$ равносилен предельному переходу $\xi \to x$, $\eta \to y$, а при этом $\varepsilon \to 0$ и $\rho \to 0$, то из последней формулы вытекает, что

$$\lim_{\substack{\xi \to x \\ \eta \to y}} f(\xi, \eta) = f(x, y),$$

а это и значит (стр. 61), что наша функция непрерывна в точке (x, y).

Из сказанного видно, что как непрерывность функции $f(x, y)$, так и существование у нее обеих частных производных в точке (x, y) являются *необходимыми* условиями дифференцируемости функции в этой точке: если хотя бы одно из этих условий не выполняется в какой-либо точке, то функция не может быть дифференцируемой в этой точке. Вскоре мы увидим, что и выполнение всех этих условий не обеспечивает дифференцируемости функции в рассматриваемой точке.

Однако практика нуждается и в достаточном критерии дифференцируемости функции; нижеследующая теорема дает один из видов достаточных условий дифференцируемости:

Если функция $u=f(x, y)$ *имеет в некоторой окрестности точки* (x, y) *обе частные производные первого порядка, а в самой этой точке частные производные непрерывны, то функция дифференцируема в точке* (x, y), т. е. допускает приближенное представление с помощью линейной функции.

В самом деле, полное приращение функции

$$\Delta u=f(x+h, y+k)-f(x, y)$$

может быть представлено в следующем виде:

$$\Delta u=\{f(x+h, y+k)-f(x, y+k)\}+\{f(x, y+k)-f(x, y)\}.$$

Каждую из фигурных скобок мы преобразуем по теореме о среднем значении дифференциального исчисления:

$$\Delta u=hf_x(x+\theta_1 h, y+k)+kf_y(x, y+\theta_2 k).$$

Так как, по условию теоремы, частные производные f_x и f_y непрерывны в точке (x, y), то мы вправе написать:

$$f_x(x+\theta_1 h, y+k)=f_x(x, y)+\varepsilon_1 \quad \text{и} \quad f_y(x, y+\theta_2 k)=f_y(x, y)+\varepsilon_2,$$

причем числа ε_1 и ε_2 стремятся к нулю вместе с h и k. Таким образом, получаем

$$\Delta u=hf_x(x,y)+kf_y(x,y)+\varepsilon_1 h+\varepsilon_2 k=$$
$$=hf_x(x,y)+kf_y(x,y)+o\left(\sqrt{h^2+k^2}\right),$$

а это и означает, что наша функция дифференцируема в точке (x, y). Выражение $hf_x(x,y)+kf_y(x,y)$ называется линейной частью приращения функции, а выражение $\varepsilon_1 h+\varepsilon_2 k=o\left(\sqrt{h^2+k^2}\right)$ называется членом высшего порядка или погрешностью линейного приближения функции.

Если допустить только существование частных производных f_x и f_y, но не их непрерывность в точке (x, y), то функция может и не быть дифференцируемой, в чем можно убедиться на следующем примере.

Рассмотрим функцию

$$u=f(x, y)=\begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}}, & \text{если } x^2+y^2\neq 0, \\ 0, & \text{если } x=y=0. \end{cases}$$

(Ход изменения этой функции станет особенно ясным, если перейти к полярным координатам $x=r\cos\theta$, $y=r\sin\theta$; тогда $u=\frac{r}{2}\sin 2\theta$.)

Первые производные по x и по y существуют всюду в окрестности начала координат и равны нулю в самом начале координат, но они имеют разрыв в этой точке, ибо, например,

$$u_x = y\left(\frac{1}{\sqrt{x^2+y^2}} - \frac{x^2}{(\sqrt{x^2+y^2})^3}\right) = \frac{y^3}{\sqrt{(x^2+y^2)^3}}.$$

Если приближаться к началу координат вдоль оси x, то u_x стремится к нулю, если же приближаться к началу по оси y, то $u_x \to +1$ при $y \to +0$ и $u_x \to -1$ при $y \to -0$. Легко убедиться, что наша функция $u = f(x, y)$ недифференцируема в начале координат. Действительно, в противном случае она имела бы своим линейным приближением $f(0, 0) + f_x(0, 0)\,h + f_y(0, 0)\,k = 0$, а разность между значением функции $f(x, y)$ и ее линейным приближением была бы $u - 0 = \frac{r}{2}\sin 2\theta$; но вдоль прямой $\theta = \frac{\pi}{4}$ имеем $\sin 2\theta = 1$ и $u = \frac{r}{2}$, так что порядок малости этой разности не выше порядка малости расстояния r вопреки тому, чего требует определение дифференцируемости функции.

Функцию с *непрерывными* первыми частными производными мы будем называть *непрерывно дифференцируемой* функцией. Если и все производные второго порядка также непрерывны, то мы будем называть функцию *дважды непрерывно дифференцируемой* и т. д.

Как и в случае одной независимой переменной, определение дифференцируемости функции двух переменных можно заменить следующим равносильным определением: функция $f(x, y)$ называется дифференцируемой в точке (x, y), если

$$f(x+h, y+k) - f(x, y) = \alpha h + \beta k,$$

где α и β зависят не только от x и y, но и от приращений h и k и являются непрерывными функциями от h и k при $h = 0$, $k = 0$.

Мы полагаем, что нет нужды специально рассказывать, как распространить определения и выводы этого номера на функции трех и большего числа аргументов.

2. Производная по заданному направлению. Дифференцируемые функции обладают тем важным свойством, что они имеют частные производные не только по x и по y или, как принято также говорить, по направлениям x и y, — их можно также *дифференцировать по любому другому направлению*. Объясним смысл этого утверждения.

Любое направление на плоскости определяется указанием угла α, образуемого этим направлением с положительной осью x, либо заданием единичного вектора $s = \{\cos\alpha, \sin\alpha\}$. Из рассматриваемой точки $P(x, y)$ проведем полупрямую по направлению вектора s, т. е. под углом α к оси x. На этой полупрямой возьмем переменную точку $Q(x+h, y+k)$. Заставим точку Q стремиться к точке P, двигаясь вдоль упомянутой полупрямой. При этом, когда расстояние ρ между неподвижной точкой P и переменной точкой Q, т. е. модуль вектора $\overline{PQ}=\{h, k\}$, стремится к нулю, то h и k тоже будут стремиться к нулю, но уже не как угодно, не независимо друг от друга, а подчиняясь формулам

$$h=\rho\cos\alpha, \qquad k=\rho\sin\alpha.$$

(Заметим, что $\rho=|\overline{PQ}|=\sqrt{h^2+k^2}$.)

Вычислим теперь, как обычно, приращение функции $f(x, y)$ при перемещении на вектор $\overline{PQ}$: $f(x+h, y+k)-f(x, y)$ и разделим его на ρ. Если эта дробь стремится к определенному пределу при $\rho\to 0$, то этот предел называется *производной* от функции $f(x, y)$ в точке $P(x, y)$ *по направлению* α (точнее, по направлению вектора $\mathbf{s}$) и обозначается символом $D_{(\alpha)}f$ или $D_{\mathbf{s}}f$.

Следовательно,

$$D_{(\alpha)}f(x, y)=D_{\mathbf{s}}f(x, y)=\lim_{\rho\to 0}\frac{f(x+h, y+k)-f(x, y)}{\rho}=$$
$$=\lim_{\rho\to 0}\frac{f(x+\rho\cos\alpha, y+\rho\sin\alpha)-f(x, y)}{\rho}$$

В частном случае, когда $\alpha=0$, то $\mathbf{s}=\mathbf{e}_1$, т. е. $k=0$ и $\rho=h$, и мы получаем $D_{(0)}f(x, y)=D_{\mathbf{e}_1}f(x, y)=\frac{\partial f}{\partial x}$; если же $\alpha=\frac{\pi}{2}$, то $\mathbf{s}=\mathbf{e}_2$, т. е. $h=0$ и $\rho=k$, и тогда $D_{\left(\frac{\pi}{2}\right)}f(x, y)=D_{\mathbf{e}_2}f(x, y)=\frac{\partial f}{\partial y}$. Нетрудно убедиться, что $D_{(\pi)}f(x, y)=D_{-\mathbf{e}_1}f(x, y)=-\frac{\partial f}{\partial x}$ и $D_{\left(-\frac{\pi}{2}\right)}f(x, y)=$ $=D_{-\mathbf{e}_2}f(x, y)=-\frac{\partial f}{\partial y}$. Вообще если функция $f(x, y)$ дифференцируема в точке (x, y), то ее приращение можно представить в следующем виде:

$$f(x+h, y+k)-f(x, y)=hf_x+kf_y+\varepsilon\rho=\rho(f_x\cos\alpha+f_y\sin\alpha+\varepsilon).$$

Когда $\rho\to 0$, то и $\varepsilon\to 0$, и для *производной по направлению* α (по направлению вектора $\mathbf{s}$) получается следующее выражение:

$$D_{(\alpha)}f(x, y)=D_{\mathbf{s}}f(x, y)=f_x\cos\alpha+f_y\sin\alpha.$$

Следовательно, *производная по направлению α равна линейной комбинации частных производных f_x и f_y с коэффициентами* $\cos\alpha$ *и* $\sin\alpha$. Этот вывод, стало быть, всегда справедлив, коль скоро известно, что производные f_x и f_y существуют и непрерывны в рассматриваемой точке (x, y).

Пример 1. Изучаемая функция равна полярному радиусу, т. е. модулю радиус-вектора: $f(x, y)=r=|\mathbf{r}|=\sqrt{x^2+y^2}$. Тогда

$$r_x=\frac{x}{\sqrt{x^2+y^2}}=\frac{x}{r}=\cos\theta, \qquad r_y=\frac{y}{\sqrt{x^2+y^2}}=\frac{y}{r}=\sin\theta,$$

где θ обозначает угол от оси x до радиус-вектора точки (x, y). Поэтому полярный радиус r имеет по направлению α производную

$$D_{(\alpha)}r = r_x \cos\alpha + r_y \sin\alpha = \cos\theta\cos\alpha + \sin\theta\sin\alpha = \cos(\theta - \alpha).$$

В частности, производная полярного радиуса по направлению радиус-вектора, т. е. $D_{(\theta)}r = D_r r = \cos 0 = 1$, а по направлению, перпендикулярному к радиус-вектору:

$$D_{\left(\theta+\frac{\pi}{2}\right)}r = D_{[e_3 r]}r = 0 \quad \text{и} \quad D_{\left(\theta-\frac{\pi}{2}\right)}r = D_{[re_3]}r = 0.$$

Пример 2. Пусть $f(x, y) = x$. Тогда $f_x = 1$, $f_y = 0$. Производная по любому направлению α есть $D_{(\alpha)}x = \cos\alpha$. Для функции $\varphi(x, y) = y$ имеем $D_{(\alpha)}y = \sin\alpha$.

В частности, производные от этих функций по направлению радиус-вектора $D_{(\theta)}x = \cos\theta$, $D_{(\theta)}y = \sin\theta$, а производные по направлению, перпендикулярному к радиус-вектору, будут $D_{\left(\theta+\frac{\pi}{2}\right)}x = -\sin\theta$, $D_{\left(\theta+\frac{\pi}{2}\right)}y = \cos\theta$ и $D_{\left(\theta-\frac{\pi}{2}\right)}x = \sin\theta$, $D_{\left(\theta-\frac{\pi}{2}\right)}y = -\cos\theta$.

Производная от функции $f(x, y)$ по направлению радиус-вектора $\boldsymbol{r}$, когда за независимые переменные приняты полярные координаты (r, θ), совпадает с частной производной $\frac{\partial f(x, y)}{\partial r}$. Для нее получается следующее часто применяемое «символическое» соотношение:

$$\frac{\partial}{\partial r} = \cos\theta\,\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial y},$$

в котором подразумевается, что за символами дифференцирования $\frac{\partial}{\partial r}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ можно поставить любую дифференцируемую функцию u от x и y.

Полезно также заметить, что если точка $Q(x+h, y+k)$ стремится к точке $P(x, y)$, двигаясь не вдоль полупрямой с направлением α, а вдоль произвольной кривой C, имеющей в точке P касательную с направлением α, то и в этом случае

$$\lim_{\rho\to 0} \frac{f(x+h, y+k) - f(x, y)}{\rho} = D_{(\alpha)}f(x, y).$$

Действительно, пусть хорда $\overline{PQ}$ кривой C образует с осью x угол β. Тогда $|\overline{PQ}| = \rho$ и $h = \rho\cos\beta$, $k = \rho\sin\beta$. Предполагая, как и раньше, что функция $f(x, y)$ дифференцируема, имеем

$$\lim_{\rho\to 0} \frac{f(x+h, y+k) - f(x, y)}{\rho} = \lim_{\rho\to 0} \frac{hf_x + kf_y + \varepsilon\rho}{\rho} =$$
$$= \lim_{\substack{\rho\to 0\\ \beta\to\alpha}} (f_x\cos\beta + f_y\sin\beta + \varepsilon) = f_x\cos\alpha + f_y\sin\alpha = D_{(\alpha)}f(x, y),$$

что и требовалось доказать. Мы воспользовались тем, что при $\rho \to 0$ также и $\varepsilon \to 0$, а угол $\beta \to \alpha$.

Совершенно аналогично строится определение производной по заданному направлению для функции $f(x, y, z)$ от трех независимых переменных. Всякое направление в пространстве можно задать с помощью единичного вектора $s=\{\cos\alpha, \cos\beta, \cos\gamma\}$, где α, β, γ — углы, образуемые этим вектором с осями координат. Наряду с точкой $P(x, y, z)$ рассматриваем переменную точку $Q(x+h, y+k, z+l)$, лежащую на полупрямой, выходящей из P по направлению вектора s. Точку P будем называть точкой наблюдения, а вектор $\overline{PQ}=\{h, k, l\}$ смещением точки наблюдения. Обозначим через ρ расстояние между точками P и Q, т. е. модуль смещения $\overline{PQ}$. Тогда

$$h=\rho\cos\alpha, \qquad k=\rho\cos\beta, \qquad l=\rho\cos\gamma \quad \text{и} \quad \rho=\sqrt{h^2+k^2+l^2}.$$

Заставим теперь точку Q стремиться к P, двигаясь вдоль упомянутой полупрямой. Тогда предел

$$\lim_{\rho\to 0}\frac{f(x+h, y+k, z+l)-f(x, y)}{\rho},$$

если он существует, называется *производной* от функции $f(x, y, z)$ в точке $P(x, y, z)$ *по направлению вектора* s и обозначается символом $D_s f(x, y, z)$. Другими словами, производной от функции $f(x, y, z)$ по направлению вектора s называется предел отношения приращения функции при смещении точки наблюдения в указанном направлении к модулю смещения, когда этот модуль стремится к нулю, конечно, при условии, что предел существует.

Если $f(x, y, z)$ — дифференцируемая функция, то приращение функции

$$f(x+h, y+k, z+l)-f(x, y, z)=hf_x+kf_y+lf_z+\varepsilon\rho=$$
$$=\rho(f_x\cos\alpha+f_y\cos\beta+f_z\cos\gamma+\varepsilon),$$

где $\varepsilon\to 0$ при $\rho\to 0$ и

$$D_s f(x, y, z)=f_x\cos\alpha+f_y\cos\beta+f_z\cos\gamma.$$

3. Геометрическое истолкование. Касательная плоскость. Сказанное в n° 2 легко истолковать геометрически для функции двух переменных $u=f(x, y)$. Вспомним, что частная производная по x дает наклон касательной к линии, по которой поверхность $u=f(x, y)$ пересекается плоскостью, проходящей через точку $(x, y, 0)$ перпендикулярно к плоскости $x0y$ и параллельно плоскости $x0u$. Подобно этому, производная по направлению α дает наклон касательной к той кривой, по которой наша поверхность пересекается плоскостью, проходящей через точку $(x, y, 0)$ перпендикулярно к плоскости $x0y$ под углом α к оси x. Формула $D_{(\alpha)}f(x, y)=f_x\cos\alpha+f_y\sin\alpha$ дает возможность вычислить наклоны касательных ко всем этим плоским сечениям, т. е. наклоны всех касательных прямых к поверхности в некоторой ее точке, через наклоны двух таких касательных.

Мы выразили нашу дифференцируемую функцию $u=f(\xi, \eta)=$ $=f(x+h, y+k)$ в окрестности точки (x, y) приближенно с помощью линейной функции

$$l(\xi, \eta)=f(x, y)+(\xi-x)f_x+(\eta-y)f_y,$$

где $\xi=x+h$, $\eta=y+k$ означают текущие координаты. Геометрическим изображением этой линейной функции является плоскость, которая, по аналогии с касательной прямой к плоской кривой, называется *касательной плоскостью* к поверхности $u=f(x,y)$ в ее точке (x, y, u). Разность между $f(\xi, \eta)$ и этой линейной функцией, т. е. погрешность приближения, стремится к нулю вместе с $h=\xi-x$ и $k=\eta-y$, и притом быстрее, чем $\rho=\sqrt{h^2+k^2}$. Согласно определению касательной к плоской кривой, это означает, что линия пересечения касательной плоскости с любой плоскостью, перпендикулярной к плоскости xOy, является касательной к линии пересечения поверхности с этой последней плоскостью. Стало быть, *все эти касательные прямые к поверхности в точке (x, y, u) лежат в одной плоскости — касательной плоскости к поверхности.*

Это свойство является геометрическим выражением дифференцируемости нашей функции в точке (x, y), которой соответствует $u=$ $=f(x, y)$. Введя обозначение $l(\xi, \eta)=\zeta$, получим уравнение касательной плоскости в текущих координатах ξ, η, ζ в следующем виде:

$$\zeta-u=(\xi-x)f_x+(\eta-y)f_y.$$

Выше (стр. 77) было доказано, что если частные производные непрерывны в данной точке, то функция непременно дифференцируема в этой точке. В противоположность положению дел у функций одной независимой переменной, одного *только существования* частных производных f_x и f_y еще *не достаточно*, чтобы обеспечить дифференцируемость функции; другими словами, если частные производные имеют разрыв в точке (x, y), то разность между $f(x+h, y+k)$ и функцией $f(x, y)+hf_x(x, y)+kf_y(x, y)$, линейной относительно h и k, может иметь порядок малости, *не превышающий* порядка малости расстояния $\rho=\sqrt{h^2+k^2}$. Геометрически это означает, что в этом случае поверхность $u=f(x, y)$ может и не иметь касательной плоскости в своей точке (x, y, u).

Поясним это на простом примере. Положим $f(x, y)=0$ вдоль прямых $x=0$ и $y=0$, $f(x, y)=|x|$ вдоль прямых $x-y=0$ и $x+y=0$.

Прямые $x=0$, $y=0$, $x-y=0$, $x+y=0$ делят плоскость xOy на восемь углов. В каждом из этих углов мы определим функцию $u=f(x, y)$ так, чтобы она изображалась в нем куском некоторой плоскости, с таким расчетом, чтобы полученная функция была всюду непрерывна. Таким образом, поверхность $u=f(x, y)$ состоит из восьми плоских углов-граней, смыкающихся друг с другом на ребрах: 1) вдоль прямой $x=0$, $u=0$; 2) вдоль прямой $y=0$, $u=0$; 3) на двух ребрах, возвышающихся над прямой $x-y=$ $=0$, $u=0$ и 4) на двух ребрах, расположенных над прямой $x+y=0$, $u=0$, Эта поверхность не имеет, очевидно, касательной плоскости в начале координат, хотя и существуют обе частные производные $f_x(0, 0) = 0$ и $f_y = (0, 0) = 0$; но эти производные имеют в начале координат разрыв. Легко убедиться, что на ребрах частные производные и вообще-то не существуют одновременно ни в одной точке, кроме начала координат.

Другим примером может служить функция

$$u = f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{при} \quad x^2+y^2 \neq 0, \\ 0 & \text{при} \quad x = y = 0, \end{cases}$$

рассмотренная на стр. 77. Мы видели, что эта функция имеет всюду обе частные производные f_x и f_y, но они терпят разрыв в начале координат. Оказалось, что функция недифференцируема в начале координат. Изображающая эту функцию поверхность $u = f(x, y)$ не имеет касательной плоскости в начале координат. Так как $f_x(0, 0) = 0$ и $f_y = (0, 0) = 0$, то формально написанное уравнение касательной плоскости в точке $x = 0$, $y = 0$, $u = 0$ было бы $\zeta = 0$, т. е. плоскость $x0y$ претендовала бы на роль касательной плоскости, но мы видели на стр. 78, что расстояние точки поверхности $u = f(x, y)$ от плоскости $x0y$ имеет в окрестности начала координат тот же порядок малости, что расстояние $r = \sqrt{x^2+y^2}$, а не выше.

4. Полный дифференциал функции. Как и у функций одной переменной, оказалось целесообразным ввести для линейной части дифференцируемой функции $u = f(x, y)$ особое название и обозначение, — эту линейную часть называют *полным дифференциалом* функции и обозначают его так:

$$du = df(x, y) = \frac{\partial f}{\partial x} h + \frac{\partial f}{\partial y} k = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Полный дифференциал представляет собой функцию от *четырех* независимых переменных: во-первых, от координат x и y рассматриваемой точки и, во-вторых, от приращений $h = dx$ и $k = dy$, которые называются дифференциалами независимых переменных. Его значение заключается в том, что du аппроксимирует линейно приращение функции $\Delta u = f(x+h, y+k) - f(x, y)$ (т. е. дает для нее линейное приближение) с ошибкой, имеющей более высокий порядок малости, чем расстояние $\rho = \sqrt{h^2+k^2}$, когда $|h|$ и $|k|$ достаточно малы. Выше (стр. 76) была доказана единственность линейного приближения дифференцируемой функции; тем самым доказана *единственность дифференциала.* Из данного там доказательства вытекает, что если каким-либо путем получено выражение для полного дифференциала, то сразу определяются все частные производные: f_x равна выражению, стоящему множителем при dx, а f_y равна множителю при dy. Таким образом, полный дифференциал объединяет в одной формуле выражения для различных частных производных.

Подчеркнем еще раз, что говорить о полном дифференциале функции $f(x, y)$ имеет смысл лишь в том случае, если эта функция дифференцируема согласно данному выше определению (для чего достаточным условием является непрерывность обеих частных производных, а отнюдь не одно только их существование).

Если функция $f(x, y)$ имеет непрерывные частные производные также и высших порядков, то можно от полного дифференциала $df(x, y)$ вычислить в свою очередь полный дифференциал, т. е. помножить его частные производные по x и по y поочередно на $h = dx$ и $k = dy$, а затем сложить эти произведения. Помня, что дифференциал $df = hf_x + kf_y$ является функцией четырех независимых переменных x, y, h и k, при вычислении его частных производных по x и y следует h и k рассматривать как постоянные. Таким путем мы получим *полный дифференциал второго порядка* (короче, второй дифференциал) функции $f(x, y)$:

$$d^2f = d(df) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k\right)h + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k\right)k =$$
$$= \frac{\partial^2 f}{\partial x^2}h^2 + 2\frac{\partial^2 f}{\partial x\,\partial y}hk + \frac{\partial^2 f}{\partial y^2}k^2 = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x\,\partial y}dx\,dy + \frac{\partial^2 f}{\partial y^2}dy^2.$$

Совершенно таким же образом составляются полные дифференциалы *высших порядков* (третьего, четвертого и т. д.):

$$d^3f = d(d^2f) = \frac{\partial^3 f}{\partial x^3}dx^3 + 3\frac{\partial^3 f}{\partial x^2\partial y}dx^2dy + 3\frac{\partial^3 f}{\partial x\,\partial y^2}dx\,dy^2 + \frac{\partial^3 f}{\partial y^3}dy^3,$$
$$d^4f = d(d^3f) = \frac{\partial^4 f}{\partial x^4}dx^4 + 4\frac{\partial^4 f}{\partial x^3\partial y}dx^3dy + 6\frac{\partial^4 f}{\partial x^2\partial y^2}dx^2dy^2 +$$
$$+ 4\frac{\partial^4 f}{\partial x\,\partial y^3}dx\,dy^3 + \frac{\partial^4 f}{\partial y^4}dy^4$$

и вообще, как легко доказать методом полной индукции:

$$d^nf = d(d^{n-1}f) = \frac{d^nf}{\partial x^n}dx^n + \binom{n}{1}\frac{\partial^n f}{\partial x^{n-1}\partial y}dx^{n-1}dy + \ldots$$
$$\ldots + \binom{n}{n-1}\frac{\partial^n f}{\partial x\,\partial y^{n-1}}dx\,dy^{n-1} + \frac{\partial^n f}{\partial y^n}dy^n.$$

Это последнее выражение можно записать символически так:

$$d^nf = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^{(n)} f.$$

Под словом «символически» имеется в виду следующее: скобки в правой части надо раскрыть по формуле бинома Ньютона, как будто бы «показатель» (n) был показателем степени, а затем полученное развернутое выражение «помножить» почленно на f, приписав f «множителем» к «числителю» ∂^n «дроби», появившейся в каждом члене.

В заключение заметим, что рассуждения и результаты этого номера непосредственно обобщаются на функции от любого числа независимых переменных.

5. Применение к исчислению ошибок. Полный дифференциал приносит практическую пользу в различных приложениях, ибо он дает удобное для вычисления приближенное значение приращения функции $\Delta f = f(x+h, y+k) - f(x, y)$. Эта польза особенно заметна в так называемом исчислении ошибок (т. I, гл. VII, § 2, n° 1). Пусть, например, требуется оценить возможную ошибку в определении удельного веса твердого тела способом взвешивания в воздухе и в воде. Пусть вес тела в воздухе $p\Gamma$, а в воде $p_1\Gamma$. По закону Архимеда, потеря в весе $p - p_1$ дает вес в граммах вытесненной этим телом воды или численно ему равный объем этой воды в $см^3$. Следовательно, и объем данного тела равен $p - p_1$ $см^3$. Поэтому удельный вес $s = \frac{p}{p - p_1} \frac{\Gamma}{см^3}$ является функцией двух независимых переменных p и p_1. Ошибка при вычислении удельного веса s, обусловленная ошибками dp и dp_1, допущенными при измерении p и p_1, выражается приближенно полным дифференциалом

$$ds = \frac{\partial s}{\partial p} dp + \frac{\partial s}{\partial p_1} dp_1 = \frac{-p_1 dp + p\, dp_1}{(p - p_1)^2}.$$

Так как знак ошибки обычно неизвестен, то оценке подлежат только абсолютные величины возможных ошибок: $|dp|$, $|dp_1|$ и $|ds|$. Поэтому приходим к формуле

$$|ds| \leqslant \frac{p_1 |dp| + p |dp_1|}{(p - p_1)^2}.$$

Пусть, например, кусок латуни весит в воздухе примерно 100 Γ с точностью до 5 $м\Gamma$, а в воде примерно 88 Γ с точностью до 8 $м\Gamma$. Тогда (в граммах веса) $p = 100$, $p_1 = 88$, $|dp| = 5 \cdot 10^{-3}$, $|dp_1| = 8 \cdot 10^{-3}$. Удельный вес куска латуни $s = \frac{25}{3} \frac{\Gamma}{см^3}$ с возможной ошибкой $|ds| \approx 8{,}6 \cdot 10^{-3} \frac{\Gamma}{см^3}$, что составляет приблизительно 1%.

§ 5. Сложные функции и введение новых независимых переменных

1. Сложные функции и их непрерывность. Часто бывает, что функция u от независимых переменных x, y задается в виде функции от промежуточных переменных ξ, η, ...:

$$u = f(\xi, \eta, \ldots),$$

где эти аргументы ξ, η, ... функции f сами являются функциями от x и y:

$$\xi = \varphi(x, y), \qquad \eta = \psi(x, y), \ldots$$

В таком случае функция

$$u = f(\xi, \eta, \ldots) = f[\varphi(x, y), \psi(x, y), \ldots] = F(x, y)$$

называется *сложной функцией* от x и y.

Например, функцию

$$u = e^{xy} \sin(x + y) = F(x, y)$$

можно рассматривать как сложную функцию

$$u = e^{\xi} \sin \eta = f(\xi, \eta), \quad \text{где} \quad \xi = xy, \qquad \eta = x + y.$$

Аналогично и функция

$$u = \ln(x^4 + y^4) \arcsin \sqrt{1 - x^2 - y^2} = F(x, y)$$

является сложной функцией от x и y, ибо ее можно записать в таком виде:

$$u = \eta \arcsin \xi = f(\xi, \eta), \text{ где } \quad \xi = \sqrt{1 - x^2 - y^2}, \ \eta = \ln(x^4 + y^4).$$

Для уточнения понятия сложной функции сделаем следующие предположения. Функции $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, ... будем считать заданными в некоторой области G независимых переменных x, y. Когда точка (x, y) пробегает область G, то точка с координатами ξ, η, ... лежит в некоторой области B пространства $\xi \eta$ ...; в этой-то области B и определена функция $u = f(\xi, \eta, \ldots)$. В конечном итоге сложная функция

$$u = f[\varphi(x, y), \quad \psi(x, y), \ldots] = F(x, y)$$

определена в области G.

Во многих случаях в таком детальном рассмотрении областей G и B нет необходимости. Так, в первом из рассмотренных только что примеров точка (x, y) может пробегать всю плоскость xy, да и функция $u = e^{\xi} \sin \eta$ определена во всей плоскости $\xi\eta$. Напротив, второй пример ясно показывает необходимость рассмотрения областей G и B при определении сложной функции. В самом деле, функции

$$\xi = \sqrt{1 - x^2 - y^2} \quad \text{и} \quad \eta = \ln(x^4 + y^4)$$

(в своей совокупности) заданы в области G, определенной неравенствами $0 < x^2 + y^2 \leqslant 1$, т. е. в круге радиуса 1 с центром в начале координат, с изъятием этого центра. В этой области G всюду $|\xi| < 1$, а η принимает все отрицательные значения и значение нуль, и в описанной таким образом области B плоскости $\xi\eta$ функция $u = \eta \arcsin \xi$ действительно определена.

Непрерывная функция от непрерывных функций сама является непрерывной функцией. Точнее:

Если функция $u = f(\xi, \eta, \ldots)$ непрерывна в области B, а промежуточные функции $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, ... непрерывны в области G, то и сложная функция $u = F(x, y)$ непрерывна в G.

Доказательство легко получается из определения непрерывности. Пусть (x_0, y_0) есть любая точка области G, а соответствующие этой точке значения промежуточных переменных ξ, η, ... обозначим через ξ_0, η_0, ... Тогда из непрерывности функции $u = f(\xi, \eta, \ldots)$ в области B вытекает, что для всякого положительного числа ε существует такое число $\delta > 0$, что

$$|f(\xi, \eta, \ldots) - f(\xi_0, \eta_0, \ldots)| < \varepsilon,$$

коль скоро одновременно $|\xi - \xi_0| < \delta$, $|\eta - \eta_0| < \delta$, ... Но, в силу непрерывности функций $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, ..., все последние неравенства выполняются, если $|x - x_0| < \gamma$ и $|y - y_0| < \gamma$, где γ — достаточно малое положительное число. Это и доказывает непрерывность сложной функции в области G.

2. Теорема о дифференцируемости сложной функции, составленной из дифференцируемых звеньев. Теперь мы докажем, что дифференцируемая функция от промежуточных аргументов, которые являются в свою очередь дифференцируемыми функциями независимых переменных, тоже представляет собой дифференцируемую функцию этих независимых переменных. Дадим сначала более точную формулировку этой теоремы:

Если $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, ... *суть дифференцируемые функции от* x *и* y *в области* G, *а* $f(\xi, \eta, \ldots)$ *есть дифференцируемая функция от своих аргументов в области* B, *то и сложная функция*

$$u = f[\varphi(x, y), \psi(x, y), \ldots] = F(x, y)$$

является дифференцируемой функцией независимых переменных x *и* y.

Для доказательства вспомним, что означает предположение, что наши функции дифференцируемы. Смысл его следующий. Дадим независимым переменным x и y приращения Δx и Δy; тогда приращения промежуточных функций (промежуточных аргументов) ξ, η, ... можно представить в следующем виде:

$$\Delta\xi = \varphi_x \Delta x + \varphi_y \Delta y + \varepsilon_1 \Delta x + \gamma_1 \Delta y,$$
$$\Delta\eta = \psi_x \Delta x + \psi_y \Delta y + \varepsilon_2 \Delta x + \gamma_2 \Delta y,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

где числа $\varepsilon_1, \varepsilon_2, \ldots; \gamma_1, \gamma_2, \ldots$ стремятся к нулю одновременно с Δx и Δy или (что то же самое) одновременно с $\sqrt{\Delta x^2 + \Delta y^2}$. Далее, если величинам ξ, η, ... сообщить приращения $\Delta\xi$, $\Delta\eta$, ..., то и функция $u = f(\xi, \eta, \ldots)$ получит приращение Δu, которое (в силу дифференцируемости последней функции) можно представить так:

$$\Delta u = f_\xi \Delta\xi + f_\eta \Delta\eta + \ldots + \delta_1 \Delta\xi + \delta_2 \Delta\eta + \ldots,$$

причем и здесь числа $\delta_1, \delta_2, \ldots$ стремятся к нулю одновременно с $\Delta\xi$, $\Delta\eta$, ... (одновременно с $\sqrt{\Delta\xi^2 + \Delta\eta^2 + \ldots}$). (Если приращения $\Delta\xi$, $\Delta\eta$, ... обращаются в нуль, то соответствующие δ_k можно принять равными нулю.)

Подставим теперь в последнее выражение для Δu, в качестве приращений $\Delta\xi$, $\Delta\eta$, ..., как раз те выражения, которые были выше получены в результате изменения x на Δx и y на Δy, и мы получим

$$\Delta u = (f_\xi \varphi_x + f_\eta \psi_x + \ldots) \Delta x + (f_\xi \varphi_y + f_\eta \psi_y + \ldots) \Delta y + \varepsilon \Delta x + \gamma \Delta y, \quad (*)$$

где через ε и γ обозначены следующие выражения:

$$\varepsilon = f_\xi \varepsilon_1 + f_\eta \varepsilon_2 + \ldots + \varphi_x \delta_1 + \psi_x \delta_2 + \ldots + \varepsilon_1 \delta_1 + \varepsilon_2 \delta_2 + \ldots,$$
$$\gamma = f_\xi \gamma_1 + f_\eta \gamma_2 + \ldots + \varphi_y \delta_1 + \psi_y \delta_2 + \ldots + \gamma_1 \delta_1 + \gamma_2 \delta_2 + \ldots,$$

В правой части каждого из этих равенств любое слагаемое содержит множителем по крайней мере одну из величин $\varepsilon_1, \varepsilon_2, \ldots, \gamma_1, \gamma_2, \ldots$ $\ldots, \delta_1, \delta_2, \ldots$ Отсюда видно, что ε и γ тоже стремятся к нулю, когда Δx и Δy стремятся к нулю.

На основании определения дифференцируемой функции отсюда вытекает, что сложная функция

$$u=f[\varphi(x, y), \psi(x, y), \ldots]$$

является дифференцируемой функцией от x и y.

Ясно, что полученный результат совершенно не зависит ни от количества независимых переменных $x, y, \ldots$, ни от числа промежуточных аргументов $\xi, \eta, \ldots$ Он, например, остается справедливым и в том случае, если промежуточные аргументы $\xi, \eta, \ldots$ зависят только от одной независимой переменнной x, так что и величина u является сложной функцией от единственной независимой переменной x. Он сохраняет также силу, если сложная функция u зависит лишь от одного промежуточного аргумента ξ, который сам является функцией от независимых переменных $x, y, \ldots$

3. Вычисление частных производных от сложной функции — правило цепочки. Из формулы (*) предыдущего номера, на основании доказанного в § 4, n° 4 предложения, получается важная формула для вычисления частных производных от сложной функции:

$$F_x=f_\xi\varphi_x+f_\eta\psi_x+\ldots,$$
$$F_y=f_\xi\varphi_y+f_\eta\psi_y+\ldots$$

или, в менее точной, но более удобной для запоминания записи:

$$u_x=u_\xi\xi_x+u_\eta\eta_x+\ldots,$$
$$u_y=u_\xi\xi_y+u_\eta\eta_y+\ldots$$

Еще одна запись этой формулы:

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x}+\frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}+\ldots,$$
$$\frac{\partial u}{\partial y}=\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial y}+\frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial y}+\ldots$$

Для того чтобы найти частную производную по x, надо вычислить частные производные от сложной функции по всем промежуточным аргументам $\xi, \eta, \ldots$, каждую из этих частных производных помножить на производную от соответствующего промежуточного аргумента по x и затем сложить все полученные произведения. Это правило является обобщением на функции многих переменных того правила цепочки, которое было выведено для функций одной переменной в т. I, гл. III, § 4, n° 1.

Ясно, что в последней системе формул должно быть столько строк (формул), сколько имеется независимых переменных $x, y, \ldots$

Если, в частности, сложная функция зависит лишь от одной независимой переменной x, то будет лишь одна формула, которая примет следующий вид:

$$\frac{du}{dx}=\frac{\partial u}{\partial \xi}\frac{d\xi}{dx}+\frac{\partial u}{\partial \eta}\frac{d\eta}{dx}+\ldots$$

Если сложная функция зависит лишь от одного промежуточного аргумента ξ, который сам зависит от x и y, то правило цепочки запишется так:

$$\frac{\partial u}{\partial x}=\frac{du}{d\xi}\frac{\partial \xi}{\partial x}, \quad \frac{\partial u}{\partial y}=\frac{du}{d\xi}\frac{d\xi}{\partial y}.$$

Для вычисления частных производных второго порядка от сложной функции надо полученные выражения для частных производных первого порядка вновь дифференцировать по x и по y, рассматривая f_ξ, $f_\eta, \ldots$ как сложные функции. Ограничиваясь, в целях простоты, тремя промежуточными аргументами ξ, η, ζ, получим:

$$u_{xx}=f_{\xi\xi}\xi_x^2+f_{\eta\eta}\eta_x^2+f_{\zeta\zeta}\zeta_x^2+2f_{\xi\eta}\xi_x\eta_x+2f_{\eta\zeta}\eta_x\zeta_x+$$
$$+2f_{\xi\zeta}\xi_x\zeta_x+f_\xi\xi_{xx}+f_\eta\eta_{xx}+f_\zeta\zeta_{xx},$$
$$u_{xy}=f_{\xi\xi}\xi_x\xi_y+f_{\eta\eta}\eta_x\eta_y+f_{\zeta\zeta}\zeta_x\zeta_y+f_{\xi\eta}(\xi_x\eta_y+\xi_y\eta_x)+$$
$$+f_{\eta\zeta}(\eta_x\zeta_y+\eta_y\zeta_x)+f_{\xi\zeta}(\xi_x\zeta_y+\xi_y\zeta_x)+f_\xi\xi_{xy}+f_\eta\eta_{xy}+f_\zeta\zeta_{xy},$$
$$u_{yy}=f_{\xi\xi}\xi_y^2+f_{\eta\eta}\eta_y^2+f_{\zeta\zeta}\zeta_y^2+2f_{\xi\eta}\xi_y\eta_y+2f_{\eta\zeta}\eta_y\zeta_y+$$
$$+2f_{\xi\zeta}\xi_y\zeta_y+f_\xi\xi_{yy}+f_\eta\eta_{yy}+f_\zeta\zeta_{yy}.$$

Дифференцируя эти формулы, можно получить выражения для частных производных третьего порядка и т. д.

Решим несколько примеров. Подчеркнем, однако, что конкретно заданные сложные функции можно продифференцировать и непосредственно, не пользуясь правилом цепочки для функций многих переменных.

1) Рассмотрим функцию $u=e^{x^2\sin^2 y+2xy\sin x\sin y+y^2}$. Положим $\xi=x^2\sin^2 y$, $\eta=2xy\sin x\sin y$, $\zeta=y^2$. Тогда $u=e^{\xi+\eta+\zeta}$, и мы получим:

$$\xi_x=2x\sin^2 y, \qquad \eta_x=2y\sin x\sin y+2xy\cos x\sin y, \qquad \zeta_x=0,$$
$$\xi_y=2x^2\sin y\cos y, \qquad \eta_y=2x\sin x\sin y+2xy\sin x\cos y, \qquad \zeta_y=2y,$$
$$u_\xi=u_\eta=u_\zeta=e^{\xi+\eta+\zeta}.$$

Следовательно,

$$u_x=2e^{x^2\sin^2 y+2xy\sin x\sin y+y^2}(x\sin^2 y+y\sin x\sin y+xy\cos x\sin y)$$

и

$$u_y=2e^{x^2\sin^2 y+2xy\sin x\sin y+y^2}(x^2\sin y\cos y+x\sin x\sin y+xy\sin x\cos y+y).$$

2) Функция $u = \sin(x^2+y^2)$. Положим $\xi = x^2+y^2$; тогда $u = \sin\xi$. Имеем:

$$u_x = 2x\cos(x^2+y^2), \qquad u_y = 2y\cos(x^2+y^2),$$
$$u_{xx} = -4x^2\sin(x^2+y^2)+2\cos(x^2+y^2),$$
$$u_{xy} = -4xy\sin(x^2+y^2),$$
$$u_{yy} = -4y^2\sin(x^2+y^2)+2\cos(x^2+y^2).$$

3) Для функции $u = \operatorname{arctg}(x^2+xy+y^2)$ введем обозначения $\xi = x^2$, $\eta = xy$, $\zeta = y^2$; тогда $u = \operatorname{arctg}(\xi+\eta+\zeta)$. Получим

$$u_x = \frac{2x+y}{1+(x^2+xy+y^2)^2}, \qquad u_y = \frac{x+2y}{1+(x^2+xy+y^2)^2}.$$

4) Рассмотрим пример сложной функции, зависящей в конечном итоге лишь от одной независимой переменной x:

$$F(x) = [\varphi(x)]^{\psi(x)}.$$

Положив $\xi = \varphi(x)$, $\eta = \psi(x)$, получим $u = F(x) = \xi^\eta$. Имеем $u_\xi = \eta\xi^{\eta-1}$, $u_\eta = \xi^\eta\ln\xi$, $\dfrac{d\xi}{dx} = \varphi'(x)$, $\dfrac{d\eta}{dx} = \psi'(x)$. По правилу цепочки

$$\frac{dF(x)}{dx} = u_\xi\frac{d\xi}{dx} + u_\eta\frac{d\eta}{dx} = \eta\xi^{\eta-1}\varphi'(x) + \xi^\eta\ln\xi\psi'(x) =$$
$$= [\varphi(x)]^{\psi(x)}\left[\psi(x)\cdot\frac{\varphi'(x)}{\varphi(x)} + \psi'(x)\ln\varphi(x)\right].$$

Этот результат был получен уже в т. I, стр. 229, с помощью несколько искусственного приема.

[4. Полный дифференциал сложной функции. Инвариантность полного дифференциала первого порядка. Рассмотрим сложную функцию

$$u = f(\xi, \eta, \ldots), \qquad \text{где} \qquad \xi = \varphi(x, y), \qquad \eta = \psi(x, y), \ldots$$

Формула (*), стр. 87, доказывает, что если промежуточные функции $f(\xi, \eta, \ldots)$, $\varphi(x, y)$, $\psi(x, y)$, ... дифференцируемы, то и сложная функция $u = f[\varphi(x, y), \psi(x, y), \ldots] = F(x, y)$ дифференцируема, и ее линейная часть, т. е. ее полный дифференциал, выражается так:

$$du = (u_\xi\xi_x + u_\eta\eta_x + \ldots)\Delta x + (u_\xi\xi_y + u_\eta\eta_y + \ldots)\Delta y =$$
$$= (u_\xi\xi_x + u_\eta\eta_x + \ldots)dx + (u_\xi\xi_y + u_\eta\eta_y + \ldots)dy.$$

Преобразуем правую часть этой формулы следующим образом: сперва раскроем скобки, а затем вынесем за скобку величины u_ξ, u_η, ... из всех членов, в которых они содержатся. В результате получим

$$du = u_\xi(\xi_x dx + \xi_y dy) + u_\eta(\eta_x dx + \eta_y dy) + \ldots$$

Теперь выражения в скобках представляют собой полные дифференциалы промежуточных аргументов ξ, η и т. д. как функций независимых переменных x и y:

$$d\xi = \xi_x dx + \xi_y dy, \qquad d\eta = \eta_x dx + \eta_y dy, \ldots$$

Следовательно,

$$du = u_\xi d\xi + u_\eta d\eta + \ldots$$

Нетрудно убедиться, что именно такое выражение получилось бы для полного дифференциала, если бы величины ξ, $\eta, \ldots$ были независимыми переменными. Стало быть, доказана следующая теорема:

Выражение для полного дифференциала функции $u = f(\xi, \eta, \ldots)$ сохраняет свой вид независимо от того, являются ли $\xi, \eta, \ldots$ независимыми переменными или только промежуточными аргументами, которые сами являются функциями от независимых переменных $x, y, \ldots$ Из хода рассуждений ясно, что этот результат справедлив не только при любом числе промежуточных аргументов $\xi, \eta, \ldots$, но и при любом числе независимых переменных $x, y, \ldots$

Эта теорема называется теоремой об *инвариантности* (неизменности) полного дифференциала. Она является обобщением свойства инвариантности дифференциала функции одной переменной (т. I, стр. 184—185).

Доказанная теорема дает возможность распространить на функции многих переменных известные формулы для дифференциалов функций одной переменной.

Пусть $\xi = \varphi(x, y, \ldots)$, $\eta = \psi(x, y, \ldots)$. Тогда для полных дифференциалов этих функций справедливы следующие формулы:

$$d(c\xi) = c\,d\xi, \quad \text{если} \quad c = \text{const},$$

$$d(\xi \pm \eta) = d\xi \pm d\eta,$$

$$d(\xi\eta) = \eta\,d\xi + \xi\,d\eta,$$

$$d\left(\frac{\xi}{\eta}\right) = \frac{\eta\,d\xi - \xi\,d\eta}{\eta^2}.$$

Докажем здесь последнюю формулу, остальные читатель сам легко докажет тем же способом. Имеем сложную функцию $u = \frac{\xi}{\eta}$, где $\xi = \varphi(x, y, \ldots)$, $\eta = \psi(x, y, \ldots)$. По свойству инвариантности полный дифференциал $du = u_\xi\,d\xi + u_\eta\,d\eta$, причем $u_\xi = \frac{1}{\eta}$, $u_\eta = -\frac{\xi}{\eta^2}$; следовательно, $du = \frac{1}{\eta}\,d\xi - \frac{\xi}{\eta^2}\,d\eta = \frac{\eta\,d\xi - \xi\,d\eta}{\eta^2}$, и формула доказана.

Стало быть, вычисления с полными дифференциалами функций многих переменных можно вести по тем же правилам, что и с дифференциалами функций одной переменной.

Например, для функции $r=\sqrt{x^2+y^2}$ имеем

$$dr=\frac{1}{2\sqrt{x^2+y^2}}d(x^2+y^2)=\frac{d(x^2)+d(y^2)}{2\sqrt{x^2+y^2}}=\frac{2x\,dx+2y\,dy}{2\sqrt{x^2+y^2}}=\frac{x\,dx+y\,dy}{\sqrt{x^2+y^2}}.$$

Теперь, из найденного полного дифференциала сразу получим обе частные производные

$$r_x=\frac{x}{r},\qquad r_y=\frac{y}{r}.$$

Точно так же и для функции $\theta=\operatorname{arctg}\frac{y}{x}$ выгодно раньше найти полный дифференциал

$$d\theta=\frac{1}{1+\frac{y^2}{x^2}}\,d\left(\frac{y}{x}\right)=\frac{1}{1+\frac{y^2}{x^2}}\cdot\frac{x\,dy-y\,dx}{x^2}=\frac{x\,dy-y\,dx}{x^2+y^2},$$

а отсюда определить частные производные

$$\theta_x=-\frac{y}{x^2+y^2},\qquad \theta_y=\frac{x}{x^2+y^2}.]$$

5. Введение новых независимых переменных. Материал, изложенный в номерах 1—4, находит особенно важное применение в вопросе о введении новых независимых переменных. Возьмем, например, функцию $u=f(\xi,\eta)$ от двух независимых переменных ξ и η, которые мы рассматриваем как прямоугольные декартовы координаты на плоскости. Повернем теперь систему координат в своей плоскости на какой-либо угол и введем новые прямоугольные координаты x, y относительно новых осей с помощью преобразования

$$\xi=\alpha_1x+\beta_1y,\qquad x=\alpha_1\xi+\alpha_2\eta,$$
$$\eta=\alpha_2x+\beta_2y,\qquad y=\beta_1\xi+\beta_2\eta.$$

(Ср. Смешанные упражнения к гл. I, упр. 3.) Тогда функция $u=f(\xi,\eta)$ преобразуется в новую функцию

$$u=f(\xi,\eta)=F(x,y),$$

и величина u будет сложной функцией по отношению к переменным x и y. В этом случае говорят, что в соотношение $u=f(\xi,\eta)$ введены новые независимые переменные x и y. С помощью правила цепочки из n° 3 находим частные производные по x и y:

$$u_x=u_\xi\alpha_1+u_\eta\alpha_2,$$
$$u_y=u_\xi\beta_1+u_\eta\beta_2,$$

причем u_x, u_y обозначают производные от функции $F(x,y)$, а u_ξ, u_η — производные от функции $f(\xi,\eta)$.

Полученные формулы показывают, что частные производные от любой функции преобразовываются при повороте системы координат по тем же формулам, что и независимые переменные (координаты точки). Этот результат справедлив и при введении новых прямоугольных (декартовых) координат в пространстве.

Другой важный пример введения новых независимых переменных дает переход от прямоугольных координат x, y к полярным координатам r, θ; этот переход выражается преобразованием

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

$$r = \sqrt{x^2+y^2}, \qquad \theta = \begin{cases} \operatorname{arctg}\dfrac{y}{x} & \text{при} \quad x \geqslant 0, \\ \operatorname{arctg}\dfrac{y}{x} + \pi & \text{при} \quad x < 0. \end{cases}$$

При введении полярных координат функция $u = f(x, y)$ преобразуется так:

$$u = f(x, y) = f(r\cos\theta, r\sin\theta) = F(r, \theta),$$

и величина u оказывается сложной функцией независимых переменных r и θ. По правилу цепочки получаем

$$u_r = u_x x_r + u_y y_r = u_x\cos\theta + u_y\sin\theta = \frac{x}{r}u_x + \frac{y}{r}u_y,$$

$$u_\theta = u_x x_\theta + u_y y_\theta = u_x(-r\sin\theta) + u_y r\cos\theta = -yu_x + xu_y.$$

Первое из этих равенств дает в сущности производную от функции $u = f(x, y)$ по направлению радиус-вектора r; мы с ней уже встретились на стр. 79—80. Второе равенство дает производную той же функции по направлению, перпендикулярному к радиус-вектору. Итак,

$$\frac{\partial u}{\partial r} = D_r u = D_{(\theta)}u, \qquad \frac{\partial u}{\partial\theta} = D_{\left(\theta+\frac{\pi}{2}\right)}u.$$

Однако часто приходится рассматривать x, y как независимые переменные, а r и θ как промежуточные аргументы. Тогда

$$u_x = u_r r_x + u_\theta\theta_x, \qquad u_y = u_r r_y + u_\theta\theta_y.$$

Внеся сюда (см. примеры в конце предыдущего номера)

$$r_x = \frac{x}{r} = \cos\theta, \qquad r_y = \frac{y}{r} = \sin\theta,$$

$$\theta_x = -\frac{y}{x^2+y^2} = -\frac{y}{r^2}, \qquad \theta_y = \frac{x}{x^2+y^2} = \frac{x}{r^2},$$

получим

$$u_x = u_r\frac{x}{r} - u_\theta\frac{y}{r^2} = u_r\cos\theta - u_\theta\frac{\sin\theta}{r},$$

$$u_y = u_r\frac{y}{r} + u_\theta\frac{x}{r^2} = u_r\sin\theta + u_\theta\frac{\cos\theta}{r}.$$

Из этих формул легко вывести еще одно часто встречающееся тождество

$$u_x^2 + u_y^2 = u_r^2 + \frac{1}{r^2}u_\theta^2.$$

Дифференцируя по правилу цепочки выражения для частных производных первого порядка, получим следующие формулы для частных производных второго порядка:

$$u_{xx}=u_{rr}\cos^2\theta+u_{\theta\theta}\frac{\sin^2\theta}{r^2}-2u_{r\theta}\frac{\cos\theta\sin\theta}{r}+u_r\frac{\sin^2\theta}{r}+2u_\theta\frac{\cos\theta\sin\theta}{r^2},$$

$$u_{xy}=u_{yx}=u_{rr}\cos\theta\sin\theta-u_{\theta\theta}\frac{\cos\theta\sin\theta}{r^2}+u_{r\theta}\frac{\cos^2\theta-\sin^2\theta}{r}+$$
$$+u_\theta\frac{\sin^2\theta-\cos^2\theta}{r^2}-u_r\frac{\sin\theta\cos\theta}{r},$$

$$u_{yy}=u_{rr}\sin^2\theta+u_{\theta\theta}\frac{\cos^2\theta}{r^2}+2u_{r\theta}\frac{\cos\theta\sin\theta}{r}+u_r\frac{\cos^2\theta}{r}-2u_\theta\frac{\cos\theta\sin\theta}{r^2}.$$

Почленное сложение выражений для u_{xx} и u_{yy} приводит к следующей формуле, дающей преобразование к полярным координатам так называемого дифференциального выражения Лапласа $u_{xx}+u_{yy}$, играющего большую роль в теории потенциала:

$$u_{xx}+u_{yy}=u_{rr}+u_{\theta\theta}\frac{1}{r^2}+u_r\frac{1}{r}=\frac{1}{r^2}\left\{r\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{\partial^2 u}{\partial\theta^2}\right\}.$$

Вообще всякий раз, когда задаются соотношения

$$u=f(\xi,\ \eta,\ \ldots),\qquad \xi=\varphi(x,y),\qquad \eta=\psi(x,y),\ \ldots,$$

определяющие сложную функцию, ее можно рассматривать как введение новых независимых переменных x, y вместо $\xi, \eta, \ldots$; при этом соответствующие друг другу системы значений переменных $\xi, \eta, \ldots$ и x, y приводят всегда к одному и тому же значению u, безразлично, рассматривается ли u как функция от одной или от другой системы переменных.

Когда идет речь о дифференцировании сложной функции, надо всегда иметь в виду следующее обстоятельство. Надо четко различать между зависимой переменной u и символом функции f, который обозначает функциональную зависимость между величиной u и аргументами $\xi, \eta, \ldots$: $u=f(\xi, \eta, \ldots)$. Символы дифференцирования u_ξ, u_η, ... приобретают смысл лишь тогда, когда конкретизирована функциональная зависимость между величиной u и независимыми переменными. Поэтому, строго говоря, для сложной функции

$$u=f(\xi,\ \eta,\ \ldots)=F(x,\ y)$$

следовало бы писать не u_ξ, u_η, ..., а также u_x и u_y, но f_ξ, f_η, ... и F_x, F_y. Однако для краткости часто пользуются более простыми обозначениями u_ξ, u_η, ..., u_x, u_y во всех случаях, когда не опасаются недоразумений.

Покажем на примере, что дифференцирование какой-либо величины получает совершенно различный смысл, смотря по характеру функциональной зависимости между нею и независимыми переменными, и результат дифференцирования по одной и той же переменной зависит от того, какие другие величины (в качестве независимых переменных) сохраняют при этом постоянные значения. Величина $u = 2\xi + \eta = f(\xi, \eta)$ при «тождественном» преобразовании $\xi = x$, $\eta = y$ переходит в функцию $u = 2x + y$; следовательно, $u_x = 2$; $u_y = 1$. Если же мы вместо этого совершим другое преобразование $\xi = x$ (как в первый раз) и $\eta = -x + t$, то $u = x + t$, откуда $u_x = 1$ и $u_t = 1$. Стало быть, дифференцирование по одной и той же независимой переменной $x = \xi$ привело к различным результатам.

Упражнения

1. Составить уравнение касательной плоскости к поверхности второго порядка $ax^2 + by^2 + cz^2 = 1$ в точке (x_0, y_0, z_0) этой поверхности.

2. Показать, что если $u = u(x, y)$ есть уравнение конической поверхности, то u удовлетворяет дифференциальному уравнению с частными производными

$$u_{xx}u_{yy} - (u_{xy})^2 = 0.$$

3. Доказать, что если функция $f(x)$ непрерывна на отрезке $[a, b]$ и имеет производную в интервале (a, b), то производная от функции

$$g(x) = \begin{vmatrix} f(x) & x & 1 \\ f(a) & a & 1 \\ f(b) & b & 1 \end{vmatrix}$$

обращается в нуль при некотором значении x, лежащем между a и b.

4. Доказать теорему: если элементы определителя третьего порядка являются функциями переменной x, то производная определителя по x равна сумме трех определителей, каждый из которых получается из данного определителя дифференцированием по x одной из его строк.

Формулировать и доказать аналогичную теорему для определителей второго и четвертого порядка.

5. Вычислить

$$\frac{d^2}{dx^2}\begin{vmatrix} a+x & b & c \\ d & e+x & f \\ g & h & k+x \end{vmatrix}.$$

6. Доказать, что

$$\frac{\partial^3}{\partial x\,\partial y\,\partial z}\begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1(y) & g_2(y) & g_3(y) \\ h_1(z) & h_2(z) & h_3(z) \end{vmatrix} = \begin{vmatrix} f_1'(x) & f_2'(x) & f_3'(x) \\ g_1'(y) & g_2'(y) & g_3'(y) \\ h_1'(z) & h_2'(z) & h_3'(z) \end{vmatrix}.$$

7. Пусть функция $f(x, y, z)$ зависит только от $r = \sqrt{x^2 + y^2 + z^2}$, т. е. $f(x, y, z) = g(r)$.

а) Вычислить $f_{xx} + f_{yy} + f_{zz}$.

б) Доказать, что если $f(x, y, z) = g(r)$ удовлетворяет уравнению Лапласа $f_{xx} + f_{yy} + f_{zz} = 0$, то $f = g(r) = \frac{a}{r} + b$, где a и b — постоянные.

8. Функция $f(x_1, x_2, \ldots, x_n) = g(r)$, где $r = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$. Вычислить

$$f_{x_1x_1} + f_{x_2x_2} + \ldots + f_{x_nx_n}.$$

9. Если $f(x_1, x_2, \ldots, x_n) = g(r)$ (см. упр. 8) удовлетворяет уравнению Лапласа

$$f_{x_1x_1} + f_{x_2x_2} + \ldots + f_{x_nx_n} = 0,$$

то $g(r)=\frac{a}{r^{n-2}}+b$ при $n>2$ и $g(r)=a\ln r+b$ при $n=2$, где a и b — постоянные (Ср. выше упр. 7б и упр. 2, стр. 73).

10*. Преобразовать дифференциальное выражение Лапласа $f_{xx}+f_{yy}+f_{zz}$ к полярным координатам в пространстве (сферическим координатам), т. е. ввести вместо x, y, z новые переменные r, θ, φ с помощью формул

$$x=r\sin\theta\cos\varphi,\qquad y=r\sin\theta\sin\varphi,\qquad z=r\cos\theta.$$

(Ср. с упр. 7а.)

11. Доказать, что выражение $f_{xx}+f_{yy}$ сохраняет свой вид при повороте системы координат на любой угол.

12. Доказать, что при линейном преобразовании

$$x=\alpha\xi+\beta\eta,\qquad y=\gamma\xi+\delta\eta$$

вторые производные $f_{xx}(x, y)$, $f_{xy}(x, y)$, $f_{yy}(x, y)$ преобразуются соответственно по тому же закону, что и коэффициенты a, b, c многочлена $ax^2+2bxy+cy^2$.

§ 6. Теорема о среднем значении и формула Тэйлора для функции многих переменных

1. Постановка задачи и предварительные замечания. В первом томе (гл. VI, § 2) мы научились аппроксимировать функцию одной переменной, имеющую производные до $n+1$-го порядка включительно, в окрестности данной точки многочленом, который мы назвали *аппроксимирующим многочленом Тэйлора*. То приближение с помощью линейной части функции, которое дает дифференциал, является лишь первым шагом к этому более точному приближению, точность которого возрастает с возрастанием n. Можно теперь попытаться и функцию многих (например, двух) независимых переменных приближенно представить в окрестности данной точки многочленом степени n. При подходящих обозначениях вопрос сводится к тому, чтобы при любом n найти для $f(x+h, y+k)$ «аппроксимирующий многочлен Тэйлора» степени n относительно приращений h и k с коэффициентами, зависящими от выбора постоянной точки (x, y).

Простая мысль дает возможность привести эту задачу к соответствующей уже решенной задаче для функции одной переменной. Введем вспомогательную переменную t и вместо функции $f(x+h, y+k)$ исследуем выражение

$$F(t)=f(x+ht, y+kt)$$

как функцию от одной независимой переменной t, считая временно x, y, h и k постоянными. При изменении t от 0 до 1 точка $(x+ht, y+kt)$ пробегает прямолинейный отрезок от точки (x, y) до точки $(x+h, y+k)$.

Сначала вычислим производные функции $F(t)$. Предполагая, что все частные производные от функции $f(x, y)$, которые понадобятся ниже, непрерывны в некоторой области, содержащей упомянутый выше отрезок, получим по правилу цепочки (§ 5, n° 3)

$$F'(t)=h\frac{\partial f}{\partial x}+k\frac{\partial f}{\partial y},$$

$$F''(t)=h^2\frac{\partial^2 f}{\partial x^2}+2hk\frac{\partial^2 f}{\partial x\,\partial y}+k^2\frac{\partial^2 f}{\partial y^2},$$

. . . .

И вообще, пользуясь методом полной индукции, получим для производной порядка n следующее выражение:

$$F^{(n)}(t)=h^n f_{x^n}+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}+\binom{n}{2}h^{n-2}k^2 f_{x^{n-2}y^2}+\ldots+k^n f_{y^n},$$

которое, как и формулу для n-го дифференциала на стр. 84, можно записать символически в следующем виде:

$$F^{(n)}(t)=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{(n)}f.$$

Скобки в правой части надо раскрыть формально по формуле бинома Ньютона, а полученное развернутое выражение «помножить» затем почленно на f. Члены полученного разложения будут содержать производные n-го порядка $\frac{\partial^n f}{\partial x^n}$, $\frac{\partial^n f}{\partial x^{n-1}\partial y}$, $\frac{\partial^n f}{\partial x^{n-2}\partial y^2}$ и т. д. Во всех этих производных следует в качестве аргументов писать вместо x и y величины $x+ht$ и $y+kt$.

2. Теорема о среднем значении. Исходным пунктом для построения искомого аппроксимирующего многочлена послужит *теорема о среднем значении*, аналогичная уже известной нам теореме о среднем значении для функции одной независимой переменной.

Эта теорема связывает приращение $f(x+h, y+k)-f(x, y)$ с частными производными f_x и f_y. При выводе формулы, выражающей эту теорему, мы определенно предполагаем, что эти частные производные непрерывны. Применим известную нам теорему о среднем значении к функции одной переменной $F(t)$ для интервала от 0 до t:

$$\frac{F(t)-F(0)}{t}=F'(\theta t),$$

где θ есть какое-то промежуточное число между 0 и 1. Подставив сюда выражения функции $F(t)$ и ее производной $F'(t)$ из n° 1, получим

$$\frac{f(x+ht, y+kt)-f(x, y)}{t}=$$

$$=hf_x(x+\theta ht, y+\theta kt)+kf_y(x+\theta ht, y+\theta kt).$$

Полагая здесь $t=1$, получим следующую формулу, выражающую *теорему о среднем значении для функции двух переменных*:

$$f(x+h, y+k)-f(x, y)=hf_x(x+\theta h, y+\theta k)+kf_y(x+\theta h, y+\theta k)=$$

$$=hf_x(\xi, \eta)+kf_y(\xi, \eta).$$

Словесно теорему можно формулировать так:

Приращение функции $f(x, y)$ при переходе от точки $P(x, y)$ к точке $Q(x+h, y+k)$ равно полному дифференциалу функции в некоторой промежуточной точке (ξ, η) прямолинейного отрезка PQ.

Подчеркиваем, что при определении промежуточных значений обеих переменных x и y как в f_x, так и в f_y все четыре раза участвует *одно и то же* значение θ $(0<\theta<1)$.

Из теоремы о среднем значении вытекает, в качестве простого следствия, такая теорема: *функция $f(x, y)$, частные производные которой f_x и f_y в некоторой области всюду существуют и равны нулю, есть постоянная.* В самом деле, при выполнении этих условий правая часть формулы, выражающей теорему о среднем значении, обращается в нуль; стало быть, $f(x+h, y+k)=f(x, y)$ при произвольных значениях h и k, так что функция имеет всюду в рассматриваемой области одно и то же значение. Обратно, если функция $f(x, y)$ сводится к постоянной, то обе ее частные производные f_x и f_y всюду существуют и равны нулю. Следовательно, обращение в нуль производных f_x и f_y является необходимым и достаточным условием постоянства функции $f(x, y)$ в данной области.

3. Формула Тэйлора для функции многих переменных. Напишем для той же функции $F(t)$ формулу Тэйлора [т. I, формула (B) на стр. 369] с остаточным членом Лагранжа (т. I, стр. 372), а затем положим $t=1$; тогда получится *формула Тэйлора для функции двух переменных*

$$f(x+h, y+k)=f(x, y)+\{hf_x(x, y)+kf_y(x, y)\}+$$
$$+\frac{1}{2!}\{h^2f_{xx}(x, y)+2hkf_{xy}(x, y)+k^2f_{yy}(x, y)\}+\ldots$$
$$\ldots+\frac{1}{n!}\left\{h^nf_{x^n}(x, y)+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}(x, y)+\ldots+k^nf_{y^n}(x, y)\right\}+R_n$$

с остаточным членом, записанным в символическом виде

$$R_n=\frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{(n+1)}f(x+\theta h, y+\theta k), \qquad 0<\theta<1.$$

Формула Тэйлора представляет приращение $f(x+h, y+k)-f(x, y)$ в виде суммы однородных многочленов первой, второй, ..., n-й и $(n+1)$-й степени относительно h и k. Первые n многочленов $\left(\text{если опустить множители } \frac{1}{1!}, \frac{1}{2!}, \ldots, \frac{1}{n!}\right)$ совпадают с дифференциалами функции $f(x, y)$ первого, второго, ..., n-го порядка в точке (x, y):

$$df=hf_x+kf_y,$$
$$d^2f=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{(2)}f(x, y)=h^2f_{xx}+2hkf_{xy}+k^2f_{yy},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$d^nf=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{(n)}f(x, y)=h^nf_{x^n}+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}+\ldots+k^nf_{y^n},$$

а многочлен $(n+1)$-й степени, фигурирующий в остаточном члене $\left(\text{без множителя } \frac{1}{(n+1)!}\right)$, есть полный дифференциал $(n+1)$-го порядка $d^{n+1}f$ в промежуточной точке (ξ, η) прямолинейного отрезка, соединяющего точки (x, y) и $(x+h, y+k)$. Поэтому можно формулу Тэйлора переписать в следующем более сжатом и легче обозримом виде:

$$f(x+h, y+k)=f(x, y)+df(x, y)+\frac{1}{2!}d^2f(x, y)+\ldots$$
$$\ldots+\frac{1}{n!}d^nf(x, y)+R_n,$$

где

$$R_n=\frac{1}{(n+1)!}d^{n+1}f(x+\theta h, y+\theta k) \qquad (0<\theta<1).$$

При этом, вообще говоря, порядок малости остаточного члена выше порядка малости последнего известного точно члена $d^nf(x, y)$, т. е. при $h\to 0$ и $k\to 0$ будет $R_n=o\left[\left(\sqrt{h^2+k^2}\right)^n\right]$.

В теории формулы Тэйлора для функции одной переменной важную роль играл предельный переход $n\to\infty$ при условии, что он влечет за собой $R_n\to 0$. В этом случае получался *бесконечный ряд Тэйлора*, который дает возможность разложить в степенной ряд обширный класс функций одной переменной. У функции многих переменных такой метод является в общем виде слишком сложным. Здесь мы еще в большей степени, чем для функции одной переменной, подчеркнем главным образом тот факт, что формула Тэйлора представляет приращение $f(x+h, y+k)-f(x, y)$ функции $f(x, y)$ в виде суммы дифференциалов различных порядков df, d^2f, ...

Упражнения

1. Найти многочлен второй степени, дающий наилучшее приближение к функции $f(x, y)=\sin x \sin y$ в окрестности начала координат.

2. Доказать, что функцию e^{-y^2+2xy} можно разложить в ряд вида

$$\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}y^n,$$

сходящийся при всех значениях x и y, и что

а) все $H_n(x)$ являются многочленами степени n (это так называемые многочлены Эрмита);

б) $H_n'(x)=2nH_{n-1}(x)$ при $n\geqslant 1$;

в) $H_{n+1}(x)-2xH_n(x)+2nH_{n-1}(x)=0$ при $n\geqslant 1$;

г) $H_n''(x)-2xH_n'(x)+2nH_n(x)=0$ при $n\geqslant 0$.

3. Найти ряды Тэйлора для следующих функций и выяснить область их пригодности:

а) $\frac{1}{1-x-y}$; б) e^{x+y}.

§ 7. Применение векторных методов

Многочисленные факты и соотношения дифференциального и интегрального исчисления функций нескольких переменных значительно выигрывают в прозрачности и простоте, если применить к ним понятия и обозначения векторного исчисления. Поэтому мы в заключение этой главы добавим еще некоторые исследования относящихся сюда вопросов.

1. Векторная и скалярная функция точки — векторное и скалярное поле. Шаг, связывающий векторное исчисление с нашей темой, заключается в следующем. Вместо того чтобы рассматривать, как в главе I, один-единственный вектор или несколько постоянных векторов, мы исследуем *векторное многообразие*, зависящее от одного или нескольких непрерывно изменяющихся параметров, [векторные и скалярные функции точки].

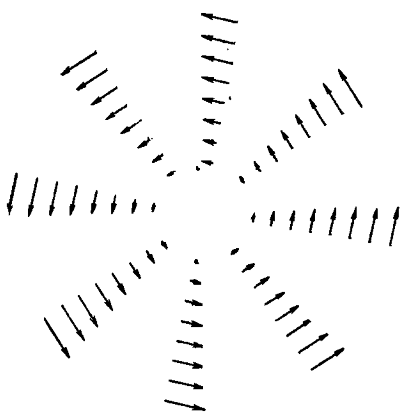


Рис. 25.

Рассмотрим, например, вещество, заполняющее некоторую область пространства и находящееся в состоянии движения. Тогда в заданный момент времени движущаяся материя имеет в каждой точке определенную скорость, представленную вектором $\boldsymbol{u}$. Принято говорить, что эти векторы $\boldsymbol{u}$ образуют *векторное поле* в рассматриваемой области. Переменный вектор $\boldsymbol{u}$ называется *вектором поля*. Он представляет собой векторную функцию от вектора положения, т. е. радиус-вектора $\boldsymbol{r}$ переменной точки области (поля) $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{r})$, а три координаты вектора поля предстают перед нами как три функции

$$u_1(x, y, z), \qquad u_2(x, y, z), \qquad u_3(x, y, z)$$

от трех координат точки наблюдения.

На рис. 25 в качестве примера векторного поля изображено поле скоростей твердого тела, вращающегося с постоянной угловой скоростью вокруг оси, перпендикулярной к плоскости чертежа.

Силы, действующие на точки движущейся в пространстве материи, тоже образуют векторное поле — силовое поле. В качестве примера силового поля можно привести поле силы притяжения, с которой материальная точка действует по закону тяготения на единичные точечные массы. Согласно закону Ньютона, все векторы этого силового поля направлены к притягивающей точке, а их модули обратно пропорциональны квадрату расстояния точки приложения от притягивающей точки.

Если путем поворота декартовой прямоугольной системы координат перейти к новой системе, то все векторы поля будут иметь новые координаты по отношению к новой системе осей. Если новые координаты точек ξ, η, ζ связаны с их старыми координатами формулами преобразования (гл. I, § 1, n° 4)

$$\begin{aligned}\xi &= \alpha_1 x + \beta_1 y + \gamma_1 z,\\ \eta &= \alpha_2 x + \beta_2 y + \gamma_2 z,\\ \zeta &= \alpha_3 x + \beta_3 y + \gamma_3 z\end{aligned}$$

и вытекающими отсюда обратными формулами

$$\begin{aligned}x &= \alpha_1\xi + \alpha_2\eta + \alpha_3\zeta,\\ y &= \beta_1\xi + \beta_2\eta + \beta_3\zeta,\\ z &= \gamma_1\xi + \gamma_2\eta + \gamma_3\zeta,\end{aligned}$$

то новые координаты вектора поля ω_1, ω_2, ω_3 связаны с его старыми координатами u_1, u_2, u_3 формулами преобразования см. гл. I, § 1, стр. 21)

$$\begin{aligned}\omega_1 &= \alpha_1 u_1 + \beta_1 u_2 + \gamma_1 u_3,\\ \omega_2 &= \alpha_2 u_1 + \beta_2 u_2 + \gamma_2 u_3,\\ \omega_3 &= \alpha_3 u_1 + \beta_3 u_2 + \gamma_3 u_3\end{aligned}$$

и, обратно,

$$\begin{aligned}u_1 &= \alpha_1\omega_1 + \alpha_2\omega_2 + \alpha_3\omega_3,\\ u_2 &= \beta_1\omega_1 + \beta_2\omega_2 + \beta_3\omega_3,\\ u_3 &= \gamma_1\omega_1 + \gamma_2\omega_2 + \gamma_3\omega_3.\end{aligned}$$

Для заданной точки поля при повороте системы координат не изменяется ни ее радиус-вектор $\boldsymbol{r}$, ни соответствующий этой точке вектор поля $\boldsymbol{u}$, причем сохраняется векторная функциональная зависимость между ними: $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{r})$, но три функции

$$u_1(x, y, z), \qquad u_2(x, y, z), \qquad u_3(x, y, z)$$

заменяются тремя новыми функциями

$$\omega_1(\xi, \eta, \zeta), \qquad \omega_2(\xi, \eta, \zeta), \qquad \omega_3(\xi, \eta, \zeta).$$

Последние три функции (в новой системе координат) получаются так: в первые три функции подставляют выражения x, y, z через ξ, η, ζ, а полученные три выражения подставляют вместо u_1, u_2, u_3 в формулы (см. выше), выражающие через них новые координаты ω_1, ω_2, ω_3 вектора поля.

В рассмотренных выше примерах векторное поле $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{r})$ задано с самого начала, а тем самым определены координаты вектора поля в любой системе декартовых прямоугольных координат. Обратно, если в некоторой выбранной системе координат x, y, z заданы три функции $u_1(x, y, z)$, $u_2(x, y, z)$, $u_3(x, y, z)$, то они определяют в этой системе координат векторное поле $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{r})$, в котором вектор поля $\boldsymbol{u}$ имеет такие координаты; координаты ω_i вектора $\boldsymbol{u}$ в другой системе координат ξ, η, ζ вычисляются по данным выше формулам преобразования.

Если в физических приложениях каждой точке некоторой пространственной области G отнесено определенное значение величины u, например плотность вещества в этой точке, так что в G задана одна-единственная функция $u=f(x, y, z)=F(r)$, причем эта величина u является не координатой вектора, а такой величиной, которая сохраняет свое значение при изменении системы координат, то эту функцию называют *скалярной функцией точки.* Говорят также, что в области G задано *скалярное поле.* Так, например, во всяком векторном поле $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{r})$ величина $|\boldsymbol{u}|^2=u_1^2+u_2^2+u_3^2$ является скалярной функцией точки: она ведь равна квадрату модуля вектора поля, а стало быть, не зависит от выбора системы координат.

2. Векторная функция скалярной переменной и ее производная. Наряду с векторными и скалярными функциями векторного аргумента (векторными и скалярными полями) рассматривают также и векторную функцию скалярной переменной, т. е. многообразие или семейство векторов, зависящее от одного параметра t; записывают это так: $\boldsymbol{u}=\boldsymbol{u}(t)$. Если перейти к координатам u_1, u_2, u_3 переменного вектора $\boldsymbol{u}$, то последняя векторная функция скалярного аргумента t равносильна системе трех скалярных функций

$$u_1=\varphi(t), \qquad u_2=\psi(t), \qquad u_3=\chi(t). \tag{1}$$

Если отождествить вектор $\boldsymbol{u}$ с радиус-вектором точки в пространстве, то при изменении t в интервале $t_1 \leqslant t \leqslant t_2$ конец радиус-вектора опишет дугу пространственной кривой, заданной параметрическими уравнениями (1).

[Для векторных функций $\boldsymbol{u}(t)$ вводят понятие предела совершенно так же, как это было сделано для обычной (скалярной) функции одной переменной:

Говорят, что переменный вектор $\boldsymbol{u}(t)$ стремится к пределу $\boldsymbol{a}$, когда t стремится к t_0, если для любого положительного числа ε существует такое положительное число δ, что неравенство $|t-t_0|<\delta$ влечет за собой неравенство $|\boldsymbol{u}(t)-\boldsymbol{a}|<\varepsilon$. Записывают это так: $\lim\limits_{t\to t_0}\boldsymbol{u}(t)=\boldsymbol{a}$ или $\boldsymbol{u}(t)\to\boldsymbol{a}$ при $t\to t_0$.

На основе этого определения нетрудно вывести правила действий с пределами, аналогичные обычным.

Так, предел постоянного вектора равен ему самому. Далее, если при $t \to t_0$ дано, что

$$\boldsymbol{u}(t) \to \boldsymbol{a}, \quad \boldsymbol{v}(t) \to \boldsymbol{b}, \quad \varphi(t) \to \alpha.$$

то

$$\boldsymbol{u}(t) + \boldsymbol{v}(t) \to \boldsymbol{a} + \boldsymbol{b}, \quad \varphi(t)\,\boldsymbol{u}(t) \to \alpha\boldsymbol{a},$$

скалярное произведение

$$\boldsymbol{u}(t)\,\boldsymbol{v}(t) \to \boldsymbol{ab},$$

векторное произведение

$$[\boldsymbol{u}(t)\,\boldsymbol{v}(t)] \to [\boldsymbol{ab}].$$

Доказательства предоставляем читателю.

Из правила вычисления предела скалярного произведения выведем следующую теорему: *предел проекции вектора $\boldsymbol{u}(t)$ на постоянную ось равен проекции предела на ту же ось.* Действительно, пусть $\boldsymbol{c}^\circ$ — единичный вектор оси, а $\lim\limits_{t \to t_0} \boldsymbol{u}(t) = \boldsymbol{a}$. Тогда при $\text{пр}_{\boldsymbol{c}^\circ} \boldsymbol{u}(t) = \boldsymbol{c}^\circ \boldsymbol{u}(t) \to \boldsymbol{c}^\circ \boldsymbol{a}$, т. е. проекция вектора стремится к проекции его предела на данную ось. Так как координаты вектора — это его проекции на оси координат, то $u_i = \boldsymbol{e}_i \boldsymbol{u}$, где $\boldsymbol{e}_i$ $(i = 1, 2, 3)$ — орты осей. Поэтому *пределы координат переменного вектора $\boldsymbol{u}(t)$ равны соответствующим координатам предела вектора,* т. е. утверждение: при $t \to t_0$

$$\boldsymbol{u}(t) = \{u_1(t),\ u_2(t),\ u_3(t)\} \to \boldsymbol{a} = \{a_1,\ a_2,\ a_3\}$$

равнозначно утверждению

$$u_1(t) \to a_1, \quad u_2(t) \to a_2, \quad u_3(t) \to a_3.$$

Векторные функции скалярной переменной t можно дифференцировать по этой переменной. Для этого вводят определение производной $\boldsymbol{u}'(t)$ от векторной функции $\boldsymbol{u}(t)$ с помощью знакомого нам предельного перехода

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{u}'(t) = \lim_{h \to 0} \frac{\boldsymbol{u}(t+h) - \boldsymbol{u}(t)}{h}$$

при том условии, конечно, что этот предел существует. Из определения видно, что производная $\boldsymbol{u}'(t)$ есть тоже вектор, являющийся функцией той же скалярной переменной t. Нетрудно доказать, что координаты производного вектора $\boldsymbol{u}'(t)$ равны производным от одноименных координат исходного вектора: если $\boldsymbol{u}(t) = \{u_1(t),\ u_2(t),\ u_3(t)\}$, то $\boldsymbol{u}'(t) = \{u_1'(t),\ u_2'(t),\ u_3'(t)\}$.

Столь же нетрудно доказать, что для векторных функций скалярной переменной справедливы знакомые нам правила дифференцирования и некоторые им аналогичные. Так, производная постоянного вектора равна нулю. Далее, для векторной суммы $\boldsymbol{w} = \boldsymbol{u}(t) + \boldsymbol{v}(t)$ имеем

$$\boldsymbol{w}' = \boldsymbol{u}'(t) + \boldsymbol{v}'(t).$$

Для произведения скалярной функции $\varphi(t)$ на векторную $\boldsymbol{u}(t)$: если $\boldsymbol{w}(t)=\varphi(t)\,\boldsymbol{u}(t)$, то

$$\boldsymbol{w}'(t)=\varphi'(t)\,\boldsymbol{u}(t)+\varphi(t)\,\boldsymbol{u}'(t)$$

или

$$(\varphi\boldsymbol{u})'=\varphi'\boldsymbol{u}+\varphi\boldsymbol{u}'.$$

Скалярное произведение двух вектор-функций $\boldsymbol{u}(t)$ и $\boldsymbol{v}(t)$ уже является обычной, т. е. скалярной функцией от t, и ее производная вычисляется по формуле

$$\frac{d(\boldsymbol{uv})}{dt}=\frac{d\boldsymbol{u}}{dt}\boldsymbol{v}+\boldsymbol{u}\frac{d\boldsymbol{v}}{dt} \quad \text{или} \quad (\boldsymbol{uv})'=\boldsymbol{u}'\boldsymbol{v}+\boldsymbol{uv}'.$$

Векторное произведение $[\boldsymbol{u}(t)\boldsymbol{v}(t)]$ дифференцируют по формуле

$$\frac{d}{dt}[\boldsymbol{uv}]=\left[\frac{d\boldsymbol{u}}{dt}\boldsymbol{v}\right]+\left[\boldsymbol{u}\frac{d\boldsymbol{v}}{dt}\right] \quad \text{или} \quad [\boldsymbol{uv}]'=[\boldsymbol{u}'\boldsymbol{v}]+[\boldsymbol{uv}'].$$

Производная от сложной функции (правило цепочки):
Если $\boldsymbol{u}=\boldsymbol{f}(t)$, где $t=\varphi(\tau)$, то $\boldsymbol{u}=\boldsymbol{f}\{\varphi(\tau)\}=\boldsymbol{F}(\tau)$ и

$$\boldsymbol{F}'(\tau)=\boldsymbol{f}'(t)\,\varphi'(\tau) \quad \text{или} \quad \boldsymbol{u}'_\tau=\boldsymbol{u}'_t\cdot t'_\tau.]$$

[3. **Длина дуги пространственной кривой. Дифференциал дуги.** Определение длины дуги плоской кривой (т. I, гл. V, § 2, n° 6) обобщают и на пространственные кривые. И здесь длиной дуги называется предел периметра вписанной в дугу ломаной, когда число ее звеньев неограниченно растет, а длина каждого звена стремится к нулю, конечно, если этот предел существует. Будем исходить из параметрического задания кривой

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

где эти три функции имеют непрерывные производные. В этих условиях упомянутый выше предел существует. Наметим ход доказательства.

Пусть начало A дуги соответствует значению α, а конец B — значению β параметра t. Интервал $\alpha\leqslant t\leqslant\beta$ разобьем на n частей точками $\alpha=t_0, t_1, t_2, \ldots, t_n=\beta$ и впишем в дугу AB ломаную с вершинами в тех точках дуги, которые соответствуют этим значениям параметра. Длина k-го звена ломаной (в понятных обозначениях) будет $\sqrt{\Delta x_k^2+\Delta y_k^2+\Delta z_k^2}=\sqrt{\left(\frac{\Delta x_k}{\Delta t_k}\right)^2+\left(\frac{\Delta y_k}{\Delta t_k}\right)^2+\left(\frac{\Delta z_k}{\Delta t_k}\right)^2}\,\Delta t_k$ $(\Delta t_k>0)$. По теореме о среднем значении дифференциального исчисления,

$$\frac{\Delta x_k}{\Delta t_k}=\dot{x}(\theta_k), \qquad \frac{\Delta y_k}{\Delta t_k}=\dot{y}(\theta'_k), \qquad \frac{\Delta z_k}{\Delta t_k}=\dot{z}(\theta''_k),$$

где точка над буквой обозначает производную по t, а θ_k, θ'_k, θ''_k — некоторые промежуточные значения между t_{k-1} и t_k. Поэтому периметр вписанной ломаной запишется в следующем виде:

$$\sum_{k=1}^{n} \sqrt{\left[\dot{x}(\theta_k)\right]^2 + \left[\dot{y}(\theta'_k)\right]^2 + \left[\dot{z}(\theta''_k)\right]^2}\, \Delta t_k.$$

На основании замечания на стр. 161 первого тома о более общей структуре сумм, имеющих своим пределом определенный интеграл, последняя сумма стремится к определенному интегралу, так что длина дуги $\overset{\frown}{AB}$

$$s(\alpha, \beta) = \int_{\alpha}^{\beta} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt.$$

Длина переменной дуги от постоянной точки A, соответствующей значению α параметра, до переменной точки P со значением параметра t есть

$$s = \int_{\alpha}^{t} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt.$$

Она является непрерывной и дифференцируемой функцией от t, и производная от длины дуги по параметру

$$\frac{ds}{dt} = \dot{s} = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} =$$

$$= \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}$$

или

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{d\boldsymbol{r}}{dt}\right)^2 = \dot{\boldsymbol{r}}^2.$$

Дифференциал дуги

$$ds = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt = \sqrt{dx^2 + dy^2 + dz^2}.]$$

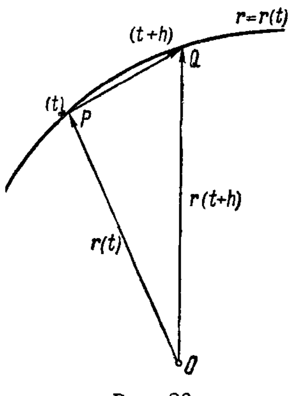


Рис. 26.

4. Кривизна пространственной кривой. Приведем некоторые простые приложения изложенных выше соображений. Пусть в пространстве x, y, z задана кривая параметрическим уравнением (в векторной записи) $\boldsymbol{r} = \boldsymbol{r}(t)$, где t — параметр, а $\boldsymbol{r}$ — радиус-вектор переменной точки кривой (рис. 26).

Тогда производный вектор $\dot{\boldsymbol{r}}(t)$, если он существует, направлен по касательной к нашей кривой в точке t. Действительно, вектор $\boldsymbol{r}(t+h) - \boldsymbol{r}(t) = \overline{PQ}$ идет по хорде от точки $P(t)$ к точке $Q(t+h)$, а следовательно, вектор

$$\frac{\boldsymbol{r}(t+h) - \boldsymbol{r}(t)}{h},$$

отличающийся от вектора $\overline{PQ}$ только скалярным множителем $\frac{1}{h}$, тоже направлен по секущей, проходящей через точки P и Q. При предельном переходе $h \to 0$ точка Q стремится вдоль кривой к точке P, и если, как было предположено, существует производная

$$\dot{\boldsymbol{r}}(t) = \lim_{h \to 0} \frac{\boldsymbol{r}(t+h) - \boldsymbol{r}(t)}{h},$$

то существует и предельное положение секущей, т. е. касательная в точке t кривой, и эта касательная имеет направление вектора $\dot{\boldsymbol{r}}(t)$, который называется поэтому *касательным вектором*. Нетрудно убедиться, что этот вектор направлен в сторону возрастания параметра t.

Примем теперь в качестве параметра длину дуги s, отсчитываемую от некоторой начальной точки на кривой; тогда в формулу для производной от длины дуги по параметру придется писать, вместо производных по t, производные по s, и мы получим

$$\frac{ds}{ds} = 1 = \sqrt{\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2}.$$

Производные по длине дуги мы будем всегда обозначать штрихами. Принимая во внимание, что

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = (x')^2 + (y')^2 + (z')^2 = \left(\frac{d\boldsymbol{r}}{ds}\right)^2 = (\boldsymbol{r}')^2,$$

последнее равенство можно записать в следующем виде:

$$(x')^2 + (y')^2 + (z')^2 = 1$$

или (в векторной записи)

$$(\boldsymbol{r}')^2 = 1.$$

Это равенство и характеризует тот факт, что параметром служит длина дуги. Так как скалярный квадрат вектора равен квадрату его модуля, то вектор $\frac{d\boldsymbol{r}}{ds} = \boldsymbol{r}'$ является *единичным касательным вектором*. Он направлен в сторону возрастания дуги s.

Дифференцируя по s равенство $(\boldsymbol{r}')^2 = 1$, получим

$$\boldsymbol{r}'\boldsymbol{r}'' = 0.$$

Это равенство устанавливает, что вектор $\boldsymbol{r}'' = \{x''(s), y''(s), z''(s)\}$ перпендикулярен к вектору $\boldsymbol{r}'$, а стало быть, *перпендикулярен к касательной*. Этот вектор $\boldsymbol{r}''$ называется *вектором кривизны* или *главным нормальным вектором*, а его модуль, т. е. длина, называется *кривизной* линии в рассматриваемой точке и обозначается буквой k. Величину $\rho = \frac{1}{k}$, обратную кривизне, называют, как и у плоской кривой, *радиусом кривизны.* Итак, кривизна

$$k=\frac{1}{\rho}=|\boldsymbol{r}''|=\sqrt{(x'')^2+(y'')^2+(z'')^2}.$$

Из рассматриваемой точки кривой отложим по направлению вектора кривизны отрезок, равный радиусу кривизны; конечная точка этого отрезка называется *центром кривизны.*

Это определение кривизны находится в согласии с определением кривизны плоской кривой, данным в первом томе, гл. V, § 2, n° 8. И здесь можно определить кривизну как абсолютную величину предела отношения угла α между направлениями касательных в двух

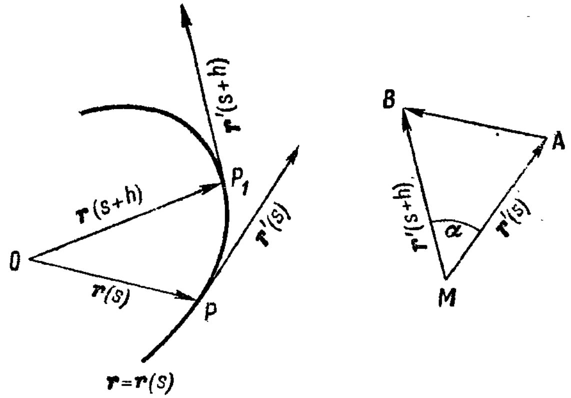


Рис. 27.

близких точках кривой к длине дуги h между этими точками: $k=\left|\lim\limits_{h\to 0}\frac{\alpha}{h}\right|$. Покажем, что такое определение кривизны совпадает с прежним.

Пусть s и $s+h$ — значения параметра (длины дуги) в близких точках P и P_1, а $\boldsymbol{r}'(s)$ и $\boldsymbol{r}'(s+h)$ — единичные касательные векторы в этих точках. Приведем эти векторы к общему началу M (рис. 27), так что $\boldsymbol{r}'(s)=\overline{MA}$ и $\boldsymbol{r}'(s+h)=\overline{MB}$. Тогда угол $\widehat{AMB}=\alpha$ и вектор $\overline{AB}=\boldsymbol{r}'(s+h)-\boldsymbol{r}'(s)$. Выражение для кривизны можно представить в следующем виде:

$$k=\left|\lim_{h\to 0}\frac{\alpha}{h}\right|=\lim_{h\to 0}\frac{\alpha}{|\overline{AB}|}\cdot\left|\lim_{h\to 0}\frac{\overline{AB}}{h}\right|.$$

Как видно из рис. 27, $|\overline{AB}|=2\sin\frac{\alpha}{2}$; поэтому

$$\lim_{h\to 0}\frac{\alpha}{|\overline{AB}|}=\lim_{\alpha\to 0}\frac{\alpha}{2\sin\frac{\alpha}{2}}=1,$$

ибо $\alpha \to 0$ при $h \to 0$. Второй предел в правой части

$$\lim_{h \to 0} \frac{\overline{AB}}{h} = \lim_{h \to 0} \frac{\boldsymbol{r}'(s+h) - \boldsymbol{r}'(s)}{h} = \boldsymbol{r}''(s).$$

Приходим, следовательно, к прежнему определению.

[Заметим, что тот вывод формулы кривизны, которым мы пользовались в первом томе в случае плоской кривой, здесь неприменим. Дело в том, что угол между касательными в двух соседних точках *плоской* кривой можно всегда представить как *приращение* угла, составляемого касательной с постоянной осью (например, осью x), в пространстве же это невозможно.]

5. Приложение к механике точки. Разложение ускорения на касательное и нормальное. Вектор кривизны (главный нормальный вектор) играет важную роль в механике точки. Представим себе, что материальная точка движется, описывая кривую $\boldsymbol{r} = \boldsymbol{r}(t)$; параметром служит время t. Тогда нетрудно убедиться, что скорость движения определяется по величине и направлению вектором $\dot{\boldsymbol{r}}(t)$, где точкой обозначена производная по времени t. Аналогично, ускорение определяется второй производной по времени $\ddot{\boldsymbol{r}}(t)$. По правилу цепочки имеем

$$\dot{\boldsymbol{r}} = \frac{d\boldsymbol{r}}{dt} = \frac{d\boldsymbol{r}}{ds}\frac{ds}{dt} = \boldsymbol{r}'\frac{ds}{dt}$$

(штрих обозначает производную по s, точка — по t).

Вектор $\dot{\boldsymbol{r}}$ является вместе с тем касательным вектором к траектории, направленным в ту ее сторону, которая соответствует возрастанию времени t. Следовательно, скорость направлена по касательной к пути движения в сторону возрастания t. Так как $\boldsymbol{r}'$ есть единичный касательный вектор, направленный в сторону возрастания s, то величина скорости измеряется производной $\frac{ds}{dt}$ от пройденного пути по времени, которая положительна, если точка движется в сторону возрастания s, и отрицательна в противном случае.

Дифференцируя вторично по t, получим, что ускорение

$$\ddot{\boldsymbol{r}} = \frac{d^2\boldsymbol{r}}{dt^2} = \boldsymbol{r}'\frac{d^2s}{dt^2} + \boldsymbol{r}''\left(\frac{ds}{dt}\right)^2.$$

Из этой формулы видно, что вектор-ускорение равен сумме двух векторов. Первый из них направлен по касательной к траектории, и его величина равна $\frac{d^2s}{dt^2}$, т. е. ускорению движущейся точки вдоль пути движения, причем движение вдоль пути является ускоренным, если $\frac{d^2s}{dt^2} > 0$, и замедленным, если $\frac{d^2s}{dt^2} < 0$. Этот вектор $\boldsymbol{r}'\frac{d^2s}{dt^2}$ называется *касательным* или *тангенциальным ускорением*. Второй составляющий вектор $r'' \left(\frac{ds}{dt}\right)^2$ направлен к центру кривизны (по главному нормальному вектору) и называется *нормальным ускорением*. Его модуль равен произведению квадрата скорости на кривизну, т. е. $\frac{v^2}{\rho}$, где v — величина скорости, ρ — радиус кривизны.

6. Градиент скалярного поля. Вернемся к рассмотрению скалярных полей. Пусть в некоторой области G пространства задана скалярная функция $f(x, y, z)$. Ее можно рассматривать как функцию точки, и в области G она определяет скалярное поле. Возьмем три частные производные

$$u_1 = f_x(x, y, z), \qquad u_2 = f_y(x, y, z), \qquad u_3 = f_z(x, y, z)$$

и будем их интерпретировать как координаты некоторого вектора u в системе координат x, y, z. Если мы перейдем, с помощью поворота осей, к новой системе прямоугольных координат ξ, η, ζ, то по формулам преобразования (n° 1) получим для вектора u новые координаты

$$\begin{aligned} \omega_1 &= \alpha_1 u_1 + \beta_1 u_2 + \gamma_1 u_3, \\ \omega_2 &= \alpha_2 u_1 + \beta_2 u_2 + \gamma_2 u_3, \\ \omega_3 &= \alpha_3 u_1 + \beta_3 u_2 + \gamma_3 u_3. \end{aligned}$$

С другой стороны, если ввести в функцию $f(x, y, z)$ в качестве новых переменных новые координаты точки ξ, η, ζ с помощью соответствующего преобразования

$$\begin{aligned} x &= \alpha_1 \xi + \alpha_2 \eta + \alpha_3 \zeta, \\ y &= \beta_1 \xi + \beta_2 \eta + \beta_3 \zeta, \\ z &= \gamma_1 \xi + \gamma_2 \eta + \gamma_3 \zeta, \end{aligned}$$

то по правилу цепочки получим

$$\begin{aligned} f_\xi &= f_x x_\xi + f_y y_\xi + f_z z_\xi = f_x \alpha_1 + f_y \beta_1 + f_z \gamma_1, \\ f_\eta &= f_x x_\eta + f_y y_\eta + f_z z_\eta = f_x \alpha_2 + f_y \beta_2 + f_z \gamma_2, \\ f_\zeta &= f_x x_\zeta + f_y y_\zeta + f_z z_\zeta = f_x \alpha_3 + f_y \beta_3 + f_z \gamma_3. \end{aligned}$$

Оказывается, таким образом, что

$$\omega_1 = f_\xi, \qquad \omega_2 = f_\eta, \qquad \omega_3 = f_\zeta.$$

Стало быть, и в новой координатной системе координаты нашего вектора u равны частным производным той же величины f по новым прямоугольным декартовым координатам. Это и означает, что всякой функции точки f в трехмерном пространстве соответствует определенный вектор, координаты которого в любой прямоугольной координатной системе равны частным производным этой функции по координатам. Этот вектор называется *градиентом функции* f или *градиентом скалярного поля* и обозначается символом

$$u = \operatorname{grad} f.$$

Для функции трех переменных, рассматриваемой как скалярная функция точки, градиент является аналогом производной от функции одной переменной. Наглядное представление о значении градиента мы получим, когда вычислим производную от функции $f(x, y, z)$ по направлению вектора c, образующего углы δ_1, δ_2, δ_3 с осями координат. На стр. 81 мы получили для этой производной выражение

$$D_c f = f_x \cos\delta_1 + f_y \cos\delta_2 + f_z \cos\delta_3.$$

Так как единичный вектор указанного направления

$$c^\circ = \{\cos\delta_1, \cos\delta_2, \cos\delta_3\},$$

то

$$D_c f = c^\circ \operatorname{grad} f.$$

Следовательно, производная функции f по направлению вектора c равна скалярному произведению градиента функции на единичный вектор указанного направления, т. е. проекции градиента на направление дифференцирования.

Именно этот факт придает понятию градиента большое значение. Поставим, например, вопрос: в каком направлении функция f быстрее всего возрастает или быстрее всего убывает? Для ответа на этот вопрос надо найти то направление, на которое проекция градиента имеет наибольшее положительное или наибольшее по абсолютной величине отрицательное значение. Первое будет, очевидно, когда вектор c имеет направление градиента, второе — когда направление c прямо противоположно направлению градиента.

Стало быть, *направление градиента есть направление наибыстрейшего возрастания функции, а направление, противоположное градиенту, есть направление наибыстрейшего ее убывания. Модуль градиента дает наибольшую скорость возрастания или убывания функции* [в различных направлениях от той точки, в которой вычислен градиент; скорость изменения функции выражена в единицах величины f, деленных на единицу длины].

Мы уже знаем, так сказать, физический смысл градиента. Но как его построить? Мы покажем здесь простой способ построения *направления* градиента. Рассмотрим сначала скалярную функцию точки $f(x, y)$ на плоскости, и речь пойдет о градиенте этой функции. Предположим, что данное скалярное поле изображено с помощью семейства линий уровня $f(x, y) = c$. Тогда в любой точке P поля производная от $f(x, y)$ по направлению линии уровня, проходящей через эту точку, равна, очевидно, нулю. Действительно, если соседняя точка Q лежит на той же линии уровня, что и P, то $f(Q) - f(P) = 0$ (смысл этой записи понятен), и при стремлении точки Q

к P вдоль линии уровня (ср. стр. 80) $\lim\limits_{\rho\to 0}\frac{f(Q)-f(P)}{\rho}=0$, где ρ — расстояние между точками P и Q. Следовательно, проекция градиента на касательную к линии уровня в точке P равна нулю. Отсюда вытекает, что *в любой точке поля градиент направлен по нормали к линии уровня, проходящей через эту точку*. Совершенно аналогично обстоит дело с градиентом скалярного поля в трехмерном пространстве. Функцию точки $f(x, y, z)$ изображают с помощью семейства ее поверхностей уровня $f(x, y, z)=c$, и тогда в любой точке P поля проекция градиента на любую касательную прямую к поверхности уровня в точке P равна нулю. Стало быть, градиент перпендикулярен касательной плоскости к поверхности уровня в точке P. Прямая, перпендикулярная касательной плоскости к поверхности и проходящая через их точку касания, называется *нормалью* к поверхности. Следовательно, в любой точке поля градиент функции $f(x, y, z)$ направлен по нормали к поверхности уровня, проходящей через эту точку. [Стрелка градиента направлена в ту сторону нормали, в которую возрастают пометки c поверхностей уровня в изображении поля. Если в какой-либо точке скалярного поля соответствующая поверхность уровня не имеет касательной плоскости, то в этой точке не существует и градиента. Это — особая точка поля.]

Градиент функции $f(x, y, z)$ является в свою очередь функцией точки, но уже векторной функцией. Для данного скалярного поля построено соответствующее ему (и характеризующее его) векторное поле его градиента. С другой стороны, во многих приложениях встречаются такие векторные поля, в которых вектор поля оказывается градиентом некоторой скалярной функции точки. Примером может служить гравитационное силовое поле (поле тяготения).

Пусть притягивающая масса M находится в точке $Q(\xi, \eta, \zeta)$, а притягиваемая масса m — в точке $P(x, y, z)$. Тогда сила тяготения Ньютона есть вектор $\boldsymbol{F}$ со следующими координатами:

$$F_1=mM\gamma\frac{\xi-x}{r^3},\qquad F_2=mM\gamma\frac{\eta-y}{r^3},\qquad F_3=mM\gamma\frac{\zeta-z}{r^3};$$

здесь $r=\sqrt{(\xi-x)^2+(\eta-y)^2+(\zeta-z)^2}$ (расстояние между обеими точками), а γ есть «постоянная тяготения». (В каждом из этих выражений можно выделить по множителю:

$$\frac{\xi-x}{r}=\cos\delta_1,\qquad \frac{\eta-y}{r}=\cos\delta_2,\qquad \frac{\zeta-z}{r}=\cos\delta_3,$$

— это направляющие косинусы вектора $\overline{PQ}$.) Нетрудно проверить прямым дифференцированием, что написанные выше составляющие силы тяготения равны частным производным функции

$$\frac{mM\gamma}{r}=\frac{mM\gamma}{\sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}}$$

по координатам x, y, z. Следовательно, наш вектор-сила является градиентом этой скалярной функции точки $P(x, y, z)$:

$$\boldsymbol{F} = \operatorname{grad} \frac{\gamma m M}{r}.$$

Если в силовом (векторном) поле сила равна градиенту некоторой скалярной функции точки, то эту скалярную функцию часто называют *силовой функцией* или *потенциалом* для данного силового поля. [В физике потенциалом обыкновенно называют силовую функцию с обратным знаком.] Это понятие мы рассмотрим с более общей точки зрения при изучении работы и энергии (гл. V, § 1, n°4 и гл. VI, §§ 1 и 5).

7. Дивергенция и ротор векторного поля. Всякой дифференцируемой скалярной функции точки мы отнесли, с помощью дифференцирования, характеризующее ее векторное поле ее градиента. Аналогично можно для векторного поля построить, с помощью некоторого процесса дифференцирования, скалярную функцию точки, называемую *дивергенцией* этого векторного поля. Исходим сначала из выбранной системы координат x, y, z и в этой системе называем дивергенцией вектор-функции точки $\boldsymbol{u}(x, y, z)$ следующую скалярную функцию:

$$\operatorname{div} \boldsymbol{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z},$$

т. е. сумму трех частных производных от координат вектора $\boldsymbol{u}$ по одноименным координатам точки. По замыслу, дивергенция должна быть скалярной функцией точки, характеризующей векторное поле $\boldsymbol{u}(x, y, z)$, и не должна зависеть от выбора системы координат. Поэтому, чтобы оправдать наше определение, надлежит доказать, что в любой другой системе координат ξ, η, ζ

$$\operatorname{div} \boldsymbol{u} = \frac{\partial \omega_1}{\partial \xi} + \frac{\partial \omega_2}{\partial \eta} + \frac{\partial \omega_3}{\partial \zeta},$$

где $\omega_1, \omega_2, \omega_3$ — координаты вектора $\boldsymbol{u}$ в системе ξ, η, ζ. И действительно, с помощью правила цепочки и формул преобразования, стр. 101, легко доказывается тождество

$$\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \equiv \frac{\partial \omega_1}{\partial \xi} + \frac{\partial \omega_2}{\partial \eta} + \frac{\partial \omega_3}{\partial \zeta}.$$

Мы здесь ограничимся формальным определением дивергенции, а ее физико-геометрическое истолкование мы рассмотрим в гл. V, § 5.

Так же формально мы введем и понятие *ротора* или *вихря* векторного поля (по-английски curl). *Ротором* векторного поля $\boldsymbol{u} = \boldsymbol{u}(x, y, z)$ называется соответствующая этому полю новая вектор-функция точки, имеющая в системе x, y, z следующие координаты:

$$\operatorname{rot} \boldsymbol{u} = \left\{\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z},\ \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x},\ \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right\}.$$

Для того чтобы оправдать это определение, следовало бы показать прямым пересчетом, что оно не зависит от выбора системы координат

нат, т. е. при переходе к любой другой системе координат (ξ, η, ζ), в которой вектор поля $\boldsymbol{u}$ будет иметь новые координаты $\{\omega_1, \omega_2, \omega_3\}$, координаты нового «вектора», обозначенного символом $\operatorname{rot} \boldsymbol{u}$, преобразуются к виду

$$\left\{\frac{\partial\omega_3}{\partial\eta} - \frac{\partial\omega_2}{\partial\zeta}, \quad \frac{\partial\omega_1}{\partial\zeta} - \frac{\partial\omega_3}{\partial\xi}, \quad \frac{\partial\omega_2}{\partial\xi} - \frac{\partial\omega_2}{\partial\eta}\right\}.$$

Однако мы не будем заниматься этим пересчетом, потому что в гл. V, § 6 естественным путем будет получено наглядное физическое истолкование ротора и вместе с тем сам собой выяснится его векторный характер.

Три дифференциальные операции, дающие градиент, дивергенцию и ротор, можно связать друг с другом с помощью одного «дифференциального оператора» [называемого также *оператором Гамильтона*], обозначаемого символом ∇ (читается: набла). Этот оператор носит характер «символического вектора»

$$\nabla = \left\{\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial z}\right\}.$$

Градиент скалярного поля есть «произведение» этого «вектора» ∇ на скалярную функцию $f(x, y, z)$:

$$\operatorname{grad} f = \nabla f = \left\{\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial z}\right\}.$$

Дивергенция векторного поля $\boldsymbol{u}$ есть «скалярное произведение» «вектора» ∇ на вектор-функцию точки $\boldsymbol{u} = \boldsymbol{u}(x, y, z)$:

$$\operatorname{div} \boldsymbol{u} = \nabla u = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}.$$

Ротор векторного поля $\boldsymbol{u}(x, y, z)$ есть «векторное произведение» оператора ∇ на вектор $\boldsymbol{u}$:

$$\operatorname{rot} \boldsymbol{u} = [\nabla \boldsymbol{u}].$$

В заключение приведем некоторые часто встречающиеся соотношения.

1) *Ротор градиента равен нулю*; в символической записи:

$$\operatorname{rot} \operatorname{grad} f = 0.$$

Действительно, это соотношение является прямым следствием теоремы о переместительности порядка дифференцирования.

2) *Дивергенция ротора равна нулю*:

$$\operatorname{div} \operatorname{rot} \boldsymbol{u} = 0.$$

Эта формула тоже сразу вытекает из определения обеих операций на основании переместительности порядка дифференцирования.

3) Чрезвычайно важное значение имеет часто встречающаяся в анализе и в его приложениях *дивергенция градиента скалярной*

функции $f(x, y, z)$. Ее можно представить как «произведение» скалярного оператора $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ (набла квадрат) на f. В результате получается *лапласиан* или *дифференциальное выражение Лапласа*, играющее большую роль в теории потенциала и издавна обозначаемое символом Δf:

$$\Delta f = \nabla^2 f = \operatorname{div} \operatorname{grad} f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

Дифференциальный символ

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

называется *оператором Лапласа* или *лапласианом*.

4) В конце гл. VI, в номере, посвященном уравнениям Максвелла, нам понадобится еще формула

$$\operatorname{rot} \operatorname{rot} \boldsymbol{u} = \operatorname{grad} \operatorname{div} \boldsymbol{u} - \nabla^2 \boldsymbol{u},$$

которую предоставляем вывести читателю. Отметим только, что здесь лапласиан берется не от скаляра, а от вектора $\boldsymbol{u} = \{u_1, u_2, u_3\}$. Это значит, что $\nabla^2 \boldsymbol{u}$ есть вектор-функция

$$\nabla^2 \boldsymbol{u} = \{\nabla^2 u_1, \nabla^2 u_2, \nabla^2 u_3\}.$$

Наконец, заметим еще, что терминология векторного анализа часто применяется и к так называемому пространству n измерений. Совокупность n чисел рассматривают как точку n-мерного пространства $(x_1, x_2, \ldots, x_n)$ или как n-мерный вектор $\{a_1, a_2, \ldots, a_n\}$. Функцию от n независимых переменных рассматривают как скалярное поле, а систему n таких функций — как векторное поле в n-мерном пространстве. Понятие скалярного произведения, а также понятие градиента легко обобщаются на n-мерное пространство, но в других отношениях положение дел там значительно сложнее, чем в трехмерном пространстве.

Упражнения

1. Найти уравнение так называемой *соприкасающейся плоскости* кривой $\boldsymbol{r} = \boldsymbol{r}(t)$ или, подробнее, $x = f(t)$, $y = \varphi(t)$, $z = \psi(t)$ в точке t_0; соприкасающейся плоскостью пространственной кривой называется предельное положение плоскости, проходящей через три точки кривой, когда эти три точки стремятся к точке с параметром t_0.

2. Показать, что касательный вектор и главный нормальный вектор оба лежат в соприкасающейся плоскости.

3*. Кривая в пространстве задана уравнением $\boldsymbol{r} = \boldsymbol{r}(s)$, где параметр s есть длина дуги, а функция $\boldsymbol{r}(s)$ имеет непрерывные производные до третьего порядка включительно. Найти центр соприкасающейся сферы, т. е. сферы, имеющей с данной кривой в точке s_0 возможно более тесное касание.

4. Дана замкнутая кривая C: $\boldsymbol{r} = \boldsymbol{r}(t)$, причем $\boldsymbol{r}(t)$ — дифференцируемая функция, и точка A, не лежащая на C. Доказать, что на кривой C существует точка B, отстоящая от A на расстоянии более коротком, чем любая

другая точка этой кривой. Доказать также, что прямая AB нормальна к кривой.

5. Кривая $\boldsymbol{r}=\boldsymbol{r}(s)$ лежит на сфере радиуса 1. Доказать, что функция $\boldsymbol{r}(s)$ удовлетворяет уравнению

$$(\boldsymbol{r}')^2(\boldsymbol{r}''')^2-(\boldsymbol{r}'\boldsymbol{r}''')^2=(\boldsymbol{r}'\boldsymbol{r}''\boldsymbol{r}''')^2,$$

где $\boldsymbol{r}'\boldsymbol{r}''\boldsymbol{r}'''$ есть смешанное произведение трех векторов.

6. Пространственная кривая задана параметрическим уравнением $\boldsymbol{r}=\boldsymbol{r}(t)$, где t — произвольный параметр. Доказать, что если вектор $\ddot{\boldsymbol{r}}(t)$ построен из той точки $\boldsymbol{r}$, для которой он рассчитан, то он лежит в соприкасающейся плоскости.

7*. *Кручением* пространственной кривой называется предел отношения двугранного угла между соприкасающимися плоскостями в двух смежных точках кривой к длине дуги Δs между этими точками при $\Delta s\to 0$ или, что то же самое, предел отношения угла между нормальными векторами этих плоскостей к Δs. Обозначим для кривой $\boldsymbol{r}=\boldsymbol{r}(s)$ единичный касательный вектор через $\boldsymbol{t}(s)$, единичный главный нормальный вектор через $\boldsymbol{n}(s)$, а вектор $[\boldsymbol{tn}]$, называемый бинормальным вектором (он тоже единичный), через $\boldsymbol{b}(s)$. Доказать формулы Френэ:

$$\frac{d\boldsymbol{t}}{ds}=k\boldsymbol{n},\qquad \frac{d\boldsymbol{n}}{ds}=-k\boldsymbol{t}+\varkappa\boldsymbol{b},\qquad \frac{d\boldsymbol{b}}{ds}=-\varkappa\boldsymbol{n},$$

где k — кривизна, а $\varkappa$ — кручение кривой $\boldsymbol{r}=\boldsymbol{r}(s)$.

8. Пользуясь векторами $\boldsymbol{t}$, $\boldsymbol{n}$, $\boldsymbol{b}$ из упр. 7 как координатными ортами, вывести выражения: а) для вектора $\boldsymbol{r}'''(s)$, б) для вектора, идущего от точки $\boldsymbol{r}(s)$ кривой к центру сферы, соприкасающейся с кривой в этой точке.

9. Показать, что кривая, во всех точках которой кручение равно нулю, является плоской кривой.

10. Пусть $z=u(x,y)$ есть уравнение поверхности, образованной семейством касательных произвольной пространственной кривой. Доказать, что а) всякая соприкасающаяся плоскость кривой является касательной плоскостью этой поверхности и б) функция $u(x,y)$ удовлетворяет уравнению $u_{xx}u_{yy}-u_{xy}^2=0$.

ДОПОЛНЕНИЯ К ГЛАВЕ II

§ 1. Принцип точки сгущения в пространстве многих измерений и его приложения

Для того чтобы заложить твердый фундамент под теорией функций многих переменных, сообщить ее понятиям надлежащую точность, освободившись от всяких ссылок на интуицию, мы пойдем тем же путем, что и для функций одной переменной. При этом достаточно будет рассмотреть эти вопросы только для случая двух независимых переменных, так как, если число этих переменных больше двух, положение дел совершенно аналогично, да и методы исследования те же.

1. Формулировка принципа точки сгущения. В основу мы опять положим *принцип точки сгущения* Больцано — Вейерштрасса. Пару чисел (x, y) мы будем называть точкой P в пространстве двух измерений и изображать, как обычно, точкой с прямоугольными координатами x и y на плоскости xy. Рассмотрим теперь ограниченное бесконечное множество таких точек, т. е. такое бесконечное множество точек, которое содержится полностью в некоторой ограниченной области, так что, например, $|x| \leqslant C$ и $|y| \leqslant C$, где C — постоянное число. Так вот, принцип точки сгущения состоит в следующем: *всякое бесконечное множество точек, лежащее в ограниченной замкнутой области, имеет в этой области по крайней мере одну точку сгущения*, т. е. в этой области существует по крайней мере одна такая точка $Q(\xi, \eta)$, что в любой сколь угодно малой ее окрестности содержится бесконечно много точек данного множества. Другими словами, при любом сколь угодно малом значении числа $\delta > 0$, в окрестности точки $Q(\xi, \eta)$, характеризуемой неравенствами

$$|x-\xi| \leqslant \delta, \qquad |y-\eta| \leqslant \delta,$$

содержится неограниченное число точек множества.

Эту же мысль можно выразить и так: *из всякого ограниченного точечного множества можно выделить последовательность точек P_1, P_2, P_3, ..., сходящуюся к предельной точке Q.* [Точка Q соответствует в одномерном случае пределу выделенной последовательности.]

Интуитивно принцип точки сгущения для случая многих измерений столь же ясен, как и для одного измерения. Однако его следует доказать аналитически, и это можно сделать тем же путем, которым мы пользовались в т. I, стр. 81, — надо только заменить интервалы прямоугольными областями. Имея в виду еще и другие приложения, мы сформулируем здесь наш принцип доказательства несколько более общо — как «принцип стягивающихся областей»:

Если у бесконечной последовательности ограниченных замкнутых областей $G_1, G_2, G_3, \ldots$ каждая область G_n содержится полностью в предшествующей области и может быть заключена в квадрат, сторона которого стремится к нулю при $n \to \infty$, то эта последовательность областей определяет одну и только одну точку Q, принадлежащую совместно всем областям G_n.

Доказательство очень просто: выберем в каждой области G_n какую-либо точку P_n; тогда абсциссы x_n и ординаты y_n этих точек (согласно принципу сходимости для одной переменной) стремятся к пределам ξ и η, которые и определяют точку $Q(\xi, \eta)$.

Принцип точки сгущения является теперь прямым следствием принципа стягивающихся областей. Действительно, пусть в области G задано бесконечное множество точек. Разделим квадрат $|x| \leqslant C$, $|y| \leqslant C$, внутри которого содержится это множество, на четыре равных квадрата со стороной C. Если к каждому из этих квадратов присоединить его границу, то по крайней мере в одном из четырех квадратов, назовем его Q_1, должно лежать бесконечно много точек заданного множества. Квадрат Q_1 мы тоже разобьем на четыре равных квадрата со стороной $\frac{C}{2}$; по крайней мере один из этих квадратов (назовем его Q_2) должен еще содержать бесконечно много точек нашего множества. К квадрату Q_2 мы применим ту же процедуру и найдем квадрат Q_3 со стороной $\frac{C}{4}$, в котором имеется неограниченное количество точек множества и т. д. Согласно принципу стягивающихся областей, последовательность квадратов $Q_1, Q_2, Q_3, \ldots$ определяет одну и только одну точку Q, которая и является точкой сгущения заданного точечного множества. Если же выбрать в каждом квадрате Q_n по точке P_n нашего множества, то из последнего будет выделена последовательность точек P_n, сходящаяся, очевидно, к предельной точке Q.

Из принципа точки сгущения для многих измерений вытекают совершенно такие же следствия, как и для одного измерения. Так как и доказательства аналогичны, то можно ограничиться только формулировкой некоторых наиболее важных положений.

В первую очередь приведем *критерий сходимости Коши*, который можно формулировать следующим образом:

Последовательность точек $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, $P_3(x_3, y_3)$, ... сходится к предельной точке в том и только в том случае, если для любого сколь угодно малого $\varepsilon > 0$ существует такой индекс $N = N(\varepsilon)$, что расстояние $\sqrt{(x_n - x_m)^2 + (y_n - y_m)^2}$ между точками P_m и P_n становится меньше чем ε, как только оба номера m и n превосходят индекс N.

Далее, совершенно таким же путем, как и для одной независимой переменной, доказываются следующие теоремы:

Функция, непрерывная в замкнутой области, принимает в этой области наибольшее и наименьшее значение.

Всякая функция $f(x, y)$, непрерывная в замкнутой области G, равномерно непрерывна в этой области, т. е. для любого положительного числа ε существует такое число $\delta = \delta(\varepsilon)$, зависящее *исключительно* от ε, но не зависящее от точки (x_0, y_0), что $|f(x, y) - f(x_0, y_0)| < \varepsilon$ во всей области, если только расстояние между точками (x_0, y_0) и (x, y), лежащими в области G, меньше чем δ.

2. Некоторые понятия теории точечных множеств. Общее понятие точки сгущения имеет фундаментальное значение для многих тонких исследований по обоснованию анализа, базирующихся на теории точечных множеств. Хотя эти вопросы и не существенны для целей этой книги, мы все же для полноты приведем здесь некоторые основные понятия и предложения.

Ограниченное множество, состоящее из бесконечно большого числа точек, называется *замкнутым*, если оно содержит все свои точки сгущения; это значит, что предельная точка любой последовательности точек, принадлежащей множеству, тоже принадлежит этому множеству. Например, все точки замкнутой кривой или поверхности образуют замкнутое множество.

Точная верхняя граница расстояний между всевозможными парами точек множества называется *поперечником* или *диаметром* этого множества. Если множество является замкнутым, то эта точная верхняя граница действительно достигается для некоторой пары точек множества.

Отметим еще следующий факт. Если точка P не принадлежит замкнутому точечному множеству M, то существует положительное *кратчайшее расстояние от этой точки до множества* M, т. е. множество M содержит такую точку Q, которая удалена от P на расстоянии меньшем (или по крайней мере не большем), чем любая другая точка множества M. Опираясь на этот факт, можно доказать, что замкнутые области, определенные на стр. 57, являются фактически замкнутыми точечными множествами в смысле данного здесь определения. Действительно, пусть C есть замкнутая кривая, а B есть та замкнутая область, которая состоит из всех точек, лежащих внутри кривой C, и из всех точек самой этой кривой. Требуется доказать, что все точки сгущения области B принадлежат ей самой. Предположим противное, а именно, что область B имеет точку сгущения P, ей не принадлежащую. Тогда точка P не лежит на кривой C и находится вне ее, так как, в противном случае, она принадлежала бы области B. Стало быть, на основании формулированного выше факта существует положительное наименьшее расстояние точки P от кривой C (ибо последняя составляет замкнутое точечное множество). Построим круг с центром в точке P радиусом меньшим, чем наименьшее расстояние точки P от кривой C. Все точки этого круга лежат вне кривой C и, следовательно, ни одна его точка не принадлежит замкнутой области B. Но мы предположили, что P является точкой сгущения области B, так что построенный нами круг должен содержать бесконечно много точек области B. Таким образом, предположение, что замкнутая область B имеет точку сгущения, ей не принадлежащую, приводит к противоречию, а это и доказывает наше утверждение. Как распространить это предложение на замкнутые области, ограниченные несколькими замкнутыми кривыми, очевидно.

Важное свойство замкнутых множеств содержится в *теореме о стягивающихся последовательностях замкнутых множеств:*

Если дана последовательность замкнутых множеств M_1, M_2, M_3, ..., каждое из которых содержится в предшествующем множестве, то существует такая точка (ξ, η), которая принадлежит всем этим множествам.

Выберем в каждом из множеств M_n по точке P_n. Последовательность точек P_n должна содержать либо бесконечное число повторений одной и той же точки, либо бесконечное число различных точек. Если точка P повторяется бесчисленное множество раз, то оно принадлежит всем нашим множествам; в самом деле, если M_n есть любое из этих множеств, то точка P несомненно принадлежит некоторому множеству M_{n_1}, номер которого $n_1 > n$, так что M_{n_1} содержится в M_n. Если же существует бесконечно много различных точек P_n, то на основании принципа точки сгущения последовательность точек P_n должна иметь точку сгущения $P(\xi, \eta)$. Эта точка P непременно принадлежит всем множествам M_n. Это следует из того, что при $m > n$ точка P_m принадлежит множеству M_n (ибо P_m есть точка множества M_m, которое само содержится в M_n). Поэтому точка $P(\xi, \eta)$ является точкой сгущения последовательности точек P_m множества M_n, а так как M_n есть замкнутое множество, то $P(\xi, \eta)$ принадлежит M_n. Стало быть, существует точка, принадлежащая всем множествам M_n, и наша теорема доказана.

Сделанное в условии теоремы предположение, что множества M_n замкнуты, существенно, как показывает следующий пример. Пусть множествами M_n являются промежутки $0 < x < \frac{1}{n}$. Каждое из этих множеств содержится в предыдущем, однако не существует точки, принадлежащей всем этим множествам. Действительно, точка $x = 0$ не принадлежит ни одному из множеств M_n, если же $x > 0$, то эта точка не принадлежит тем множествам M_n, для которых $\frac{1}{n} < x$.

Множество называется *открытым*, если для каждой его точки существует некоторый круг с центром в этой точке, содержащийся полностью в данном множестве.

Множество называется *связным*, если любые две точки этого множества можно соединить ломаной линией, все точки которой принадлежат множеству.

Связное открытое множество называется *открытой областью*. Примером открытой области может служить множество всех точек, лежащих внутри любой замкнутой кривой, или множество, состоящее из всех внутренних точек круга, за исключением всех точек одного из его радиусов. Точки сгущения открытой области, которые сами не принадлежат этой области, называются ее *краевыми* или *граничными точками*. Множество краевых точек открытой области G называется ее *границей* или *контуром*. *Граница C области G является замкнутым множеством.* Наметим доказательство этого утверждения. Всякая точка сгущения P границы C не принадлежит открытой области G, так как любая точка последней лежит в круге, состоящем только из точек области G и не имеющем поэтому общих точек с контуром C. Но точка P является также точкой сгущения области G, так как на C можно найти точку Q, сколь угодно близкую к P, а в G можно найти точки, сколь угодно близкие к Q. Следовательно, P принадлежит границе C.

Если к открытой области присоединить все ее краевые точки, то получится замкнутое множество. В самом деле, всякая точка сгущения полученного множества является либо точкой сгущения контура C и принадлежит этому контуру, либо точкой сгущения для G и принадлежит области G или ее контуру C. Такие замкнутые множества называются *замкнутыми областями*; для наших целей они особенно полезны.

В заключение определим *окрестность* точки P как любую открытую область, содержащую эту точку. Самыми простыми видами окрестности точки $P(\xi, \eta)$ являются круговая окрестность, состоящая из всех точек (x, y), удовлетворяющих неравенству

$$(x-\xi)^2+(y-\eta)^2<\delta^2,$$

и квадратная окрестность, состоящая из всех точек (x, y), удовлетворяющих неравенствам

$$|x-\xi|<\delta \quad \text{и} \quad |y-\eta|<\delta.$$

3. Теорема Гейне — Бореля о покрытии. Еще одно следствие принципа точки сгущения представляет теорема о покрытии Гейне — Бореля, полезная для многих доказательств и более тонких исследований.

Если каждой точке ограниченного замкнутого множества M отнесена некоторая окрестность этой точки, например круг или квадрат, то из всех этих окрестностей можно выделить конечное их число таким образом, что они полностью покроют множество M. Это утверждение означает, что всякая точка множества M принадлежит хотя бы одной из конечного числа выделенных окрестностей.

Доказательство получается почти непосредственно из теоремы о стягивающейся последовательности замкнутых множеств путем умозаключения от противного. Предположим, что теорема не верна, т. е. множество M не покрывается конечным числом наших окрестностей. Множество M как ограниченное лежит в некотором квадрате Q. Квадрат Q мы разобьем на четыре равных квадрата. По крайней мере в одном из этих равных квадратов и на его границе лежит такая часть множества M, которая не может быть покрыта конечным числом окрестностей, ибо если бы каждая из четырех частей могла быть покрыта таким путем, то и само множество M оказалось бы покрытым. Эту часть множества M обозначим через M_1, и сразу видно, что M_1 — замкнутое множество. Теперь разделим квадрат, содержащий M_1, на четыре равных квадрата. Тем же рассуждением обнаружим, что часть M_2 множества M_1, лежащая в одном из этих квадратов или на его границе, не может быть покрыта конечным числом окрестностей. Продолжая этот процесс, получим последовательность вложенных друг в друга замкнутых множеств M_1, M_2, M_3, ...; каждое из них содержится в квадрате, сторона которого стремится к нулю, и ни одно из них не может быть покрыто конечным числом наших окрестностей. По теореме о стягивающейся последовательности замкнутых множеств, существует точка $Q(\xi, \eta)$, принадлежащая всем этим множествам, а следовательно и множеству M. Точке Q соответствует, согласно условию теоремы, одна из наших окрестностей, которая, очевидно, содержит малый квадрат, окружающий точку Q. Так как любое M_n содержит точку Q и само содержится в квадрате со стороной, стремящейся к нулю, то все M_n, начиная с некоторого номера, полностью лежат в малом квадрате вокруг точки Q и, стало быть, покрыты одной лишь окрестностью из всего их множества, а это противоречит полученному выводу, что ни одно из множеств M_n не может быть покрыто конечным числом этих окрестностей. Итак, предположение, что теорема ложна, приводит к противоречию, а следовательно, теорема о покрытии доказана.

Упражнения

1. Выпуклую область G можно определить как ограниченную замкнутую область, обладающую тем свойством, что если какие-либо две точки A и B принадлежат области G, то и весь отрезок AB принадлежит G. Доказать следующие утверждения:

а)* Если точка P не принадлежит выпуклой области G, то существует прямая линия, проходящая через P и не имеющая ни одной общей точки с G.

б)* Через всякую точку P границы выпуклой области G проходит такая прямая l (так называемая «опорная прямая»), что все точки области G лежат по одну и ту же сторону от прямой l или на самой этой прямой.

в) Если точка M лежит на той же стороне от любой опорной прямой, что и точки выпуклой области G, то M тоже является точкой области G.

г) Центр массы выпуклой области G принадлежит этой области.

д) Если замкнутая кривая не имеет более двух общих точек с любой прямой, то она является границей выпуклой области.

е)* Если кривизна $\frac{d\varphi}{ds}$ замкнутой кривой положительна во всех ее точках, то эта кривая образует границу выпуклой области. (Предполагается, что при обходе всей кривой ее касательная совершает полный оборот.)

2. а) Если S есть произвольное замкнутое ограниченное точечное множество, то существует «наименьшее выпуклое множество» E, содержащее S, т. е. такое множество, которое

1) содержит все точки множества S,

2) содержится полностью во всех выпуклых множествах, содержащих множество S, и

3) само является выпуклым.

б) Это множество E можно также описать и следующим путем:

Любая точка P плоскости принадлежит множеству E в том и только в том случае, если для любой прямой, оставляющей все точки исходного множества S по одну сторону от себя, точка P лежит на той же стороне.

в) Центр массы множества S принадлежит E.

§ 2. Более подробное исследование понятия предела функции многих переменных

Полезно сопоставить между собой различные процессы предельного перехода для функции многих переменных и точнее их охарактеризовать с единой точки зрения. Здесь мы также ограничимся типичным случаем двух переменных.

1. Двойные последовательности и их пределы. В случае одной переменной мы начали с изучения последовательности чисел a_n, где номер n пробегает всю натуральную последовательность чисел. Для многих переменных столь же важны *двойные последовательности*, т. е. множества чисел a_{nm}, снабженных двумя номерами n и m, которые пробегают, независимо друг от друга, все числа натурального ряда, так что двойная последовательность начинается, например, так:

$$a_{11},\ a_{12},\ a_{21},\ a_{13},\ a_{22},\ a_{31},\ a_{14},\ a_{23},\ \ldots$$

Вот примеры таких последовательностей:

$$a_{nm}=\frac{1}{n+m},\qquad a_{nm}=\frac{1}{n^2+m^2},\qquad a_{nm}=\frac{n}{n+m},$$

Введем следующее определение:

Двойная последовательность a_{nm} стремится или сходится при $n\to\infty$ и $m\to\infty$ к пределу l (двойному пределу), если абсолютная величина разности $|a_{nm}-l|$ становится меньше любого сколь угодно малого наперед заданного положительного числа ε, коль скоро оба индекса n и m выбраны достаточно большими, например, оба превышают некоторое число N, зависящее только от ε. Этот факт записывают так:

$$\lim_{\substack{n\to\infty\\ m\to\infty}} a_{nm}=l.$$

Так, например,

$$\lim_{\substack{n\to\infty\\ m\to\infty}}\frac{1}{n+m}=0,\qquad \lim_{\substack{n\to\infty\\ m\to\infty}}\frac{m+n^2}{mn^2}=\lim_{\substack{n\to\infty\\ m\to\infty}}\left(\frac{1}{n^2}+\frac{1}{m}\right)=0.$$

И здесь *критерий сходимости Коши* дает возможность судить о том, сходится ли последовательность или нет, исходя из ее определения и не нуждаясь для этого в какой-либо информации о пределе. Для двойных последовательностей критерий Коши формулируется так:

Числовая последовательность a_{nm} сходится в том и только в том случае, если для любого числа $\varepsilon>0$ существует такое число $N=N(\varepsilon)$, что $|a_{nm}-a_{n'm'}|<\varepsilon$, если только все четыре номера n, m, n', m' превосходят N.

Многочисленные вопросы анализа, касающиеся функций многих переменных, сводятся к возможности разложить такой двойной предельный переход на два последовательных обыкновенных предельных перехода. Другими словами, вместо того чтобы заставить n и m одновременно расти неограниченно, пытаются сначала, оставляя неизменным один из индексов, например m, выполнить предельный переход $n\to\infty$. Полученный таким путем предел (если он существует), будет еще в общем случае зависеть от m; обозначим его через l_m. Теперь заставим m расти неограниченно и постараемся выяснить, сходится ли l_m при $m\to\infty$. Если $\lim\limits_{m\to\infty} l_m$ существует, то возникнет вопрос, совпадает ли предел последовательности l_m с искомым двойным пределом, и если да, то при каких условиях. И другой вопрос: при разложении двойного предела на два последовательных простых предельных перехода зависит ли результат от того, в каком порядке совершаются составляющие предельные переходы, т. е. можно ли, наоборот, сначала найти $\lim\limits_{m\to\infty} a_{nm}=\lambda_n$, а затем $\lim\limits_{n\to\infty}\lambda_n$, не рискуя получить другой результат.

Для того чтобы ориентироваться в положении дела, рассмотрим сначала несколько примеров.

1) Для двойной последовательности $a_{nm}=\dfrac{1}{n+m}$, сохраняя m постоянным, получим $\lim\limits_{n\to\infty} a_{nm}=l_m=0$, а затем и $\lim\limits_{m\to\infty} l_m=0$; тот же результат получится, если совершить предельные переходы в обратном порядке.

2) Иначе обстоит дело у последовательности $a_{nm}=\dfrac{n}{n+m}=\dfrac{1}{1+\dfrac{m}{n}}$. Теперь $\lim\limits_{n\to\infty} a_{nm}=l_m=1$, так что и $\lim\limits_{m\to\infty} l_m=1$. Если же выполним предельные переходы в обратном порядке, то сначала получим $\lim\limits_{m\to\infty} a_{nm}=\lambda_n=0$, а затем $\lim\limits_{n\to\infty}\lambda_n=0$. Стало быть, в этом случае результат двух последовательных переходов к пределу *зависит* от порядка, в котором они выполнены:

$$\lim_{m\to\infty}(\lim_{n\to\infty} a_{nm})\neq\lim_{n\to\infty}(\lim_{m\to\infty} a_{nm}).$$

К тому же здесь при одновременном переходе $n\to\infty$ и $m\to\infty$ *двойной* предел вообще не существует. Действительно, если бы этот двойной предел существовал, то он должен был бы равняться нулю, так как числа a_{nm} можно сделать сколь угодно близкими к нулю, выбирая n достаточно большим и $m=n^2$. С другой стороны, если положить $n=m$, то $a_{nm}=\frac{1}{2}$, сколь бы большим ни выбрать n. Сочетание этих двух фактов противоречит допущению, что двойной предел существует.

3) Пример последовательности $a_{nm}=\dfrac{1}{n-m+\dfrac{1}{2}}$ показывает, что двойной предел может не существовать даже и в том случае, когда $\lim\limits_{m\to\infty}(\lim\limits_{n\to\infty} a_{nm})=\lim\limits_{n\to\infty}(\lim\limits_{m\to\infty} a_{nm})$.

4) Интересный пример представляет двойная последовательность $a_{nm}=\dfrac{\sin n}{m}$. Здесь двойной предел $\lim\limits_{\substack{n\to\infty\\ m\to\infty}} a_{nm}$ существует и равен нулю, так как абсолютная величина числителя не может превзойти единицу, тогда как знаменатель неограниченно возрастает. Этот же предел получается, если сперва устремить $m\to\infty$, так что $\lim\limits_{m\to\infty} a_{nm}=\lambda_n=0$, а затем совершить предельный переход $n\to\infty$: $\lim\limits_{n\to\infty}\lambda_n=0$. Но если бы мы пожелали выполнить предельные переходы в обратном порядке, оставляя сначала m неизменным и увеличивая неограниченно n, то встретились бы с затруднением: $\lim\limits_{n\to\infty}\sin n$ вообще не существует. В данном случае разложение двойного предела на два последовательных простых предельных перехода возможно только одним из двух способов.

Существующую здесь закономерность можно формулировать в виде следующих двух теорем.

Первая теорема. *Если существует двойной предел* $\lim\limits_{\substack{n\to\infty\\ m\to\infty}} a_{nm}=l$, *а простой предел* $\lim\limits_{n\to\infty} a_{nm}=l_m$ *существует при всяком значении* m, *то существует также и предел* l_m *при* $m\to\infty$, *и притом* $\lim\limits_{m\to\infty} l_m=l$. *Если, кроме того, существует также и* $\lim\limits_{m\to\infty} a_{nm}=\lambda_n$ *при любом значении* n, *то и* λ_n *при* $n\to\infty$ *стремится к тому же пределу* l.

Это можно записать в виде одной формулы

$$l=\lim_{\substack{n\to\infty\\ m\to\infty}} a_{nm}=\lim_{m\to\infty}(\lim_{n\to\infty} a_{nm})=\lim_{n\to\infty}(\lim_{m\to\infty} a_{nm}),$$

так что двойной предел можно разложить на простые предельные переходы, и результат не зависит от порядка составляющих предельных переходов.

Доказательство вытекает почти сразу из определения этих пределов. Из существования двойного предела l следует, что для всякого $\varepsilon>0$ имеется такое число $N=N(\varepsilon)$, что $|a_{nm}-l|<\varepsilon$, как только n и m становятся больше чем N. Оставляя в этом неравенстве m неизменным и неограниченно увеличивая n, находим, что $|\lim\limits_{n\to\infty} a_{nm}-l|=|l_m-l|\leqslant\varepsilon$. Это неравенство выполняется при любом положительном ε, при том лишь условии, что $m>N(\varepsilon)$; стало быть, оно равносильно утверждению, что $\lim\limits_{m\to\infty}(\lim\limits_{n\to\infty} a_{nm})=l$. Точно таким же способом доказывается и вторая часть теоремы.

Вторая теорема является в некотором смысле обратной по отношению к первой. Она дает достаточное условие эквивалентности двух последовательных предельных переходов двойному пределу. Эта теорема основывается на понятии *равномерной сходимости*, которое определяется следующим образом:

Говорят, что *последовательность* a_{nm} *при* $n\to\infty$ *сходится* (*или стремится*) *к пределу* l_m *равномерно относительно* m, *если не только существует при всяком* m *предел* $\lim\limits_{n\to\infty} a_{nm}=l_m$, *но, сверх того, для любого сколь угодно малого* $\varepsilon>0$ *можно найти такое число* $N=N(\varepsilon)$, *зависящее только от* ε, *но не зависящее от* m, *что* $|l_m-a_{nm}|<\varepsilon$, *коль скоро* $n>N$.

Так, например, последовательность $a_{nm}=\dfrac{n}{m(n+m)}=\dfrac{1}{m}-\dfrac{1}{n+m}$ сходится равномерно относительно m к пределу $l_m=\dfrac{1}{m}$, как это видно из оценки

$$\left| a_{nm} - \frac{1}{m} \right| = \frac{1}{n+m} < \frac{1}{n},$$

— стоит только положить $N \geqslant \frac{1}{\varepsilon}$.

Напротив, последовательность $a_{nm} = \frac{m}{n+m}$ при $n \to \infty$ стремится к своему пределу неравномерно (относительно m). Хотя при любом постоянном значении m предел $\lim\limits_{n\to\infty} a_{nm} = l_m = 0$, однако сходимость эта неравномерна. Действительно, при любом наперед заданном ε, как бы ни было велико значение n, всегда можно найти такие значения m, для которых $|a_{nm} - l_m| = a_{nm} > \varepsilon$; такие значения m найдем, решая уравнение $\frac{m}{n+m} > \varepsilon$; окажется, что достаточно выбрать $m > \frac{\varepsilon n}{1-\varepsilon}$, чтобы a_{nm} отличалось от своего предела $l_m = 0$ больше чем на ε.

Вторая теорема. *Если последовательность a_{nm} при постоянном m стремится к пределу $\lim\limits_{n\to\infty} a_{nm} = l_m$ равномерно относительно m и если к тому же существует $\lim\limits_{m\to\infty} l_m = l$, тогда существует и двойной предел $\lim\limits_{\substack{n\to\infty\\ m\to\infty}} a_{nm}$ и имеет то же самое значение l*:

$$\lim_{m\to\infty} \left(\lim_{n\to\infty} a_{nm} \right) = \lim_{\substack{n\to\infty\\ m\to\infty}} a_{nm}.$$

Если при этом известно, что и $\lim\limits_{m\to\infty} a_{nm} = \lambda_n$ существует, то (на основании первой теоремы) порядок предельных переходов можно изменить.

Доказательство приводится тем же путем, что и для первой теоремы, с помощью неравенства $|a_{nm} - l| \leqslant |a_{nm} - l_m| + |l_m - l|$; мы предоставляем его читателю.

2. Двойной предел в случае непрерывно изменяющихся независимых переменных. Во многих случаях встречаются такие предельные переходы, в которых участвуют одновременно как целочисленные индексы, так и одна или несколько непрерывных переменных, причем индексы, например n, неограниченно возрастают, а непрерывные переменные $x, y, \ldots$ стремятся к определенным пределам $\xi, \eta, \ldots$; встречаются и такие процессы, в которых нет индексов, а участвуют в них в качестве аргументов только лишь непрерывные переменные, стремящиеся к своим пределам. Ко всем таким случаям принципиально вполне приложимы известные уже нам рассуждения и результаты без существенных изменений. Прежде всего отметим, что понятие предела последовательности функций $f_n(x)$ или $f_n(x, y)$ при неограниченном возрастании n является как раз примером первого типа только что перечисленных предельных переходов. Действительно, величина $f_n(x)$ или $f_n(x, y)$ зависит не только от переменной x или от (x, y), непрерывно изменяющихся в некоторой области, но еще и от целочисленного аргумента — индекса n. Мы видели в первом томе (гл. VIII, § 4), что если последовательность непрерывных функций $f_n(x)$ сходится равномерно, то и предельная функция $f(x)$ также непрерывна. Эта теорема приводит к следующим равенствам:

$$f(\xi) = \lim_{x \to \xi} f(x) = \lim_{x \to \xi} [\lim_{n \to \infty} f_n(x)] = \lim_{n \to \infty} f_n(\xi) = \lim_{n \to \infty} [\lim_{x \to \xi} f_n(x)],$$

выражающим переместительность предельных переходов $n \to \infty$ и $x \to \xi$. Эта теорема справедлива и в том случае, когда члены последовательности $f_n(x, y, \ldots)$ зависят от нескольких переменных, — при этом ничто не изменяется ни в определениях, ни в доказательствах.

С другими примерами, показывающими, какое большое значение имеет в анализе вопрос о переместительности предельных переходов, мы уже встречались не раз в прошлом, например в теореме о перемене порядка дифференцирования при вычислении частных производных, и еще встретимся в будущем. Здесь мы рассмотрим еще только пример функции

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}.$$

При постоянном $y \neq 0$ и $x \to 0$ получается $\lim\limits_{x \to 0} f(x, y) = -1$, а при постоянном $x \neq 0$ и $y \to 0$ другой предел $\lim\limits_{y \to 0} f(x, y) = 1$. Стало быть, у этой функции $\lim\limits_{y \to 0} [\lim\limits_{x \to 0} f(x, y)] \neq \lim\limits_{x \to 0} [\lim\limits_{y \to 0} f(x, y)]$ и результат зависит от порядка предельных переходов. Это связано, очевидно, с наличием разрыва функции в начале координат.

Заметим в заключение, что *в случае непрерывных переменных вопрос о разложении двойного предела на последовательные простые предельные переходы и о переместительности порядка составляющих простых предельных переходов подчиняется таким же точно теоремам, какие были только что доказаны в* n° 1 *для двойных последовательностей.*

3. Теорема Дини о равномерной сходимости монотонных последовательностей функций. Во многих более тонких аналитических исследованиях применяется с пользой одна общая теорема о равномерной сходимости. Мы уже знаем (т. I, стр. 449—450), что последовательность непрерывных функций может сходиться к непрерывной предельной функции, несмотря на то, что эта сходимость неравномерна. Однако существует важный случай, когда из непрерывности предельной функции вытекает равномерность сходимости. Это тот случай, когда последовательность функций монотонна, т. е. при любых фиксированных значениях x и y значение функции $f_n(x, y)$, с возрастанием n, либо никогда не убывает, либо никогда не возрастает. Теорема формулируется так:

Если в замкнутой области G последовательность непрерывных функций $f_n(x, y)$ сходится к непрерывной предельной функции $f(x, y)$ и если во всех точках области $f_{n+1}(x, y) \geqslant f_n(x, y)$ либо во всех точках области $f_{n+1}(x, y) \leqslant f_n(x, y)$, то сходимость равномерна в области G.

При доказательстве можно ограничиться случаем, когда $f_n(x, y)$ монотонно возрастает или монотонно не убывает (первое неравенство), — при втором предположении рассуждение аналогично. Приводимое здесь доказательство от противного представляет собой типичный пример применения принципа точки сгущения. Если бы сходимость была неравномерной, то существовало бы такое положительное число α, что для некоторой бесконечной последовательности номеров $n_1, n_2, \ldots$ нашлось бы в области G по точке $P_{n_1}, P_{n_2}, \ldots$, в которой значение функции $f_{n_k}(P_{n_k})$ отличалось бы от значения предельной функции $f(P_{n_k})$ больше чем на α. Последовательность точек $P_{n_1}, P_{n_2}, \ldots$ имеет по крайней мере одну точку сгущения Q, а так как область G замкнутая, то точка Q должна принадлежать этой области. Но для всякой точки P области G и всякого натурального числа μ должно быть

$$f(P) = f_\mu(P) + R_\mu(P),$$

где $f_\mu(P)$ и «остаток» $R_\mu(P)$ являются непрерывными функциями точки P. Так как последовательность функций $f_n(P)$ монотонно возрастает (или по крайней мере не убывает), то $R_\mu(P) \geqslant R_n(P)$ при $n > \mu$. В частности, в каждой из точек P_{n_k} при $n_k > \mu$ будет

$$R_\mu(P_{n_k}) \geqslant R_{n_k}(P_{n_k}) \geqslant \alpha.$$

Из последовательности точек $P_{n_1}, P_{n_2}, P_{n_3}, \ldots$ можно выделить подпоследовательность, стремящуюся в пределе к точке Q; в силу непрерывности $R_\mu(P)$ при любом фиксированном значении μ будет также $R_\mu(Q) \geqslant \alpha$. Так как в нашем предельном переходе номер n_k неограниченно возрастает, то индекс μ можно взять сколь угодно большим, ибо, как бы велико ни было μ, всегда еще будет бесчисленное множество номеров n_k, превышающих μ. Но неравенство $R_\mu(Q) \geqslant \alpha$ при сколь угодно большом μ противоречит тому факту, что $R_\mu(Q)$ стремится к нулю при возрастании μ. Следовательно, предположение, что сходимость неравномерна, приводит к противоречию, что и доказывает нашу теорему.

Упражнения

1. Выяснить, существуют ли следующие двойные пределы:

а) $\lim\limits_{\substack{n \to \infty \\ m \to \infty}} \dfrac{(\ln n)^2 - (\ln m)^2}{(\ln n)^2 + (\ln m)^2}$;

б) $\lim\limits_{\substack{n\to\infty\\ m\to\infty}} \dfrac{\operatorname{tg} n+\operatorname{tg} m}{1-\operatorname{tg} n \operatorname{tg} m}$;

в) $\lim\limits_{\substack{n\to\infty\\ m\to\infty}} \dfrac{1}{m^2}\sum\limits_{\nu=1}^{n} \cos\dfrac{\nu}{m}$.

2. Доказать, что функция $f(x, y)$ непрерывна, если а) при фиксированном y функция f является непрерывной функцией от x; б) при фиксированном x функция f равномерно непрерывна относительно y в том смысле, что для всякого $\varepsilon>0$ существует такое $\delta>0$ (не зависящее от x и y), что $|f(x, y_1)-f(x, y)|\leqslant\varepsilon$, коль скоро $|y_1-y|\leqslant\delta$.

3. Доказать, что функция $f(x, y)$ непрерывна в точке $x=0$, $y=0$, если функция $\Phi(t, \varphi)=f(t\cos\varphi, t\sin\varphi)$

а) является непрерывной функцией от t при постоянном φ и

б) равномерно непрерывна относительно φ при постоянном t, так что для всякого $\varepsilon>0$ существует такое $\delta>0$ (не зависящее от t и φ), что $|\Phi(t, \varphi_1)-\Phi(t, \varphi)|\leqslant\varepsilon$, если только $|\varphi_1-\varphi|\leqslant\delta$.

4. Доказать, что дополнительное множество для замкнутого множества S (т. е. множество всех точек, не принадлежащих S) является открытым множеством.

§ 3. Однородные функции

В заключение сообщим некоторые сведения об *однородных функциях*. Простейшими однородными функциями являются однородные целые рациональные функции (короче — многочлены) от многих переменных. Многочлен $ax+by$ называется однородной функцией первой степени относительно x и y, многочлен $ax^2+bxy+cy^2$ — однородной функцией второй степени и вообще *целая рациональная функция от x и y или от большего числа переменных называется однородной функцией степени n, если сумма показателей независимых переменных имеет во всех членах одно и то же значение n*, так что в случае двух переменных x и y эта функция имеет следующий вид:

$$a_0x^n+a_1x^{n-1}y+a_2x^{n-2}y^2+\ldots+a_ny^n,$$

где $a_0, a_1, \ldots, a_n$ — постоянные коэффициенты. Эти однородные многочлены обладают тем свойством, что при любом значении t выполняется тождество

$$f(tx, ty)=t^n f(x, y).$$

Это свойство однородных многочленов кладут в основу общего определения однородной функции:

Функция $f(x, y, \ldots)$ называется однородной функцией измерения (или степени) h, если она удовлетворяет равенству

$$f(tx, ty, \ldots)=t^h f(x, y, \ldots).$$

Приведем несколько примеров однородных функций, не являющихся многочленами:

1) $$\operatorname{tg}\left(\frac{y}{x}\right), \qquad h=0.$$

2) $$x^2 \sin\frac{x}{y} + y\sqrt{x^2+y^2}\ln\frac{x+y}{x}, \qquad h=2.$$

3) $$\frac{x^3+3xy^2+z^3}{xy+y^2+xz}, \qquad h=1.$$

4) Косинус угла между двумя векторами $\{x, y, z\}$ и $\{u, v, w\}$

$$\frac{xu+yv+zw}{\sqrt{x^2+y^2+z^2}\cdot\sqrt{u^2+v^2+w^2}}$$

является однородной функцией шести переменных измерения $h=0$.

5) Длина вектора $\{x, y, z\}$, т. е. $\sqrt{x^2+y^2+z^2}$, является примером функции *положительно* однородной первого измерения, т. е. тождество, служащее определением однородных функций, выполняется здесь лишь при $t \geqslant 0$.

Положительно однородной является и функция 2).

Если однородная функция дифференцируема, то она удовлетворяет *соотношению однородности Эйлера*

$$xf_x + yf_y + zf_z + \ldots = hf(x, y, z),$$

которое представляет собой дифференциальное уравнение с частными производными первого порядка. Для доказательства продифференцируем по t определяющее тождество $f(tx, ty, tz, \ldots) = t^h f(x, y, z, \ldots)$. Применяя к левой части правило цепочки, получим

$$xf_x(tx, ty, \ldots) + yf_y(tx, ty, \ldots) + \ldots = ht^{h-1} f(x, y, \ldots).$$

При $t=1$ отсюда вытекает доказываемое соотношение.

Соотношение Эйлера вполне характеризует однородные функции; это значит, что не только соотношение Эйлера является следствием однородности функции, но и, обратно, всякая функция, удовлетворяющая соотношению Эйлера, есть однородная функция, так что *соотношение Эйлера служит необходимым и достаточным условием однородности функции.*

Прежде чем приступить к доказательству упомянутой обратной теоремы, заметим, что определяющее свойство однородных функций равносильно следующему свойству: если разделить функцию на x^h, то полученное частное зависит только от отношений $\frac{y}{x}, \frac{z}{x}, \ldots$

[Действительно, подставив $t=\frac{1}{x}$ в тождество $t^h f(x, y, z, \ldots) = f(tx, ty, tz, \ldots)$, получим $\frac{f(x, y, z, \ldots)}{x^h} = f\left(1, \frac{y}{x}, \frac{z}{x}, \ldots\right) = \varphi\left(\frac{y}{x}, \frac{z}{x}, \ldots\right)$.

Обратно, если $\frac{f(x, y, z, \ldots)}{x^h} = \varphi\left(\frac{y}{x}, \frac{z}{x}, \ldots\right)$, то $f(x, y, z, \ldots) = x^h\varphi\left(\frac{y}{x}, \frac{z}{x}, \ldots\right)$. Подставив сюда вместо $x, y, z, \ldots$ величины $tx, ty, tz, \ldots$, получим $f(tx, ty, tz, \ldots) = t^h x^h \varphi\left(\frac{y}{x}, \frac{z}{x}, \ldots\right) = t^h f(x, y, z, \ldots)$.]

Наша цель — доказать, что функция, удовлетворяющая соотношению Эйлера, является однородной функцией степени h. На основании сказанного, для этого достаточно показать, что если вместо переменных y, z и т. д. ввести новые переменные $\eta = \frac{y}{x}$, $\zeta = \frac{z}{x}$, ..., так что $y = \eta x$, $z = \zeta x$, ... (переменную x сохраняем), то функция

$$\frac{1}{x^h} f(x, y, z, \ldots) = \frac{1}{x^h} f(x, \eta x, \zeta x, \ldots) = g(x, \eta, \zeta, \ldots)$$

в действительности уже не зависит от переменной x, так что $g_x = 0$ тождественно. А это легко проверить, дифференцируя последнее равенство по x по правилу цепочки:

$$g_x = \frac{1}{x^h}(f_x + \eta f_y + \zeta f_z + \ldots) - \frac{h}{x^{h+1}} f(x, y, z, \ldots) =$$

$$= \frac{1}{x^{h+1}}(xf_x + yf_y + zf_z + \ldots) - \frac{h}{x^{h+1}} f(x, y, z, \ldots).$$

Последнее выражение обращается в нуль в силу соотношения Эйлера, и тем самым обратная теорема доказана.

Приведем еще другое доказательство, более изящное, хотя и менее прямое. Наша цель — вывести из соотношения Эйлера, что функция

$$g(t) = f(tx, ty, \ldots) - t^h f(x, y, \ldots)$$

равна нулю при всех значениях t. Сразу видно, что $g(1) = 0$. Далее,

$$g'(t) = xf_x(tx, ty, \ldots) + yf_y(tx, ty, \ldots) + \ldots - ht^{h-1} f(x, y, \ldots).$$

Подставляя в соотношение Эйлера вместо $x, y, \ldots$ соответственно $tx, ty, \ldots$, получим

$$txf_x(tx, ty, \ldots) + tyf_y(tx, ty, \ldots) + \ldots = hf(tx, ty, \ldots).$$

С помощью этого равенства выражение для $g'(t)$ принимает следующий вид:

$$g'(t) = \frac{h}{t} g(t).$$

Это — дифференциальное уравнение с отделяющимися переменными, которому должна удовлетворять функция $g(t)$. Отделяя переменные:

$$\frac{dg(t)}{g(t)} = \frac{h}{t} dt$$

и интегрируя, получим

$$\ln|g(t)| = h\ln|t| + \ln|C|, \qquad \text{откуда} \qquad |g(t)| = |C|\cdot|t^h|,$$

и окончательно $g(t) = Ct^h$. Постоянную интегрирования C определяем из найденного выше условия $g(1) = 0$, которое дает $C = 0$, так что $g(t) = 0$ при всех значениях t, что и требовалось доказать.

Упражнения

1. Доказать, что если $f(x, y, z, \ldots)$ есть однородная функция степени h, то ее производная k-го порядка является однородной функцией степени $h - k$.

2. Доказать, что однородная функция f первой степени удовлетворяет дифференциальному уравнению

$$f_{xx} + f_{yy} + f_{zz} + \ldots + 2f_{xy} + 2f_{xz} + \ldots = 0.$$

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ II

1. а) Показать, что функция u вида $u(x, y) = f(x)g(y)$ удовлетворяет дифференциальному уравнению с частными производными $uu_{xy} - u_x u_y = 0$.

б) Доказать обратное предложение.

2. Доказать, что

$$u(x, y, z) = \frac{f(t+r)}{r} + \frac{g(t-r)}{r}, \qquad \text{где} \qquad r^2 = x^2 + y^2 + z^2,$$

удовлетворяет дифференциальному уравнению $\nabla^2 u = u_{tt}$.

3. Показать, что если частные производные функции $u(x, y)$ удовлетворяют соотношению вида $F(u_x, u_y) = 0$, то функция u удовлетворяет дифференциальному уравнению $u_{xx}u_{yy} - u_{xy}^2 = 0$.

4*. Доказать, что поверхность $u = \varphi(x, y)$, образуемая прямыми линиями, пересекающими ось u, или, что то же, поверхность, пересекающая все плоскости семейства $y = cx$ по прямым линиям, удовлетворяет дифференциальному уравнению

$$x^2 u_{xx} + 2xyu_{xy} + y^2 u_{yy} = 0.$$

5. Найти для заданного ε соответствующее $\delta = \delta(\varepsilon, x, y)$ для следующих непрерывных функций (ср. гл. II, § 2, n° 1):

а) $f(x, y) = \sqrt{1 + x^2 + 2y^2}$; б) $f(x, y) = \sqrt{1 + e^{xy}}$.

6. Показать, что функции

$$f(x, y) = \frac{x^4y^4}{(x^2+y^4)^3}, \qquad g(x, y) = \frac{x^2}{x^2 + y^2 - x}$$

стремятся к нулю, когда точка (x, y) приближается к началу координат вдоль любой прямой, проходящей через начало, но эти функции имеют разрыв в начале координат.

7. Дана гладкая плоская кривая с непрерывно вращающейся касательной. Обозначим через d расстояние между двумя точками P и Q кривой, т. е. длину хорды PQ, а через l длину дуги $\widehat{PQ}$. Доказать, что $l - d = o(d)$, когда $d \to 0$.

8. Вычислить $S=\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\frac{(a+b)!}{a!\,b!}\,\frac{a}{x^a y^b}$ при условии, что $\frac{1}{x}+\frac{1}{y}<1$, $x>0$, $y>0$.

9. Показать, пользуясь соотношением Эйлера (стр. 129), что однородная функция $S_n(x, y, z)$ степени n, удовлетворяющая уравнению Лапласа $\nabla^2 S_n=0$, удовлетворяет также соотношению

$$\nabla^2(r^{2m}S_n)=2m(2n+2m+1)\,r^{2m-2}S_n,$$

где $r^2=x^2+y^2+z^2$.

10. Вывести следующую формулу для кривизны пространственной кривой $\boldsymbol{r}=\boldsymbol{r}(t)$, где t — произвольный параметр:

$$k=\sqrt{\frac{\dot{\boldsymbol{r}}^2\ddot{\boldsymbol{r}}^2-(\dot{\boldsymbol{r}}\ddot{\boldsymbol{r}})^2}{(\dot{\boldsymbol{r}}^2)^3}}=\frac{|[\dot{\boldsymbol{r}}\ddot{\boldsymbol{r}}]|}{|\dot{\boldsymbol{r}}|^3}$$

(точки над символом вектора означают дифференцирование по t).

11*. Пространственная кривая C задана уравнениями $x=x(s)$, $y=y(s)$, $z=as$, параметром служит длина дуги s плоской кривой $x=x(s)$, $y=y(s)$. Доказать, что соприкасающаяся плоскость кривой C в ее точке P (ср. упр. 1, стр. 114) содержит нормаль к цилиндрической поверхности $x=x(s)$, $y=y(s)$ в точке P. Показать, что кривизна и кручение кривой C (см. упр. 7, стр. 115) определяются по формулам

$$k=\frac{|\dot{x}\ddot{y}-\ddot{x}\dot{y}|}{1+a^2},\qquad \varkappa=\frac{a(\dot{x}\ddot{y}-\ddot{x}\dot{y})}{1+a^2}.$$

12. Составить уравнение соприкасающейся плоскости (см. упр. 1, стр. 114, и его решение) в любой точке кривой $x=\cos\theta$, $y=\sin\theta$, $z=f(\theta)$. Показать, что если $f(\theta)=\frac{1}{A}\operatorname{ch}A\theta$, то любая соприкасающаяся плоскость кривой касается сферы с центром в начале координат и радиусом, равным $\sqrt{1+\frac{1}{A^2}}$.

13. На цилиндре $x^2+y^2=a^2$ начерчена кривая, обладающая тем свойством, что угол между осью z и касательной к этой кривой в любой ее точке P равен углу между осью y и касательной плоскостью к цилиндру в той же точке P. Доказать, что координаты любой точки P кривой можно выразить через некоторый параметр θ уравнениями

$$x=a\cos\theta,\qquad y=a\sin\theta,\qquad z=c\pm a\ln\sin\theta$$

и что кривизна кривой равна $\frac{1}{a}\sin\theta\sqrt{1+\sin^2\theta}$.

14. Дана кривая $x=\frac{1}{3}at^3$, $y=\frac{1}{2}bt^2$, $z=ct$.

а) Доказать, что уравнение плоскости, проходящей через три точки t_1, t_2, t_3 этой кривой, есть

$$\frac{3x}{a}-2(t_1+t_2+t_3)\frac{y}{b}+(t_2t_3+t_3t_1+t_1t_2)\frac{z}{c}-t_1t_2t_3=0.$$

б) Показать, что точка пересечения соприкасающихся плоскостей кривой в точках t_1, t_2, t_3 лежит в этой плоскости.

15. Пусть a, b, c — стороны треугольника, A, B, C — противолежащие им углы, S — площадь треугольника, R — радиус описанной окружности. Показать, что

$$dS = R(\cos A da + \cos B db + \cos C dc).$$

16. Движение точки P относительно некоторой системы координат считается известным, т. е. задан ее радиус-вектор $\boldsymbol{r} = \boldsymbol{r}(t)$ как функция времени t. Тогда скорость точки P равна вектору $\dot{\boldsymbol{r}}$. Доказать, что

$$\frac{dr}{dt} = \boldsymbol{r}^{\circ}\dot{\boldsymbol{r}} = \boldsymbol{r}^{\circ}\frac{d\boldsymbol{r}}{dt},$$

где $r = |\boldsymbol{r}|$ есть полярный радиус точки P, а $\boldsymbol{r}^{\circ}$ есть единичный вектор радиус-вектора. Словесно: (скалярная) скорость изменения полярного радиуса точки равна проекции ее скорости на направление ее радиус-вектора.

ГЛАВА III

ПОСТРОЕНИЕ ДИФФЕРЕНЦИАЛЬНОГО ИСЧИСЛЕНИЯ И ЕГО ПРИЛОЖЕНИЯ

§ 1. Неявные функции

1. Общие замечания. В аналитической геометрии кривые часто задаются не уравнением вида $y=f(x)$ или $x=\varphi(y)$, а уравнением вида $F(x, y)=0$. Например, прямую можно представить общим уравнением $ax+by+c=0$, а эллипс — уравнением $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. Для того чтобы привести такое уравнение кривой к виду $y=f(x)$, надо решить уравнение $F(x, y)=0$ относительно y.

В первом томе мы рассмотрели задачу о нахождении обратной функции $y=f(x)$ или, что то же самое, о решении уравнения $F(x, y)\equiv y-f(x)=0$ относительно переменной x. Все эти примеры показывают, как важно изучение вопроса о решении уравнения вида $F(x, y)=0$ относительно x или y. Мы теперь займемся этим исследованием, а полученные здесь результаты обобщим в § 3 на функции многих переменных.

В простейших случаях, вроде приведенных выше примеров, явное решение можно выразить через элементарные функции. В других случаях результат можно получить приближенно с любой желательной точностью. Однако для многих целей целесообразнее работать не с разрешенным явным видом уравнения (точным или приближенным), а вместо этого исследовать решение путем изучения самой функции $F(x, y)$, в которой не отдается предпочтения ни одной из переменных x и y.

Было бы ошибочно думать, что при любой функции $F(x, y)$ уравнение $F(x, y)=0$ определяет в неявном виде функцию $y=f(x)$ или $x=\varphi(y)$. Напротив, легко привести примеры таких функций $F(x, y)$, для которых уравнение $F(x, y)=0$ не имеет решения в виде функции одной переменной. Так, например, уравнению $x^2+y^2=0$ удовлетворяет единственная пара значений $x=y=0$, тогда как уравнению $x^2+y^2+1=0$ не удовлетворяет ни одна (действительная) пара значений (x, y). Необходимо поэтому внимательно исследовать, в каких случаях уравнение $F(x, y)=0$ определяет в неявном виде функцию $y=f(x)$ и какими свойствами обладает эта функция.

2. Геометрическое истолкование [1]. Для того чтобы уяснить себе положение дел, изобразим функцию $u=F(x, y)$ поверхностью в трех-

[1]) Ср. также т. I, гл. X, § 5.

мерном пространстве. Интересующие нас решения уравнения $F(x, y)=0$ являются общими решениями двух уравнений $u=F(x, y)$ и $u=0$. Геометрически наша задача состоит в том, чтобы выяснить, существуют ли кривые $y=f(x)$ или $x=\varphi(y)$, по которым поверхность $u=F(x, y)$ пересекает плоскость xy. (Как *далеко* может простираться такая линия пересечения, нас здесь не интересует.)

Первая возможность, которая приходит на ум: поверхность и плоскость могут не иметь ни одной общей точки. Например, параболоид $u=F(x, y)=x^2+y^2+1$ весь лежит выше плоскости $x0y$. В этом случае, очевидно, не существует линии пересечения. Поэтому подлежат рассмотрению лишь те случаи, в которых существует точка (x_0, y_0), в которой $F(x_0, y_0)=0$; пару значений x_0, y_0 мы будем называть *начальным решением.*

Если начальное решение существует, то остаются две возможности: либо плоскость $u=0$ касается поверхности в точке (x_0, y_0), либо нет. В первом случае легко показать на примерах, что решение $y=f(x)$ или $x=\varphi(y)$ может и не существовать. Например, параболоид $u=x^2+y^2$ имеет начальное решение $x=0$, $y=0$, но не имеет других общих точек с плоскостью $x0y$. Еще пример: гиперболический параболоид $u=xy$ имеет начальное решение $x=y=0$

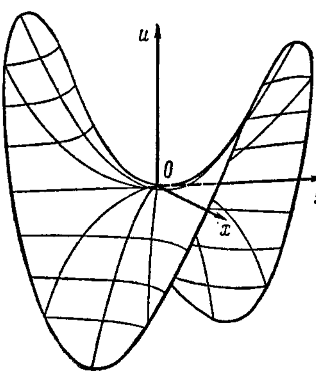


Рис. 28.

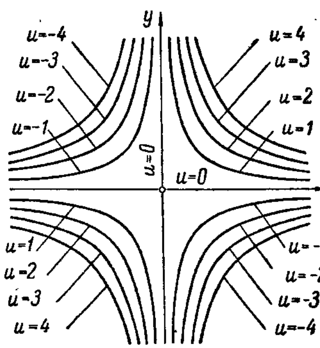


Рис. 29.

и действительно пересекает плоскость $x0y$ по прямым $x=0$ и $y=0$ (ср. рис. 28 и 29). Но ни в какой окрестности начала координат невозможно представить *всю* линию пересечения с помощью функции вида $y=f(x)$ или вида $x=\varphi(y)$. С другой стороны, и в этом случае, когда плоскость $u=0$ является касательной плоскостью к поверхности в точке, служащей начальным решением, не исключена возможность, что уравнение $F(x, y)=0$ имеет решение, как показывает пример $F(x, y)\equiv(y-x)^4=0$. Следовательно, в том случае (его естественно считать исключительным), когда плоскость $u=0$ является касательной плоскостью к поверхности $u=F(x, y)$ в точке

(x_0, y_0), не представляется возможным высказать заранее определенное общее заключение.

Остается еще возможность, что в точке (x_0, y_0) — начальном решении — касательная плоскость существует, но не совпадает с плоскостью $x0y$. Тогда наглядное представление показывает, грубо говоря, что поверхность $u = F(x, y)$ не может изгибаться настолько быстро, чтобы избежать пересечения с плоскостью $x0y$ в окрестности точки (x_0, y_0) по единственной вполне определенной кривой, и что часть этой кривой вблизи начального решения может быть выражена уравнением вида $y = f(x)$ или $x = \varphi(y)$. Но утверждение, что касательная плоскость в точке (x_0, y_0) плоскости $x0y$ не совпадает с этой плоскостью, равносильно утверждению, что частные производные $F_x(x_0, y_0)$ и $F_y(x_0, y_0)$ не равны одновременно нулю. Этот-то случай мы и рассмотрим аналитически в следующем номере.

3. Теорема существования неявной функции и правило ее дифференцирования. Общая теорема, устанавливающая достаточные условия существования неявной функции и дающая в то же время правило ее дифференцирования, формулируется так:

Если функция $F(x, y)$ имеет непрерывные производные F_x и F_y и если точка (x_0, y_0) из области ее определения удовлетворяет уравнению $F(x_0, y_0) = 0$, между тем как $F_y(x_0, y_0) \neq 0$, то вокруг точки (x_0, y_0) можно выделить такой прямоугольник $x_1 \leqslant x \leqslant x_2$, $y_1 \leqslant y \leqslant y_2$, что в интервале $x_1 \leqslant x \leqslant x_2$ уравнением $F(x, y) = 0$ определяется однозначная непрерывная функция $y = f(x)$, значения которой лежат в интервале $y_1 \leqslant y \leqslant y_2$. Эта функция принимает при $x = x_0$ значение $y_0 = f(x_0)$ и в каждой точке x интервала $[x_1, x_2]$ удовлетворяет уравнению

$$F[x, f(x)] = 0.$$

Функция $y = f(x)$ дифференцируема, а ее производная и дифференциал выражаются следующими формулами:

$$y' = f'(x) = -\frac{F_x}{F_y} \quad и \quad dy = df(x) = -\frac{F_x}{F_y}\,dx.$$

Эти формулы выражают *правило дифференцирования неявной функции.*

В этом номере мы будем предполагать, что первая часть теоремы, говорящая о существовании и непрерывности неявно заданной функции, уже доказана, и ограничимся здесь лишь доказательством дифференцируемости функции $y = f(x)$ и выводом правила дифференцирования. Аналитическое доказательство существования и непрерывности неявной функции мы отложим до n° 6.

Если бы можно было дифференцировать тождество $F[x, f(x)] = 0$ по правилу цепочки, то сразу получили бы формулы дифференцирования (ср. т. I, стр. 577). Но так как дифференцируемость функции $f(x)$

требуется еще доказать, то мы вынуждены воспользоваться более сложным рассуждением.

Согласно условию теоремы, производные F_x и F_y непрерывны. Следовательно, функция $F(x, y)$ дифференцируема, и можно написать

$$F(x+h, y+k)=F(x, y)+hF_x(x, y)+kF_y(x, y)+\varepsilon_1 h+\varepsilon_2 k,$$

где ε_1 и ε_2 — два числа, стремящихся к нулю вместе с h и k или, что то же, вместе с $\rho=\sqrt{h^2+k^2}$. Будем рассматривать только такие пары значений (x, y) и $(x+h, y+k)$, для которых x и $x+h$ лежат в интервале $x_1 \leqslant x \leqslant x_2$, причем $y=f(x)$ и $y+k=f(x+h)$. Для таких пар значений будет $F(x, y)=0$ и $F(x+h, y+k)=0$, так что последнее равенство примет следующий вид:

$$0=hF_x+kF_y+\varepsilon_1 h+\varepsilon_2 k.$$

Мы уже условились считать здесь доказанной непрерывность функции $f(x)$; поэтому при $h \to 0$ будет и $k \to 0$, а вместе с ними также $\varepsilon_1 \to 0$ и $\varepsilon_2 \to 0$. По условию, $F_y \neq 0$, стало быть, можно разделить полученное равенство на hF_y, что дает

$$\left(1+\frac{\varepsilon_2}{F_y}\right)\frac{k}{h}+\frac{F_x}{F_y}+\frac{\varepsilon_1}{F_y}=0,$$

а отсюда, переходя к пределу при $h \to 0$, получим

$$\lim_{h \to 0}\frac{k}{h}+\frac{F_x}{F_y}=0.$$

Это равенство и доказывает существование предела

$$\lim_{h \to 0}\frac{k}{h}=\lim_{h \to 0}\frac{f(x+h)-f(x)}{h},$$

т. е. дифференцируемость функции $f(x)$ и вместе с тем дает требуемое правило дифференцирования

$$y'=f'(x)=-\frac{F_x}{F_y}.$$

Это правило можно записать и в таком виде:

$$F_x+F_y y'=0,$$

а также в виде

$$dF=F_x\,dx+F_y\,dy=0.$$

Это равенство, устанавливающее, что в силу уравнения $F(x, y)=0$ полный дифференциал функции $F(x, y)$ обращается в нуль, показывает, что по той же причине дифференциалы dx и dy нельзя выбирать независимо друг от друга.

Неявную функцию обычно проще дифференцировать по этому правилу, чем путем предварительного нахождения явного выражения функции и прямого дифференцирования этого выражения. Это правило всегда приводит к цели, когда неявная функция (на основании теоремы этого номера) существует, даже в тех нередких случаях, когда явное выражение функции через обычные (рациональные, тригонометрические и т. д.) практически получить очень сложно, а порою и невозможно.

Предположим, что функция $F(x, y)$ имеет непрерывные частные производные второго порядка. Тогда можно равенство $y' = -\frac{F_x}{F_y}$, правая часть которого представляет собой сложную функцию от x, дифференцировать по x, пользуясь правилом дифференцирования дроби и правилом цепочки:

$$y'' = -\frac{F_y(F_{xx} + F_{xy}y') - F_x(F_{yx} + F_{yy}y')}{F_y^2}.$$

Подставив сюда вместо y' ее значение $-\frac{F_x}{F_y}$, получим формулу для второй производной от функции $y = f(x)$:

$$y'' = -\frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_y^3}.$$

Таким же путем последовательного дифференцирования можно вычислять и дальнейшие производные от функции $y = f(x)$, заданной в неявном виде.

4. Примеры. 1) Из уравнения *эллипса*

$$F(x, y) \equiv \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

находим для определенной этим уравнением функции $y = f(x)$ ее производную

$$y' = -\frac{F_x}{F_y} = -\frac{b^2x}{a^2y}.$$

Проверим этот результат. Решая уравнение эллипса относительно y, получим две однозначные функции: $y = \frac{b}{a}\sqrt{a^2 - x^2}$ для верхней половины эллипса и $y = -\frac{b}{a}\sqrt{a^2 - x^2}$ для нижней половины. Прямое дифференцирование дает для верхней половины эллипса $y' = -\frac{bx}{a\sqrt{a^2 - x^2}}$, а для нижней половины $y' = \frac{bx}{a\sqrt{a^2 - x^2}}$. Для обеих ветвей $y' = -\frac{b^2x}{a^2y}$.

2) Из уравнения *лемнискаты* (т. I, стр. 94)

$$F(x, y) \equiv (x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0$$

нелегко получить явное выражение для y. Имеем

$$F_x = 4x(x^2 + y^2) - 4a^2x, \qquad F_y = 4y(x^2 + y^2) + 4a^2y.$$

При $x = 0$ и $y = 0$ оказывается $F = 0$, $F_x = 0$ и $F_y = 0$. Следовательно, в начале координат теорема существования не действует, и правило дифференцирования нельзя применить; это можно было ожидать, так как через начало проходят две различные ветви лемнискаты. Однако для всех точек кривой, в которых $y \neq 0$, имеем $F_y \neq 0$, и наше правило приложимо. Для производной от функции $y = f(x)$ получаем

$$y' = -\frac{F_x}{F_y} = -\frac{4x(x^2+y^2) - 4a^2x}{4y(x^2+y^2) + 4a^2y}.$$

Из этой формулы можно извлечь важную информацию о течении кривой, и не подставляя в нее явного выражения y через x. Например, максимумы или минимумы функция $y = f(x)$ может иметь лишь в тех точках, где $y' = 0$, стало быть, либо при $x = 0$, либо там, где $x^2 + y^2 = a^2$. Из уравнения лемнискаты видно, что при $x = 0$ также и $y = 0$; но в начале координат, как показывает, например, рис. 26 (т. I, стр. 94), экстремума нет. Подставляя же в уравнение лемнискаты $x^2 + y^2 = a^2$, получим еще одно уравнение $x^2 - y^2 = \frac{a^2}{2}$. Из этих двух уравнений определяются четыре точки $\left(\pm \frac{a}{2}\sqrt{3}, \pm \frac{a}{2}\right)$, которые действительно дают экстремумы неявно заданной функции $y = f(x)$.

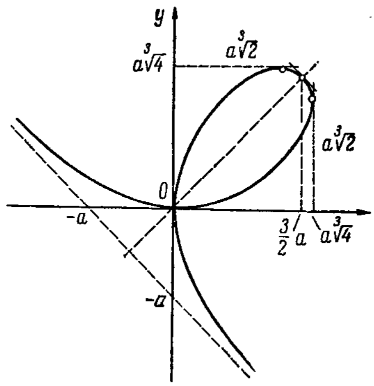


Рис. 30.

3) У *декартова листа* $F(x, y) \equiv x^3 + y^3 - 3axy = 0$ (рис. 30) явное выражение функции было бы чрезвычайно неудобным. В начале координат, где кривая сама себя пересекает, наше правило опять неприменимо, так как в этой точке $F = F_x = F_y = 0$, главное, $F_y = 0$. Во всех точках, в которых $y^2 \neq ax$, имеем

$$y' = -\frac{F_x}{F_y} = -\frac{x^2 - ay}{y^2 - ax}.$$

Отсюда видно, что производная обращается в нуль в тех точках, где $x^2 - ay = 0$. Это уравнение совместно с уравнением кривой дает одну точку: $x = a\sqrt[3]{2}$, $y = a\sqrt[3]{4}$. Испытание этой точки (выяснением знака второй производной y'') обнаруживает в ней максимум функции $y = f(x)$.

5. Теорема существования неявной функции нескольких переменных. Общую теорему о неявных функциях можно обобщить на случай многих независимых переменных следующим образом:

Пусть $F(x, y, \ldots, z, u)$ есть непрерывная функция своих аргументов $x, y, \ldots, z, u$, обладающая непрерывными частными производными $F_x, F_y, \ldots, F_z, F_u$. Пусть система значений $x_0, y_0, \ldots, z_0, u_0$ удовлетворяет уравнению $F = 0$, так что $F(x_0, y_0, \ldots, z_0, u_0) = 0$, а

$$F_u(x_0, y_0, \ldots, z_0, u_0) \neq 0.$$

Тогда существует такой интервал $u_1 \leqslant u \leqslant u_2$ вокруг u_0 и такая область G, содержащая внутри себя $(x_0, y_0, \ldots, z_0)$, что уравнением

$F(x, y, \ldots, z, u)=0$ *определяется в области G однозначная функция* $u=f(x, y, \ldots, z)$, *значения которой лежат в интервале* $u_1 \leqslant u \leqslant u_2$; *эта функция удовлетворяет в каждой точке области G* уравнению

$$F[x, y, \ldots, z, f(x, y, \ldots, z)]=0,$$

и для нее

$$u_0=f(x_0, y_0, \ldots, z_0).$$

Кроме того, функция $u=f(x, y, \ldots, z)$ *является непрерывной функцией независимых переменных* $x, y, \ldots, z$ *и имеет по ним непрерывные частные производные, которые определяются из уравнений:*

$$F_x+F_u f_x=0,$$
$$F_y+F_u f_y=0,$$
$$\cdots\cdots\cdots$$
$$F_z+F_u f_z=0.$$

По поводу доказательства существования и непрерывности мы отошлем читателя к следующему номеру. Что касается формул дифференцирования, то они получаются из правила дифференцирования неявной функции одной независимой переменной. Для этого надо приписать всем независимым переменным, кроме одной, постоянные значения и взять производную по оставшейся переменной; если, например, фиксировать значения $y, \ldots, z$, то получим первую из записанных выше формул, для вычисления f_x.

Помножим первую из формул для производных на dx, вторую — на $dy, \ldots$, последнюю — на dz; тогда если учесть, что $f_x dx+f_y dy+\ldots$ $\ldots+f_z dz=df=du$, все формулы дифференцирования объединятся в одном уравнении

$$F_x dx+F_y dy+\ldots+F_z dz+F_u du=0,$$

которое выражает следующее: *если аргументы $x, y, \ldots, z$, u функции $F(x, y, \ldots, z, u)$ не изменяются независимо друг от друга, а подчинены условию $F=0$, то и их дифференциалы $dx, dy, \ldots$ $\ldots, dz, du$ тоже не являются независимыми, — они связаны линейным уравнением*

$$dF=F_x dx+F_y dy+\ldots+F_z dz+F_u du=0.$$

Обратно, из этого уравнения можно вывести исходные формулы для производных. Для этого надо заменить du его выражением $u_x dx+u_y dy+\ldots+u_z dz$, после чего уравнение примет следующий вид:

$$(F_x+F_u u_x)\,dx+(F_y+F_u u_y)\,dy+\ldots+(Fz+F_u u_z)\,dz=0.$$

Так как дифференциалы $dx, dy, \ldots, dz$ уже взаимно независимы, то можно последовательно положить один из них не равным нулю, а остальные приравнять нулю, и тогда получатся все формулы для производных от функции $u=f(x, y, \ldots, z, u)$.

Понятие неявной функции позволяет дать общее определение понятия *алгебраической функции*. Функция $u=f(x, y, \ldots)$ называется алгебраической функцией независимых переменных $x, y, \ldots$, если она может быть определена неявно уравнением $F(x, y, \ldots, u)=0$, где F есть целая рациональная функция своих аргументов $x, y, \ldots, u$; выражаясь короче, если u «удовлетворяет алгебраическому уравнению». Все функции, не удовлетворяющие алгебраическому уравнению, называются *трансцендентными*.

В качестве примера на дифференцирование неявной функции многих переменных рассмотрим уравнение шаровой поверхности $x^2+y^2+u^2-1=0$.

Для частных производных первого порядка получаем

$$u_x=-\frac{x}{u}, \quad u_y=-\frac{y}{u}.$$

Дальнейшее дифференцирование дает:

$$u_{xx}=-\frac{1}{u}+\frac{x}{u^2}u_x=-\frac{x^2+u^2}{u^3},$$

$$u_{xy}=\frac{x}{u^2}u_y=-\frac{xy}{u^3},$$

$$u_{yy}=-\frac{1}{u}+\frac{y}{u^2}u_y=-\frac{y^2+u^2}{u^3}.$$

6. Доказательство существования и непрерывности неявной функции. Во многих конкретных случаях существование и непрерывность неявной функции вытекает из фактической возможности выразить из уравнения $F(x, y)=0$ неизвестную y через x с помощью элементарных функций. Тем не менее совершенно необходимо дать общее аналитическое доказательство теоремы существования, сформулированной в n° 3.

Прежде всего выделим прямоугольник

$$x_1 \leqslant x \leqslant x_2, \quad y_1 \leqslant y \leqslant y_2,$$

в котором уравнением $F(x, y)=0$ однозначно определяется функция $y=f(x)$. При этом мы отнюдь не будем пытаться найти наибольший прямоугольник, обладающий этим свойством; мы намерены лишь показать, что такой прямоугольник *существует*.

Так как производная $F_y(x, y)$ непрерывна, а $F_y(x_0, y_0) \neq 0$, то можно найти прямоугольник R с центром в точке $P(x_0, y_0)$ настолько малый, что во всем этом прямоугольнике производная F_y отлична от нуля и сохраняет постоянный знак. Не теряя общности, мы вправе-допустить, что этот знак положительный, т. е. что $F_y > 0$ во всем прямоугольнике R; действительно, если бы было $F_y < 0$, то можно было бы заменить функцию F на $-F$, не изменяя уравнения $F(x, y) = 0$. В силу $F_y > 0$, на отрезке любой прямой $x = \mathrm{const}$, параллельной оси y, лежащем в R, функция $F(x, y)$, рассматриваемая как функция одного лишь y, монотонно возрастает. Так как $F(x_0, y_0) = 0$, то в некоторой точке $A(x_0, y_1)$, лежащей в R на вертикали, проходящей через точку P, причем $y_1 < y_0$, значение функции $F(x_0, y_1)$ отрицательно, а в некоторой точке $B(x_0, y_2)$, $y_2 > y_0$, значение $F(x_0, y_2)$ положительно (рис. 31). Вследствие непрерывности функции $F(x, y)$ отсюда вытекает, что $F(x, y)$ принимает отрицательные значения вдоль некоторого лежащего в R отрезка горизонтальной прямой $y = y_1$, проходящей через A, и положительные значения на лежащем в R отрезке прямой $y = y_2$, проходящей через B. Стало быть, вокруг x_0 можно отметить интервал $x_1 \leqslant x \leqslant x_2$ настолько малый, что для значений x из этого интервала функция $F(x, y)$ отрицательна на горизонтальном отрезке, проходящем через A, и положительна на горизонтальном отрезке, проходящем через B. Другими словами, во всех точках интервала $x_1 \leqslant x \leqslant x_2$ выполняются неравенства $F(x, y_1) < 0$ и $F(x, y_2) > 0$.

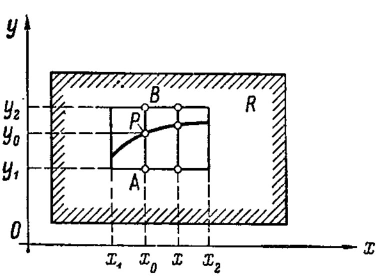


Рис. 31.

Выберем теперь любое значение x из интервала $x_1 \leqslant x \leqslant x_2$ и, оставляя его неизменным, заставим y возрастать от y_1 до y_2. Тогда точка (x, y) остается в прямоугольнике

$$x_1 \leqslant x \leqslant x_2, \qquad y_1 \leqslant y \leqslant y_2,$$

который мы можем считать лежащим целиком внутри R. Так как $F_y(x, y) > 0$, то значение функции $F(x, y)$ монотонно и непрерывно возрастает от отрицательного значения до положительного и в двух точках с одинаковой абсциссой не может принимать одно и то же значение. Следовательно, каждому значению x из интервала $x_1 \leqslant x \leqslant x_2$ соответствует *однозначно определенное* значение y, которое удовлетворяет уравнению $F(x, y) = 0$. Это значение y является поэтому функцией от x; тем самым мы доказали существование однозначной неявной функции $y = f(x)$, определяемой уравнением $F(x, y) = 0$. При этом ясно обнаруживается и роль условия $F_y \neq 0$. Если бы это условие не было выполнено, то значения функции $F(x, y)$ на горизонтальных прямых, проходящих через A и B, могли бы и не иметь противоположных знаков, так что $F(x, y)$ могла бы и не проходить через нуль на вертикальных отрезках. Если же значения функции на горизонталях, проходящих через A и B, даже имели бы различные знаки, то производная F_y могла бы менять свой знак, так что функция $F(x, y)$ не изменялась бы монотонно на вертикальном отрезке (т. е. при постоянном значении x) и могла бы поэтому обратиться в нуль несколько раз, что сделало бы неявную функцию многозначной.

Отметим еще важность ограничения $y_1 \leqslant y \leqslant y_2$. Если его опустить, то однозначность функции $y = f(x)$ не была бы обеспечена. Пусть, например, $F(x, y) = x^2 + y^2 - 1$ и $x_0 = 0$, $y_0 = 1$. Тогда в интервале $-\frac{1}{2} \leqslant x \leqslant \frac{1}{2}$ уравнением $x^2 + y^2 - 1 = 0$ определяется единственное решение $y = f(x)$, если ограничить y интервалом $0 \leqslant y \leqslant 2$; если же не ограничивать значений y, то получится два решения: $y = \sqrt{1 - x^2}$ и $y = -\sqrt{1 - x^2}$.

Наше доказательство только лишь устанавливает тот факт, что функция $y = f(x)$ существует. Оно является образцом чистого *«доказательства существования»*, совершенно не затрагивающего вопроса о практических возможностях вычисления функции. Этот отказ от нахождения практических методов является порою существенно важным шагом для упрощения доказательства.

Остается доказать непрерывность функции $f(x)$, но она является непосредственным следствием наших рассуждений. В самом деле, пусть прямоугольник $R'(x'_1 \leqslant x \leqslant x'_2,\ y'_1 \leqslant y \leqslant y'_2)$ лежит полностью внутри прямоугольника $x_1 \leqslant x \leqslant x_2$, $y_1 \leqslant y \leqslant y_2$. Для этого меньшего прямоугольника можно выполнить точно тот же самый процесс построения решения $y = f(x)$ уравнения $F(x, y) = 0$. Так как в объемлющем, бо́льшем, прямоугольнике это решение было однозначно определенным, то вновь найденная функция $y = f(x)$ совпадает с прежней. Для того чтобы доказать непрерывность функции $y = f(x)$ в какой-либо точке $x = x_0$, надлежит показать, что при любом сколь угодно малом наперед заданном числе $\varepsilon > 0$ можно сделать $|f(x) - f(x_0)| < \varepsilon$, если только выбрать x достаточно близким к x_0. Для этой цели положим

$$y'_1 = y_0 - \varepsilon \quad \text{и} \quad y'_2 = y_0 + \varepsilon,$$

где $y_0 = f(x_0)$, и для этих значений y'_1 и y'_2 определим соответствующий интервал для x: $x'_1 \leqslant x \leqslant x'_2$. Тогда, по данному выше построению, для всякого x из этого интервала соответствующее значение функции $y = f(x)$ лежит между границами y'_1 и y'_2, и отличается поэтому от $y_0 = f(x_0)$ меньше чем на ε. Это-то и выражает непрерывность функции $f(x)$ в точке x_0. Так как это рассуждение применимо к любой точке интервала $x_1 \leqslant x \leqslant x_2$, то непрерывность функции доказана для всех точек этого интервала.

Доказательство общей теоремы о существовании и непрерывности неявной функции $u = f(x, y, \ldots, z)$, определяемой уравнением $F(x, y, \ldots, z, u) = 0$ с бо́льшим числом независимых переменных, проводится по совершенно такому же образцу, как только что выполненное доказательство, и не встречается ни с какими новыми затруднениями.

Упражнения

1. Доказать, что каждое из данных ниже уравнений определяет в неявном виде однозначную функцию $y=f(x)$ в окрестности указанных точек:

а) $x^2+xy+y^2=7$ в окрестности точки $(2;\ 1)$;

б) $x\cos(x,y)=0$ в окрестности точки $\left(1;\ \frac{\pi}{2}\right)$;

в) $xy+\ln(xy)=1$ в окрестности точки $(1;\ 1)$;

г) $x^5+y^5+xy=3$ в окрестности точки $(1;\ 1)$.

2. Найти первые производные от функций, заданных неявно в упр. 1, в указанных там точках.

3. Вычислить вторые производные от функций, заданных неявно в упр. 1, в указанных там точках.

4. Определить максимумы и минимумы функции $y=f(x)$, определенной неявно уравнением $x^2+xy+y^2=27$.

5. Функция $z=f(x,y)$ определена в неявном виде уравнением

$$x^3+y^3+z^3-3xyz=0.$$

Выразить z_x и z_y через x, y, z.

6. Показать, что уравнение $x+y+u=\sin(xyu)$ определяет в окрестности точки $(0,\ 0,\ 0)$ однозначную функцию $u=f(x,\ y)$, принимающую при $x=y=0$ значение $u=0$. Найти частные производные этой функции.

§ 2. Неявное задание плоских кривых и неявное задание поверхностей

1. Неявное задание плоской кривой. До сих пор мы задавали плоскую кривую несимметрическим уравнением вида $y=f(x)$, в котором одной из координат отдается предпочтение. В такой же несимметрической форме писалось уравнение касательной к этой кривой

$$\eta=y+(\xi-x)f'(x)$$

и уравнение ее нормали

$$\eta=y-(\xi-x)\frac{1}{f'(x)},$$

где ξ и η — текущие координаты касательной или нормали, а x и y — координаты точки касания. Для явного уравнения кривой мы вывели также выражение для кривизны и признаки точек перегиба. Теперь мы выведем соответствующие формулы для кривой, заданной неявным уравнением $F(x,\ y)=0$. Мы будем при этом предполагать, что функция $F(x,y)$ дифференцируема и что в рассматриваемой точке F_x и F_y не обращаются одновременно в нуль, так что в этой точке $F_x^2+F_y^2\neq 0$.

Допустим, например, что $F_y\neq 0$. Тогда в уравнениях касательной и нормали в точке (x,y) кривой можно заменить y' ее выражением $-\frac{F_x}{F_y}$, и сразу получится уравнение касательной

$$(\xi-x)F_x+(\eta-y)F_y=0$$

и уравнение нормали

$$(\xi - x)F_y - (\eta - y)F_x = 0$$

в текущих координатах ξ и η.

Нынешней постановке вопроса более соответствовал бы вывод уравнения касательной прямым путем без использования явного уравнения кривой. Это можно выполнить, если дать касательной несколько иное определение. Мы теперь будем называть касательной к кривой в ее точке P такую прямую, проходящую через эту точку P, что расстояние от любой близкой точки P_1 кривой до этой прямой имеет порядок малости более высокий, чем расстояние PP_1, когда это последнее расстояние стремится к нулю.

Предоставляем читателю доказать, что двух различных прямых, удовлетворяющих этому условию, быть не может, так что новое определение касательной однозначно.

Пусть теперь $P(x, y)$ — любая точка кривой, т. е. $F(x, y) = 0$. Уравнение всякой прямой, проходящей через точку P, можно записать в следующем виде:

$$a(\xi - x) + b(\eta - y) = 0, \tag{1}$$

где ξ и η — текущие координаты этой прямой. Для того чтобы эта прямая была касательной в точке P кривой, мы должны так подобрать коэффициенты a и b, чтобы расстояние от переменной точки P_1 кривой с координатами $x_1 = x + h$ и $y_1 = y + k$ до этой прямой имело более высокий порядок малости, чем $\rho = \sqrt{h^2 + k^2}$, когда расстояние ρ между точками P и P_1 стремится к нулю. В силу дифференцируемости функции $F(x, y)$ имеем

$$F(x + h, y + k) = F(x, y) + hF_x + kF_y + \varepsilon\rho,$$

где $\varepsilon \to 0$ при $\rho \to 0$. Так как обе точки P и P_1 лежат на кривой, то

$$hF_x + kF_y = -\varepsilon\rho.$$

Для того чтобы выразить расстояние от точки P_1 до прямой (1), надо привести уравнение (1) к нормальному виду, подставить в левую часть полученного уравнения вместо ξ, η координаты $x + h$, $y + k$ точки P_1 и взять абсолютную величину результата. Следовательно, искомое расстояние равно $\left|\frac{ah + bk}{\sqrt{a^2 + b^2}}\right|$.

Теперь сразу ясно, что если положить $a = F_x(x, y)$ и $b = F_y(x, y)$, то расстояние от точки P_1 до прямой будет равно (вспомним, что $F_x^2 + F_y^2 \neq 0$)

$$\left|\frac{hF_x + kF_y}{\sqrt{F_x^2 + F_y^2}}\right| = \frac{\varepsilon}{\sqrt{F_x^2 + F_y^2}}\rho = \varepsilon_1\rho,$$

где $\varepsilon_1 \to 0$ при $\rho \to 0$. Стало быть, при $\rho \to 0$ это расстояние тоже стремится к нулю, имея более высокий порядок малости, чем ρ. Отсюда следует, что из нового определения касательной получается для нее то же самое уравнение, которое было выведено ранее.

Из уравнения касательной видим сразу, что вектор $\{F_x, F_y\}$ является ее нормальным вектором, т. е. *нормальным вектором кривой.*

Из уравнения нормали видно, что за направляющий вектор касательной (касательный вектор) можно принять вектор $\{F_y, -F_x\}$.

Заметим, что если вместо кривой $F(x, y) = 0$ будем рассматривать кривую $F(x, y) = C$, где C — произвольная постоянная, то ничего не изменится ни в наших рассуждениях, ни в вычислениях. Придется лишь заменить функцию $F(x, y)$ функцией $F(x, y) - C$, имеющей те же частные производные, что и первая функция; поэтому уравнения касательной и нормали имеют при любом C формально точно тот же вид, какой выведен выше.

Множество всех кривых $F(x, y) = C$, которые получаются, когда C пробегает все значения в некотором интервале, называется *семейством кривых.* Градиент функции $F(x, y)$, т. е. плоский вектор $\{F_x, F_y\}$, является в каждой точке плоскости нормальным вектором кривой семейства, проходящей через эту точку, другими словами, перпендикулярен к этой кривой, как мы уже видели в гл. II, § 7, n° 6. Этот факт дает еще один способ для вывода уравнения касательной — вектор $\{\xi - x, \eta - y\}$ направлен по касательной, стало быть, он перпендикулярен к градиенту, и скалярное произведение этих двух векторов должно равняться нулю:

$$(\xi - x) F_x + (\eta - y) F_y = 0,$$

а это и есть уравнение касательной.

В т. I, стр. 188, мы вывели для кривой, заданной явным уравнением $y = f(x)$, *необходимое условие точки перегиба*: $f''(x) = 0$. Заменяя $f''(x)$ выражением

$$f''(x) = -\frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_y^3}$$

из предыдущего параграфа, мы получим необходимое условие точки перегиба

$$F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2 = 0 \qquad (2)$$

для кривой, заданной неявным уравнением $F(x, y) = 0$. В этой форме необходимого условия точки перегиба обе переменные x и y уже равноправны; эта форма условия совершенно симметрична и можно показать, что она уже не связана с допущением, что $F_y \neq 0$.

Выражение для кривизны

$$k = \frac{y''}{(1 + y'^2)^{3/2}},$$

выведенное в т. I, стр. 325, тоже можно преобразовать к симметричному виду, пригодному для кривой, заданной неявным уравнением $F(x, y)=0$. Вспомним, что по этой формуле кривизна $k>0$, если кривая в данной точке обращена вогнутостью вверх, и $k<0$, если она обращена вогнутостью вниз. Заменив y' через $-\frac{F_x}{F_y}$, а $y''=f''(x)$ выражением, написанным выше, получим

$$\left(1+y'^2\right)^{3/2}=\frac{(F_x^2+F_y^2)^{3/2}}{|F_y|^3}$$

и, наконец, новую формулу для *кривизны*

$$k=-\operatorname{sgn} F_y\cdot\frac{F_{xx}F_y^2-2F_{xy}F_xF_y+F_{yy}F_x^2}{(F_x^2+F_y^2)^{3/2}}, \tag{3}$$

где

$$\operatorname{sgn} a=\begin{cases}+1 & \text{при } a>0,\\ -1 & \text{при } a<0 \text{ (ср. т. I, стр. 69).}\end{cases}$$

Знак кривизны по этой формуле имеет тот же смысл, что и в формуле, из которой она выведена (см. выше).

[Если $F_y=0$, то множитель сигнум надо просто опустить, и абсолютная величина кривизны получится правильно. Однако для случая $F_x\neq 0$ можно вывести соответствующую формулу для кривизны из формулы $k=\frac{x''}{(1+x'^2)^{3/2}}$, пригодной для кривой, заданной уравнением $x=\varphi(y)$. Тогда получится формула

$$k=-\operatorname{sgn} F_x\cdot\frac{F_{xx}F_y^2-2F_{xy}F_xF_y+F_{yy}F_x^2}{(F_x^2+F_y^2)^{3/2}}. \tag{4}$$

По обеим этим формулам кривизна получается положительной, если кривая обращена вогнутостью вправо, и отрицательной, если она обращена вогнутостью влево.]

Для координат (ξ, η) центра кривизны получаются следующие выражения $\left(\text{в которых } \rho=\frac{1}{k}\right)$:

$$\xi=x+\rho\frac{|F_x|}{\sqrt{F_x^2+F_y^2}}=x-\frac{F_x(F_x^2+F_y^2)}{F_{xx}F_y^2-2F_{xy}F_xF_y+F_{yy}F_x^2},$$

где $\rho=\frac{1}{k}$ определено по формуле (4) и

$$\eta=y+\rho\frac{|F_y|}{\sqrt{F_x^2+F_y^2}}=y-\frac{F_y(F_x^2+F_y^2)}{F_{xx}F_y^2-2F_{xy}F_xF_y+F_{yy}F_x^2},$$

где $\rho=\frac{1}{k}$ определено по формуле (3).

Если две кривые $F(x, y)=0$ и $G(x, y)=0$ пересекаются в точке (x, y), то *углом между этими кривыми* называется угол ω между их касательными или между нормалями в точке их пересечения. Возьмем, например, нормальные векторы $N=\{F_x, F_y\}$ и $N_1=\{G_x, G_y\}$. Тогда

$$\cos\omega=\frac{NN_1}{|N|\cdot|N_1|}=\frac{F_xG_x+F_yG_y}{\sqrt{F_x^2+F_y^2}\sqrt{G_x^2+G_y^2}}.$$

Полагая в этой формуле $\omega=\frac{\pi}{2}$, получаем *условие перпендикулярности* (условие ортогональности) двух кривых:

$$NN_1=F_xG_x+F_yG_y=0.$$

Условие касания двух кривых в их общей точке можно получить либо из коллинеарности нормальных (или касательных) векторов обеих кривых

$$F_x:G_x=F_y:G_y,$$

либо из того факта, что отношение дифференциалов $dy:dx$ должно быть одинаковым для обеих кривых:

$$dy:dx=-F_x:F_y=-G_x:G_y.$$

Это условие касания можно записать и в следующем виде:

$$\begin{vmatrix} F_x & F_y \\ G_x & G_y \end{vmatrix}=0.$$

В качестве примера рассмотрим семейство *софокусных парабол*

$$y^2-2p\left(x+\frac{p}{2}\right)=0,$$

имеющих общий фокус в начале координат (сделать чертеж; ср. рис. 36, стр. 157). Если $p_1>0$, а $p_2<0$, то две параболы

$$F(x, y)\equiv y^2-2p_1\left(x+\frac{p_1}{2}\right)=0 \quad \text{и} \quad G(x, y)\equiv y^2-2p_2\left(x+\frac{p_2}{2}\right)=0$$

пересекаются, и притом под прямым углом, т. е. они ортогональны, так как в точке их пересечения

$$F_xG_x+F_yG_y=4(p_1p_2+y^2)=4\frac{p_2F-p_1G}{p_2-p_1}=0,$$

ибо $p_2-p_1\neq 0$, а $F=G=0$ в точке пересечения.

В качестве второго примера рассмотрим эллипс $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. Уравнение касательной в точке (x, y): $\frac{x}{a^2}(\xi-x)+\frac{y}{b^2}(\eta-y)=0$; его можно привести к виду

$$\frac{x\xi}{a^2}+\frac{y\eta}{b^2}=1,$$

известному из аналитической геометрии.

Кривизна $k = \frac{-\operatorname{sgn} y \cdot a^4 b^4}{\sqrt{(b^4x^2 + a^4y^2)^3}}$, так как $\operatorname{sgn} F_y = \operatorname{sgn} \frac{2y}{b^2} = \operatorname{sgn} y$. Если $a > b$, то абсолютная величина кривизны достигает наибольшего значения $\frac{a}{b^2}$ в вершинах $y = 0$, $x = \pm a$, а наименьшего значения — в двух других вершинах $x = 0$, $y = \pm b$; это наименьшее значение есть $|k| = \frac{b}{a^2}$.

2. Особые точки плоской кривой. Добавим несколько слов об *особых точках плоской кривой.* Мы здесь удовольствуемся тем, что приведем несколько типичных примеров, назначение которых показать необходимость исследования этого вопроса; само исследование читатель найдет в Дополнениях к этой главе.

В формулах, выведенных в n° 1, часто встречается выражение $F_x^2 + F_y^2$ в знаменателе дроби. Стало быть, обращение этой величины в нуль, т. е. одновременное обращение в нуль частных производных F_x и F_y в какой-либо точке кривой позволяет ожидать чего-то необычного, какой-то особенности кривой в этой точке. Прежде всего это проявляется в том, что в такой точке выражение $y' = -\frac{F_x}{F_y}$ для наклона касательной теряет смысл.

Точка кривой называется *регулярной* или *обыкновенной* точкой, если в окрестности этой точки можно представить ординату y в виде непрерывно дифференцируемой функции от x или абсциссу x в виде непрерывно дифференцируемой функции от y. В обоих случаях кривая имеет в такой точке касательную, и в окрестности этой точки мало отклоняется от своей касательной. Все же точки, не обладающие этим свойством, называются *особыми точками* или *особенностями* кривой.

Из теории неявных функций известно, что если в какой-либо точке кривой $F(x, y) = 0$ производная $F_y \neq 0$, то эта точка является обыкновенной точкой кривой, так как при этом условии из уравнения $F(x, y) = 0$ можно выразить y в виде однозначной дифференцируемой функции $y = f(x)$. По аналогии ясно, что точка, в которой $F_x \neq 0$, тоже будет обыкновенной точкой кривой. Следовательно, особые точки кривой надо искать среди тех точек, которые удовлетворяют не только уравнению кривой $F(x, y) = 0$, но еще и уравнениям

$$F_x = 0 \quad \text{и} \quad F_y = 0.$$

Первый по важности тип особой точки представляет узловая или кратная точка. Точка кривой называется *кратной* или *узловой*, если через нее проходят две или большее число различных ветвей кривой. Пример узловой точки представляет начало координат для лемнискаты

$$(x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0.$$

В окрестности такой точки невозможно однозначно представить уравнение кривой в виде $y=f(x)$ или $x=\varphi(y)$.

Другой тип особой точки представляет так называемая *точка заострения или возврата.* В такой точке сходятся две ветви кривой, имея в ней общую касательную.

Пример такой точки дает кривая $y^3-x^2=0$, изображенная на рис. 32. Она имеет точку возврата в начале координат. Другой пример представляет кривая $x^3-y^2=0$, у которой тоже точка возврата в начале координат. Кстати, вторая кривая получается из первой поворотом на 90° вокруг начала координат.

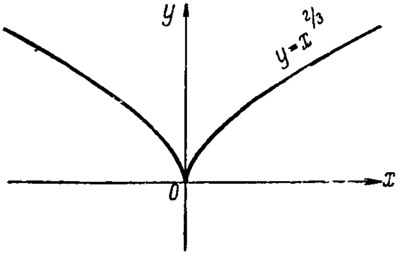


Рис. 32.

Во всех этих примерах действительно $F_x=0$ и $F_y=0$ в особой точке. Однако совмещение условий $F_x=0$ и $F_y=0$ в какой-либо точке кривой $F(x, y)=0$ необходимо, но отнюдь не достаточно для того, чтобы эта точка была особой в смысле данного выше определения. Можно сказать, что точка, удовлетворяющая этим условиям, является особой точкой *уравнения* кривой, но не обязательно особой точкой самой кривой. Точка кривой может оказаться регулярной несмотря на то, что в ней обращаются в нуль обе частные производные F_x и F_y.

Пример такой точки мы находим у кривой $y^3-x^4=0$ в начале координат, где $F_x(0, 0)=F_y(0, 0)=0$. Сразу видно, что эта кривая симметрична относительно оси y, так как при замене x на $(-x)$ уравнение кривой не изменяется. Уравнение можно привести к явному виду $y=x^{4/3}$, откуда $y'=\frac{4}{3}x^{1/3}$. Ясно, что начало координат является регулярной точкой кривой. Далее, $y(0)=0$ и $y'(0)=0$, так что кривая касается в начале координат оси x. (Общий облик кривой напоминает обыкновенную параболу.)

Однако начало координат все же чем-то выделяется среди других точек кривой: в нем вторая производная обращается в бесконечность. Следовательно, кривизна бесконечна в начале координат, в то время как направление касательной в этой точке не обнаруживает никаких особенностей.

Другой пример представляет линия $(y-x)^2=0$, во всех точках которой $F_x=0$ и $F_y=0$. Однако же это — прямая, у которой все точки, конечно, обыкновенные.

Из приведенных здесь рассуждений и примеров вытекает следующий вывод: при отыскании особых точек кривой недостаточно найти все ее точки, удовлетворяющие условиям $F_x=0$ и $F_y=0$. Каждая такая точка должна быть подвергнута особому исследованию.

3. Неявное задание поверхности. До сих пор мы обычно изображали функцию $z=f(x, y)$ (будем теперь писать z вместо u) поверхностью в пространстве, рассматривая x, y, z как прямоугольные координаты точки. Если же первоначально задана не функция, а поверхность в пространстве, то предпочтение, оказываемое при таком явном задании координате z, может привести к таким же неудобствам, к каким приводило задание плоской кривой уравнением $y=f(x)$. Более естественным и более общим является задание поверхности в пространстве уравнением вида $F(x, y, z)=0$ или $F(x, y, z)=\text{const}$; так, например, шаровую поверхность чаще удобнее представить уравнением $x^2+y^2+z^2-a^2=0$, чем явным уравнением $z=\pm\sqrt{a^2-x^2-y^2}$. Уравнение $z-f(x, y)=0$ можно трактовать как частный случай общего неявного уравнения поверхности.

Для того чтобы составить уравнение касательной плоскости к поверхности $F(x, y, z)=0$ в ее точке (x, y, z), допустим, что в этой точке $F_x^2+F_y^2+F_z^2 \neq 0$[1]), т. е. что по крайней мере одна из частных производных не равна нулю. Пусть, например, $F_z \neq 0$. Тогда из уравнения поверхности можно принципиально выразить z как однозначную явную функцию от x и y. В уравнение касательной плоскости (гл. II, § 4, n° 3)

$$\zeta-z=(\xi-x) z_x+(\eta-y) z_y$$

подставим вместо z_x и z_y их выражения $z_x=-\frac{F_x}{F_z}$ и $z_y=-\frac{F_y}{F_z}$, и тогда уравнение касательной плоскости приведется к симметричному виду

$$(\xi-x) F_x+(\eta-y) F_y+(\zeta-z) F_z=0,$$

где ξ, η и ζ — текущие координаты.

Уравнение касательной плоскости можно вывести и непосредственно из неявного уравнения поверхности так, как было выведено уравнение касательной к плоской кривой. Для этого надо по-иному поставить задачу: найти такую плоскость, проходящую через точку поверхности $P(x, y, z)$, что расстояние от переменной близкой точки поверхности $Q(x+h, y+k, z+l)$ до этой плоскости будем иметь более высокий порядок малости, чем расстояние $\rho=\sqrt{h^2+k^2+l^2}$ между точками P и Q, когда $\rho \to 0$.

Нормальный вектор касательной плоскости $N=\{F_x, F_y, F_z\}$ называется, как мы уже знаем, нормальным вектором поверхности в точке касания (x, y, z). Следовательно, направляющие косинусы нормали выражаются следующими формулами:

$$\cos\alpha=\frac{F_x}{\sqrt{F_x^2+F_y^2+F_z^2}}, \qquad \cos\beta=\frac{F_y}{\sqrt{F_x^2+F_y^2+F_z^2}},$$

$$\cos\gamma=\frac{F_z}{\sqrt{F_x^2+F_y^2+F_z^2}}.$$

[1]) Обращение этого выражения в нуль указывает на возможность появления особенностей; мы не будем здесь заниматься этим вопросом.

Если две поверхности $F(x, y, z)=0$ и $G(x, y, z)=0$ пересекаются в некоторой точке, то *углом* ω *между поверхностями* называется двугранный угол между их касательными плоскостями в этой точке или численно равный ему угол между нормальными векторами N и N_1 обеих поверхностей. Ясно, что

$$\cos\omega = \frac{NN_1}{|N|\cdot|N_1|}.$$

Отсюда получается условие ортогональности (перпендикулярности) двух поверхностей:

$$NN_1 = 0 \quad \text{или} \quad F_xG_x + F_yG_y + F_zG_z = 0.$$

Вместо отдельной поверхности $F(x, y, z)=0$ можно рассматривать целое семейство поверхностей $F(x, y, z)=C$, где C — постоянная, различная для разных поверхностей семейства. Мы будем предполагать, что через каждую точку пространства или по крайней мере через всякую точку некоторой пространственной области проходит одна и только одна поверхность семейства или, как принято говорить, что семейство покрывает эту область пространства *однократно*. Отдельные поверхности семейства называются тогда *поверхностями уровня* функции $F(x, y, z)$. В гл. II, § 7, n° 6 мы ввели понятие градиента этой функции: $\operatorname{grad} F = \{F_x, F_y, F_z\}$. Из сказанного выше явствует, что *градиент* функции в точке (x, y, z) является *нормальным вектором поверхности уровня, проходящей через эту точку*. Можно стать на другую точку зрения и считать этот факт известным из гл. II, § 7, n° 6; тогда, зная нормальный вектор поверхности в точке касания, можно составить уравнение касательной плоскости, и это будет третий и самый простой способ вывода ее уравнения.

В качестве примера рассмотрим *шаровую поверхность*

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

Имеем $F = x^2 + y^2 + z^2 - a^2$, $\operatorname{grad} F = \{2x, 2y, 2z\}$, за нормальный вектор можно принять $N = \{x, y, z\}$ и уравнение касательной плоскости будет $x(\xi - x) + y(\eta - y) + z(\zeta - z) = 0$ или

$$x\xi + y\eta + z\zeta - a^2 = 0.$$

Нормальный вектор $N = \{x, y, z\} = r$; следовательно, нормаль шаровой поверхности в любой ее точке проходит по радиусу, идущему к этой точке.

Для *эллипсоида* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ уравнение касательной плоскости можно привести к следующему виду:

$$\frac{x\xi}{a^2} + \frac{y\eta}{b^2} + \frac{z\zeta}{c^2} - 1 = 0,$$

Упражнения

1. Составить уравнение касательной плоскости
а) к поверхности $x^3+2xy^2-7z^3+3y+1=0$ в точке $(1;\ 1;\ 1)$;
б) к поверхности $(x^2+y^2)^2+x^2-y^2+7xy+3x+z^4-z=14$ в точке $(1;\ 1;\ 1)$;
в) к поверхности $\sin^2 x+\cos(y+z)=\frac{3}{4}$ в точке $\left(\frac{\pi}{6},\ \frac{\pi}{3},\ 0\right)$.

2. Вычислить кривизну плоской кривой $\sin x+\cos y=1$ в начале координат.

3. Из трех семейств поверхностей

$$xy=Cz, \qquad \sqrt{x^2+z^2}+\sqrt{y^2+z^2}=C_1, \qquad \sqrt{x^2+z^2}-\sqrt{y^2+z^2}=C_2$$

можно выделить три поверхности (по одной от каждого семейства), имеющих общую особую точку. Доказать, что эти три поверхности попарно ортогональны.

4. Точки A и B движутся равномерно с одинаковой (скалярной) скоростью. Точка A движется по оси z, точка B — по прямой, параллельной оси y. В начальный момент A находилась в начале координат, а B — в точке $(a,\ 0,\ 0)$. Найти поверхность, являющуюся геометрическим местом прямых, соединяющих обе движущиеся точки.

5. Доказать, что точки пересечения кривой

$$(x+y-a)^3+27axy=0 \qquad (a\neq 0)$$

с прямой $x+y=a$ являются точками перегиба кривой.

6. Определить a и b так, чтобы кривые второго порядка

$$4x^2+4xy+y^2-10x-10y+11=0,$$
$$(y+bx-1-b)^2-a(by-x+1-b)=0$$

пересекались в точке $(1;\ 1)$ под прямым углом и имели в этой точке одинаковую по абсолютной величине кривизну.

7. Поверхность задана уравнением $F(x,\ y,\ z)=1$, где F есть однородная функция измерения h. Показать, что касательная плоскость этой поверхности в ее точке $(x,\ y,\ z)$ приводится к виду

$$\xi F_x+\eta F_y+\zeta F_z=h.$$

8. K_1 и K_2 — две окружности, имеющие две общие точки A и B. Доказать, что если окружность K ортогональна к обеим окружностям K_1 и K_2, то она также ортогональна и ко всякой окружности, проходящей через точки A и B.

§ 3. Системы функций, преобразования и отображения

1. Первая интерпретация системы функций: преобразование и отображение. Сведения, полученные о неявных функциях, дают теперь возможность изучать системы функций, т. е. несколько функций, рассматриваемых совокупно. В этом параграфе мы рассмотрим особенно важный случай таких систем, в которых число функций равно числу независимых переменных. Выясним сперва смысл таких систем функций на случае двух независимых переменных. Пусть даны две функции

$$\xi=\varphi(x,\ y) \quad \text{и} \quad \eta=\psi(x,\ y),$$

дифференцируемые в некоторой области G плоскости xy. Эту систему можно истолковать двумя различными способами. В этом номере мы изложим первую интерпретацию, рассматривающую систему функций как отображение или преобразование (а вторая будет дана в n° 2). Точке $P(x, y)$ одной плоскости соответствует, в качестве ее *изображения*, точка $\Pi(\xi, \eta)$ другой плоскости. Точка $P(x, y)$ называется *оригиналом* своего изображения $\Pi(\xi, \eta)$.

Примером такого отображения является аффинное отображение или преобразование (гл. I, § 4)

$$\xi = a_1x + b_1y,$$
$$\eta = a_2x + b_2y,$$

где a_1, b_1, a_2 и b_2 — постоянные числа.

Точки (x, y) и (ξ, η) часто рассматривают как точки одной и той же плоскости; тогда говорят об *отображении плоскости* (x, y) *на самое себя* или же о *деформации* плоскости.

Одну функцию одной независимой переменной $\xi = f(x)$ тоже можно интерпретировать как отображение, считая, что эта функция точке x на оси x приводит в соответствие точку ξ на оси ξ; это точечное соответствие отображает ось x или ее часть на ось ξ или часть этой оси. Равномерная шкала равноотстоящих точек на оси x превращается в шкалу точек ξ (вообще говоря, неравномерную) на оси ξ, на которой точки то сгущаются, то разрежаются. Шкалу на оси ξ можно использовать для наглядного изображения функции $\xi = f(x)$; такая точка зрения часто применяется с успехом в приложениях (например, в *номографии*).

Основной задачей, которая возникает в связи с преобразованием, является задача об его *обращении*, т. е. вопрос о том, дают ли уравнения преобразования $\xi = \varphi(x, y)$ и $\eta = \psi(x, y)$ возможность рассматривать обратно x и y как функции от ξ и η, и если да, то как дифференцировать эти обратные функции.

Когда точка (x, y) пробегает область G, пусть точка (ξ, η), которая служит ее изображением, пробегает область Γ плоскости $\xi\eta$; тогда эта область Γ называется изображением области G. *Если* при этом *двум различным точкам области G всегда соответствуют две различные точки области* Γ, то для всякой точки области Γ существует единственная точка области G, изображением которой она является. Стало быть, можно обратно каждой точке области Γ отнести исходную точку области G, ее оригинал, т. е. наше преобразование можно обратить и выразить переменные x, y однозначно как обратные функции

$$x = g(\xi, \eta), \qquad y = h(\xi, \eta),$$

определенные в области Γ. В этом случае отображение называется *однозначно обратимым*, *взаимно однозначным* или *одно-однозначным*, а система функций $x = g(\xi, \eta)$, $y = h(\xi, \eta)$ называется преобразованием или отображением, *обратным* первоначальному.

Если точка $P(x, y)$ пробегает в области G некоторую кривую, то изображение этой точки опишет в области Γ тоже кривую, которая называется изображением первоначальной кривой. Так, например, линии $x = c$, т. е. прямой, параллельной оси y, соответствует на плоскости $\xi\eta$ кривая, выражаемая параметрическими уравнениями

$$\xi = \varphi(c, y), \qquad \eta = \psi(c, y),$$

в которых параметром служит y. Аналогично, прямой $y = k$ соответствует изображающая ее кривая

$$\xi = \varphi(x, k), \qquad \eta = \psi(x, k).$$

Припишем постоянным c и k последовательно различные значения $c_1, c_2, c_3, \ldots$ и $k_1, k_2, k_3, \ldots$; тогда прямоугольной *координатной*

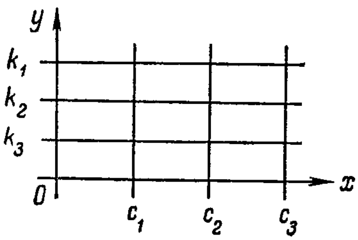


Рис. 33.

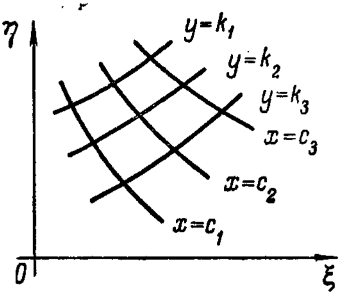


Рис. 34.

сетке прямых $x = \text{const}$ и $y = \text{const}$ (например, сети прямых на миллиметровой бумаге) соответствует на плоскости $\xi\eta$ *сетка линий*, вообще говоря, *криволинейная* (рис. 33 и 34). Имея в своем распоряжении выражение обратного преобразования

$$x = g(\xi, \eta), \qquad y = h(\xi, \eta),$$

можно оба семейства кривых этой криволинейной сетки представить и в неявном виде уравнениями

$$g(\xi, \eta) = c, \qquad h(\xi, \eta) = k.$$

Подобным же образом двум семействам прямых на плоскости изображения

$$\xi = \gamma \quad \text{и} \quad \eta = \varkappa$$

соответствуют на исходной плоскости xy два семейства кривых

$$\varphi(x, y) = \gamma \quad \text{и} \quad \psi(x, y) = \varkappa.$$

В качестве примера однозначно обратимого преобразования рассмотрим *инверсию*, иначе называемую *отображением* с помощью *обратных радиусов* или *симметрией относительно единичной окружности*. Это преобразование задается уравнениями

$$\xi = \frac{x}{x^2 + y^2}, \qquad \eta = \frac{y}{x^2 + y^2}.$$

Радиус-вектор точки P есть $\overline{OP}=\mathbf{r}=\{x, y\}$, а радиус-вектор ее изображения, точки П, $\overline{O\Pi}=\frac{\mathbf{r}}{r^2}=\frac{\mathbf{r}^0}{r}$. Таким образом, точка P и ее изображение П лежат на одной полупрямой, выходящей из начала координат (центра окружности $x^2+y^2=1$), и полярный радиус точки П равен числу, обратному полярному радиусу точки $P: O\Pi=\frac{1}{OP}$. Точки, лежащие внутри единичной окружности отображаются на точки, лежащие вне ее, и обратно.

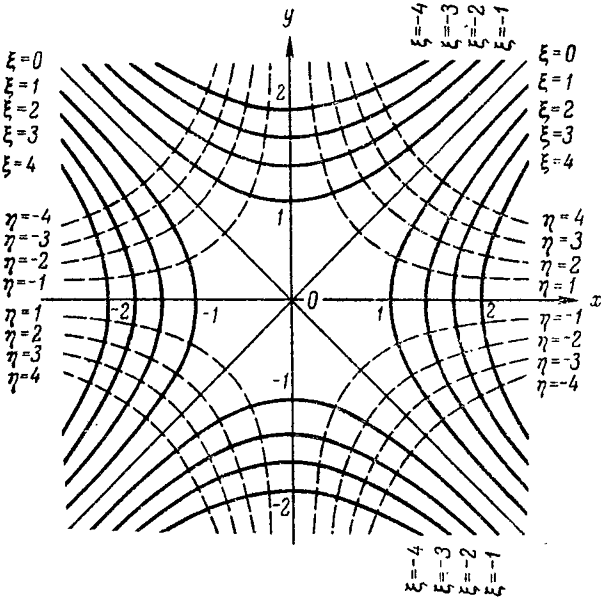


Рис. 35.

Так как $\xi^2+\eta^2=\frac{1}{x^2+y^2}$, то *обратное преобразование* есть

$$x=\frac{\xi}{\xi^2+\eta^2}, \qquad y=\frac{\eta}{\xi^2+\eta^2},$$

т. е. тоже *инверсия*.

За область G можно принять всю плоскость xy, за исключением начала координат, а за область Γ — всю плоскость ξ, η, за исключением начала координат. Прямым $\xi=c$ и $\eta=k$ плоскости $\xi\eta$ соответствуют на плоскости xy окружности $x^2+y^2-\frac{1}{c}x=0$ и $x^2+y^2-\frac{1}{k}y=0$; в начале координат первая из этих окружностей касается оси y, а вторая—оси x. Аналогично, координатная сетка $x=c$ и $y=k$ плоскости xy изображается на плоскости $\xi\eta$ криволинейной сеткой, состоящей из двух пучков окружностей, касающихся в начале координат оси η (первый пучок) и оси ξ (второй пучок).

В качестве второго примера рассмотрим отображение

$$\xi=x^2-y^2, \qquad \eta=2xy. \qquad (A)$$

Прямым $\xi = \text{const}$ соответствуют на плоскости xy равнобочные гиперболы $x^2 - y^2 = \text{const}$, имеющие асимптотами прямые $y = x$ и $y = -x$. Прямым $\eta = \text{const}$ соответствует другое семейство равнобочных гипербол $xy = \text{const}$, имеющее асимптотами оси координат. Гиперболы каждого семейства пересекают гиперболы другого семейства под прямым углом (рис. 35).

Прямым плоскости xy, параллельным осям координат, соответствуют на плоскости $\xi\eta$ два семейства парабол (рис. 36), а именно: прямым $x = c$ соответствуют параболы $\eta^2 = -4c^2(\xi - c^2)$, а прямым $y = k$ — параболы $\eta^2 = 4k^2(\xi + k^2)$. Оба семейства парабол имеют общий фокус в начале координат и общую ось — ось ξ (объединенное семейство софокусных и соосных парабол, рис. 36).

[Каждой точке плоскости xy соответствует одна и только одна точка плоскости $\xi\eta$. Однако на всей плоскости xy это преобразование не является взаимно однозначным. Формулы обратного преобразования будут

$$x = \pm\sqrt{\frac{\sqrt{\xi^2+\eta^2}+\xi}{2}},$$

$$y = \pm\sqrt{\frac{\sqrt{\xi^2+\eta^2}-\xi}{2}},$$

причем, в силу равенства $\eta = 2xy$, при $\eta > 0$ надо в обеих формулах брать одинаковые знаки, а при $\eta < 0$ — различные знаки. Отсюда видно, что на всю верхнюю полуплоскость $\eta > 0$ плоскости $\xi\eta$ отображается как первая, так и третья четверть плоскости xy; а на нижнюю полуплоскость $\eta < 0$ плоскости $\xi\eta$ отображается как вторая, так и четвертая четверть плоскости xy. Поэтому, чтобы отображение (A) было однозначно обратимым, можно за область Г принять всю плоскость $\xi\eta$, а за область G выбрать одну из четырех возможностей: либо полуплоскость $y > 0$, либо полуплоскость $y < 0$, либо полуплоскость $x > 0$, либо полуплоскость $x < 0$.]

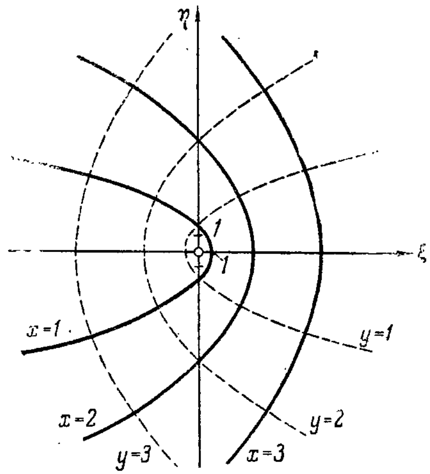


Рис. 36.

Обратимые преобразования можно наглядно представить себе как деформации или перемещения непрерывно распределенного вещества, например жидкости. Они и применяются для изучения таких деформаций или перемещений. Пусть такое вещество в некоторый момент времени распределено непрерывно по двумерной области G, а затем подвергается деформации, так что вещество, которое располагалось раньше на области G, уже будет покрывать область Г, которая может не совпадать с областью G. Каждую частицу вещества можно отметить в начале движения ее координатами (x, y) в области G, а после деформации—ее координатами (ξ, η) в области Г. Взаимно однозначный характер преобразования, устанавливающего соответствие между положениями частицы (x, y) и (ξ, η) до и после деформации, есть просто математическое выражение того физически очевидного факта, что каждая частица сохраняет свою индивидуальность и после деформации, так что *различные* частицы остаются *различными*.

2. Вторая интерпретация системы функций: введение новых, криволинейных координат. С первой интерпретацией (как отображения) тесно связана вторая интерпретация системы функций $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$ как *преобразования координат* на плоскости. Однако если φ и ψ не являются линейными функциями, это уже будет не аффинное преобразование, а *преобразование к общим криволинейным координатам.*

Полагаем по-прежнему, что когда точка (x, y) пробегает область G плоскости xy, то соответствующая ей точка (ξ, η) пробегает область Γ на плоскости $\xi\eta$, и что, обратно, для каждой точки области Γ соответствующая пара значений (x, y) тоже определяется однозначно; другими словами, что преобразование взаимно однозначно. Обратное преобразование обозначим по-прежнему через $x = g(\xi, \eta), y = h(\xi, \eta)$.

Координатами точки P области G можно считать любую пару чисел, пригодную для однозначного указания положения точки P в области G. Прямоугольные координаты представляют простейший пример координат, эффективных на всей плоскости. Другим типичным примером является полярная система координат, которая вводится на плоскости xy с помощью формул:

$$\xi = r = \sqrt{x^2 + y^2}, \quad \underset{(x^2+y^2 \neq 0)}{\eta = \theta =} \begin{cases} \operatorname{arctg} \dfrac{y}{x} & \text{при } x \geqslant 0, \\ \operatorname{arctg} \dfrac{y}{x} + \pi & \text{при } x \leqslant 0,\ y \geqslant 0, \\ \operatorname{arctg} \dfrac{y}{x} - \pi & \text{при } x \leqslant 0,\ y < 0, \end{cases}$$

так что $-\pi < \theta \leqslant \pi$.

И вообще, когда задана система функций $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, то для каждой точки $P(x, y)$ можно рассматривать соответствующую пару чисел (ξ, η) как ее новые координаты. Действительно, всякая пара чисел (ξ, η), принадлежащая области Γ, однозначно определяет пару значений (x, y) и тем самым однозначно указывает положение точки P в области G; поэтому мы вправе называть ξ, η координатами точки P.

В любой системе координат *координатными линиями* называются такие линии, вдоль которых одна из координат сохраняет постоянное значение. Например, в декартовой прямоугольной системе координатными линиями являются прямые $x = c$ и $y = k$, параллельные осям координат. На этой же плоскости xy мы имеем в новой системе координат ξ, η два семейства координатных линий, $\xi = c = \text{const}$ и $\eta = k = \text{const}$, которые определяются в неявном виде уравнениями $\varphi(x, y) = c$ и $\psi(x, y) = k$. Эти координатные линии покрывают область G координатной сеткой, как правило, криволинейной; поэтому новые координаты (ξ, η) называются также *криволинейными координатами* в области G. Стало быть, декартова система координат, как прямоугольная, так и косоугольная, является прямолинейной системой координат, а полярные координаты принадлежат к числу криволинейных.

Подчеркнем еще раз, что оба различных толкования системы функций тесно связаны между собой. В плоскости xy имеется система прямоугольных декартовых координат (x, y), в плоскости $\xi\eta$ — такая же система координат (ξ, η). С точки зрения первой интерпретации, функции $\xi=\varphi(x, y)$, $\eta=\psi(x, y)$ отображают область G первой плоскости на область Γ второй, причем прямым $x=c$, $y=k$ области G соответствуют два семейства кривых в области Γ плоскости $\xi\eta$. С точки зрения второй интерпретации обратная система функций $x=g(\xi, \eta)$, $y=h(\xi, \eta)$ вводит в области Γ плоскости $\xi\eta$ новую, криволинейную систему координат (x, y) и ее координатными линиями являются как раз только что упомянутые два семейства кривых, служащих изображениями прямых $x=c$, $y=k$.

Обратно, координатными линиями криволинейной системы координат $\xi=\varphi(x, y)$, $\eta=\psi(x, y)$ для области G плоскости xy являются кривые, отображающие на эту область прямые $\xi=c$, $\eta=k$ плоскости $\xi\eta$.

Даже при истолковании пар (ξ, η) как криволинейных координат на плоскости xy целесообразно, для достижения полной ясности, рассматривать одновременно еще и плоскость $\xi\eta$ и на ней область Γ, которую пробегает переменная точка (ξ, η). Различие главным образом в точке зрения[1]). Если нас интересует главным образом область G плоскости xy, то мы рассматриваем ξ, η только как новый способ координации точек в области G, а область Γ плоскости $\xi\eta$ будет играть лишь вспомогательную роль. Если же нас равно интересуют обе области: G в плоскости xy и Γ в плоскости $\xi\eta$, то лучше видеть в системе функций установление соответствия между обеими областями. Однако всегда желательно иметь в виду одновременно оба толкования — отображение и преобразование координат.

Пусть, например, в плоскости xy вводятся полярные координаты (r, θ); тогда полезно ввести в рассмотрение вспомогательную плоскость $r\theta$, в которой толкуем r и θ как прямоугольные декартовы координаты. Окружности $r=\text{const}$ и полупрямые $\theta=\text{const}$ данной плоскости отобразятся в виде прямых, параллельных осям, на вспомогательной плоскости $r\theta$. Если область G плоскости xy есть круг $x^2+y^2\leqslant 1$, то соответствующая область Γ второй плоскости будет прямоугольник $0\leqslant r\leqslant 1$, $0\leqslant\theta<2\pi$; при этом сторону $\theta=2\pi$ прямоугольника мы не присоединяем к области Γ, а вся сторона прямоугольника $r=0$ служит изображением одной точки — начала координат $x=0$, $y=0$.

Рассмотрим другой пример криволинейных координат — *систему параболических координат*. Мы придем к ним, рассматривая в плоскости

---

[1]) Однако существует и реальное различие, состоящее в том, что система равенств $\xi=\varphi(x, y)$, $\eta=\psi(x, y)$ всегда определяет *отображение*, независимо от того, сколько точек соответствует одной точке (x, y), между тем как *переход к криволинейным координатам* она задает лишь в том случае, если соответствие взаимно однозначно.

xy семейство парабол (сделать чертеж; ср. стр. 148, и стр. 157, рис. 36)

$$y^2 = 2p\left(x + \frac{p}{2}\right),$$

имеющих общий фокус в начале координат и общую ось — ось x. Через каждую точку плоскости проходят две параболы семейства, из которых одна соответствует отрицательному значению параметра $p_1 = \xi < 0$, а вторая — положительному значению $p_2 = \eta > 0$. Эти два значения получаем, решая относительно p квадратное уравнение $p^2 + 2px - y^2 = 0$, в котором рассматриваем x и y как координаты конкретно заданной точки:

$$\xi = -x - \sqrt{x^2 + y^2}, \qquad \eta = -x + \sqrt{x^2 + y^2}. \tag{1}$$

Этими двумя функциями мы введем криволинейные координаты на плоскости xy. Новыми координатами (ξ, η) точки (x, y) являются, таким образом, параметры $p_1 = \xi$ и $p_2 = \eta$ тех двух парабол, которые в ней пересекаются; отсюда и название *параболические координаты*.

Однако двум точкам $P(x, y)$ и $Q(x, -y)$, расположенным симметрично относительно оси x, соответствует одна и та же пара координат (ξ, η), а стало быть, одна и та же точка плоскости $\xi\eta$: эти точки P и Q являются точками пересечения одной и той же пары парабол нашего семейства. Так как $\xi + \eta = p_1 + p_2 = -2x$ и $\xi\eta = p_1 p_2 = -y^2$, то обратная система функций будет

$$x = -\frac{\xi + \eta}{2}, \qquad y = \pm\sqrt{-\xi\eta}. \tag{2}$$

Для того чтобы соответствие (1) было взаимно однозначным (при введении криволинейных координат это необходимо), надо ограничиться либо верхней половиной плоскости xy, либо ее нижней половиной. Если выбрать за область G верхнюю полуплоскость $y \geqslant 0$, то во второй из формул (2) надо при квадратном корне взять знак плюс.

Будем теперь рассматривать формулы (1) как отображение; тогда, при сделанном выборе области G, они дают отображение верхней половины плоскости xy $(y \geqslant 0)$ на вторую четверть плоскости $\xi\eta$ (это область Γ). Изображениями полупрямых $x = c$, лежащих в области G, являются лежащие в области Γ части параллельных прямых $\eta = -\xi - 2c$, а изображениями прямых $y = k$ служат лежащие в Γ ветви гипербол $\xi\eta = -k^2$.

[Все семейство парабол в целом совпадает с семейством парабол на рис. 36, только название осей различное: там ξ, η, а здесь x, y. Существенное различие состоит в том, что пометки (надписи) на каждой конкретно выбранной параболе в обоих отображениях различны.]

3. Система трех функций от трех независимых переменных. В случае трех или большего числа независимых переменных дело обстоит совершенно аналогично. Например, систему трех непрерывно дифференцируемых функций

$$\xi = \varphi(x, y, z), \qquad \eta = \psi(x, y, z), \qquad \zeta = \chi(x, y, z),$$

определенных в некоторой области G пространства xyz, можно рассматривать как отображение области G на некоторую область Γ пространства $\xi\eta\zeta$. Допустим, что это отображение области G на область Γ взаимно однозначно, так что и, обратно, для каждой точки (ξ, η, ζ) из области Γ координаты x, y, z исходной точки P в G выражаются однозначно с помощью функций

$$x = g(\xi, \eta, \zeta), \quad y = h(\xi, \eta, \zeta), \quad z = l(\xi, \eta, \zeta);$$

тогда ξ, η и ζ можно также рассматривать как общие криволинейные координаты точки P в области G. Поверхности $\xi = \text{const}$, $\eta = \text{const}$ и $\zeta = \text{const}$ или

$$\varphi(x, y, z) = \text{const}, \quad \psi(x, y, z) = \text{const}, \quad \chi(x, y, z) = \text{const}$$

образуют систему трех семейств поверхностей, покрывающих область G; эти поверхности (как правило, кривые) называются координатными поверхностями.

Так же как и в случае двух переменных, взаимно однозначные преобразования трех переменных можно наглядно иллюстрировать как деформации вещества, непрерывно распределенного в некоторой области трехмерного пространства.

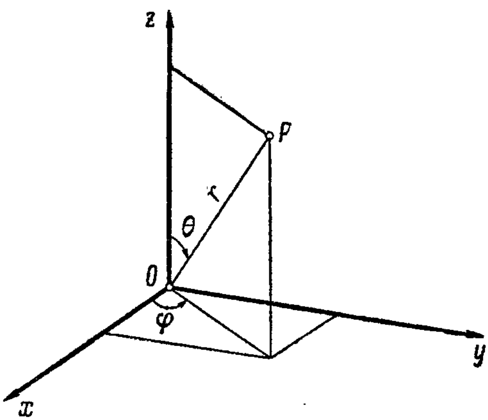


Рис. 37.

Важнейшим примером введения криволинейных координат является преобразование к *сферическим* или *пространственным полярным координатам*. В этой системе координат положение точки P в пространстве задается с помощью следующих трех чисел: 1) расстояния $r = \sqrt{x^2 + y^2 + z^2}$ точки P от начала координат, 2) географической долготы φ, т. е. угла от плоскости $x0z$ до плоскости, проходящей через точку P и ось z, и 3) полярного угла θ, т. е. угла от положительного направления оси z (полярной оси) до радиус-вектора $\overline{OP}$. Из рис. 37 сразу находим формулы преобразования, выражающие прямоугольные координаты через сферические:

$$x = r \sin\theta \cos\varphi, \quad y = r \sin\theta \sin\varphi, \quad z = r \cos\theta.$$

Решая эти уравнения относительно r, θ, φ как неизвестных, получим формулы, выражающие сферические координаты через декартовы прямоугольные:

$$r=\sqrt{x^2+y^2+z^2},\quad \theta=\arccos\frac{z}{\sqrt{x^2+y^2+z^2}},$$

$$\varphi=\begin{cases}\arccos\dfrac{x}{\sqrt{x^2+y^2}} & \text{при} \quad y\geqslant 0,\\ -\arccos\dfrac{x}{\sqrt{x^2+y^2}} & \text{при} \quad y<0.\end{cases}$$

Таким образом, $r\geqslant 0$, $0\leqslant\theta\leqslant\pi$, $-\pi<\varphi\leqslant\pi$.

Для полярных координат на плоскости полюс (начало координат) является исключительной точкой, в которой нарушается однозначность, так как полярный угол делается неопределенным. В полярной системе координат в пространстве вся ось z занимает исключительное положение, ибо во всех ее точках долгота φ неопределенна. В начале координат становится неопределенным еще и полярный угол θ.

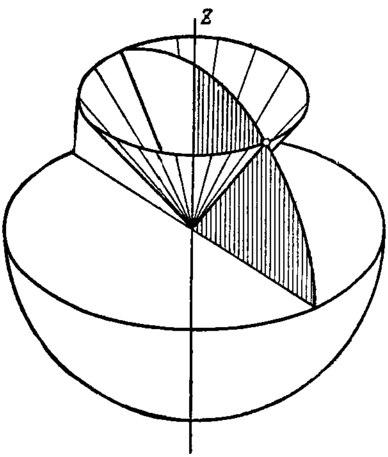


Рис. 38.

Сферическая система координат имеет следующие три семейства координатных поверхностей: 1) при различных постоянных значениях r — концентрические шаровые поверхности с центром в начале координат, 2) при постоянных значениях φ — семейство полуплоскостей, проходящих через ось z, и 3) при различных постоянных значениях θ — конические поверхности вращения вокруг оси z с вершиной в начале координат (точнее, каждая такая координатная поверхность есть только половина, т. е. одна полость конической поверхности (рис. 38)).

Для координации точек в пространстве часто также пользуются и *цилиндрической системой координат*; ее получают, вводя на плоскости xy полярные координаты ρ, φ и удерживая z в качестве третьей координаты. Формулы, выражающие прямоугольные координаты через цилиндрические, таковы:

$$x=\rho\cos\varphi,\qquad y=\rho\sin\varphi,\qquad z=z;$$

отсюда обратное преобразование

$$\rho=\sqrt{x^2+y^2},\qquad \varphi=\begin{cases}\arccos\dfrac{x}{\sqrt{x^2+y^2}} & \text{при} \quad y\geqslant 0;\\ -\arccos\dfrac{x}{\sqrt{x^2+y^2}} & \text{при} \quad y<0,\end{cases}\qquad z=z.$$

Координатные поверхности $\rho = \text{const}$ представляют собой круговые цилиндры с осью z, пересекающие плоскость xy по концентрическим окружностям с центром в начале координат. Поверхности $\varphi = \text{const}$ суть полуплоскости, проходящие через ось z, а координатные поверхности $z = \text{const}$ — плоскости, параллельные плоскости xy.

4. Формулы дифференцирования обратных функций. Во многих практически важных случаях возможно решить заданную систему уравнений $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, рассматривая x и y как неизвестные (так было в рассмотренных выше примерах), и убедиться, что обратные функции непрерывны и имеют непрерывные производные. В связи с этим, мы покамест будем предполагать, что система функций обратима и что обратные функции дифференцируемы. При этом предположении можно вычислить производные от обратных функций, не нуждаясь в предварительном явном решении системы уравнений. Для этого представим себе, что в данные уравнения $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$ подставлены обратные функции $x = g(\xi, \eta)$, $y = h(\xi, \eta)$. Эти уравнения превратятся тогда в тождества

$$\xi = \varphi[g(\xi, \eta), h(\xi, \eta)], \qquad \eta = \psi[g(\xi, \eta), h(\xi, \eta)],$$

в которых правые части являются сложными функциями от ξ и η. Эти тождества мы продифференцируем по ξ и по η, рассматривая их как независимые переменные. В правых частях мы воспользуемся правилом цепочки для дифференцирования сложных функций, и в итоге получим следующие четыре равенства, справедливые при всех значениях ξ и η, т. е. четыре тождества относительно ξ и η:

$$1 = \varphi_x g_\xi + \varphi_y h_\xi, \qquad 0 = \varphi_x g_\eta + \varphi_y h_\eta,$$
$$0 = \psi_x g_\xi + \psi_y h_\xi, \qquad 1 = \psi_x g_\eta + \psi_y h_\eta.$$

Но эти равенства можно рассматривать как две системы уравнений для определения искомых частных производных. Обе системы уравнений имеют общий определитель $D = \varphi_x\psi_y - \varphi_y\psi_x$, который мы предполагаем не равным нулю. Из первой системы находим g_ξ и h_ξ, из второй — g_η и h_η:

$$g_\xi = \frac{\psi_y}{D}, \qquad g_\eta = -\frac{\varphi_y}{D}, \qquad h_\xi = -\frac{\psi_x}{D}, \qquad h_\eta = \frac{\varphi_x}{D}$$

или

$$x_\xi = \frac{\eta_y}{D}, \qquad x_\eta = -\frac{\xi_y}{D}, \qquad y_\xi = -\frac{\eta_x}{D}, \qquad y_\eta = \frac{\xi_x}{D}.$$

Таким образом, частные производные обратных функций $x = g(\xi, \eta)$ и $y = h(\xi, \eta)$ по ξ и по η выражены через частные производные исходных функций $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$ по x и y. Выражение

$$D = \xi_x\eta_y - \xi_y\eta_x = \begin{vmatrix} \dfrac{\partial\xi}{\partial x} & \dfrac{\partial\xi}{\partial y} \\ \dfrac{\partial\eta}{\partial x} & \dfrac{\partial\eta}{\partial y} \end{vmatrix}$$

называется *якобианом* или *функциональным определителем* системы функций $\xi = \varphi(x, y)$ и $\eta = \psi(x, y)$ по переменным x и y.

Выше, как порою бывало и раньше, мы пользовались краткой записью $\xi(x, y)$ и $\eta(x, y)$ вместо подробного обозначения $\xi = \varphi(x, y)$ и $\eta = \psi(x, y)$, в котором делается различие между самими *величинами* ξ и η, с одной стороны, и *законами функционального соответствия* $\varphi(x, y)$ и $\psi(x, y)$ — с другой. Таким несколько вольным, но зато более кратким и наглядным обозначением мы будем пользоваться и в дальнейшем, если не будет оснований опасаться недоразумений.

В качестве примера рассмотрим выражение полярных координат на плоскости через прямоугольные координаты

$$\xi = r = \sqrt{x^2 + y^2}, \qquad \underset{(x^2+y^2\neq 0)}{\eta = \theta =} \begin{cases} \operatorname{arctg}\dfrac{y}{x} & \text{при} \quad x \geqslant 0, \\ \operatorname{arctg}\dfrac{y}{x} + \pi & \text{при} \quad x \leqslant 0, \quad y \geqslant 0, \\ \operatorname{arctg}\dfrac{y}{x} - \pi & \text{при} \quad x \leqslant 0, \quad y < 0. \end{cases}$$

Частные производные будут

$$r_x = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r}, \qquad r_y = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{r},$$

$$\theta_x = -\frac{y}{x^2+y^2} = -\frac{y}{r^2}, \qquad \theta_y = \frac{x}{x^2+y^2} = \frac{x}{r^2}.$$

Следовательно, якобиан

$$D = \begin{vmatrix} r_x & r_y \\ \theta_x & \theta_y \end{vmatrix} = \frac{x}{r}\cdot\frac{x}{r^2} - \frac{y}{r}\left(-\frac{y}{r^2}\right) = \frac{1}{r},$$

и частные производные от обратных функций, выражающих прямоугольные координаты через полярные, будут

$$x_r = \frac{\theta_y}{D} = \frac{x}{r}, \quad x_\theta = -\frac{r_y}{D} = -y, \quad y_r = -\frac{\theta_x}{D} = \frac{y}{r}, \quad y_\theta = \frac{r_x}{D} = x.$$

В этом примере известны формулы обратного преобразования $x = r\cos\theta$, $y = r\sin\theta$, и, конечно, те же результаты легче получить прямым дифференцированием этих формул.

Функциональный определитель (якобиан) D столь часто встречается, что для него ввели специальный символ

$$D = \frac{\partial(\xi, \eta)}{\partial(x, y)},$$

в целесообразности которого мы сейчас убедимся. Из формул

$$x_\xi = \frac{\eta_y}{D}, \qquad x_\eta = -\frac{\xi_y}{D},$$

$$y_\xi = -\frac{\eta_x}{D}, \quad y_\eta = \frac{\xi_x}{D},$$

выражающих частные производные от обратных функций, находим якобиан системы функций $x=x(\xi, \eta)$, $y=y(\xi, \eta)$ по переменным ξ и η:

$$\frac{\partial(x, y)}{\partial(\xi, \eta)}=x_\xi y_\eta - x_\eta y_\xi = \frac{\xi_x\eta_y-\xi_y\eta_x}{D^2}=\frac{1}{D}=1:\frac{\partial(\xi, \eta)}{\partial(x, y)}.$$

Якобиан обратной системы функций равен обратной величине якобиана исходной системы.

Таким же путем можно выразить и производные второго порядка от обратных функций через первые и вторые производные исходных функций. Для этого надо только продифференцировать по ξ и по η полученные выше (стр. 163) тождества, линейные относительно g_ξ, h_ξ, g_η и h_η, пользуясь правилом цепочки. (Мы, естественно, предполагаем, что исходные функции имеют непрерывные производные второго порядка.) Полученные равенства, тождественные относительно ξ и η, будут линейными уравнениями относительно искомых производных второго порядка, которые мы и вычислим из этих уравнений.

Например, для вычисления производных $\frac{\partial^2 x}{\partial\xi^2}=g_{\xi\xi}$ и $\frac{\partial^2 y}{\partial\xi^2}=h_{\xi\xi}$ продифференцируем равенства

$$1=\xi_x x_\xi+\xi_y y_\xi, \qquad 0=\eta_x x_\xi+\eta_y y_\xi$$

еще раз по ξ; пользуясь правилом цепочки, получим

$$0=\xi_{xx}x_\xi^2+2\xi_{xy}x_\xi y_\xi+\xi_{yy}y_\xi^2+\xi_x x_{\xi\xi}+\xi_y y_{\xi\xi},$$
$$0=\eta_{xx}x_\xi^2+2\eta_{xy}x_\xi y_\xi+\eta_{yy}y_\xi^2+\eta_x x_{\xi\xi}+\eta_y y_{\xi\xi}.$$

Это система линейных уравнений относительно неизвестных $x_{\xi\xi}$ и $y_{\xi\xi}$, причем определитель системы опять равен D и, стало быть, по условию, отличен от нуля. Решив эту систему и подставив затем вместо x_ξ и y_ξ найденные для них выше значения, получим

$$x_{\xi\xi}=-\frac{1}{D^3}\begin{vmatrix}\xi_{xx}\eta_y^2-2\xi_{xy}\eta_x\eta_y+\xi_{yy}\eta_x^2 & \xi_y\\ \eta_{xx}\eta_y^2-2\eta_{xy}\eta_x\eta_y+\eta_{yy}\eta_x^2 & \eta_y\end{vmatrix},$$
$$y_{\xi\xi}=\frac{1}{D^3}\begin{vmatrix}\xi_{xx}\eta_y^2-2\xi_{xy}\eta_x\eta_y+\xi_{yy}\eta_x^2 & \xi_x\\ \eta_{xx}\eta_y^2-2\eta_{xy}\eta_x\eta_y+\eta_{yy}\eta_x^2 & \eta_x\end{vmatrix}.$$

Тем же способом находят и производные третьего порядка и выше, дифференцируя повторно получающиеся линейные равенства, и на каждой ступени появляется система линейных уравнений с отличным от нуля определителем D.

5. Умножение отображений и преобразований. Если преобразование

$$\xi=\varphi(x, y), \qquad \eta=\psi(x, y)$$

отображает взаимно однозначно переменную точку (x, y) области G на точку (ξ, η) области Γ плоскости $\xi\eta$ и если затем эта область Γ

в свою очередь отображается взаимно однозначно преобразованием

$$u = \Phi(\xi, \eta), \qquad v = \Psi(\xi, \eta)$$

на область B плоскости uv, то тем самым одновременно возникает взаимно однозначное отображение области G на область B. Это *результирующее* отображение или преобразование называется *произведением* данных двух отображений или преобразований и выражается формулами

$$u = \Phi[\varphi(x, y), \psi(x, y)], \qquad v = \Psi[\varphi(x, y), \psi(x, y)].$$

Частные производные от u и v по x и по y легко вычисляются по правилу дифференцирования сложных функций:

$$\frac{\partial u}{\partial x} = \Phi_\xi \varphi_x + \Phi_\eta \psi_x,$$

$$\frac{\partial u}{\partial y} = \Phi_\xi \varphi_y + \Phi_\eta \psi_y,$$

$$\frac{\partial v}{\partial x} = \Psi_\xi \varphi_x + \Psi_\eta \psi_x,$$

$$\frac{\partial v}{\partial y} = \Psi_\xi \varphi_y + \Psi_\eta \psi_y.$$

Сопоставляя эти формулы с законом умножения определителей (в конце гл. I), находим, что якобиан системы функций u и v по x и y:

$$\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \Phi_\xi & \Phi_\eta \\ \Psi_\xi & \Psi_\eta \end{vmatrix} \cdot \begin{vmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{vmatrix},$$

т. е. *якобиан результирующего преобразования равен произведению якобианов составляющих преобразований.* Другими словами: при умножении преобразований перемножаются и их якобианы. Символическая запись этого правила

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial(u, v)}{\partial(\xi, \eta)} \cdot \frac{\partial(\xi, \eta)}{\partial(x, y)} \qquad (*)$$

показывает целесообразность нашего обозначения якобиана.

В этой записи *правило вычисления якобиана результирующего преобразования напоминает правило вычисления производной от сложной функции по производным составляющих функций одной переменной.*

Если якобианы составляющих преобразований отличны от нуля, то и якобиан их произведения отличен от нуля.

В частности, если в качестве второго преобразования взять преобразование $x = g(\xi, \eta)$, $y = h(\xi, \eta)$, обратное первому, то произведение будет тождественным преобразованием $u = x$, $v = y$ с якобианом, равным единице; тогда из формулы (*) вытекает вновь соотношение

$$\frac{\partial(\xi, \eta)}{\partial(x, y)} \cdot \frac{\partial(x, y)}{\partial(\xi, \eta)} = 1,$$

выведенное на стр. 165. Отсюда, кстати, следует, что ни один из этих двух якобианов не может обратиться в нуль.

6. Разложение произвольного преобразования на примитивные. В главе I мы видели, что всякое аффинное преобразование плоскости можно разложить на простые или, как говорят, *примитивные* преобразования, из которых первое деформирует плоскость только в одном направлении, а второе подвергает уже деформированную плоскость новой деформации в другом направлении. В каждом из этих последовательных преобразований вводится фактически только *одна* новая переменная.

Такого же результата можно добиться и для любых преобразований.

Рассмотрим сначала следующее примитивное преобразование:

$$\xi = \varphi(x, y), \quad \eta = y.$$

Предположим, что якобиан этого преобразования $D = \varphi_x$ отличен от нуля во всей области G; пусть, например, $\varphi_x > 0$ во всей области. Наше преобразование деформирует область G в область Γ, причем каждая точка области G смещается вдоль прямой, параллельной оси

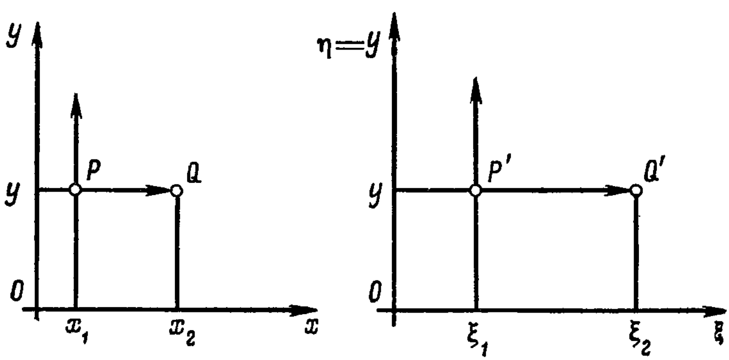


Рис. 39.

x; все точки с одной и той же ординатой y сохраняют и после деформации неизменную ординату $\eta = y$. В результате преобразования точка (x, y) приобретает новую абсциссу ξ, зависящую как от x, так и от y. Из условия $\varphi_x > 0$ вытекает, что при постоянном y абсцисса ξ монотонно возрастает вместе с x. Это условие обеспечивает взаимную однозначность соответствия между положениями точек прямой $y = \text{const}$ до и после деформации; действительно, две точки $P(x_1, y)$ и $Q(x_2, y)$ с одинаковой ординатой y при $x_2 > x_1$ переходят в две точки $P'(\xi_1, \eta)$ и $Q'(\xi_2, \eta)$, имеющие ту же ординату $\eta=y$ и новые абсциссы, удовлетворяющие неравенству $\xi_2>\xi_1$ (рис. 39). Вместе с тем этот факт показывает, что наше преобразование не изменяет направления вращения.

Если предположить, что $\varphi_x<0$, то две точки $P(x_1, y)$ и $Q(x_2, y)$ с одинаковой ординатой, у которых $x_2>x_1$, преобразуются в точки

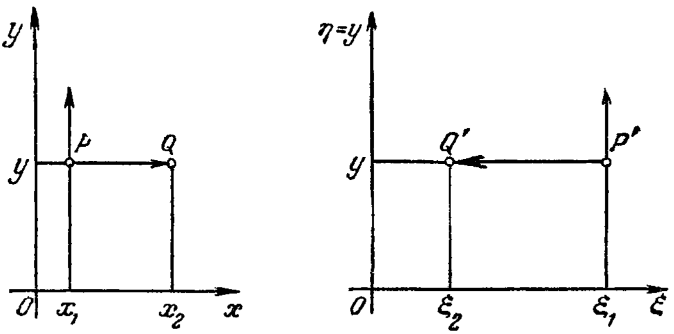


Рис. 40.

$P'(\xi_1, y)$ и $Q'(\xi_2, y)$, но на этот раз $\xi_1>\xi_2$ (рис. 40). При этом направление вращения изменится на обратное, подобно тому, что мы наблюдали в гл. I, стр. 48, на простом примере аффинного преобразования с отрицательным определителем.

Если примитивное преобразование

$$\xi=\varphi(x, y), \qquad \eta=y$$

непрерывно дифференцируемо, а его якобиан φ_x отличен от нуля в некоторой точке $P(x_0, y_0)$, то найдется такая окрестность точки P, в которой преобразование однозначно обратимо, причем обратное преобразование будет примитивным того же типа.

Действительно, в силу условия $\varphi_x \neq 0$, можно применить теорему о неявной функции (§ 1, n°3), согласно которой уравнение $\varphi(x,y)-\xi=0$ определяет в окрестности точки (x_0, y_0) величину x однозначно, как непрерывно дифференцируемую функцию $x=g(\xi, y)$ от ξ и y. Стало быть, формулы

$$x=g(\xi, \eta), \qquad y=\eta$$

дают обратное преобразование, якобиан которого $g_\xi=\dfrac{1}{\varphi_x}\neq 0$.

Пусть теперь область Γ плоскости $\xi\eta$ отображается в свою очередь на область B плоскости uv преобразованием

$$u=\xi, \qquad v=\Psi(\xi, \eta),$$

причем $\Psi_\eta \neq 0$; тогда положение дел такое же, как и на первом этапе, только смещение происходит теперь параллельно другой оси координат; да и направление вращения либо сохраняется, либо изменяется на противоположное, смотря по тому, каков знак Ψ_η — положительный или отрицательный.

Комбинируя эти два преобразования, составляем результирующее преобразование

$$u = \varphi(x, y), \qquad v = \Psi[\varphi(x, y), y] = \psi(x, y)$$

и, согласно теореме о якобиане произведения двух преобразований,

$$\frac{\partial(\varphi, \psi)}{\partial(x, y)} = \varphi_x \Psi_\eta.$$

Теперь мы утверждаем, что всякое взаимно однозначное непрерывно дифференцируемое[1]) преобразование

$$u = \varphi(x, y), \qquad v = \psi(x, y),$$

отображающее область G плоскости xy на область B плоскости uv, можно в окрестности любой внутренней точки области G разложить на примитивные непрерывно дифференцируемые преобразования, если только во всей области G якобиан

$$\frac{\partial(u, v)}{\partial(x, y)} = \varphi_x\psi_y - \psi_x\varphi_y$$

отличен от нуля.

В самом деле, из того, что якобиан не обращается в нуль, вытекает, что ни в одной точке φ_x и φ_y не могут одновременно обратиться в нуль. Рассмотрим точку $P(x_0, y_0)$ и допустим сначала, что в этой точке $\varphi_x \neq 0$. Тогда (согласно основной теореме § 1, n°5) можно отграничить квадрат $x_1 \leqslant x \leqslant x_2$, $y_1 \leqslant y \leqslant y_2$ вокруг (x_0, y_0) и интервал $u_1 \leqslant u \leqslant u_2$ вокруг $u_0 = \varphi(x_0, y_0)$ таким образом, что в этих границах уравнение $u = \varphi(x, y)$ однозначно разрешимо относительно x и определяет $x = g(u, y)$ как непрерывно дифференцируемую функцию от u и y. Вообразим, что это выражение подставлено в функцию $v = \psi(x, y)$, тогда $v = \psi[g(u, y), y] = \Psi(u, y)$. Стало быть, в указанной выше окрестности точки (x_0, y_0) наше преобразование $u = \varphi(x, y)$, $v = \psi(x, y)$ можно рассматривать как произведение двух примитивных преобразований

$$\xi = \varphi(x, y), \quad \eta = y \quad \text{и} \quad u = \xi, \quad v = \Psi(\xi, \eta).$$

Аналогично, если в точке (x_0, y_0) $\varphi_y \neq 0$, то в некоторой ее окрестности можно данное преобразование разложить на два примитивных преобразования вида

$$\xi = x, \quad \eta = \varphi(x, y) \quad \text{и} \quad u = \eta, \quad v = \psi[x, h(u, x)] = \Psi_1(\xi, \eta),$$

где $y = h(u, x)$ есть явное выражение y из уравнения $u = \varphi(x, y)$.

---

[1]) То есть функции u и v имеют непрерывные частные производные.

Нет оснований ожидать, что возможно единое разложение данного преобразования на примитивные во всей области G. Однако, так как разложение одного из двух указанных видов можно выполнить в окрестности любой внутренней точки области G, то область G или, во всяком случае, всякую замкнутую область, лежащую полностью внутри G, можно разбить на конечное число замкнутых частичных областей [1]) так, что в каждой из них выполнимо разложение одного из двух видов.

Из возможности такого разложения можно вывести интересное следствие. Мы знаем, что примитивное преобразование сохраняет направление вращения или изменяет его на противоположное, смотря по тому, каков знак якобиана — положительный или отрицательный. Из сказанного выше вытекает, что и *общее преобразование оставляет направление вращения неизменным или, напротив, изменяет его на противоположное, смотря по тому, имеет ли якобиан положительный или отрицательный знак.* Действительно, если якобиан данного преобразования имеет знак плюс, то якобианы его составляющих примитивных преобразований либо оба положительны, либо оба отрицательны. В первом случае сразу очевидно, что направление вращения сохраняется; во втором случае это вытекает из того факта, что двукратное изменение направления вращения на противоположное приводит к первоначальному направлению. Если же якобиан нашего преобразования имеет знак минус, то отрицательный якобиан будет иметь одно и только одно из примитивных преобразований, которое и приведет к изменению направления вращения, а другое примитивное преобразование не изменит его.

7. Общая теорема об обращении преобразования и о системах неявных функций. *Если функции $\varphi(x, y)$ и $\psi(x, y)$ непрерывно дифференцируемы в окрестности точки (x_0, y_0) и принимают в этой точке значения $u_0=\varphi(x_0, y_0)$ и $v_0=\psi(x_0, y_0)$ и если к тому же якобиан $D=\varphi_x\psi_y-\varphi_y\psi_x$ не равен нулю в (x_0, y_0), то система уравнений $u=\varphi(x, y)$, $v=\psi(x, y)$ однозначно разрешима в некоторой окрестности точки (x_0, y_0),* т. е. существует однозначно определенная обратная пара функций $x=g(u, v)$, $y=h(u, v)$ такого рода, что $x_0=g(u_0, v_0)$, $y_0=h(u_0, v_0)$, и в какой-то окрестности точки (u_0, v_0) выполняются тождества

$$u=\varphi[g(u, v), h(u, v)] \quad \text{и} \quad v=\psi[g(u, v), h(u, v)].$$

В окрестности точки (u_0, v_0) обратные функции $x=g(u, v)$, $y=h(u, v)$ имеют непрерывные частные производные, которые можно вычислить по следующим формулам:

$$\frac{\partial x}{\partial u}=\frac{1}{D}\frac{\partial v}{\partial y}, \qquad \frac{\partial x}{\partial v}=-\frac{1}{D}\frac{\partial u}{\partial y},$$
$$\frac{\partial y}{\partial u}=-\frac{1}{D}\frac{\partial v}{\partial x}, \qquad \frac{\partial y}{\partial v}=\frac{1}{D}\frac{\partial u}{\partial x}.$$

[1]) Это вытекает из теоремы о покрытии (стр. 120).

Доказательство вытекает из результатов предшествующего n° 6. Действительно, в достаточно малой окрестности точки (x_0, y_0) данное преобразование $u=\varphi(x, y)$, $v=\psi(x, y)$ можно разложить на непрерывно дифференцируемые примитивные преобразования; каждое из этих примитивных преобразований имеет однозначно определенное обратное преобразование, тоже примитивное и непрерывно дифференцируемое. Результирующее этих двух обратных преобразований и будет преобразованием, обратным данному; оно дифференцируемо, так как составляющие преобразования дифференцируемы. Формулы же дифференцирования обратных функций уже выведены в n° 4.

Приведенная здесь теорема обращения преобразований является частным случаем более общей теоремы, которую можно рассматривать как распространение теоремы о неявных функциях на системы функций. Эта теорема (§ 1, n° 5) говорит о возможности решения одного уравнения относительно одной из переменных. Общая теорема формулируется так:

Пусть $\varphi(x, y, u, v, \ldots, w)$ *и* $\psi(x, y, u, v, \ldots, w)$ — *две непрерывно дифференцируемые функции своих аргументов* $x, y, u, v, \ldots, w$, *и пусть уравнениям*

$$\varphi(x, y, u, v, \ldots, w)=0 \quad \text{и} \quad \psi(x, y, u, v, \ldots, w)=0$$

удовлетворяет некоторая система значений $x_0, y_0, u_0, v_0, \ldots, w_0$. *Если к тому же якобиан* $D=\varphi_x\psi_y-\varphi_y\psi_x$ *системы функций* φ *и* ψ *по* x *и* y *не обращается в нуль в этой точке, т. е. при указанной системе значений всех аргументов, то в окрестности этой точки можно однозначно решить уравнения* $\varphi=0$ *и* $\psi=0$ *относительно* x *и* y *как неизвестных, и это решение дает* x *и* y *как непрерывно дифференцируемые функции от* $u, v, \ldots, w$.

Доказательство этой теоремы аналогично доказательству теоремы обращения, приведенной в начале этого номера. Из условия $D \neq 0$ можно заключить, что в рассматриваемой точке φ_x и φ_y не равны одновременно нулю. Пусть, например, $\varphi_x \neq 0$. Тогда, по теореме о неявной функции (§ 1, n° 5), если ограничить аргументы $x, y, u, v, \ldots$ $\ldots, w$ достаточно малыми интервалами вокруг $x_0, y_0, u_0, v_0, \ldots, w_0$ соответственно, то уравнение $\varphi(x, y, u, v, \ldots, w)=0$ однозначно определяет $x=g(y, u, v, \ldots, w)$ как непрерывно дифференцируемую функцию остальных переменных, имеющую частную производную $g_y=-\dfrac{\varphi_y}{\varphi_x}$. Подставив эту функцию $x=g(y, u, v, \ldots, w)$ в $\psi(x, y, u, v, \ldots, w)$, мы получим функцию

$$\psi[g(y, u, v, \ldots, w), y, u, v, \ldots, w]=\Psi(y, u, v, \ldots, w),$$

причем

$$\Psi_y=\psi_x g_y+\psi_y=-\psi_x\frac{\varphi_y}{\varphi_x}+\psi_y=\frac{D}{\varphi_x}.$$

Так как, по условию, $D \neq 0$, то производная $\Psi_y \neq 0$. Поэтому, если ограничить переменные $y, u, v, \ldots, w$ некоторыми интервалами, окружающими $y_0, u_0, v_0, \ldots, w_0$ (они будут теперь уже тех интервалов, которыми мы их ранее ограничили), то уравнение $\Psi(y, u, v, \ldots, w) = 0$ определяет y как непрерывно дифференцируемую функцию переменных $u, v, \ldots, w$. Подставив это выражение для y в равенство $x = g(y, u, v, \ldots, w)$, получим и x как функцию от $u, v, \ldots$ $\ldots, w$. Итак, мы доказали существование однозначного решения в виде пары непрерывно дифференцируемых функций x и y, причем $x, y, u, v, \ldots, w$ ограничены достаточно малыми промежутками вокруг $x_0, y_0, u_0, v_0, \ldots, w_0$.

Производные этих функций x и y по $u, v, \ldots, w$ вычисляются из тождеств

$$\begin{cases} \varphi_x x_u + \varphi_y y_u + \varphi_u = 0, \\ \psi_x x_u + \psi_y y_u + \psi_u = 0, \end{cases}$$

$$\begin{cases} \varphi_x x_v + \varphi_y y_v + \varphi_v = 0, \\ \psi_x x_v + \psi_y y_v + \psi_v = 0, \end{cases}$$

$$\cdots\cdots\cdots\cdots$$

$$\begin{cases} \varphi_x x_w + \varphi_y y_w + \varphi_w = 0, \\ \psi_x x_w + \psi_y y_w + \psi_w = 0, \end{cases}$$

получаемых по правилу цепочки и рассматриваемых как системы линейных уравнений для пар неизвестных $x_u, y_u; \ldots; x_w, y_w$.

8. Взаимная зависимость функций. Необходимо подчеркнуть, что если якобиан D обращается в нуль в некоторой точке (x_0, y_0), то по вопросу о существовании обратной системы функций в окрестности этой точки невозможно высказать наперед никаких общих утверждений. Однако, если обратные функции существуют, то они не могут быть дифференцируемыми в соответствующей точке, так как в противном случае произведение якобианов $\frac{\partial(\xi, \eta)}{\partial(x, y)} \frac{\partial(x, y)}{\partial(\xi, \eta)}$ обратилось бы в нуль, между тем как, согласно n° 5, оно должно равняться единице.

Например, система функций $u = x^3$, $v = y$, якобиан которой обращается в нуль на оси y, имеет все же однозначную обратную систему $x = \sqrt[3]{u}$, $y = v$; однако функция $\sqrt[3]{u}$ недифференцируема при $u = 0$.

С другой стороны, система функций

$$u = x^2 - y^2, \qquad v = 2xy$$

не может иметь однозначной обратной системы функций в окрестности начала координат, так как двум различным точкам (x, y) и $(-x, -y)$ плоскости xy соответствует одна и та же точка плоскости uv.

Если же якобиан равен нулю *тождественно*, т. е. не только в одной точке (x_0, y_0), но и во всех точках некоторой ее окрестности, то преобразование называется ***вырожденным***. В этом случае говорят, что функции $u=\varphi(x, y)$ и $v=\psi(x, y)$ ***взаимно зависимы*** или просто ***зависимы***.

Рассмотрим сначала частный случай, являющийся почти тривиальным, когда $\varphi_x=0$ и $\varphi_y=0$ во всех точках рассматриваемой окрестности, так что в ней функция $\varphi(x, y)=\text{const}$. Ясно, что когда точка (x, y) пробегает некоторую двумерную область, ее изображение, точка (u, v), всегда остается на прямой $u=\text{const}$. Стало быть, наша область отображается уже не на область, а на отрезок прямой, и о взаимно однозначном отображении двух двумерных областей друг на друга не может быть и речи.

Подобным же образом обстоит дело и в общем случае, когда из частных производных φ_x и φ_y по крайней мере одна не равна тождественно нулю, между тем как $D=0$ тождественно. Допустим, например, что в точке (x_0, y_0) рассматриваемой области $\varphi_x \neq 0$. Согласно n° 6, наше преобразование можно разложить на два примитивных преобразования $\xi=\varphi(x, y)$, $\eta=y$ и $u=\xi$, $v=\Psi(\xi, \eta)$, ибо мы там пользовались только допущением, что $\varphi_x \neq 0$. Так как якобиан $D=\varphi_x\Psi_\eta \equiv 0$, то в окрестности точки (x_0, y_0), в которой φ_x не обращается в нуль, производная Ψ_η должна быть тождественно равна нулю; стало быть, функция $v=\Psi$ вовсе не зависит от η и является функцией лишь одной переменной $\xi=u$: $v=\Psi(u)$. Итак, мы приходим к следующему выводу:

Если якобиан преобразования тождественно равен нулю, то двумерная область плоскости xy отобразится не на двумерную область плоскости uv, а на кривую линию этой плоскости, так как каждому значению u (в некотором интервале) соответствует лишь одно значение v. Следовательно, *если якобиан системы функций $u=\varphi(x, y)$, $v=\psi(x, y)$ тождественно равен нулю, то эти функции взаимно зависимы, т. е. существует соотношение вида*

$$F(\varphi, \psi)=0,$$

которому удовлетворяют все пары значений (x, y) в упомянутой выше области. Действительно, если $F(u, v)=0$ есть уравнение той кривой плоскости uv, на которую отображается упомянутая выше область плоскости xy, то

$$F[\varphi(x, y), \psi(x, y)]=0$$

для всех точек этой области, т. е. последнее равенство есть тождество относительно x и y.

Тот случай, с которого мы начали рассмотрение этого вопроса ($\varphi_x=0$ и $\varphi_y=0$), очевидно, содержится в этом предложении как частный случай. Кривая, на которую отображается область плоскости xy, является в этом случае просто прямой $u=\text{const}$, параллельной оси v.

Приведем пример вырожденного преобразования

$$\xi = x + y, \qquad \eta = (x+y)^2.$$

Это преобразование переводит все точки плоскости xy в точки параболы $\eta = \xi^2$ плоскости $\xi\eta$. Что обращение этого преобразования невозможно, видно уже из того, что все точки прямой $x + y = c$ отображаются на одну точку $\xi = c$, $\eta = c^2$. Легко проверить, что якобиан этого преобразования тождественно равен нулю. Соотношение $F(\xi, \eta) = 0$, которое, согласно общей теореме, должно связывать функции ξ и η, есть здесь $\xi^2 - \eta = 0$.

Заключительное замечание. Во всем этом параграфе мы видели, что общее преобразование во многом напоминает аффинное преобразование и что якобиан играет здесь такую же роль, какую играет определитель аффинного преобразования. Следующие соображения помогут нам лучше уяснить себе это обстоятельство. Так как функции $\xi = \varphi(x, y)$ и $\eta = \psi(x, y)$ дифференцируемы в окрестности точки (x_0, y_0), то их можно представить в следующем виде:

$$\xi - \xi_0 = (x - x_0)\varphi_x(x_0, y_0) + (y - y_0)\varphi_y(x_0, y_0) + \varepsilon\sqrt{(x-x_0)^2 + (y-y_0)^2}$$

$$\eta - \eta_0 = (x - x_0)\psi_x(x_0, y_0) + (y - y_0)\psi_y(x_0, y_0) + \delta\sqrt{(x-x_0)^2 + (y-y_0)^2},$$

где ε и δ стремятся к нулю одновременно с расстоянием $\rho = \sqrt{(x-x_0)^2 + (y-y_0)^2}$. Отсюда видно, что при достаточно малых значениях $|x - x_0|$ и $|y - y_0|$ наше преобразование можно рассматривать в первом приближении как аффинное, т. е. в первом приближении его можно заменить аффинным преобразованием

$$\xi = \xi_0 + (x - x_0)\varphi_x(x_0, y_0) + (y - y_0)\varphi_y(x_0, y_0),$$

$$\eta = \eta_0 + (x - x_0)\psi_x(x_0, y_0) + (y - y_0)\psi_y(x_0, y_0),$$

определитель которого есть значение якобиана исходного общего преобразования в точке (x_0, y_0).

9. Несколько слов о преобразованиях в пространстве n измерений. Обобщение изложенной теории на три или большее число функций такого же числа независимых переменных не встречает особых затруднений. Главное различие состоит в том, что вместо определителя D второго порядка якобианом будет определитель третьего или более высокого порядка.

Для преобразований в трехмерном пространстве

$$\xi = \varphi(x, y, z), \qquad \eta = \psi(x, y, z), \qquad \zeta = \chi(x, y, z),$$

$$x = g(\xi, \eta, \zeta), \qquad y = h(\xi, \eta, \zeta), \qquad z = l(\xi, \eta, \zeta)$$

якобиан будет определителем третьего порядка

$$D = \frac{\partial(\xi, \eta, \zeta)}{\partial(x, y, z)} = \begin{vmatrix} \varphi_x & \psi_x & \chi_x \\ \varphi_y & \psi_y & \chi_y \\ \varphi_z & \psi_z & \chi_z \end{vmatrix}$$

Для преобразований в n-мерном пространстве

$$\left.\begin{aligned} \xi_i &= \varphi_i(x_1, x_2, \ldots, x_n), \\ x_i &= g_i(\xi_1, \xi_2, \ldots, \xi_n) \end{aligned}\right\}$$
$$(i = 1, 2, \ldots, n)$$

якобиан будет определителем n-го порядка

$$\frac{\partial(\xi_1, \xi_2, \ldots, \xi_n)}{\partial(x_1, x_2, \ldots, x_n)} = \begin{vmatrix} \frac{\partial\varphi_1}{\partial x_1} & \frac{\partial\varphi_2}{\partial x_1} & \ldots & \frac{\partial\varphi_n}{\partial x_1} \\ \frac{\partial\varphi_1}{\partial x_2} & \frac{\partial\varphi_2}{\partial x_2} & \ldots & \frac{\partial\varphi_n}{\partial x_2} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial\varphi_1}{\partial x_n} & \frac{\partial\varphi_2}{\partial x_n} & \ldots & \frac{\partial\varphi_n}{\partial x_n} \end{vmatrix}.$$

При любом числе измерений якобиан результирующего преобразования равен произведению якобианов составляющих преобразований

$$\frac{\partial(\xi_1, \xi_2, \ldots, \xi_n)}{\partial(\eta_1, \eta_2, \ldots, \eta_n)} \cdot \frac{\partial(\eta_1, \eta_2, \ldots, \eta_n)}{\partial(x_1, x_2, \ldots, x_n)} = \frac{\partial(\xi_1, \xi_2, \ldots, \xi_n)}{\partial(x_1, x_2, \ldots, x_n)}.$$

И здесь из последней формулы вытекает, что якобиан обратного преобразования равен единице, деленной на якобиан исходного преобразования.

Теоремы, касающиеся умножения и разложения преобразований, обращения преобразования, а также взаимной зависимости преобразований остаются в силе при любом числе измерений. Да и доказательства совершенно аналогичны случаю $n = 2$, и нет нужды их здесь повторять.

Упражнения

1. Если $f(x)$ есть непрерывно дифференцируемая функция, то преобразование

$$u = f(x), \qquad v = -y + xf(x)$$

имеет единственное обратное преобразование в любой области плоскости xy, в которой $f'(x) \neq 0$. Это обратное преобразование имеет следующий вид:

$$x = g(u), \qquad y = -v + ug(u).$$

2. Преобразование называется *конформным* (см. дальше § 4, n° 3), если оно сохраняет угол между любыми двумя кривыми.

а) Доказать, что инверсия

$$\xi = \frac{x}{x^2 + y^2}, \qquad \eta = \frac{y}{x^2 + y^2}$$

является конформным преобразованием.

б) Доказать, что инверсия переводит любую окружность в другую окружность или в прямую.

в) Вычислить якобиан инверсии.

3. Доказать, что в криволинейном треугольнике, образованном тремя окружностями, проходящими через одну точку O, сумма углов равна π.

4. Дано преобразование плоскости $u=\varphi(x, y)$, $v=\psi(x, y)$.

Доказать, что если функции φ и ψ удовлетворяют тождествам $\varphi_x=\psi_y$ и $\varphi_y=-\psi_x$, то это преобразование является конформным (см. упр. 2).

5. Уравнение $\dfrac{x^2}{a-t}+\dfrac{y^2}{b-t}=1$ $(a>b)$ определяет два значения t, зависящих от x и y:

$$t_1=\lambda(x, y), \qquad t_2=\mu(x, y).$$

а) Доказать, что кривые $t_1=\text{const}$ и $t_2=\text{const}$ являются все эллипсами и гиперболами с общими фокусами (софокусные кривые второго порядка).

б) Доказать, что кривые $t_1=\text{const}$ ортогональны кривым $t_2=\text{const}$.

в) Пары чисел t_1 и t_2 можно использовать в качестве криволинейных координат (так называемых *бифокальных* координат). Выразить x и y через эти координаты.

г) Выразить якобиан $\dfrac{\partial(t_1, t_2)}{\partial(x, y)}$ через x и y.

д) Найти условие, при котором две кривые, передставленные в бифокальной системе координат параметрическими уравнениями

$$t_1=f_1(\lambda), \quad t_2=f_2(\lambda) \quad \text{и} \quad t_1=g_1(\mu), \quad t_2=g_2(\mu),$$

взаимно ортогональны.

6. а) Доказать, что уравнение для неизвестной t

$$\frac{x^2}{a-t}+\frac{y^2}{b-t}+\frac{z^2}{c-t}=1 \qquad (a>b>c)$$

имеет три различных действительных корня t_1, t_2, t_3, лежащих в промежутках: $-\infty<t_1<c$, $c<t_2<b$, $b<t_3<a$, при условии, что точка (x, y, z) не лежит на какой-либо координатной плоскости.

б) Доказать, что три поверхности $t_1=\text{const}$, $t_2=\text{const}$ и $t_3=\text{const}$ попарно ортогональны.

в) Выразить x, y, z через t_1, t_2, t_3, которые можно принять за криволинейные координаты в пространстве. Систему t_1, t_2, t_3 можно назвать системой «фокальных координат».

7. Доказать, что преобразование плоскости xy, заданное уравнениями

$$\xi=\frac{1}{2}\left(x+\frac{x}{x^2+y^2}\right), \qquad \eta=\frac{1}{2}\left(y-\frac{y}{x^2+y^2}\right),$$

а) конформно;

б) преобразует прямые, проходящие через начало, и окружности с центром в начале координат плоскости xy в софокусные конические сечения (кривые второго порядка) $t=\text{const}$, определяемые уравнением

$$\frac{\xi^2}{t+\frac{1}{2}}+\frac{\eta^2}{t-\frac{1}{2}}=1.$$

8. Инверсия в трехмерном пространстве определяется формулами

$$\xi=\frac{x}{x^2+y^2+z^2}, \qquad \eta=\frac{y}{x^2+y^2+z^2}, \qquad \zeta=\frac{z}{x^2+y^2+z^2}.$$

Доказать, что а) угол между любыми двумя поверхностями не изменяется при инверсии; б) шаровые поверхности переходят либо в шаровые поверхности, либо в плоскости.

9. Доказать, что, если нормали к поверхности $z=u(x, y)$ встречают ось z, то эта поверхность является поверхностью вращения.

§ 4. Приложения

1. Параметрическое задание поверхности. Для исследования поверхностей, как и для изучения кривых, часто самым удобным оказывается параметрический способ задания, но при этом требуется уже не один, а два параметра, которые мы обозначим через u и v. Таким образом, параметрическое представление поверхности имеет следующий вид:

$$x=\varphi(u, v), \qquad y=\psi(u, v), \qquad z=\chi(u, v)$$

или, в векторной записи: $\boldsymbol{r}=\boldsymbol{r}(u, v)=\{\varphi(u, v), \psi(u, v), \chi(u, v)\}$, где φ, ψ и χ — заданные функции параметров u и v, а точка (u, v) пробегает некоторую заданную область G плоскости uv. Соответствующая точка (x, y, z) описывает некоторый геометрический образ в пространстве xyz. Этот образ, вообще говоря, является поверхностью, которую можно представить также уравнением вида $z=f(x, y)$. Действительно, можно попытаться выразить u и v из каких-нибудь двух уравнений из числа заданных трех через соответствующие две прямоугольные координаты; подставив затем найденные выражения для u и v в третье параметрическое уравнение, получим уравнение поверхности в несимметричной форме, например $z=f(x, y)$. (Заметим, кстати, что параметрическое представление содержит этот явный вид задания как частный случай, что сразу видно, если положить $x=u$, $y=v$.) Итак, для того чтобы заданная система параметрических уравнений действительно представляла поверхность, должна быть обеспечена возможность описанного только что исключения параметров u и v, а для этого достаточно предположить, что три якобиана

$$\begin{vmatrix} \varphi_u & \varphi_v \\ \psi_u & \psi_v \end{vmatrix}, \qquad \begin{vmatrix} \psi_u & \psi_v \\ \chi_u & \chi_v \end{vmatrix}, \qquad \begin{vmatrix} \chi_u & \chi_v \\ \varphi_u & \varphi_v \end{vmatrix}$$

не обращаются одновременно в нуль; эти три условия можно объединить в одном предположении, что

$$(\varphi_u\psi_v-\varphi_v\psi_u)^2+(\psi_u\chi_v-\psi_v\chi_u)^2+(\chi_u\varphi_v-\chi_v\varphi_u)^2>0. \qquad (*)$$

При этом предположении действительно возможно в окрестности любой точки пространства, представляемой тремя заданными уравнениями, выразить однозначно одну из координат x, y, z через две другие.

Пример 1. Рассмотрим следующее параметрическое представление шаровой поверхности $x^2+y^2+z^2=a^2$ радиуса a:

$$x=a\cos u\sin v, \qquad y=a\sin u\sin v, \qquad z=a\cos v$$
$$(0\leqslant u<2\pi, \quad 0\leqslant v\leqslant\pi),$$

в котором параметрами служат полярный угол $u=\varphi$ и географическая долгота $v=\theta$ точки сферы.

На этом примере видно одно из преимуществ параметрического задания. Прямоугольные координаты заданы как явные и притом однозначные функции от u и v. При изменении $v=0$ от 0 до $\frac{\pi}{2}$ получается верхняя полусфера, а при $\frac{\pi}{2}\leqslant v\leqslant\pi$ получается нижняя полусфера. Стало быть, для получения всей шаровой поверхности при параметрическом ее задании не приходится рассматривать две однозначные ветви $z=\pm\sqrt{a^2-x^2-y^2}$ двузначной функции z, определяемой неявным уравнением сферы $x^2+y^2+z^2-a^2=0$.

Можно получить и другое параметрическое представление шаровой поверхности с помощью *стереографической проекции*. Принимаем *северный полюс* $N(0, 0, a)$ за центр проекций, и из N проектируем все точки сферы на *плоскость экватора* $z=0$. Для этого соединяем прямой линией каждую точку $P(x, y, z)$ сферы с полюсом N; точка пересечения Q прямой NP с плоскостью экватора называется стереографическим изображением точки P сферы (рис. 41). Таким образом, между точками сферы (кроме северного полюса N) и точками экваториальной плоскости устанавливается взаимно однозначное соответствие. Обозначим прямоугольные координаты изображающей точки Q в плоскости экватора (относительно осей x и y) через u, v. Тогда из коллинеарности векторов $\overline{NP}=\{x, y, z-a\}$ и $\overline{NQ}=\{u, v, -a\}$

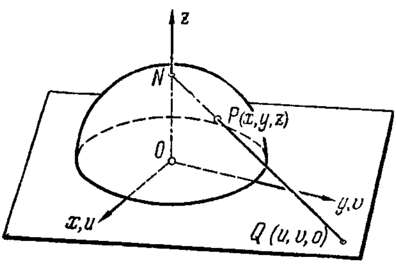


Рис. 41.

вытекает, что $\frac{x}{u}=\frac{y}{v}=\frac{z-a}{-a}$. Решая эти уравнения совместно с уравнением сферы $x^2+y^2+z^2=a^2$, получим следующие формулы соответствия:

$$x=\frac{2a^2u}{u^2+v^2+a^2},\qquad y=\frac{2a^2v}{u^2+v^2+a^2},\qquad z=\frac{(u^2+v^2-a^2)\,a}{u^2+v^2+a^2}.$$

Эти уравнения можно, очевидно, рассматривать и как параметрическое представление шаровой поверхности, причем параметрами служат прямоугольные координаты u, v изображающей точки Q в плоскости экватора.

Пример 2. *Однополостный гиперболоид* $\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$ (рис. 42) можно представить параметрическими уравнениями

$$x=a\cos u\,\mathrm{ch}\,v,\qquad y=b\sin u\,\mathrm{ch}\,v,\qquad z=c\,\mathrm{sh}\,v$$
$$(0\leqslant u<2\pi,\quad -\infty<v<+\infty).$$

Пример 3. *Двуполостный гиперболоид* $-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ (рис. 43) имеет следующее параметрическое представление:

$$x=a\cos u\,\mathrm{sh}\,v,\qquad y=b\sin u\,\mathrm{sh}\,v,\qquad z=\pm c\,\mathrm{ch}\,v$$
$$(0\leqslant u<2\pi,\quad -\infty<v<+\infty).$$

Параметрическое представление поверхности можно вообще рассматривать как *отображение области G плоскости uv на соответствующий кусок поверхности*, причем под словом отображение разумеют, как всегда, точечное соответствие. Каждой точке области G

плоскости uv соответствует одна точка поверхности, и обратное утверждение часто тоже справедливо [1]).

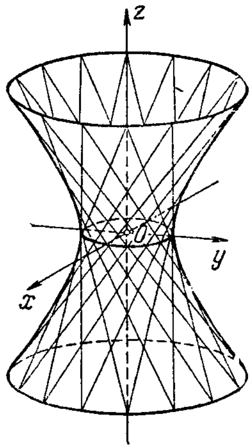


Рис. 42.

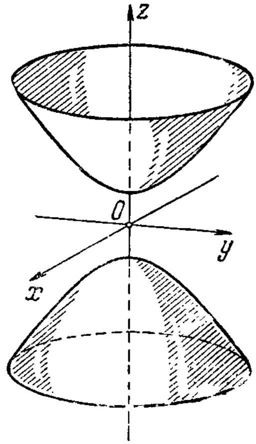


Рис. 43.

Кривой $u = u(t)$, $v = v(t)$ плоскости uv соответствует на данной поверхности кривая

$$x = \varphi(u(t), v(t)) = x(t),$$
$$y = \psi(u(t), v(t)) = y(t),$$
$$z = \chi(u(t), v(t)) = z(t)$$

или, в векторной записи,

$$\boldsymbol{r} = \boldsymbol{r}(t) = \{x(t),\ y(t),\ z(t)\}.$$

Например, в параметрическом представлении сферы с помощью географических координат (пример 1) прямым, параллельным оси v: $u = c = \text{const}$, $v = t$ соответствуют на сфере меридианы

$$x = a\cos c\sin t, \qquad y = a\sin c\sin t, \qquad z = a\cos t;$$

прямым, параллельным оси u: $u = \tau$, $v = k = \text{const}$ соответствуют на сфере параллели широт

$$x = a\sin k\cos\tau, \qquad y = a\sin k\sin\tau, \qquad z = a\cos k.$$

Стало быть, прямым, параллельным осям плоскости uv, соответствует сетка меридианов и параллелей на сфере.

---

[1]) Правда, не всегда. Бывает и так, что отдельным точкам поверхности соответствуют целые дуги кривых в плоскости uv; например, в параметрическом задании сферы с помощью географических координат ее полюсам соответствуют отрезки прямых $v = \theta = 0$ и $v = \theta = \pi$.

2. Линейный элемент поверхности. Одним из важнейших методов для исследования поверхности является изучение кривых, лежащих на этой поверхности. Мы здесь выведем только выражение для длины дуги s такой кривой на поверхности. На основании гл. II, § 7, n° 3

$$\left(\frac{ds}{dt}\right)^2=\left(\frac{d\boldsymbol{r}}{dt}\right)^2=\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2.$$

Так как $\frac{d\boldsymbol{r}}{dt}=\frac{\partial\boldsymbol{r}}{\partial u}\frac{du}{dt}+\frac{\partial\boldsymbol{r}}{\partial v}\frac{dv}{dt}$, то

$$\left(\frac{ds}{dt}\right)^2=E\left(\frac{du}{dt}\right)^2+2F\frac{du}{dt}\frac{dv}{dt}+G\left(\frac{dv}{dt}\right)^2, \tag{1}$$

причем, для большей сжатости, введены так называемые *гауссовы коэффициенты*

$$E=\left(\frac{\partial\boldsymbol{r}}{\partial u}\right)^2=\left(\frac{\partial x}{\partial u}\right)^2+\left(\frac{\partial y}{\partial u}\right)^2+\left(\frac{\partial z}{\partial u}\right)^2,$$
$$F=\frac{\partial\boldsymbol{r}}{\partial u}\frac{\partial\boldsymbol{r}}{\partial v}=\frac{\partial x}{\partial u}\frac{\partial x}{\partial v}+\frac{\partial y}{\partial u}\frac{\partial y}{\partial v}+\frac{\partial z}{\partial u}\frac{\partial z}{\partial v},$$
$$G=\left(\frac{\partial\boldsymbol{r}}{\partial v}\right)^2=\left(\frac{\partial x}{\partial v}\right)^2+\left(\frac{\partial y}{\partial v}\right)^2+\left(\frac{\partial z}{\partial v}\right)^2.$$

Эти величины не зависят от выбора кривой на поверхности; для заданной поверхности они зависят только от выбора ее параметрического представления и от выбора точки (u, v) на поверхности.

Умножив (1) на dt^2, получим для *дифференциала дуги* ds или, как принято говорить, для *линейного элемента поверхности* следующую *квадратичную дифференциальную форму* от дифференциалов du и dv криволинейных координат:

$$ds^2=E\,du^2+2F\,du\,dv+G\,dv^2.$$

[Квадратичной формой называется вообще однородный многочлен второй степени.]

Мы уже знаем (стр. 151), что нормальный вектор поверхности, заданной неявным уравнением $\Phi(x, y, z)=0$, имеет направление градиента функции Φ:

$$\operatorname{grad}\Phi=\{\Phi_x,\ \Phi_y,\ \Phi_z\}.$$

Для того чтобы получить нормальный вектор поверхности, заданной в параметрической форме

$$x=\varphi(u, v),\qquad y=\psi(u, v),\qquad z=\chi(u, v)$$

или (что то же самое)

$$\boldsymbol{r}=\boldsymbol{r}(u, v)=\{\varphi(u, v),\ \psi(u, v),\ \chi(u, v)\},$$

предположим, что эта же поверхность имеет неявное уравнение $\Phi(x, y, z)=0$. Подставив в это неявное уравнение выражения координат x, y, z через параметры u и v, получим соотношение

$$\Phi(\varphi(u, v), \psi(u, v), \chi(u, v)) = 0,$$

тождественное относительно u и v. Дифференцируя это соотношение по u и по v, получим два новых тождества

$$\Phi_x\varphi_u + \Phi_y\psi_u + \Phi_z\chi_u = 0 \quad \text{и} \quad \Phi_x\varphi_v + \Phi_y\psi_v + \Phi_z\chi_v = 0$$

или, в векторной записи,

$$\frac{\partial \boldsymbol{r}}{\partial u}\operatorname{grad}\Phi = 0 \quad \text{и} \quad \frac{\partial \boldsymbol{r}}{\partial v}\operatorname{grad}\Phi = 0.$$

Из равенства нулю этих двух скалярных произведений вытекает, что $\operatorname{grad}\Phi$ перпендикулярен к каждому из векторов $\frac{\partial \boldsymbol{r}}{\partial u}$ и $\frac{\partial \boldsymbol{r}}{\partial v}$, а стало быть, коллинеарен их векторному произведению. Следовательно, за нормальный вектор поверхности можно принять вектор $[\boldsymbol{r}_u\ \boldsymbol{r}_v]$. Так как

$$\boldsymbol{r}_u = \{\varphi_u, \psi_u, \chi_u\},$$
$$\boldsymbol{r}_v = \{\varphi_v, \psi_v, \chi_v\},$$

то нормальный вектор нашей поверхности

$$\boldsymbol{N} = [\boldsymbol{r}_u\boldsymbol{r}_v] = \{\psi_u\chi_v - \chi_u\psi_v,\ \chi_u\varphi_v - \varphi_u\chi_v,\ \varphi_u\psi_v - \psi_u\varphi_v\}.$$

В силу сделанного в самом начале n° 1 допущения (*), этот вектор не равен нулю. Для вычисления направляющих косинусов нормального вектора понадобится его модуль. Согласно тождеству Лагранжа (стр. 33, упр. 7), имеем

$$|\boldsymbol{N}|^2 = [\boldsymbol{r}_u\boldsymbol{r}_v]^2 = \boldsymbol{r}_u^2\boldsymbol{r}_v^2 - (\boldsymbol{r}_u\boldsymbol{r}_v)^2 = EG - F^2.$$

Следовательно, модуль нормального вектора поверхности $|\boldsymbol{N}| = \sqrt{EG - F^2}$.

В конце n° 1 мы видели, что уравнения $u = g(t)$, $v = h(t)$ определяют кривую на поверхности. Касательный вектор $\boldsymbol{T}$ этой кривой дается производной от радиус-вектора по параметру t. По правилу цепочки

$$\boldsymbol{T} = \frac{d\boldsymbol{r}}{dt} = \boldsymbol{r}_u\frac{dg}{dt} + \boldsymbol{r}_v\frac{dh}{dt} =$$
$$= \{x_u\dot{g}(t) + x_v\dot{h}(t),\ y_u\dot{g}(t) + y_v\dot{h}(t),\ z_u\dot{g}(t) + z_v\dot{h}(t)\},$$

а *единичный* касательный вектор получим делением вектора $\boldsymbol{T}$ на его модуль

$$|\boldsymbol{T}| = \left|\frac{d\boldsymbol{r}}{dt}\right| = \frac{ds}{dt} = \sqrt{E\dot{g}^2 + 2F\dot{g}\dot{h} + G\dot{h}^2}$$

$\left(\text{см. формулу (1), где теперь } \frac{du}{dt} = \dot{g}, \quad \frac{dv}{dt} = \dot{h}\right)$.

Рассмотрим на нашей поверхности еще другую кривую $u = g_1(\tau)$, $v = h_1(\tau)$, пересекающую первую в точке P, соответствующей некоторым значениям параметра t первой кривой и параметра τ второй. Обозначим касательные векторы обеих кривых в их общей точке через $\boldsymbol{T}$ и $\boldsymbol{T}_1$. Тогда угол ω между обеими кривыми как угол между их касательными векторами определяется по формуле

$$\cos\omega = \frac{\boldsymbol{T}\boldsymbol{T}_1}{|\boldsymbol{T}|\cdot|\boldsymbol{T}_1|} = \frac{E\dot{g}\dot{g}_1 + F(\dot{g}\dot{h}_1 + \dot{g}_1\dot{h}) + G\dot{h}\dot{h}_1}{\sqrt{E\dot{g}^2 + 2F\dot{g}\dot{h} + G\dot{h}^2}\cdot\sqrt{E\dot{g}_1^2 + 2F\dot{g}_1\dot{h}_1 + G\dot{h}_1^2}},$$

где $\dot{g} = \frac{dg}{dt}$, $\dot{h} = \frac{dh}{dt}$, $\dot{g}_1 = \frac{dg_1}{d\tau}$, $\dot{h}_1 = \frac{dh_1}{d\tau}$; значения всех этих производных должны быть взяты в точке пересечения обеих кривых.

В частности, можно рассматривать те кривые на нашей поверхности, которые определяются уравнениями вида $u = \text{const}$ или $v = \text{const}$. Подставляя в параметрические уравнения поверхности вместо u постоянное значение $u = c$, получим лежащую на поверхности пространственную кривую, для которой переменным параметром служит v; если же подставим постоянное значение $v = k$, то получится кривая на поверхности с переменным параметром u. Эти кривые $u = c$ и $v = k$ называются *параметрическими кривыми* на поверхности; мы уже встречались с ними. Сетка параметрических кривых соответствует сетке прямых, параллельных осям, в плоскости uv (рис. 44).

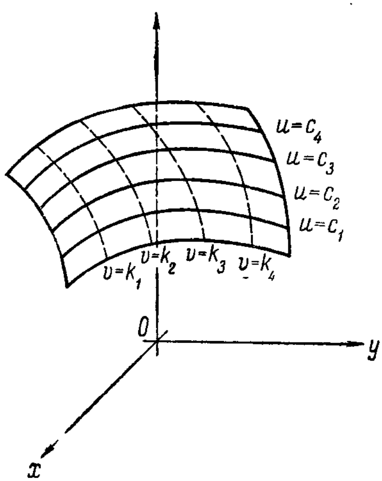


Рис. 44.

Отображение плоской области на другую плоскую область можно рассматривать как частный случай параметрического представления поверхности. Действительно, пусть $z = \chi(u, v) = 0$ во всей заданной области плоскости uv. Когда точка (u, v) пробегает свою область изменения, то при этом соответствующая точка (x, y, z) опишет некоторую область плоскости (x, y). Стало быть, наши уравнения дают просто отображение области плоскости uv на область плоскости xy. Если же мы предпочитаем точку зрения преобразования координат, то можно сказать, что те же уравнения определяют систему криволинейных координат в плоской области uv; обратные функции (если они существуют) определяют в свою очередь криволинейную систему координат u, v в плоской области xy. Линейный элемент ds плоскости xy выражается через дифференциалы криволинейных координат u, v следующей формулой:

$$ds^2 = E du^2 + 2F du\, dv + G dv^2,$$

где

$$E = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2, \quad F = \frac{\partial x}{\partial v}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v}, \quad G = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2.$$

Приведем еще один пример параметрического задания поверхности, а именно выведем параметрические уравнения *тора*. Эту поверхность описывает окружность при своем вращении вокруг оси, лежащей в плоскости окружности и ее не пересекающей (рис. 45). Примем ось вращения за ось z, а ось x проведем через центр окружности; абсциссу этого центра обозначим через a. Если радиус окружности $r < |a|$, то сразу получим параметрические уравнения окружности (в ее начальном положении) в плоскости xz:

$$x - a = r\cos\theta, \qquad y = 0,$$
$$z = r\sin\theta \qquad (0 \leqslant \theta < 2\pi).$$

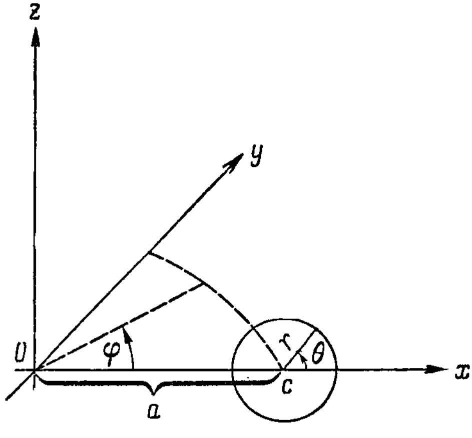


Рис. 45.

При вращении этой окружности вокруг оси z расстояние

$$\rho = \sqrt{x^2 + y^2} = a + r\cos\theta$$

каждой ее фиксированной точки от оси вращения остается постоянным. Таким образом, обозначив через φ угол поворота вокруг оси z, получим следующие параметрические уравнения тора:

$$x = (a + r\cos\theta)\cos\varphi, \qquad y = (a + r\cos\theta)\sin\varphi, \qquad z = r\sin\theta$$
$$(0 \leqslant \varphi < 2\pi, \qquad 0 \leqslant \theta < 2\pi)$$

с параметрами θ и φ. При этом поверхность тора представлена как однозначное отображение квадрата со стороной 2π в плоскости $\theta\varphi$. Исключение составляют точки периметра квадрата: любая пара точек периметра, лежащая на одной и той же прямой $\theta = \text{const}$ или $\varphi = \text{const}$, соответствует лишь одной точке поверхности тора, да и все четыре вершины квадрата соответствуют одной и той же точке тора.

Нетрудно вычислить выражение для линейного элемента поверхности тора

$$ds^2 = r^2 d\theta^2 + (a + r\cos\theta)^2 d\varphi^2.$$

3. Понятие о конформном отображении. Преобразование

$$\xi = \varphi(x, y), \qquad \eta = \psi(x, y) \tag{1}$$

называется *конформным*, если оно отображает любые две пересекающиеся кривые в такие две другие, которые составляют между собой тот же угол, что и исходная пара кривых.

Теорема. *Для того чтобы* (*непрерывно дифференцируемое*), *преобразование* (1) *было конформным, необходимо и достаточно* *чтобы оно удовлетворяло условиям или уравнениям Коши — Римана*

$$\varphi_x - \psi_y = 0, \quad \varphi_y + \psi_x = 0 \tag{2}$$

либо

$$\varphi_x + \psi_y = 0, \quad \varphi_y - \psi_x = 0.$$

В первом случае направление отсчета углов сохраняется, во втором же случае оно изменяется на противоположное; это последнее утверждение сразу вытекает из результатов § 3, n° 6 о знаке якобиана $\varphi_x\psi_y - \varphi_y\psi_x$.

Доказательство необходимости. Предположим, что преобразование (1) конформно. Линии $\xi = \text{const}$ и $\eta = \text{const}$ плоскости $\xi\eta$ ортогональны; следовательно, соответствующие им в плоскости xy кривые $\varphi(x, y) = \text{const}$ и $\psi(x, y) = \text{const}$ тоже должны быть ортогональны. Стало быть,

$$\varphi_x\psi_x + \varphi_y\psi_y = 0$$

(см. условие ортогональности двух кривых, стр. 148). Отсюда вытекает, что

$$\varphi_x = \lambda\psi_y, \qquad \varphi_y = -\lambda\psi_x. \tag{*}$$

С другой стороны, кривые, соответствующие в плоскости xy взаимно перпендикулярным прямым $\xi + \eta = \text{const}$ и $\xi - \eta = \text{const}$, тоже должны быть ортогональны. Это дает

$$(\varphi_x + \psi_x)(\varphi_x - \psi_x) + (\varphi_y + \psi_y)(\varphi_y - \psi_y) = 0,$$

откуда

$$\varphi_x^2 + \varphi_y^2 = \psi_x^2 + \psi_y^2.$$

Подставив в это равенство выражения (*), получим $\lambda^2 = 1$, а это значит, что функции $\varphi(x, y)$ и $\psi(x, y)$ удовлетворяют либо первой, либо второй системе уравнений Коши — Римана.

Доказательство достаточности. Предположим, что условия Коши — Римана выполнены. Пусть в плоскости xy даны две кривые $F(x, y) = 0$ и $G(x, y) = 0$, и пусть, в силу нашего преобразования, $F(x, y) = \Phi(\xi, \eta)$ и $G(x, y) = \Gamma(\xi, \eta)$. С помощью правила цепочки находим

$$F_x = \Phi_\xi\varphi_x + \Phi_\eta\psi_x, \qquad G_x = \Gamma_\xi\varphi_x + \Gamma_\eta\psi_x,$$
$$F_y = \Phi_\xi\varphi_y + \Phi_\eta\psi_y, \qquad G_y = \Gamma_\xi\varphi_y + \Gamma_\eta\psi_y.$$

А теперь, воспользовавшись условиями Коши — Римана, после несложных выкладок получим

$$F_x^2 + F_y^2 = (\Phi_\xi^2 + \Phi_\eta^2)(\varphi_x^2 + \varphi_y^2),$$
$$G_x^2 + G_y^2 = (\Gamma_\xi^2 + \Gamma_\eta^2)(\varphi_x^2 + \varphi_y^2),$$
$$F_xG_x + F_yG_y = (\Phi_\xi\Gamma_\xi + \Phi_\eta\Gamma_\eta)(\varphi_x^2 + \varphi_y^2);$$

следовательно,

$$\frac{F_xG_x+F_yG_y}{\sqrt{F_x^2+F_y^2}\,\sqrt{G_x^2+G_y^2}}=\frac{\Phi_\xi\Gamma_\xi+\Phi_\eta\Gamma_\eta}{\sqrt{\Phi_\xi^2+\Phi_\eta^2}\,\sqrt{\Gamma_\xi^2+\Gamma_\eta^2}}.$$

В левой части этого равенства стоит косинус угла между кривыми $F=0$ и $G=0$, в правой — косинус угла между их изображениями $\Phi=0$ и $\Gamma=0$. Стало быть, угол между двумя кривыми плоскости xy равен углу между их изображениями в плоскости $\xi\eta$. Итак, наше преобразование конформно, и достаточность условий Коши — Римана доказана.

Упражнения

1. а) Доказать, что стереографическая проекция сферы на плоскость является конформным отображением.

б) Доказать, что стереографическая проекция преобразует окружности на сфере в окружности или в прямые линии на плоскости.

в) Доказать, что отражение сферической поверхности в плоскости экватора (как в зеркале) соответствует, при стереографической проекции, инверсии в плоскости uv.

г) Найти выражение для линейного элемента сферы в стереографических параметрах u, v.

2. Вычислить линейный элемент

а) на сфере $x=a\sin\theta\cos\varphi$, $y=a\sin\theta\sin\varphi$, $z=a\cos\theta$;

б) на однополостном гиперболоиде

$$x=a\cos u\operatorname{ch}v,\qquad y=b\sin u\operatorname{ch}v,\qquad z=c\operatorname{sh}v;$$

в) на поверхности, описанной вращением вокруг оси z кривой $x=f(z)$ плоскости xz. Полагаем $f(z)>0$, так что кривая не пересекает оси вращения. В качестве криволинейных координат на поверхности вращения принять цилиндрические координаты z и θ;

г)* на поверхности $t_s=\text{const}$ семейства софокусных поверхностей второго порядка, заданного уравнением

$$\frac{x^2}{a-t}+\frac{y^2}{b-t}+\frac{z^2}{c-t}=1,$$

приняв t_1 и t_2 за криволинейные координаты на изучаемой поверхности (ср. упр. 6, стр. 176).

3. На поверхности с параметрами u, v вводится новая система криволинейных координат r, s с помощью уравнений

$$u=u(r,s),\qquad v=v(r,s).$$

Доказать, что

$$E'G'-F'^2=(EG-F^2)\left(\frac{\partial(u,v)}{\partial(r,s)}\right)^2,$$

где E', F', G' обозначают гауссовы коэффициенты на поверхности в системе параметров r, s и E, F, G — гауссовы коэффициенты в системе u, v.

4. Обозначим через t касательную к поверхности S в ее точке P и рассмотрим линии сечения поверхности S всеми плоскостями, содержащими касательную t. Доказать, что центры кривизны всех этих линий лежат на окружности.

5. Как и в упр. 4, пусть t — какая-либо касательная к поверхности S в ее точке P. Построим нормальное плоское сечение, проходящее через

касательную t (т. е. линию сечения поверхности S плоскостью, проходящей через t и через нормаль к S в точке P). Обозначим кривизну этого нормального сечения в точке P через k (она зависит от направления касательной t). На каждой касательной в точке P отложим от P вектор $\overline{PQ}$ длиною $\frac{1}{\sqrt{k}}$. Доказать, что конечные точки Q этих векторов лежат на кривой второго порядка.

6*. Кривая задана как сечение двух поверхностей

$$x^2+y^2+z^2=1 \quad \text{и} \quad ax^2+by^2+cz^2=0.$$

Составить а) уравнения касательной и б) уравнение соприкасающейся плоскости в любой точке кривой.

§ 5. Семейства кривых и семейства поверхностей; их огибающие

1. Понятие семейства кривых и семейства поверхностей. Нам уже неоднократно приходилось рассматривать кривые (или поверхности) не как единичные фигуры, а как члены *семейства* кривых (или поверхностей); таково семейство кривых на плоскости $f(x, y) = c$, где каждому значению постоянной c соответствует своя кривая семейства. Число c, постоянное для каждой кривой, но изменяющееся при переходе от одной кривой семейства к другой, называется *параметром* семейства.

Так, например, все прямые плоскости xy, параллельные оси y, т. е. прямые $x = c$, составляют семейство. Множество концентрических окружностей $x^2+y^2=c^2$ с центром в начале координат тоже образует семейство кривых; каждому значению c соответствует одна окружность этого семейства, а именно окружность радиуса c. Мы уже встречались с семейством равнобочных гипербол $xy=c$, изображенным на рис. 29 (стр. 135); в этом семействе частное значение $c=0$ соответствует вырожденной гиперболе, состоящей из обеих координатных осей.

Интересный пример семейства линий представляет множество всех нормалей какой-либо данной кривой. Если эта кривая задана параметрическими уравнениями $\xi=\varphi(t)$, $\eta=\psi(t)$, то семейство нормалей определяется уравнением

$$[x-\varphi(t)]\,\dot\varphi(t)+[y-\psi(t)]\,\dot\psi(t)=0;$$

в этом примере параметром семейства нормалей служит параметр t исходной кривой (вместо прежнего c).

Теперь дадим аналитическое определение понятия семейства кривых на плоскости xy. Пусть $f(x, y, c)$ обозначает непрерывно дифференцируемую функцию двух независимых переменных x и y и *параметра* c, изменяющегося в некотором заданном интервале. (Стало быть, параметр c есть фактически третья независимая переменная, которой дано другое название только потому, что она играет другую роль, чем x и y.) Так вот, если уравнение

$$f(x, y, c)=0$$

представляет при всяком значении параметра c кривую, то совокупность этих кривых, которые получаются, когда c пробегает свой интервал, называется (*однопараметрическим*) *семейством кривых*, зависящим от параметра c.

Кривые такого семейства могут быть заданы и в параметрическом виде

$$x=\varphi(t,\ c),\qquad y=\psi(t,\ c),$$

где c есть *параметр семейства.* Если подставить вместо c определенное значение, то эти уравнения дают параметрическое представление конкретной кривой семейства с *параметром t кривой*; изменение параметра t (при постоянном c) вызывает движение точки (x, y) вдоль этой кривой.

Например, рассмотренное выше семейство концентрических окружностей $x^2+y^2=c^2$ можно представить и в параметрическом виде

$$x=c\cos t,\qquad y=c\sin t.$$

Приведенное выше семейство равнобочных гипербол тоже можно задать в параметрическом виде $x=ct,\ y=\frac{1}{t}$.

Порою приходится рассматривать и такие семейства плоских кривых, которые зависят не от одного параметра c, а от нескольких параметров. Например, множество всех окружностей на плоскости $(x-a)^2+(y-b)^2=c^2$ есть семейство кривых, зависящее от трех параметров a, b, c. В дальнейшем мы под семейством кривых будем всегда разуметь *однопараметрическое* семейство, если не оговорено противное. Там же, где они понадобятся, мы для отличия скажем, что речь идет о двупараметрическом, трехпараметрическом или многопараметрическом семействе кривых.

Совершенно таким же путем вводится понятие семейства поверхностей в пространстве. Если дана непрерывно дифференцируемая функция $f(x, y, z, c)$ своих четырех аргументов и если, при всяком значении c в некотором интервале, уравнение

$$f(x,y,z,c)=0$$

представляет поверхность в пространстве xyz, то совокупность поверхностей, которые получаются, когда c пробегает свой интервал, называется (*однопараметрическим*) *семейством поверхностей* с параметром c.

Существуют также и семейства поверхностей, зависящие от двух или нескольких параметров.

Например, множество концентрических сфер $x^2+y^2+z^2=c^2$ с центром в начале координат является однопараметрическим семейством поверхностей, а совокупность плоскостей, определяемая уравнением

$$ax+by+\sqrt{1-a^2-b^2}\,z+1=0,$$

является двупараметрическим семейством с параметрами a и b, пробегающими область $a^2+b^2 \leqslant 1$. Это семейство всех плоскостей, отстоящих от начала координат на расстоянии, равном единице.

Иногда говорят, что однопараметрическое семейство содержит ∞^1 элементов (в данном случае, поверхностей), двупараметрическое — ∞^2 элементов и т. д.

2. Огибающая и дискриминантная кривая однопараметрического семейства плоских линий. Если семейство прямых тождественно с совокупностью касательных к некоторой кривой, то эта кривая называется *огибающей* семейства прямых. Так, например, семейство нормалей плоской кривой C совпадает с семейством касательных к ее эволюте E; стало быть, эволюта E является огибающей семейства нормалей исходной кривой C (ср. т. I, стр. 327 и 352).

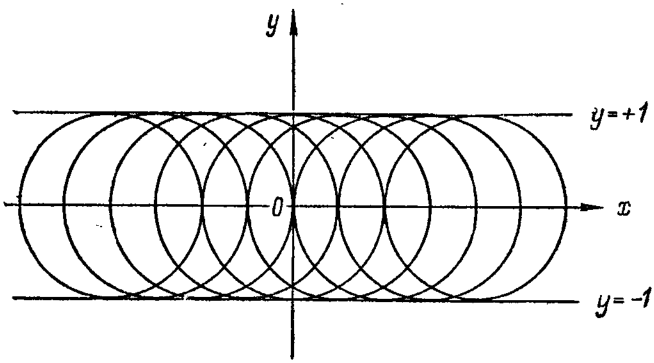


Рис. 46.

Рассмотрим аналогичный пример семейства окружностей $(x-c)^2+$ $+y^2-1=0$, т. е. окружностей радиуса 1, центры которых лежат на оси x (рис. 46). Пара прямых $y=1$ и $y=-1$ касается всех окружностей семейства; в соответствии с этим говорят, что эта пара прямых служит огибающей нашего семейства окружностей.

Как получить в этих примерах точку касания огибающей с линией огибаемого семейства? Найдем точку пересечения двух смежных кривых семейства со значениями параметра c и $c+h$ и заставим затем h стремиться к нулю. Естественно ожидать, что предельное положение точки пересечения и будет точкой касания огибающей с кривой со значением параметра c. Эту мысль выражают кратко так: огибающая является геометрическим местом точек пересечения бесконечно близких кривых семейства.

Встречаются и другие семейства линий $f(x, y, c)=0$, для которых существует такая кривая E, которая в каждой из своих точек имеет касание с одной из линий семейства, причем в разных своих точках кривая E касается разных линий семейства. Тогда кривая E называется *огибающей* семейства линий $f(x, y, c)=0$.

Естественно возникает вопрос: как найти огибающую данного семейства кривых $f(x, y, c)=0$? Начнем с допущения, что огибающая действительно существует и что она является, как в приведенных выше примерах, геометрическим местом точек пересечения бесконечно близких кривых семейства[1]).

В соответствии с этим мы получим аналитически точку касания кривой $f(x, y, c)=0$ с огибающей E следующим путем. Наряду с указанной кривой семейства рассмотрим соседнюю кривую $f(x, y, c+h)=0$, найдем точку пересечения этих двух кривых, а затем заставим h стремиться к нулю. Тогда точка пересечения должна стремиться к искомой точке касания. Координаты точки пересечения удовлетворяют системе уравнений $f(x, y, c)=0$ и $f(x, y, c+h)=0$, а следовательно, и уравнению

$$\frac{f(x, y, c+h)-f(x, y, c)}{h}=0.$$

В последнем уравнении легко выполнить предельный переход $h \to 0$. Так как мы с самого начала предположили, что существует частная производная f_c, то для точки касания кривой $f(x, y, c)=0$ с огибающей E получаем систему уравнений

$$f(x, y, c)=0 \quad \text{и} \quad f_c(x, y, c)=0. \tag{*}$$

Если возможно выразить из этих уравнений x и y как функции от c, то получится параметрическое представление некоторой кривой, с параметром c, и естественно ожидать, что эта кривая и есть огибающая. Но можно получить уравнение этой кривой и в неявном виде $g(x, y)=0$, исключив из системы (*) параметр c. Уравнение $g(x, y)=0$ называется *дискриминантным уравнением*, а кривая, представляемая этим уравнением, *дискриминантной кривой* семейства.

Итак, мы пришли к следующему правилу: *для того чтобы найти дискриминантную кривую семейства кривых $f(x, y, c)=0$, надо рассматривать совместно два уравнения $f(x, y, c)=0$ и $f_c(x, y, c)=0$ и стараться либо определить из этой системы x и y как функции от c, либо исключить из нее параметр c.*

Наглядные соображения делают правдоподобным, что дискриминантная кривая и есть огибающая.

Это эвристическое рассуждение мы теперь заменим более общим и полноценным выводом, основанным на самом определении огибающей как кривой, касающейся всех кривых данного семейства. Вместе с тем выяснится, при каких условиях дискриминантная кривая действительно дает огибающую, и какие другие возможности содержит наше правило.

---

[1]) На примерах выяснится, что это допущение является слишком узким; поэтому мы вскоре дадим более общий и полноценный вывод.

Сначала допустим, что существует огибающая E, которая может быть представлена в параметрическом виде уравнениями

$$x = x(c), \qquad y = y(c),$$

причем функции $x(c)$ и $y(c)$ непрерывно дифференцируемы и $\left(\frac{dx}{dc}\right)^2 + \left(\frac{dy}{dc}\right)^2 \neq 0$, и что в своей точке с некоторым значением параметра c она касается той кривой семейства, которая соответствует тому же значению параметра c. Докажем, что функции $x = x(c)$ и $y = y(c)$ удовлетворяют системе уравнений (*).

Для доказательства заметим, что при всяком значении c координаты соответствующей точки касания, с одной стороны, получаются из параметрических уравнений огибающей, а с другой стороны, удовлетворяют уравнению $f(x, y, c) = 0$. Поэтому если подставить в последнее уравнение выражения $x = x(c)$ и $y = y(c)$, то оно удовлетворится тождественно при всех значениях параметра c из интервала его изменения. Дифференцируя это тождество по c, получим

$$f_x \frac{dx}{dc} + f_y \frac{dy}{dc} + f_c = 0.$$

Но факт касания огибающей E и кривой семейства $f(x, y, c) = 0$ выражается условием

$$f_x \frac{dx}{dc} + f_y \frac{dy}{dc} = 0,$$

так как вектор $\left\{\frac{dx}{dc}, \frac{dy}{dc}\right\}$ есть касательный вектор огибающей E, а вектор $\{f_x, f_y\}$ есть нормальный вектор кривой семейства, а эти векторы должны быть взаимно перпендикулярны, и их скалярное произведение равно нулю. Отсюда вытекает, что огибающая удовлетворяет также и уравнению $f_c = 0$. Стало быть, полученное выше правило дает *необходимое условие* для огибающей, т. е. огибающая, если она существует, входит в состав дискриминантной кривой.

Для того чтобы выяснить, в какой мере это условие является *достаточным*, предположим, что некоторая кривая E, заданная параметрическими уравнениями $x = x(c)$, $y = y(c)$, удовлетворяет системе уравнений $f(x, y, c) = 0$ и $f_c(x, y, c) = 0$. Это значит, что если подставить в них $x = x(c)$ и $y = y(c)$, то получатся тождества относительно c. Дифференцируя первое из них по c и учитывая, что $f_c = 0$, получим соотношение

$$f_x \frac{dx}{dc} + f_y \frac{dy}{dc} = 0,$$

которое, стало быть, выполняется во всех точках кривой E.

Если в какой-либо точке кривой E оказывается $f_x^2 + f_y^2 \neq 0$ и $\left(\frac{dx}{dc}\right)^2 + \left(\frac{dy}{dc}\right)^2 \neq 0$, то в этой точке линия E и кривая семейства имеют каждая определенную касательную, и полученное соотношение устанавливает, что кривая E и кривая семейства касаются друг друга в этой точке. Следовательно, при дополнительном условии $f_x^2+f_y^2\neq 0$ наше правило дает не только необходимое, но и *достаточное* условие для огибающей. Если же f_x и f_y обращаются одновременно в нуль, то рассматриваемая точка кривой E может оказаться особой точкой проходящей через нее кривой семейства (см. § 2, n° 2). В этом случае без дополнительного исследования невозможно делать какие-либо заключения о наличии или отсутствии касания обеих кривых в этой точке. Совокупность особых точек кривых семейства может образовать целую линию, входящую в состав дискриминантной кривой; возможен и такой случай, когда вся дискриминантная кривая является геометрическим местом особых точек кривых семейства (когда $f_x=f_y=0$ во всех точках кривой E). Что касается того случая, когда $\frac{dx}{dc}$ и $\frac{dy}{dc}$ обращаются одновременно в нуль, то он менее интересен: это может иметь место лишь в отдельных точках кривой E, которые могут оказаться ее особыми точками.

Таким образом, после того, как дискриминантная кривая найдена, необходимо еще всегда специальным исследованием выяснить, является ли она действительно огибающей или же геометрическим местом особых точек семейства кривых, либо, может быть, содержит и то и другое.

В заключение приведем уравнения для определения дискриминантной кривой семейства, заданного в параметрическом виде

$$x=\varphi(t, c), \qquad y=\psi(t, c), \tag{A}$$

где c — параметр семейства, а t — параметр на кривой. Условие $f_c=0$ заменяется здесь условием

$$\varphi_t\psi_c-\varphi_c\psi_t=0. \tag{B}$$

Это условие можно вывести путем перехода к прежнему виду уравнения семейства $f(x, y, c)=0$ (с помощью исключения параметра t).

Если удастся получить из (B) выражение t через c, то, подставив его в (A), найдем параметрическое представление дискриминантной кривой: $x=x(c)$, $y=y(c)$. Если же из трех уравнений (A) и (B) исключить оба параметра t и c, то уравнение дискриминантной кривой получится в виде $g(x, y)=0$.

3. Примеры. 1) Семейство окружностей радиуса 1, с центрами на оси x, мы уже рассмотрели наглядно в n° 2, рис. 46. Уравнение этого семейства $(x-c)^2+y^2=1$. Мы видели, что огибающая должна состоять из двух прямых $y=1$ и $y=-1$. К этому результату приводит и наше правило (*), ибо здесь $f(x, y, c)\equiv(x-c)^2+y^2-1=0$ и $f_c\equiv-2(x-c)=0$, откуда для дискриминантной кривой E получается уравнение $y^2=1$. Так как во всех точках линии E производная $f_y=2y$ принимает значения ± 2, а не нуль, то пара прямых $y^2=1$ и является огибающей.

2) Семейство окружностей радиуса 1, проходящих через начало координат. Ясно, что центры этих окружностей лежат на единичной окружности с центром в начале, и уравнение семейства есть

$$(x-\cos c)^2+(y-\sin c)^2=1$$

или

$$f(x, y, c) \equiv x^2+y^2-2x\cos c-2y\sin c=0,$$

откуда

$$f_c(x, y, c) \equiv 2x\sin c-2y\cos c=0.$$

Эта система уравнений имеет очевидное решение $x=y=0$. Если же $x^2+y^2 \neq 0$, то из наших уравнений вытекает, что $x=2\cos c$, $y=2\sin c$. Исключив c, получим $x^2+y^2=4$. Нетрудно проверить, что в точках этой окружности f_x и f_y не равны одновременно нулю. Итак, дискриминантная кривая состоит из окружности радиуса 2 с центром в начале координат

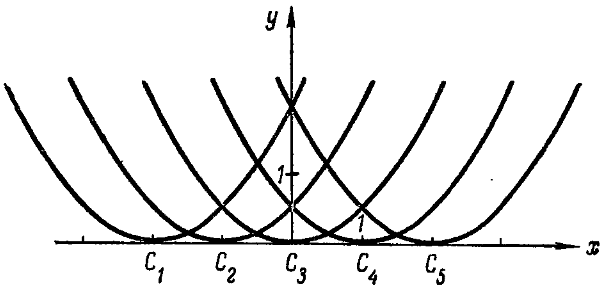


Рис. 47.

(это и есть огибающая, что соответствует наглядному представлению) и из изолированной точки (начала координат), являющейся общей точкой всех кривых семейства.

[Легко понять, что если все кривые семейства имеют общую точку A, то эта точка должна входить в состав дискриминантной кривой. Действительно, такая точка будет точкой пересечения любых двух бесконечно близких кривых семейства, а стало быть, либо геометрическим местом этих точек будет сама точка A, т. е. дискриминантная кривая вырождается в точку, либо точка A должна входить в состав дискриминантной кривой.

3) Найдем дискриминантную кривую семейства линий

$$u(x, y)+cv(x, y)=0.$$

Приравнивая нулю производную от левой части по c, получим $v(x, y)=0$, и для нахождения дискриминантной кривой имеем систему уравнений

$$u(x, y)=0, \qquad v(x, y)=0.$$

Если эта система совместна, то она определяет несколько отдельных точек — точек пересечения линий $u(x, y)=0$ и $v(x, y)=0$. Данное семейство есть пучок кривых, проходящих через общие точки основных линий пучка: $u=0$ и $v=0$. Итак, дискриминантная кривая пучка линий состоит из вершин этого пучка (ср. заключительное замечание предыдущего примера).

4) Семейство кривых может вообще не иметь дискриминантной кривой. Это будет в том случае, когда определяющие ее уравнения $f(x, y, c)=0$ и $f_c(x, y, c)=0$ несовместны.

Рассмотрим, например, линии уровня функции $F(x, y)$, т. е. семейство кривых $F(x, y) = c$. Тогда

$$f(x, y, c) = F(x, y) - c \quad \text{и} \quad f_c(x, y, c) \equiv -1,$$

а следовательно, f_c не может равняться нулю. Несуществование дискриминантной кривой вполне понятно, так как любые две различные кривые семейства $F(x, y) = c$ не имеют общих точек.]

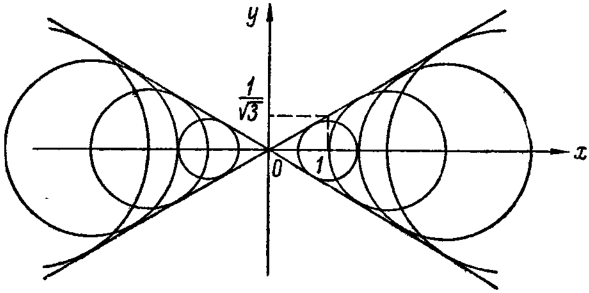


Рис. 48.

5) Для семейства парабол $(x-c)^2 - 2y = 0$ (рис. 47) получается дискриминантная кривая $y = 0$, т. е. ось x. Так как $f_y = -2 \neq 0$, то огибающей является ось x, что и геометрически ясно.

6) Рассмотрим семейство окружностей $f(x, y, c) \equiv \equiv (x-2c)^2 + y^2 - c^2 = 0$ (рис. 48). Уравнение $f_c = 0$ приводится к виду $2x - 3c = 0$. Подставив $c = \frac{2}{3}x$ в уравнение семейства, получим дискриминантную кривую $E: y^2 = \frac{x^2}{3}$. Так как $f_y = 2y$ на E обращается в нуль только в начале координат, то E есть огибающая, состоящая из двух прямых $y = \frac{x}{\sqrt{3}}$ и $y = -\frac{x}{\sqrt{3}}$, причем начало координат представляет исключение в том смысле, что в нем нет касания.

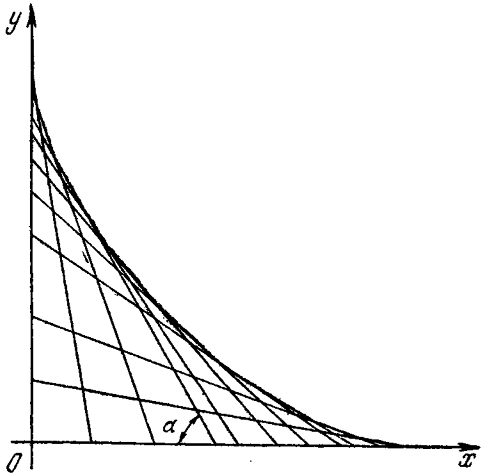


Рис. 49.

7) Рассмотрим семейство прямых линий, на которых оси x и y отсекают отрезки длиной 1. Роль параметра будет играть угол α, отмеченный на рис. 49. Уравнение семейства прямых запишется тогда так:

$$f(x, y, \alpha) \equiv \frac{x}{\cos\alpha} + \frac{y}{\sin\alpha} - 1 = 0,$$

откуда

$$f_\alpha \equiv \frac{\sin\alpha}{\cos^2\alpha}x - \frac{\cos\alpha}{\sin^2\alpha}y = 0.$$

Из этой системы уравнений получаются параметрические уравнения дискриминантной кривой

$$x = \cos^3\alpha, \qquad y = \sin^3\alpha,$$

а отсюда ее неявное уравнение $x^{2/3} + y^{2/3} = 1$. Эта кривая называется *астроидой* (ср. т. I, стр. 311, упр. 6). Она состоит из четырех симметричных

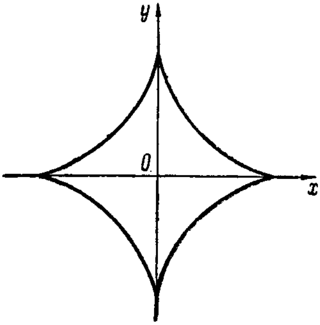


Рис. 50.

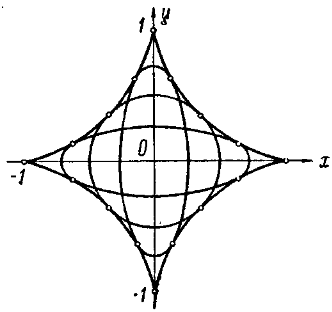


Рис. 51.

ветвей, смыкающихся в четырех точках заострения (рис. 49 и 50). Так как $f_x = \frac{1}{\cos\alpha}$ и $f_y = \frac{1}{\sin\alpha}$ не обращаются в нуль ни при каком значении α, то эта астроида является огибающей нашего семейства прямых.

8) Рассмотрим семейство эллипсов $\frac{x^2}{c^2} + \frac{y^2}{(1-c)^2} = 1$ $(0 < c < 1)$, полуоси которых имеют постоянную сумму, равную 1. Это семейство имеет своей огибающей ту же астроиду $x^{2/3} + y^{2/3} = 1$ (рис. 51).

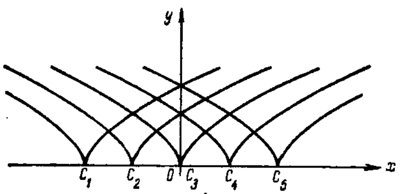


Рис. 52.

9) Семейство кривых $(x-c)^2 - y^3 = 0$ представляет собой пример семейства, не имеющего огибающей. Наше правило дает дискриминантную кривую $y = 0$, т. е. ось x. Но во всех ее точках f_x и f_y обращаются в нуль. Рис. 52 показывает, что ось x является не огибающей (она не касается кривых семейства), а геометрическим местом особых точек — точек заострения кривых семейства.

10) Для семейства $(x-c)^3 - y^2 = 0$ (рис. 53) в качестве дискриминантной кривой тоже получается ось x. Она и здесь является геометрическим местом точек заострения, но она касается каждой кривой семейства, а стало быть, ее следует считать огибающей.

11) Семейство *строфоид* (рис. 54):

$$[x^2 + (y-c)^2](x-2) + x = 0$$

интересно тем, что его дискриминантная кривая состоит из двух линий: огибающей и геометрического места узловых (двойных) точек. Все кривые семейства конгруэнтны и получаются перемещением одной из них параллельно оси y. Дифференцирование дает $f_c \equiv -2(y-c)(x-2)=0$, откуда

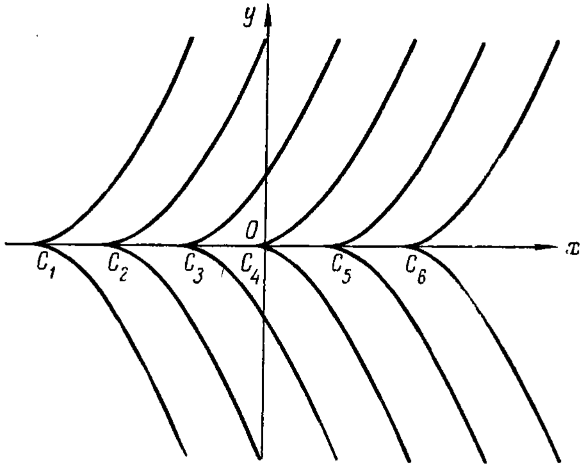


Рис. 53.

либо $x=2$, либо $y=c$. Прямая $x=2$ в счет не идет, так как при $x=2$ не получается конечного значения y. Остается принять $y=c$, и уравнение дискриминантной кривой будет $x^2(x-2)+x=0$ или $x(x-1)^2=0$. Эта кривая распадается на две прямых: $x=0$ и $x=1$. Из рис. 54 видно, что первая из них, $x=0$, является огибающей, а прямая $x=1$ есть геометрическое место двойных точек кривых семейства.

12) Огибающая не обязательно является геометрическим местом точек пересечения бесконечно близких кривых семейства. Это видно на примере семейства конгруэнтных кубических парабол $y-(x-c)^3=0$. Все эти параболы получаются

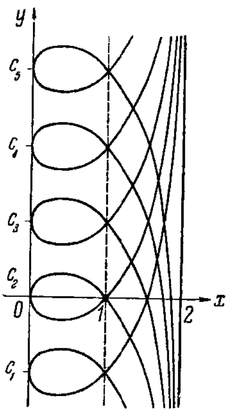


Рис. 54.

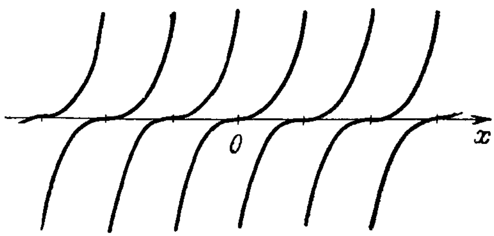


Рис. 55.

перемещением одной из них параллельно оси x, и никакие две из них не пересекаются (рис. 55). Наше правило дает уравнение $f_c \equiv 3(x-c)^2=0$, так что дискриминантной кривой является $y=0$, т. е. ось x. Она и является огибающей, так как касается всех кривых семейства.

13) Понятие огибающей предоставляет возможность дать новое определение эволюты плоской кривой (ср. т. I, стр. 327, 351 и след.). *Эволюту* E кривой C мы определим теперь как *огибающую семейства нормалей* кривой C. Пусть эта кривая задана уравнениями $x=\varphi(t)$, $y=\psi(t)$. Тогда семейство нормалей кривой C имеет уравнение (параметром семейства является t)

$$f(x, y, t)\equiv[x-\varphi(t)]\varphi'(t)+[y-\psi(t)]\psi'(t)=0.$$

Для нахождения дискриминантной кривой дифференцируем это уравнение по t и приравниваем f_t нулю:

$$f_t\equiv[x-\varphi(t)]\varphi''(t)+[y-\psi(t)]\psi''(t)-\varphi'^2(t)-\psi'^2(t)=0.$$

Из этого уравнения и из уравнения семейства нормалей находим параметрические уравнения дискриминантной кривой

$$x=\varphi(t)-\psi'(t)\frac{\varphi'^2+\psi'^2}{\psi''\varphi'-\varphi''\psi'}=\varphi-\frac{\psi'\rho}{\sqrt{\varphi'^2+\psi'^2}},$$

$$y=\psi(t)+\varphi'(t)\frac{\varphi'^2+\psi'^2}{\psi''\varphi'+\varphi''\psi'}=\psi+\frac{\varphi'\rho}{\sqrt{\varphi'^2+\psi'^2}},$$

где

$$\rho=\frac{(\varphi'^2+\psi'^2)^{3/2}}{\psi''\varphi'-\varphi''\psi'}$$

есть радиус кривизны кривой C (т. I, стр. 326). Так как $f_x=\varphi'(t)$ и $f_y=\psi'(t)$ не равны одновременно нулю при всех значениях t, соответствующих заведомо обыкновенным точкам кривой C, то дискриминантная кривая и является огибающей семейства нормалей, т. е. эволютой кривой C. Полученные здесь уравнения совпадают с уравнениями эволюты, выведенными в т. I, стр. 327.

14) Кривая C задана уравнениями $x=\varphi(t)$, $y=\psi(t)$. Найдем огибающую E семейства окружностей, проходящих через начало O и имеющих центры на C. Уравнение семейства есть

$$x^2+y^2-2x\varphi(t)-2y\psi(t)=0.$$

Так как окружности не имеют особых точек, то дискриминантная кривая должна состоять из огибающей E и начала координат [общей точки всех линий семейства — см. примеры 2 и 3]. Для ее получения надо к уравнению семейства присоединить уравнение

$$x\varphi'(t)+y\psi'(t)=0.$$

Это уравнение при каждом значении t представляет прямую, проходящую через начало O (общую точку всех окружностей), и ее вторая точка пересечения с окружностью семейства с тем же значением t есть как раз точка касания Q огибающей с этой окружностью. Указанная прямая, очевидно, перпендикулярна касательной кривой C в точке P этой кривой, служащей центром той же окружности. Ясно, что $PQ=PO$ как радиусы одной окружности; следовательно, PO и PQ образуют равные углы с касательной к кривой C в точке P.

Представим себе, что O есть светящаяся точка, а C — отражающая кривая (зеркало); тогда QP дает направление отраженного луча, соответствующего падающему лучу OP. Огибающая отраженных лучей называется *каустической кривой* или *каустикой* относительно точки O. Эта *каустическая кривая совпадает с эволютой кривой* E, т. е. огибающая отраженных лучей является эволютой огибающей E нашего семейства окружностей. Действительно, отраженный луч QP есть нормаль кривой E как радиус окружности с центром P, проведенный в точку касания этой окружности с кривой E, а из предшествующего примера мы знаем, что огибающая нормалей линии E является ее эволютой.

Пусть, например, кривая C (линия центров наших окружностей) есть тоже окружность, проходящая через начало O. Тогда огибающая E пучка окружностей есть путь, описываемый точкой O' окружности C', конгруэнтной с C и катящейся по C извне, причем в начале движения точка O' совпадает с O. Действительно, все время движения точка O' (как и точка Q в предыдущем рассмотрении) симметрична с O относительно общей касательной обеих окружностей. Следовательно, E является одной из эпициклоид, а именно кардиоидой (ср. т. I, стр. 310, упр. 2 и 3). Так как эволюта эпициклоиды есть тоже эпициклоида, подобная данной (см. т. I, стр. 357, упр. 1), то каустика окружности C относительно ее точки O есть кардиоида.

4. Огибающая семейства поверхностей. Все рассуждения, относящиеся к огибающей семейства плоских кривых, приложимы, лишь с небольшими изменениями, к семейству поверхностей. Рассмотрим сначала однопараметрическое семейство поверхностей $f(x, y, z, c)=0$ в некотором определенном интервале значений параметра c. Поверхность E называется *огибающей семейства*, если она касается каждой поверхности семейства вдоль целой кривой, причем множество этих кривых касания составляет семейство кривых, покрывающее полностью огибающую поверхность E.

Примером является семейство всех шаровых поверхностей радиуса 1 с центрами на оси z. Наглядно ясно, что огибающей будет поверхность кругового цилиндра $x^2+y^2-1=0$ радиуса 1, симметричная относительно оси z. Семейство кривых касания состоит из окружностей радиуса 1, плоскости которых параллельны плоскости $x0y$, а центры лежат на оси z. Заметим, кстати, что огибающая поверхность семейства сфер постоянного радиуса, центры которых лежат на какой-либо заданной кривой, называется *трубчатой поверхностью*.

Для нахождения огибающей поверхности, в предположении, что она существует, можно применить совершенно такой же эвристический метод, как в n° 2. Начинаем с рассмотрения двух поверхностей семейства $f(x, y, z, c)=0$ и $f(x, y, z, c+h)=0$. Эта пара уравнений определяет совместно линию пересечения обеих поверхностей, существование которой мы определенно предполагаем. Второе уравнение этой системы мы заменим уравнением

$$\frac{f(x, y, z, c+h)-f(x, y, z, c)}{h}=0,$$

которое является линейной комбинацией обоих уравнений, причем новая система равносильна прежней. Заставим теперь h стремиться к нулю; тогда линия пересечения будет стремиться к некоторому предельному положению, и эта предельная кривая, которую образно называют линией пересечения двух бесконечно близких поверхностей семейства, определяется следующей системой уравнений:

$$f(x, y, z, c)=0, \qquad f_c(x, y, z, c)=0. \qquad (*)$$

Она зависит еще от параметра c, так что все эти кривые пересечения образуют однопараметрическое семейство пространственных кривых.

Если из двух уравнений ($*$) исключить параметр c, то получим так называемое *дискриминантное уравнение* вида $g(x, y, z) = 0$. Представляемая этим уравнением поверхность называется *дискриминантной поверхностью* системы. Совершенно таким же способом, как в n° 2, можно показать, что огибающая поверхность, если она существует, входит в состав дискриминантной поверхности и что, обратно, дискриминантная поверхность имеет со всякой поверхностью семейства в любой их общей точке общую касательную плоскость, если только $f_x^2 + f_y^2 + f_z^2 \neq 0$ в этой общей точке, т. е. во всяком случае, если эта точка является обыкновенной точкой семейства. В итоге мы приходим к аналогичному выводу: дискриминантная поверхность состоит из огибающей и из геометрического места особых точек поверхностей семейства.

Пример. Найдем огибающую семейства $f(x, y, z, c) \equiv x^2 + y^2 + (z - c)^2 - 1 = 0$, подвергнутого выше наглядному рассмотрению. Согласно правилу ($*$), присоединяем равенство $f_c(x, y, z, c) \equiv -2(z - c) = 0$. При всяком фиксированном значении c оба уравнения, взятые совместно, определяют окружность радиуса 1, лежащую в плоскости, параллельной плоскости xOy на уровне $z = c$. Исключив из этих двух уравнений параметр c, получим уравнение дискриминантной поверхности $x^2 + y^2 - 1 = 0$. Эта поверхность и будет огибающей, так как шаровые поверхности не имеют особых точек; она представляет собой цилиндрическую поверхность, имеющую направляющей горизонтальную окружность радиуса 1, а осью — ось z.

Задача о нахождении огибающей семейства кривых имеет смысл лишь в том случае, если это семейство зависит только от *одного* параметра. Для семейства же *поверхностей* эта задача имеет смысл и в том случае, если это семейство $f(x, y, z, \alpha, \beta) = 0$ зависит от *двух* параметров α и β. Таково, например, семейство всех шаровых поверхностей радиуса 1, центры которых лежат в плоскости xOy. Уравнение этого семейства

$$(x - \alpha)^2 + (y - \beta)^2 + z^2 - 1 = 0.$$

Наглядно ясно, что обе плоскости $z = 1$ и $z = -1$ касаются в каждой своей точке одной из сфер семейства.

Вообще поверхность E называется огибающей двупараметрического семейства поверхностей, если она в каждой своей точке P касается одной из поверхностей семейства и если при этом выполняются следующие два условия: 1) когда точка P пробегает поверхность E, то значения параметров (α, β), соответствующие той поверхности семейства, которая касается огибающей E в точке P, должны пробегать некоторую область плоскости α, β, и 2) различным точкам (α, β) должны соответствовать различные точки поверхности E. Стало быть, теперь поверхность семейства касается огибающей, вообще говоря, в одной точке, а не вдоль целой кривой, как это было у однопараметрического семейства поверхностей.

При допущениях, подобных сделанным выше в теории огибающей плоских кривых, найдем, что точка касания поверхности семейства с огибающей, если последняя существует, должна удовлетворять системе уравнений

$$f(x, y, z, \alpha, \beta)=0, \qquad f_\alpha(x, y, z, \alpha, \beta)=0, \qquad f_\beta(x, y, z, \alpha, \beta)=0. \quad \text{(I)}$$

Если из этой системы уравнений удается выразить x, y, z через параметры α и β:

$$x=x(\alpha, \beta), \qquad y=y(\alpha, \beta), \qquad z=z(\alpha, \beta),$$

то получим параметрические уравнения так называемой *дискриминантной поверхности* семейства; исключив из уравнений (I) параметры α и β, получим уравнение этой поверхности в неявном виде $g(x, y, z)=0$. И здесь дискриминантная поверхность объединяет в себе огибающую и геометрическое место особых точек поверхностей семейства.

Если огибающая поверхность существует, то из трех уравнений (I) можно найти координаты точки касания огибающей с любой поверхностью семейства; надо только подставить в эти уравнения значения параметров α и β, соответствующие интересующей нас поверхности семейства.

Найдем, например, огибающую семейства всех шаровых поверхностей радиуса 1 с центрами в плоскости xOy:

$$f(x, y, z, \alpha, \beta) \equiv (x-\alpha)^2+(y-\beta)^2+z^2-1=0.$$

Это семейство с двумя параметрами α и β. Согласно (I), надо добавить еще два уравнения

$$f_\alpha \equiv -2(x-\alpha)=0, \qquad f_\beta \equiv -2(y-\beta)=0.$$

Стало быть, дискриминантное уравнение будет $z^2-1=0$, а так как шаровые поверхности не имеют особых точек, то наше семейство имеет огибающую, состоящую из двух плоскостей $z=1$ и $z=-1$. Этот результат мы предвидели выше из соображений наглядности.

Упражнения

1. Пусть $z=u(x, y)$ есть уравнение трубчатой поверхности, т. е. огибающей семейства сфер радиуса 1, центры которых лежат на кривой $y=f(x)$ плоскости xOy. Доказать, что

$$u^2(u_x^2+u_y^2+1)=1.$$

2. а) Найти огибающую двупараметрического семейства плоскостей, для которых $OP+OQ+OR=1$, где O — начало координат, а буквы P, Q, R обозначают точки пересечения плоскостей семейства с осями координат.

б) Найти огибающую семейства плоскостей, для которых

$$OP^2+OQ^2+OR^2=1.$$

3. Пусть C — произвольная кривая на плоскости. Рассмотрим окружности этой плоскости радиуса ρ, центры которых лежат на C. Доказать, что огибающая этого семейства окружностей состоит из двух кривых, параллельных

кривой C и отстоящих от нее на расстоянии p (определение параллельных кривых см. т. I, стр. 334, упр. 21).

4*. Семейство прямых в пространстве задано как сечение двух плоскостей, зависящих от параметра t:

$$a(t)x+b(t)y+c(t)z=1,$$
$$A(t)x+B(t)y+C(t)z=1.$$

Доказать, что если все эти прямые являются касательными некоторой кривой, другими словами, если это семейство прямых имеет огибающую, то

$$\begin{vmatrix} a-A & b-B & c-C \\ \dot{a} & \dot{b} & \dot{c} \\ \dot{A} & \dot{B} & \dot{C} \end{vmatrix}=0.$$

5*. Семейство плоскостей задано уравнением

$$x\cos t+y\sin t+z=t,$$

где t — параметр семейства.

а) Найти уравнение огибающей поверхности в цилиндрических координатах ρ, θ, z.

б) Доказать, что эта огибающая образована касательными к некоторой кривой.

6. Тело бросают вверх из заданного начального положения, с заданной начальной скоростью, в одной и той же вертикальной плоскости, но под различными углами к горизонту. Доказать, что траектории движения тела образуют семейство парабол и что огибающая этого семейства есть тоже парабола.

7*. Найти огибающую семейства шаровых поверхностей, касающихся трех данных сфер:

$$S_1:\ \left(x-\frac{3}{2}\right)^2+y^2+z^2=\frac{9}{4},$$
$$S_2:\ x^2+\left(y-\frac{3}{2}\right)^2+z^2=\frac{9}{4},$$
$$S_3:\ x^2+y^2+\left(z-\frac{3}{2}\right)^2=\frac{9}{4}.$$

8. Если плоская кривая C задана уравнениями $x=f(t)$, $y=g(t)$, то ее *взаимно полярной кривой* C' называется огибающая семейства прямых

$$\xi f(t)+\eta g(t)=1,$$

где (ξ, η) — текущие координаты на C'.

а) Доказать, что и кривая C является взаимно полярной для C'.

б) Найти взаимно полярную кривую окружности $(x-a)^2+(y-b)^2=1$.

в) Найти взаимно полярную кривую эллипса $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$.

§ 6. Максимумы и минимумы

1. Определение. Теория максимумов и минимумов является и у функций многих переменных одним из важнейших приложений дифференцирования.

Рассмотрим сначала функцию двух независимых переменных $u=f(x, y)$ и изобразим ее наглядно поверхностью в пространстве xyu.

Мы будем говорить, что величина u имеет максимум в точке (x_0, y_0), если все другие значения функции в некоторой окрестности этой точки меньше значения $u_0 = f(x_0, y_0)$. Геометрически такому максимуму соответствует вершина (верхушка) на поверхности. Аналогично мы будем говорить, что u имеет в точке (x_0, y_0) минимум, если все другие значения функции в некоторой окрестности этой точки больше значения $u_0 = f(x_0, y_0)$. Как и у функций одной переменной, эти понятия всегда относятся лишь к достаточно малой окрестности рассматриваемой точки. Весь кусок поверхности, изображающий функцию $u = f(x, y)$ в области ее задания, может, конечно, содержать точки, лежащие выше максимума или ниже минимума.

Дадим аналитическую формулировку определения максимума и минимума в общем виде для функции любого числа независимых переменных:

Функция $u = f(x, y, \ldots)$ имеет в точке $(x_0, y_0, \ldots)$ максимум (минимум), если она в некоторой окрестности этой точки принимает всюду значения меньшие (большие), чем в самой точке $(x_0, y_0, \ldots)$.

Если в окрестности точки $(x_0, y_0, \ldots)$ функция принимает значения, только не превышающие значения этой функции в самой точке $(x_0, y_0, \ldots)$, т. е. если $u(x, y, \ldots) \leqslant u(x_0, y_0, \ldots)$, то говорят, что функция имеет *несобственный* или *нестрогий максимум* в этой точке. Аналогично определяется и *несобственный минимум*.

Мы вновь подчеркиваем, что наше определение относится к надлежащим образом выбранной окрестности, окружающей рассматриваемую точку со всех сторон [т. е. оно носит локальный (местный) характер]. Мы уже знаем, что функция, непрерывная в замкнутой области, всегда принимает в некоторой точке области наибольшее, и в некоторой ее точке — наименьшее значение. Так вот, вполне может случиться, что значение максимума будет ниже наибольшего значения функции в замкнутой области, а значение минимума может оказаться выше наименьшего значения функции в той же области. Если функция принимает свое наибольшее (наименьшее) значение в точке P_0 границы области, то это значение не обязательно будет максимумом (минимумом) в определенном выше смысле; вспомним, что так именно обстоит дело и у функций одной переменной. Действительно, если функция определена *только* в замкнутой области, то *полной окрестности* точки P_0, в которой бы функция была определена, вообще *не* существует; если же функция определена в некоторой более обширной области, содержащей упомянутую замкнутую область, то в этой более обширной области функция может и не иметь максимума (минимума) в точке P_0. Это можно наблюдать на следующем примере. Возьмем функцию $u = -x - y$, определенную на всей плоскости xy, но будем ее рассматривать только на квадрате $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. В этой замкнутой области функция u принимает свое наибольшее значение нуль в начале координат, но это наибольшее значение не является максимумом. В самом деле, если рассматривать полную (всестороннюю) окрестность начала координат, то в ней функция $u = -x - y$ принимает и значения, большие нуля.

Однако практически важно отметить следующий факт: если функция принимает свое наибольшее (наименьшее) значение во *внутренней* точке области, то это значение обязательно будет максимумом (минимумом) функции в смысле нашего определения [собственным или несобственным].

2. Необходимые условия экстремума. Как и у функций одной переменной, мы наряду со словами *максимум* и *минимум* будем пользоваться термином *экстремум* или *экстремальное значение*, объединяющим эти два понятия. Сообщим теперь *необходимые условия* существования экстремума, т. е. такие условия, которые непременно должны быть выполнены в точке $(x_0, y_0, \ldots)$, если функция имеет в этой точке экстремум.

Для того чтобы дифференцируемая функция $u = f(x, y, z, \ldots)$ имела экстремум в точке $P_0(x_0, y_0, z_0, \ldots)$, необходимо, чтобы выполнялись следующие равенства:

$$\left.\begin{array}{l} f_x(x_0, y_0, z_0, \ldots) = 0, \\ f_y(x_0, y_0, z_0, \ldots) = 0, \\ f_z(x_0, y_0, z_0, \ldots) = 0, \\ \ldots\ldots\ldots\ldots \end{array}\right\} \qquad \text{(I)}$$

Эти условия легко вывести из уже известного необходимого условия экстремума дифференцируемой функции одной переменной. В самом деле, припишем, например, переменным $y, z, \ldots$ постоянные значения $y = y_0$, $z = z_0, \ldots$ и будем рассматривать нашу функцию в окрестности точки P_0, как функцию $f(x, y_0, z_0, \ldots)$, зависящую только от x. Тогда эта функция одной переменной x должна иметь экстремум при $x = x_0$, а необходимым условием такого экстремума является $f_x(x_0, y_0, z_0, \ldots) = 0$. Итак, необходимым условием существования в точке $P_0(x_0, y_0, z_0, \ldots)$ экстремума дифференцируемой функции $u = f(x, y, z, \ldots)$ является обращение в этой точке в нуль всех частных производных $f_x, f_y, f_z, \ldots$

В случае дифференцируемой функции двух переменных $u = f(x, y)$ это необходимое условие имеет простой геометрический смысл: функция может иметь в точке $P_0(x_0, y_0)$ экстремум лишь в том случае, если поверхность $u = f(x, y)$ имеет в P_0 касательную плоскость, параллельную плоскости xOy.

Для многих целей удобнее объединить формально наши условия (I) в одно условие, налагаемое на полный дифференциал функции:

$$df(x_0, y_0, z_0, \ldots) = f_x(x_0, y_0, z_0, \ldots)\,dx + f_y(x_0, y_0, z_0, \ldots)\,dy + \\ + f_z(x_0, y_0, z_0, \ldots)\,dz + \ldots = 0. \qquad \text{(II)}$$

Этот вид необходимого условия можно выразить словами так: *функция $u = f(x, y, z, \ldots)$ может иметь экстремум лишь в таких точках, в которых полный дифференциал (линейное приращение функции) обращается в нуль при любых значениях дифференциалов $dx, dy, dz, \ldots$ независимых переменных $x, y, z, \ldots$*

Не только из условий (I) вытекает условие (II), но и обратно, из (II) можно вывести уравнения (I). Для этого достаточно положить в (II) все независимые дифференциалы $dx, dy, dz, \ldots$, кроме одного, равными нулю.

В системе уравнений (I) число неизвестных x_0, y_0, z_0, ... равно числу уравнений. Стало быть, из этой системы можно, вообще говоря, вычислить координаты точек, в которых есть основание ожидать экстремума. Однако в любой полученной таким путем точке функция *не* обязательно имеет экстремум.

Рассмотрим, например, функцию $u = xy$. Условия (I) (их теперь два) дают сразу $x = 0$, $y = 0$. Однако в окрестности этой точки функция принимает как положительные, так и отрицательные значения, смотря по тому, в какой четверти берется точка. Стало быть, в точке (0, 0) функция экстремума не имеет. Изображающий эту функцию гиперболический параболоид $u = xy$ (рис. 28, стр. 135) имеет в начале координат так называемую *точку перевала* или *седловину*.

Полезно иметь специальное название для точек, удовлетворяющих уравнениям (I), независимо от того, дают ли они экстремум функции или нет. Всякую точку $(x_0, y_0, z_0, \ldots)$, в которой обращаются в нуль все частные производные f_x, f_y, f_z, ..., мы будем называть *стационарной точкой* функции $u = f(x, y, z, \ldots)$, а значение $f(x_0, y_0, z_0, \ldots)$ в такой точке — *стационарным значением* функции. В такой точке функция имеет, так сказать, замедленное изменение. Ясно, что точки экстремума надо искать среди стационарных точек функции.

Для решения вопроса, дает ли найденная стационарная точка экстремум функции или нет, требуется дополнительное исследование. Однако во многих практических задачах положение дел ясно с самого начала. В частности, если известно, что функция принимает свое наибольшее или наименьшее значение во всей области в некоторой ее *внутренней* точке, а система уравнений (I) определяет единственную стационарную точку $P_0(x_0, y_0, z_0, \ldots)$, то функция имеет в P_0 экстремум, а именно в первом случае — максимум, а во втором — минимум. Если же условие задачи не дает оснований для подобных соображений, то каждую стационарную точку надо отдельно исследовать. В дополнениях к этой главе будут выведены достаточные условия максимума и минимума для функции двух переменных.

3. Примеры. Поясним полученные результаты на нескольких примерах.

1) Частные производные функции $u = x^2 + y^2$ обращаются в нуль только при $x = y = 0$. Следовательно, имеется лишь одна стационарная точка — начало координат. В этой точке функция действительно имеет не только минимум, но и наименьшее значение, так как во всех точках, кроме начала координат, функция как сумма квадратов принимает лишь положительные значения.

2) Функция $u = \sqrt{1 - x^2 - y^2}$ $(x^2 + y^2 \leq 1)$ имеет частные производные $u_x = -\dfrac{x}{\sqrt{1 - x^2 - y^2}}$, $u_y = -\dfrac{y}{\sqrt{1 - x^2 - y^2}}$. Они обращаются в нуль только при $x = y = 0$. Единственная стационарная точка (начало координат) дает не только максимум, но и наибольшее значение функции в ее области задания, так как подкоренное выражение $1 - x^2 - y^2$ имеет, очевидно, во всех точках области, отличных от начала координат, меньшие значения, чем в самом

начале. Свое наименьшее значение наша функция принимает на окружности $x^2+y^2=1$, являющейся границей области.

3) Требуется построить треугольник, у которого произведение синусов всех его углов принимает наибольшее значение. Задача сводится к нахождению тех значений углов α и β, при которых функция

$$f(\alpha, \beta)=\sin\alpha\sin\beta\sin(\alpha+\beta)$$

достигает своего наибольшего значения в треугольной области $0\leqslant\alpha\leqslant\pi$, $0\leqslant\beta\leqslant\pi$, $0\leqslant\alpha+\beta\leqslant\pi$ (рис. 56). Внутри области функция положительна, а на ее границе $f(\alpha, \beta)=0$. Следовательно, наша функция принимает наименьшее значение (нуль) на границе области, а наибольшее значение (положительное) — в некоторой точке внутри треугольной области.

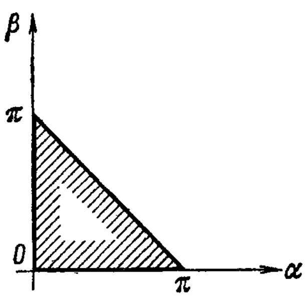


Рис. 56.

Приравнивая частные производные нулю, получаем следующую систему уравнений:

$$\cos\alpha\sin\beta\sin(\alpha+\beta)+\sin\alpha\sin\beta\cos(\alpha+\beta)=0,$$
$$\sin\alpha\cos\beta\sin(\alpha+\beta)+\sin\alpha\sin\beta\cos(\alpha+\beta)=0.$$

Из этих уравнений вытекает, что $\operatorname{tg}\beta=\operatorname{tg}\alpha$, а так как внутри области $0<\alpha<\pi$, $0<\beta<\pi$ и $0<\alpha+\beta<\pi$, то $\beta=\alpha$. Подставив в первое уравнение системы $\beta=\alpha$, получим соотношение $\sin 3\alpha=0$. Отсюда $\alpha=\frac{\pi}{3}$ и $\beta=\frac{\pi}{3}$; это единственная стационарная точка внутри области. Следовательно, искомый треугольник — равносторонний.

4) Три точки $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, $P_3(x_3, y_3)$ являются вершинами остроугольного треугольника. Требуется найти такую четвертую точку $P(x, y)$, чтобы ее расстояния от данных точек P_1, P_2 и P_3 имели наименьшую сумму. Сумма расстояний $P_1P+P_2P+P_3P$ является непрерывной функцией от координат x и y точки P, и эта функция обязательно имеет наименьшее значение в некоторой точке P_0 очень большого круга, содержащего внутри себя данный треугольник. Точка P_0 не может быть вершиной этого треугольника, так как основание перпендикуляра, опущенного из какой-либо другой вершины на противоположную сторону треугольника, имело бы заведомо меньшую сумму расстояний. Но P_0 не может лежать и на окружности круга, если точки этой окружности достаточно удалены от вершин треугольника. Итак, требуется найти наименьшее значение функции $f(x, y)=R_1+R_2+R_3$, где $R_k=P_kP=\sqrt{(x-x_k)^2+(y-y_k)^2}$, $k=1, 2, 3$. Эта функция дифференцируема везде, кроме точек P_1, P_2, P_3. Мы знаем, что в искомой точке P_0 должны обратиться в нуль частные производные f_x и f_y. Таким образом получается система уравнений

$$\frac{x-x_1}{R_1}+\frac{x-x_2}{R_2}+\frac{x-x_3}{R_3}=0,$$

$$\frac{y-y_1}{R_1}+\frac{y-y_2}{R_2}+\frac{y-y_3}{R_3}=0.$$

Из этих уравнений видно, что три единичных вектора плоскости xy

$$R_1^\circ=\left\{\frac{x-x_1}{R_1}, \frac{y-y_1}{R_1}\right\}, \quad R_2^\circ=\left\{\frac{x-x_2}{R_2}, \frac{y-y_2}{R_2}\right\}, \quad R_3^\circ=\left\{\frac{x-x_3}{R_3}, \frac{y-y_3}{R_3}\right\}$$

имеют (векторную) сумму, равную нулю, так что при построении их суммы получается равносторонний треугольник (рис. 57). Отсюда вытекает, что направление каждого из этих векторов переходит в направление следующего при повороте на угол $2\pi/3$. Так как наши три единичных вектора имеют направления векторов $\overline{P_1P_0}$, $\overline{P_2P_0}$ и $\overline{P_3P_0}$, то каждая сторона данного треугольника должна быть видна из точки P_0 под одним и тем же углом $2\pi/3$.

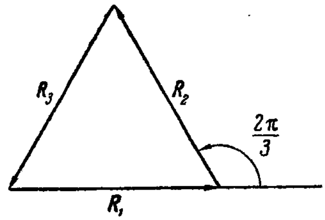


Рис. 57.

[Может случиться, что частные производные функции $u=f(x, y, \ldots)$ в отдельных точках обращаются в бесконечность или вообще не существуют. Следующие ниже примеры показывают, что в таких точках тоже может быть экстремум; поэтому каждую такую точку, как и каждую стационарную точку, надо проверить, нет ли в ней экстремума.

5) Функция $u=\sqrt[3]{x^2+y^2}$ имеет частные производные $u_x=\frac{2x}{3\sqrt[3]{(x^2+y^2)^2}}$, $u_y=\frac{2y}{3\sqrt[3]{(x^2+y^2)^2}}$. Единственная «подозрительная» на экстремум точка — начало координат, где выражения для u_x и u_y становятся неопределенными. Поэтому вычисляем отдельно частные производные в точке (0, 0):

$$u_x(0, 0)=\lim_{x\to 0}\frac{\sqrt[3]{x^2}-0}{x-0}=\infty, \quad u_y(0, 0)=\lim_{y\to 0}\frac{\sqrt[3]{y^2}-0}{y-0}=\infty.$$

Вместе с тем ясно, что в начале координат функция имеет минимум, так как $u(0, 0)=0$, а во всех других точках $u>0$.

Нетрудно убедиться, что поверхность $u=\sqrt[3]{x^2+y^2}$ образована вращением вокруг оси u полукубической параболы $u=\sqrt[3]{x^2}$. В начале координат поверхность имеет единственную касательную прямую — ось u. Такой минимум (и аналогичный максимум) естественно называть *острым*, а экстремум, имеющий место в стационарной точке — *гладким*.

Читатель легко сам обнаружит, что функция $u=-\sqrt[3]{\frac{x^2}{a^2}+\frac{y^2}{b^2}}$ имеет единственный экстремум в начале координат (острый максимум).

6) Исследуем на экстремум функцию $u=\sqrt{x^2+y^2}$. Частные производные нигде не обращаются в нуль, а в начале координат их выражения становятся неопределенными. Отдельное вычисление дает

$$\lim_{x\to 0}\frac{u(x, 0)-u(0, 0)}{x-0}=\lim_{x\to 0}\frac{\sqrt{x^2}-0}{x-0}=\begin{cases}-1 \text{ при } x\to -0,\\ +1 \text{ при } x\to +0\end{cases}$$

и аналогично

$$\lim_{y\to 0}\frac{u(0, y)-u(0, 0)}{y-0}=\begin{cases}-1 \text{ при } y\to -0,\\ +1 \text{ при } y\to +0.\end{cases}$$

Итак, в начале координат частные производные не существуют. Однако наша функция имеет в этой точке очевидный минимум. Изображающая ее поверхность — верхняя полость кругового конуса $x^2+y^2-u^2=0$, имеющего осью ось u.

7) Так же точно решается задача об экстремумах функции $u = e^{\sqrt{x^2+y^2}}$. Стационарных точек нет. В начале координат частные производные не существуют (как и в предыдущем примере, здесь тоже требуется отдельное вычисление). Однако ясно, что эта функция имеет при $x = y = 0$ минимум $u = 1$. Геометрическое изображение — поверхность, образуемая вращением вокруг оси u дуги кривой $u = e^x$, $x \geqslant 0$, плоскости xu. Ясно, что в своей точке $(0, 0, 1)$ поверхность имеет, вместо касательной плоскости, касательный конус (круговой) с вершиной в этой точке. Такую особую точку поверхности можно назвать *конической* точкой.

Читатель сам найдет, что функция $u = -e^{\sqrt{\frac{x^2}{a^2}+\frac{y^2}{b^2}}}$ имеет при $x = y = 0$ максимум $u(0, 0) = -1$. Точка $(0, 0, -1)$ соответствующей поверхности тоже коническая точка, но касательный конус уже не круговой.

8) Найдем наибольшее и наименьшее значения функции $u = \sin\sqrt{x^2+y^2}$ в замкнутой области, ограниченной окружностью $x^2 + y^2 = \pi^2$, а также экстремумы функции, лежащие в этой области. Наглядное рассмотрение сразу дает ответ: функция принимает свое наименьшее значение нуль в начале координат и во всех точках граничной окружности, а наибольшее значение, единицу, на окружности $x^2 + y^2 = \frac{\pi^2}{4}$. Вычисление показывает, что все точки этой последней окружности являются стационарными; в каждой из этих точек функция имеет несобственный максимум. В начале координат частные производные не существуют; однако функция имеет в этой точке минимум. Изображающая функцию поверхность образована вращением вокруг оси u дуги синусоиды $u = \sin x$, $0 \leqslant x \leqslant \pi$, и имеет в начале координат коническую точку.]

4. Условные экстремумы. Задача на отыскание экстремумов функций многих переменных часто возникает в форме, отличной от изложенной выше. Пусть, например, требуется найти на данной поверхности $\varphi(x, y, z) = 0$ точку, ближайшую к началу координат. Для решения этой задачи придется определить минимумы функции

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2},$$

в которой, однако, координаты x, y, z уже не являются все независимыми переменными, а связаны между собой дополнительным условием: они должны удовлетворять уравнению поверхности $\varphi(x, y, z) = 0$.

Экстремумы функции $u = f(x, y, \ldots)$, аргументы которой удовлетворяют еще некоторым добавочным уравнениям (число которых, конечно, меньше числа аргументов), называются *условными* или *относительными* экстремумами. В отличие от них, рассмотренные выше экстремумы (без дополнительных условий) называются *свободными* или *безусловными* экстремумами. Задача отыскания условных максимумов и минимумов не является, правда, принципиально новой проблемой. Так, в примере, с которого мы начали, надо только выразить из уравнения поверхности одну из переменных, например z, через остальные, и подставить полученное выражение в исследуемую функцию $f = \sqrt{x^2 + y^2 + z^2}$; тогда мы придем к обычной задаче на экстремум функции двух независимых переменных x и y. Но при таком решении переменные x, y, z перестают быть равноправными, двум из них (x и y) отдается предпочтение перед z.

Представляется более удобным и более изящным найти симметричную запись необходимого условия относительного экстремума, в которой ни одна из переменных не пользуется преимуществом перед другими.

Рассмотрим сначала более простую, но типичную задачу об условном экстремуме функции *двух* переменных: *требуется найти точки экстремума функции $f(x, y)$, если аргументы x и y не являются независимыми переменными, а связаны добавочным условием* $\varphi(x, y)=0$.

Аналитическому решению мы предпошлем геометрическое рассмотрение. Допустим, что уравнение $\varphi(x, y)=0$ изображается кривой без особых точек (рис. 58) и что семейство кривых $f(x, y)=c$ покрывает часть плоскости, как показано на том же рисунке. Задача состоит в том, чтобы среди кривых этого семейства, пересекающих кривую $\varphi(x, y)=0$, найти ту, для которой постоянная c имеет максимум или минимум. Двигаясь вдоль кривой $\varphi=0$, мы пересекаем последовательно кривые семейства $f(x, y)=c$, и при этом параметр c семейства изменяется, вообще говоря, монотонно.

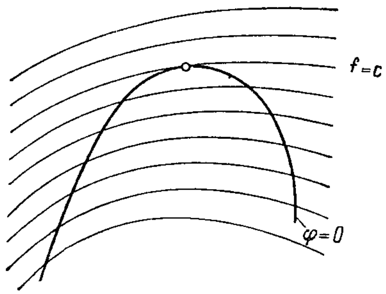


Рис. 58.

Экстремума параметра c следует ожидать в такой точке кривой $\varphi=0$, при прохождении которой изменяется направление пробега шкалы c. Из рис. 58 видно, что это может произойти в такой точке кривой $\varphi=0$, в которой она как раз касается кривой семейства, соответствующей экстремальному значению c. Координаты точки касания и будут искомыми значениями $x=\xi$, $y=\eta$, которым соответствует экстремум функции $f(x, y)$. Но касание кривых $\varphi=0$ и $f=c$ означает, что эти линии имеют в точке (ξ, η) общую касательную; следовательно, в точке касания нормальные векторы обеих кривых $\{f_x, f_y\}$ и $\{\varphi_x, \varphi_y\}$ коллинеарны, откуда

$$f_x : \varphi_x = f_y : \varphi_y = -\lambda,$$

где через $(-\lambda)$ обозначено общее значение равных отношений. Стало быть, координаты (ξ, η) точки касания удовлетворяют системе трех уравнений

$$f_x + \lambda\varphi_x = 0, \qquad f_y + \lambda\varphi_y = 0, \qquad \varphi(x, y) = 0,$$

из которой надо определить координаты (ξ, η) точки касания и коэффициент пропорциональности λ. Это и будут необходимые условия относительного экстремума.

Мы не зря предположили, что кривая $\varphi(x, y)=0$ не имеет особых точек. Если отбросить это предположение, то может случиться, что точка (ξ, η), в которой сливаются две точки пересечения кривой $\varphi=0$ с кривой семейства $f=c$, является особой точкой линии $\varphi=0$, и тогда наши рассуждения неприменимы. Но в этом случае $\varphi_x(\xi, \eta)=0$ и $\varphi_y(\xi, \eta)=0$. Найдя особые точки линии $\varphi=0$, их можно отдельно исследовать. На рис. 59 показан случай, когда кривая $\varphi=0$ имеет особую точку — точку заострения, а экстремум существует именно в этой точке.

Итак, мы пришли наглядным путем к следующей теореме:

Для того чтобы функция $f(x, y)$ имела в точке $x=\xi$, $y=\eta$ экстремум, при условии $\varphi(x, y)=0$, необходимо, чтобы существовал такой множитель λ, что три числа ξ, η, λ удовлетворяют системе трех уравнений

$$f_x(\xi, \eta)+\lambda\varphi_x(\xi, \eta)=0,$$
$$f_y(\xi, \eta)+\lambda\varphi_y(\xi, \eta)=0,$$
$$\varphi(\xi, \eta)=0.$$

При этом предполагается, что частные производные $\varphi_x(\xi, \eta)$ и $\varphi_y(\xi, \eta)$ не обращаются одновременно в нуль, так что точка (ξ, η) заведомо не является особой точкой кривой $\varphi(x, y)=0$.

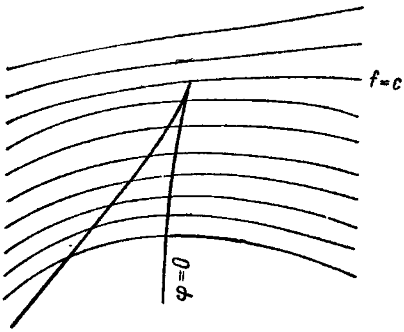


Рис. 59.

Доказательство этой теоремы будет дано в n° 5.

Это необходимое условие дает для определения ξ, η и λ столько же уравнений, сколько имеется неизвестных. Прибавилось еще одно неизвестное λ, но зато мы выиграли в том отношении, что получили вполне симметричную формулировку задачи. В итоге мы имеем следующее удобное практическое правило для отыскания условных экстремумов:

Для того чтобы найти экстремумы функции $f(x, y)$, подчиненные дополнительному условию $\varphi(x, y)=0$, надо сложить исследуемую функцию $f(x, y)$ с произведением функции $\varphi(x, y)$ на неизвестный множитель λ, не зависящий от x и y, а затем написать известные необходимые условия экстремума для вспомогательной функции $F=f(x, y)+\lambda\varphi(x, y)$, а именно:

$$f_x+\lambda\varphi_x=0, \qquad f_y+\lambda\varphi_y=0.$$

Эти уравнения вместе с дополнительным условием $\varphi(x, y)=0$ служат для определения координат точек, «подозрительных» на экстремум, и множителя λ. Множитель λ называется *множителем Лагранжа*, а сформулированное только что правило называется *правилом* или *методом неопределенных множителей Лагранжа.*

Прежде чем перейти к доказательству правила неопределенных множителей, покажем его применение на простом примере. Требуется найти экстремумы функции $u=xy$ на окружности радиуса 1 с центром в начале координат, т. е. при дополнительном условии

$$x^2+y^2-1=0.$$

Согласно правилу Лагранжа, строим вспомогательную функцию

$$F(x, y)=xy+\lambda(x^2+y^2-1)$$

и пишем для нее необходимые условия экстремума:

$$F_x\equiv y+2\lambda x=0, \qquad F_y\equiv x+2\lambda y=0;$$

вместе с дополнительным условием $x^2+y^2-1=0$ они составляют систему уравнений для определения неизвестных x, y, λ.

Решив эту систему, получим четыре стационарные точки:

$$\xi=\frac{1}{2}\sqrt{2}, \qquad \eta=\frac{1}{2}\sqrt{2},$$

$$\xi=-\frac{1}{2}\sqrt{2}, \quad \eta=-\frac{1}{2}\sqrt{2},$$

$$\xi=\frac{1}{2}\sqrt{2}, \qquad \eta=-\frac{1}{2}\sqrt{2},$$

$$\xi=-\frac{1}{2}\sqrt{2}, \quad \eta=\frac{1}{2}\sqrt{2}.$$

В первых двух точках функция $u=xy$ принимает значение $u=\frac{1}{2}$, в последних двух точках — значение $u=-\frac{1}{2}$.

Наша функция обязана принять в некоторых точках окружности свое наибольшее и свое наименьшее значение, а так как окружность, будучи замкнутой линией, не имеет граничных точек, то эти крайние значения функция может принять только в своих стационарных точках. Отсюда вытекает, что в первых двух найденных точках функция имеет максимум и вместе с тем свое наибольшее значение $u=\frac{1}{2}$ на окружности, а в последних двух точках — минимум и наименьшее значение $u=-\frac{1}{2}$.

5. Доказательство правила неопределенных множителей для условного экстремума функции двух переменных. Аналитическое доказательство правила Лагранжа мы получим естественным путем, приведя задачу к уже знакомому случаю свободных экстремумов. Допустим, что в точке экстремума обе частные производные $\varphi_x(\xi, \eta)$ и $\varphi_y(\xi, \eta)$ не обращаются одновременно в нуль; пусть, для определенности, $\varphi_y(\xi, \eta)\neq 0$. Тогда на основании § 1, n° 3 уравнение $\varphi(x, y)=0$ определяет в окрестности этой точки однозначным образом величину $y=g(x)$ как непрерывно дифференцируемую функцию от x. Подставим это выражение в функцию $f(x, y)$; тогда функция

$$f[x, g(x)]=u(x)$$

должна иметь при $x=\xi$ свободный экстремум, а для этого должно выполняться при $x=\xi$ необходимое условие

$$u'(x)\equiv f_x+f_y g'(x)=0.$$

Кроме того, для функции $y=g(x)$, определенной неявно уравнением $\varphi(x, y)=0$, удовлетворяется тождественно соотношение

$$\varphi_x+\varphi_y g'(x)=0.$$

Из последних двух уравнений исключим $g'(x)$ следующим способом: сложим первое из них со вторым уравнением, помноженным на $\lambda=-\frac{f_y}{\varphi_y}$, в результате чего получится уравнение

$$f_x+\lambda\varphi_x=0.$$

С другой стороны, из определения множителя λ вытекает уравнение

$$f_y+\lambda\varphi_y=0.$$

Тем самым правило неопределенных множителей доказано.

Из хода доказательства ясно, насколько существенно предположение, что в точке (ξ, η) не обращаются одновременно в нуль обе частные производные φ_x и φ_y. Да это видно и из самого правила, которое теряет силу в этом случае. Действительно, если функция $f(x, y)$ имеет в точке (ξ, η) экстремум при условии $\varphi(x, y)=0$, причем $\varphi_x(\xi, \eta)=\varphi_y(\xi, \eta)=0$, то уравнения $f_x+\lambda\varphi_x=0$ и $f_y+\lambda\varphi_y=0$ принимают вид $f_x=0$ и $f_y=0$, которым точка (ξ, η) отнюдь не обязана удовлетворять.

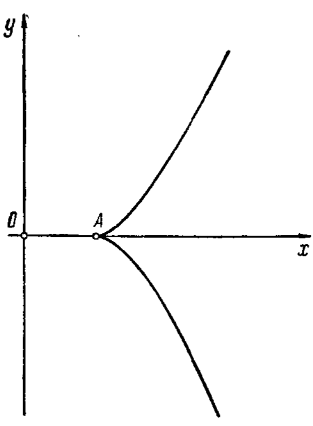


Рис. 60.

Поясним это на следующей задаче. Требуется найти на кривой $(x-1)^3-y^2=0$ точку, наиболее близкую к началу координат. Для ее решения надо определить точки минимума функции

$$f(x, y)=x^2+y^2,$$

подчиненные условию

$$\varphi(x, y)\equiv(x-1)^3-y^2=0.$$

Из рис. 60 видно, что ближайшей к началу точкой этой полукубической параболы является ее точка заострения A (1; 0) (легко доказать, что окружность радиуса 1 с центром в начале координат не имеет других общих точек с нашей кривой). Координаты точки A, т. е. $\xi=1$, $\eta=0$, удовлетворяют, очевидно, уравнению $f_y+\lambda\varphi_y=0$ и даже при любом значении λ, но

$$f_x(\xi, \eta)+\lambda\varphi_x(\xi, \eta)=2\xi+3\lambda(\xi-1)^2=2\neq 0.$$

Доказательству правила неопределенных множителей можно придать несколько иной вид, который особенно удобен для обобщения, Этот новый вид доказательства основан на другой формулировке необходимого условия экстремума. Мы знаем, что необходимым условием наличия в некоторой точке экстремума функции одной или многих переменных является обращение в нуль дифференциала функции в этой точке. И в нынешней проблеме условного экстремума можно дать соответствующей теореме следующую аналогичную формулировку:

Для того чтобы функция $f(x, y)$ имела в точке (ξ, η) экстремум, подчиненный условию $\varphi(x, y)=0$, необходимо, чтобы дифференциал df обращался в нуль в этой точке, когда дифференциалы dx и dy уже не являются независимыми друг от друга, а связаны соотношением

$$d\varphi \equiv \varphi_x\,dx + \varphi_y\,dy = 0,$$

вытекающим из условия $\varphi(x, y)=0$. Таким образом, если числа dx и dy удовлетворяют условию $d\varphi=0$, то они в точке (ξ, η) экстремума непременно должны удовлетворять уравнению

$$df \equiv f_x(\xi, \eta)\,dx + f_y(\xi, \eta)\,dy = 0.$$

Сложим это уравнение с предыдущим соотношением, помноженным на неопределенный пока множитель λ; в итоге получится

$$(f_x + \lambda\varphi_x)\,dx + (f_y + \lambda\varphi_y)\,dy = 0.$$

А теперь выберем λ так, чтобы было

$$f_y + \lambda\varphi_y = 0$$

(это выполнимо в силу предположения, что $\varphi_y \neq 0$); тогда обязательно будет $(f_x + \lambda\varphi_x)\,dx = 0$. Но дифференциал dx можно выбрать произвольно, например равным 1, и следовательно, также

$$f_x + \lambda\varphi_x = 0.$$

Итак, новый вывод правила Лагранжа завершен.

6. Обобщение метода неопределенных множителей. Правило неопределенных множителей можно обобщить на любое число переменных и на любое (но, конечно, меньшее) число дополнительных условий. Мы здесь рассмотрим только частный случай, но он содержит в себе все существенные черты общей проблемы. Пусть требуется определить экстремумы функции

$$u = f(x, y, z, t)$$

при том условии, что переменные x, y, z, t удовлетворяют двум уравнениям

$$\varphi(x, y, z, t) = 0, \qquad \psi(x, y, z, t) = 0.$$

(Эти условные уравнения называются также *уравнениями связей.*) Допустим, что в точке $P(\xi, \eta, \zeta, \tau)$ функция u действительно имеет определенного вида экстремум по сравнению со значениями функции во всех соседних точках, удовлетворяющих двум уравнениям связей. Допустим также, что в точке P якобиан

$$\frac{\partial(\varphi, \psi)}{\partial(z, t)} = \varphi_z\psi_t - \varphi_t\psi_z$$

не обращается в нуль. В силу этого последнего допущения (см. § 3, n° 7), переменные z и t можно выразить из двух условных уравнений в виде дифференцируемых функций от двух других переменных x и y: $z = g(x, y)$ и $t = h(x, y)$. Подставим эти два выражения вместо z и t в исследуемую функцию $u = f(x, y, z, t)$. Тогда $f(x, y, z, t)$ превратится в функцию двух *независимых* переменных x и y, и эта функция должна иметь в точке $x = \xi$, $y = \eta$ свободный экстремум. Стало быть, в этой точке должны обратиться в нуль обе частные производные этой функции:

$$\left.\begin{aligned} f_x + f_z \frac{\partial z}{\partial x} + f_t \frac{\partial t}{\partial x} = 0, \\ f_y + f_z \frac{\partial z}{\partial y} + f_t \frac{\partial t}{\partial y} = 0. \end{aligned}\right\} \qquad (1)$$

Участвующие в этих уравнениях четыре производные $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, $\frac{\partial t}{\partial x}$ и $\frac{\partial t}{\partial y}$ надо выразить из уравнений связей, дифференцированием которых можно получить следующие две системы уравнений:

$$\varphi_x + \varphi_z \frac{\partial z}{\partial x} + \varphi_t \frac{\partial t}{\partial x} = 0,$$
$$\psi_x + \psi_z \frac{\partial z}{\partial x} + \psi_t \frac{\partial t}{\partial x} = 0$$

$$\varphi_y + \varphi_z \frac{\partial z}{\partial y} + \varphi_t \frac{\partial t}{\partial y} = 0,$$
$$\psi_y + \psi_z \frac{\partial z}{\partial y} + \psi_t \frac{\partial t}{\partial y} = 0.$$

Общим определителем этих двух систем уравнений является якобиан $\frac{\partial(\varphi, \psi)}{\partial(z, t)}$, который, согласно допущению, не обращается в нуль. Поэтому из последних двух систем уравнений можно определить искомые величины $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, $\frac{\partial t}{\partial x}$ и $\frac{\partial t}{\partial y}$. Подставив их выражения в уравнения (1), мы получили бы вместе с уравнениями $\varphi = 0$ и $\psi = 0$ систему четырех уравнений с четырьмя неизвестными x, y, z, t для определения стационарных точек, и задача была бы решена.

Однако, вместо этого пути, мы для достижения формальной симметрии и, пожалуй, большей ясности предпочитаем действовать следующим образом. Найдем такие два числа λ и μ, чтобы в точке экстремума удовлетворялись два уравнения

$$f_z + \lambda\varphi_z + \mu\psi_z = 0,$$
$$f_t + \lambda\varphi_t + \mu\psi_t = 0.$$

Определение этих двух чисел λ и μ всегда возможно, в силу допущения, что якобиан $\frac{\partial(\varphi, \psi)}{\partial(z, t)}$ не обращается в нуль.

Помножим теперь оба уравнения

$$\varphi_x + \varphi_z \frac{\partial z}{\partial x} + \varphi_t \frac{\partial t}{\partial x} = 0 \quad \text{и} \quad \psi_x + \psi_z \frac{\partial z}{\partial x} + \psi_t \frac{\partial t}{\partial x} = 0$$

на λ и μ и сложим с уравнением

$$f_x + f_z \frac{\partial z}{\partial x} + f_t \frac{\partial t}{\partial x} = 0,$$

в результате чего получится

$$f_x + \lambda\varphi_x + \mu\psi_x + (f_z + \lambda\varphi_z + \mu\psi_z) \frac{\partial z}{\partial x} + (f_t + \lambda\varphi_t + \mu\psi_t) \frac{\partial t}{\partial x} = 0,$$

а отсюда, в силу определения λ и μ, вытекает

$$f_x + \lambda\varphi_x + \mu\psi_x = 0.$$

Действуя аналогично, сложим уравнение

$$f_y + f_z \frac{\partial z}{\partial y} + f_t \frac{\partial t}{\partial y} = 0$$

с уравнениями

$$\varphi_y + \varphi_z \frac{\partial z}{\partial y} + \varphi_t \frac{\partial t}{\partial y} = 0 \quad \text{и} \quad \psi_y + \psi_z \frac{\partial z}{\partial y} + \psi_t \frac{\partial t}{\partial y} = 0,$$

помноженными предварительно на λ и μ соответственно, и получим

$$f_y + \lambda\varphi_y + \mu\psi_y + (f_z + \lambda\varphi_z + \mu\psi_z) \frac{\partial z}{\partial y} + (f_t + \lambda\varphi_t + \mu\psi_t) \frac{\partial t}{\partial y} = 0.$$

Отсюда, в силу определения λ и μ, имеем

$$f_y + \lambda\varphi_y + \mu\psi_y = 0.$$

Итак, мы пришли к следующему результату:

Если функция $f(x, y, z, t)$ имеет экстремум в точке $P(\xi, \eta, \zeta, \tau)$, подчиненной уравнениям связей

$$\varphi(x, y, z, t) = 0 \quad \textit{и} \quad \psi(x, y, z, t) = 0,$$

причем в этой точке якобиан $\frac{\partial(\varphi, \psi)}{\partial(z, t)} \neq 0$, то существуют такие *два числа λ и μ, что в точке P, кроме двух уравнений связей, удовлетворяются еще и следующие четыре уравнения:*

$$f_x+\lambda\varphi_x+\mu\psi_x=0,$$
$$f_y+\lambda\varphi_y+\mu\psi_y=0,$$
$$f_z+\lambda\varphi_z+\mu\psi_z=0,$$
$$f_t+\lambda\varphi_t+\mu\psi_t=0.$$

Последние четыре уравнения совершенно симметричны. Переменные x и y утратили свое привилегированное положение, и мы пришли бы к этим же четырем уравнениям, если бы исходили из допущения, что не обращается в нуль какой-либо один (безразлично какой) из якобианов $\frac{\partial(\varphi,\psi)}{\partial(x,y)}, \frac{\partial(\varphi,\psi)}{\partial(x,z)}, \ldots, \frac{\partial(\varphi,\psi)}{\partial(z,t)}$, а не обязательно последний из них. Из этого допущения вытекало бы, что какая-то пара из четырех величин x, y, z и t (возможно, вовсе не z и t) может быть выражена через оставшуюся пару в окрестности точки P. Правда, эта симметрия достигнута ценой введения еще двух неизвестных, множителей Лагранжа λ и μ, в дополнение к неизвестным ξ, η, ζ, τ. Вместо четырех неизвестных мы имеем теперь шесть, и для их определения располагаем таким же числом уравнений.

И здесь можно провести доказательство более изящно, пользуясь аппаратом дифференциалов. Тогда необходимое условие наличия в точке P экстремума надо записать в таком виде:

$$df(x, y, z, t)\equiv f_x dx+f_y dy+f_z dz+f_t\,dt=0.$$

Дифференциалы dx, dy, dz, dt, участвующие в этом уравнении, связаны двумя соотношениями

$$d\varphi\equiv\varphi_x dx+\varphi_y dy+\varphi_z dz+\varphi_t dt=0,$$
$$d\psi\equiv\psi_x dx+\psi_y dy+\psi_z dz+\psi_t dt=0,$$

которые получены дифференцированием уравнений связей $\varphi=0$ и $\psi=0$. Допустим, что *не* все определители второго порядка, составленные из коэффициентов при dx, dy, dz, dt в этих уравнениях, обращаются в нуль в точке $P(\xi, \eta, \zeta, \tau)$. Пусть, например, якобиан $\frac{\partial(\varphi,\psi)}{\partial(z,t)}\neq 0$ в P. Тогда существуют такие два числа λ и μ, которые удовлетворяют системе уравнений

$$f_z+\lambda\varphi_z+\mu\psi_z=0,$$
$$f_t+\lambda\varphi_t+\mu\psi_t=0.$$

Помножим теперь уравнение $d\varphi=0$ на λ, уравнение $d\psi=0$ на μ и полученные два уравнения сложим с уравнением $df=0$; в итоге получится

$$(f_x+\lambda\varphi_x+\mu\psi_x)\,dx+(f_y+\lambda\varphi_y+\mu\psi_y)\,dy+(f_z+\lambda\varphi_z+\mu\psi_z)\,dz+$$
$$+(f_t+\lambda\varphi_t+\mu\psi_t)\,dt=0\cdot$$

В силу равенств, определяющих λ и μ, это уравнение принимает следующий вид:

$$(f_x+\lambda\varphi_x+\mu\psi_x)\,dx+(f_y+\lambda\varphi_y+\mu\psi_y)\,dy=0.$$

Так как dx и dy являются здесь независимыми дифференциалами (т. е. произвольными числами), то отсюда вытекает, что числа λ и μ удовлетворяют еще и следующим двум уравнениям:

$$f_x+\lambda\varphi_x+\mu\psi_x=0,$$
$$f_y+\lambda\varphi_y+\mu\psi_y=0.$$

Стало быть, мы вновь пришли к правилу неопределенных множителей:

При любом числе переменных и любом (но, конечно, меньшем) числе условных уравнений правило неопределенных множителей формулируется совершенно аналогично и доказывается тем же путем. Приводим это правило в его общем виде.

Пусть требуется найти экстремумы функции

$$u=f(x_1,\ x_2,\ \ldots,\ x_n),$$

причем ее n аргументов не являются независимыми, а подчинены m условиям связей ($m<n$):

$$\varphi_1(x_1,\ x_2,\ \ldots,\ x_n)=0,$$
$$\varphi_2(x_1,\ x_2,\ \ldots,\ x_n)=0,$$
$$\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots$$
$$\varphi_m(x_1,\ x_2,\ \ldots,\ x_n)=0.$$

Тогда надо ввести m неопределенных множителей $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_m$ и построить вспомогательную функцию

$$F=f+\lambda_1\varphi_1+\lambda_2\varphi_2+\ldots+\lambda_m\varphi_m,$$

для которой следует записать n необходимых условий свободного экстремума

$$\frac{\partial F}{\partial x_1}=0,\quad \frac{\partial F}{\partial x_2}=0,\ \ldots,\ \frac{\partial F}{\partial x_n}=0$$

(при вычислении частных производных считать $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_m$ постоянными). Полученные n уравнений вместе с m данными уравнениями связей

$$\varphi_1=0,\ \varphi_2=0,\ \ldots,\ \varphi_m=0$$

образуют систему $m+n$ уравнений с $m+n$ неизвестными $x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m$. Этим уравнениям обязательно удовлетворяют координаты любой точки экстремума данной функции f, если только не имеет места тот исключительный случай, когда в этой точке обращаются в нуль все якобианы m функций $\varphi_1, \varphi_2, \ldots, \varphi_m$ по m из числа переменных $x_1, x_2, \ldots, x_n$.

По поводу этого метода неопределенных множителей важно подчеркнуть следующее. Сформулированное выше правило дает лишь формальный способ для определения точек, в которых только и возможны экстремумы функции; стало быть, оно дает лишь *необходимое* условие. Естественно возникает вопрос, дают ли точки, найденные с помощью правила множителей, действительно максимум или минимум функции и при каких обстоятельствах, т. е. вопрос о достаточных условиях максимума и минимума. Этого вопроса мы здесь и не затрагиваем, его исследование завело бы нас слишком далеко. Как и в случае свободного экстремума, в практических приложениях обыкновенно заранее известно, что экстремум *существует*. Если при этом правило множителей определяет точку P однозначно и в этой точке не возникает исключительный случай (равенства нулю всех якобианов), то нет сомнения, что найдена именно точка экстремума.

7. Примеры. Поясним практику пользования методом неопределенных множителей на нескольких примерах.

1) Найдем наибольшее значение функции $f(x, y, z) = x^2y^2z^2$ на сфере $x^2+y^2+z^2=c^2$. На этой поверхности функция должна принять свое наибольшее значение, а так как сфера не имеет краевых точек, то это наибольшее значение должно быть максимумом. Стало быть, надо найти максимумы функции $f=x^2y^2z^2$ при дополнительном условии $x^2+y^2+z^2=c^2$. Строим вспомогательную функцию

$$F = x^2y^2z^2 + \lambda(x^2+y^2+z^2-c^2)$$

и ее частные производные по x, y, z приравниваем нулю:

$$2xy^2z^2 + 2\lambda x = 0,$$
$$2x^2yz^2 + 2\lambda y = 0,$$
$$2x^2y^2z + 2\lambda z = 0.$$

Те решения, в которых x, либо y, либо z равен нулю, можно опустить, так как в этих точках функция f принимает свое наименьшее значение — нуль. Остальные решения удовлетворяют уравнениям

$$x^2 = y^2 = z^2, \quad \lambda = -x^4 \quad \text{и} \quad x^2+y^2+z^2=c^2,$$

откуда легко находим искомые значения координат

$$x = \pm\frac{c}{\sqrt{3}}, \qquad y = \pm\frac{c}{\sqrt{3}}, \qquad z = \pm\frac{c}{\sqrt{3}}.$$

Во всех этих восьми точках функция принимает одно и то же значение $\frac{c^6}{27}$, которое и является поэтому ее наибольшим значением на заданной шаровой поверхности.

Отсюда, кстати, вытекает, что для любой тройки чисел (x, y, z) выполняется соотношение

$$\sqrt[3]{x^2y^2z^2} \leqslant \frac{c^2}{3} = \frac{x^2+y^2+z^2}{3},$$

т. е. среднее геометрическое трех положительных чисел x^2, y^2, z^2 никогда не превосходит их среднего арифметического.

Справедлива и общая теорема, что среднее геометрическое любого количества n положительных чисел не превышает их среднего арифметического. Ее можно доказать аналогичным путем, в результате решения задачи об определении наибольшего значения функции $f = x_1^2x_2^2 \ldots x_n^2$ при дополнительном условии

$$x_1^2 + x_2^2 + \ldots + x_n^2 = c^2.$$

По поводу другого доказательства этой теоремы см. т. I, стр. 197, упр. 18.

2) Требуется найти треугольник (со сторонами x, y, z), который при заданном периметре $2p$ имел бы наибольшую площадь.

По формуле Герона, квадрат площади треугольника выражается следующей функцией его сторон:

$$f(x, y, z) = p(p-x)(p-y)(p-z).$$

Стало быть, требуется найти наибольшее значение этой функции при дополнительном условии

$$\varphi(x, y, z) \equiv x+y+z-2p = 0;$$

к тому же переменные x, y, z ограничены неравенствами

$$x \geqslant 0, \quad y \geqslant 0, \quad z \geqslant 0, \quad x+y \geqslant z, \quad x+z \geqslant y, \quad y+z \geqslant x.$$

Эти неравенства определяют замкнутую область изменения переменных x, y, z. На границе этой области, т. е. всякий раз, как одно из неравенств превращается в равенство, функция f везде принимает значение нуль, которое, очевидно, является ее наименьшим значением. Следовательно, свое наибольшее значение функция принимает внутри области, и именно в точке локального максимума. Строим вспомогательную функцию

$$F(x, y, z) = p(p-x)(p-y)(p-z) + \lambda(x+y+z-2p),$$

дифференцированием которой получим три уравнения

$$-p(p-y)(p-z)+\lambda = 0, \quad -p(p-x)(p-z)+\lambda = 0,$$
$$-p(p-x)(p-y)+\lambda = 0.$$

Вместе с уравнением связи $x+y+z-2p=0$ они дают

$$x = y = z = \frac{2p}{3}.$$

Стало быть, при данном периметре наибольшую площадь имеет равносторонний треугольник.

3) *Неравенство Гёльдера* (*Hölder*). Докажем сначала следующую т е о р е м у: *неравенство*

$$uv \leqslant \frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta$$

справедливо при всех $u \geqslant 0$, $v \geqslant 0$ *и всех* $\alpha > 0$, $\beta > 0$, *для которых* $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Это неравенство, очевидно, выполняется, если либо u, либо v равно нулю. Следовательно, можно ограничиться рассмотрением таких значений u и v, произведение которых $uv \neq 0$. Далее, нетрудно убедиться, что если неравенство справедливо для пары чисел (u, v), то оно справедливо и для пар чисел вида $\left(ut^{\frac{1}{\alpha}}, vt^{\frac{1}{\beta}}\right)$, где t — любое положительное число. Поэтому достаточно рассматривать лишь такие пары чисел (u, v), для которых $uv = 1$. Стало быть, требуется доказать, что для всех пар положительных чисел (u, v), для которых $uv = 1$, выполняется неравенство

$$\frac{1}{\alpha} u^\alpha + \frac{1}{\beta} v^\beta \geqslant 1.$$

Для доказательства мы поставим и решим задачу на определение наименьшего значения функции $f(u, v) = \frac{1}{\alpha} u^\alpha + \frac{1}{\beta} v^\beta$ при дополнительном условии $uv = 1$. Это наименьшее значение, очевидно, существует и достигается функцией f в некоторой точке (u, v), где $u \neq 0$ и $v \neq 0$. Стало быть, существует такой множитель $(-\lambda)$, который удовлетворяет в этой точке уравнениям

$$u^{\alpha-1} - \lambda v = 0, \qquad v^{\beta-1} - \lambda u = 0.$$

Помножив первое уравнение на u, а второе на v, получим $u^\alpha = \lambda$ и $v^\beta = \lambda$, а так как $uv = 1$, то окончательно $u = v = 1$. Таким образом, наименьшее значение функции $f = \frac{1}{\alpha} u^\alpha + \frac{1}{\beta} v^\beta$ равно $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, и наше утверждение, что $\frac{1}{\alpha} u^\alpha + \frac{1}{\beta} u^\beta \geqslant 1$, коль скоро $uv = 1$, доказано.

Тем самым доказано (см. выше) и неравенство

$$uv \leqslant \frac{1}{\alpha} u^\alpha + \frac{1}{\beta} v^\beta$$

при условиях, перечисленных в условии теоремы.

Подставим теперь в это неравенство вместо u и v следующие выражения:

$$u = \frac{u_i}{\left(\sum_{i=1}^{n} u_i^\alpha\right)^{\frac{1}{\alpha}}} \quad \text{и} \quad v = \frac{v_i}{\left(\sum_{i=1}^{n} v_i^\beta\right)^{\frac{1}{\beta}}},$$

где $u_1, u_2, \ldots, u_n$ и $v_1, v_2, \ldots, v_n$ — любые не отрицательные числа, среди которых по крайней мере одно из u_i и по крайней мере одно из v_i не равно нулю. Просуммируем полученные n неравенств по i от 1 до n; тем самым будет доказано новое неравенство:

$$\sum_{i=1}^{n} u_i v_i \leqslant \left(\sum_{i=1}^{n} u_i^\alpha\right)^{\frac{1}{\alpha}} \left(\sum_{i=1}^{n} v_i^\beta\right)^{\frac{1}{\beta}},$$

которое называется *неравенством Гёльдера*. Оно справедливо для любых $2n$ чисел (u_i, v_i), где $u_i \geqslant 0$, $v_i \geqslant 0$ $(i = 1, 2, \ldots, n)$ при показателях $\alpha > 0$ и $\beta > 0$, связанных условием $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Кроме того, предполагается, что по крайней мере одно u_i и одно v_i не равны нулю.

4) Наконец, будем искать такие точки поверхности

$$\varphi(x, y, z) = 0,$$

в которых расстояние от данной точки пространства $A(x_0, y_0, z_0)$ до поверхности имеет экстремум. Так как одновременно с расстоянием имеет экстремум и его квадрат, то задача приводится к нахождению экстремумов функции

$$f(x, y, z) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$$

при дополнительном условии $\varphi(x, y, z) = 0$. Составляем вспомогательную функцию

$$F(x, y, z) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 + \lambda\varphi(x, y, z).$$

Для определения неизвестных x, y, z и λ получаем (в дополнение к уравнению $\varphi = 0$) еще три уравнения

$$2(x - x_0) + \lambda\varphi_x = 0, \quad 2(y - y_0) + \lambda\varphi_y = 0, \quad 2(z - z_0) + \lambda\varphi_z = 0,$$

откуда

$$\frac{x - x_0}{\varphi_x} = \frac{y - y_0}{\varphi_y} = \frac{z - z_0}{\varphi_z}.$$

Это значит, что вектор $\overline{AP_0}$, идущий от данной точки A до той точки P_0 поверхности, которая дает экстремум расстояния AP, является нормалью к поверхности в ее точке P_0. Следовательно, для того чтобы прямолинейный путь от данной точки A до (дифференцируемой) поверхности был короче (или, напротив, длиннее) любого соседнего достаточно близкого прямолинейного пути до нее, надо двигаться вдоль нормали к поверхности. Конечно, каждую найденную точку надо еще подвергнуть специальному исследованию для выяснения вопроса, дает ли она действительно экстремум, и если да, то какой: минимум или максимум.

Для конкретности рассмотрим, например, точку A внутри шара. Точки, дающие экстремум расстояния от точки до его поверхности, лежат на концах диаметра, проходящего через A. Расстояние от A до одного конца диаметра есть минимум, до другого конца — максимум. Здесь минимум дает вместе с тем и наименьшее расстояние от A до поверхности, а максимум — наибольшее.

Упражнения

1. Найти наибольшее и наименьшее расстояние от точки эллипса $\frac{x^2}{4} + \frac{y^2}{1} = 1$ до прямой $x + y - 4 = 0$.

2. Сумма длин двенадцати ребер прямоугольного параллелепипеда равна a; сумма площадей его шести граней равна $\frac{a^2}{25}$. Вычислить, каковы должны быть длины ребер параллелепипеда, чтобы разность между его объемом и объемом куба, ребро которого равно самому короткому ребру параллелепипеда, имело наибольшее значение.

3. Определить максимумы и минимумы функции

$$(ax^2 + by^2)e^{-x^2 - y^2} \qquad (0 < a < b).$$

4. Показать, что максимум выражения

$$\frac{ax^2 + 2bxy + cy^2}{ex^2 + 2fxy + gy^2} \qquad (eg - f^2 > 0)$$

равен большему из корней квадратного уравнения относительно λ:

$$(ac-b^2)-\lambda(ag-2bf+ec)+\lambda^2(eg-f^2)=0.$$

5. Вычислить максимумы следующих выражений:

$$\text{а) } \frac{x^2+6xy+3y^2}{x^2-xy+y^2}; \qquad \text{б) } \frac{x^4+2x^3y}{x^4+y^4}.$$

6. Определить стационарные точки функции

$$f(x, y)=y^2\left(\sin x-\frac{x}{2}\right)$$

и выяснить их характер.

7*. Найти полуоси a и b эллипса $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ наименьшей площади, содержащего внутри себя окружность $(x-1)^2+y^2=1$.

8. Найти четырехугольник с заданными сторонами a, b, c и d, имеющий наибольшую площадь.

9. Какая точка шаровой поверхности $x^2+y^2+z^2=1$ наиболее удалена от точки $A(1, 2, 3)$?

10. Дан выпуклый четырехугольник $P_1P_2P_3P_4$. Найти такую точку O, расстояния которой от вершин P_1, P_2, P_3, P_4 имеют наименьшую сумму.

11. Дан эллипсоид $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$. Обозначим через A, B, C отрезки, отсекаемые на осях координат касательной плоскостью эллипсоида в его точке (x, y, z), $(x>0,\ y>0,\ z>0)$. Найти на данном эллипсоиде такую точку (x, y, z), для которой величина

$$\text{а) } A+B+C; \qquad \text{б) } \sqrt{A^2+B^2+C^2}$$

имеет наименьшее значение.

12. Среди прямоугольных параллелепипедов, вписанных в эллипсоид $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$, найти тот, который имеет наибольший объем.

13. Среди прямоугольников, вписанных в эллипс $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, найти такой, который имеет наибольший периметр.

14. На эллипсе $5x^2-6xy+5y^2=4$ найти такую точку, в которой касательная имеет наибольшее расстояние от начала координат.

15*. Доказать, что длина l самой большой оси эллипсоида

$$Ax^2+By^2+Cz^2+2Dxy+2Exz+2Gyz=1$$

равна наибольшему действительному корню уравнения

$$\begin{vmatrix} A-\frac{1}{l^2} & D & E \\ D & B-\frac{1}{l^2} & G \\ E & G & C-\frac{1}{l^2} \end{vmatrix}=0.$$

ДОПОЛНЕНИЯ К ГЛАВЕ III

§ 1. Достаточные условия экстремума функции двух переменных

1. Постановка вопроса. В основном тексте этой главы, в теории максимумов и минимумов, мы ограничились выводом *необходимых* условий экстремума. Во многих случаях специальный характер решаемой практической задачи позволяет судить о природе найденной стационарной точки и выяснить, дает ли она экстремум и какой. Однако же важно иметь общие *достаточные* условия наличия максимума или минимума в стационарной точке. Такие достаточные условия мы здесь выведем для типичного случая двух назависимых переменных.

Рассмотрим стационарную точку (x_0, y_0) функции $f(x, y)$, т. е. точку, в которой обращаются в нуль обе частные производные первого порядка f_x и f_y. Наличие или отсутствие в этой точке экстремума зависит от того, сохраняет ли полное приращение

$$f(x_0+h,\ y_0+k)-f(x_0,\ y_0)$$

один и тот же знак при всех достаточно малых значениях h и k или нет. Разложим это выражение по формуле Тэйлора (гл. II, § 6) с остаточным членом R_2. Так как $f_x(x_0, y_0)=0$ и $f_y(x_0, y_0)=0$, то требуемое разложение принимает следующий вид:

$$f(x_0+h,\ y_0+k)-f(x_0,\ y_0)=$$
$$=\frac{1}{2}[h^2f_{xx}(x_0,\ y_0)+2hkf_{xy}(x_0,\ y_0)+k^2f_{yy}(x_0,\ y_0)]+\varepsilon\rho^2,$$

где $\rho^2=h^2+k^2$, а $\varepsilon\to 0$, когда расстояние $\rho\to 0$.

Отсюда видно, что в достаточно малой окрестности точки (x_0, y_0) поведение полного приращения функции

$$f(x_0+h,\ y_0+k)-f(x_0,\ y_0)$$

определяется свойствами выражения

$$Q(h,\ k)=ah^2+2bhk+ck^2,$$

где, для сокращения записи, введены следующие обозначения:

$$f_{xx}(x_0,\ y_0)=a,\qquad f_{xy}(x_0,\ y_0)=b,\qquad f_{yy}(x_0,\ y_0)=c.$$

Выражение $Q(h, k)$ представляет собой однородный многочлен второй степени относительно h и k или, как принято говорить, *квадратичную форму*. Для решения вопроса о наличии в точке (x_0, y_0) экстремума надо исследовать квадратичную форму $Q(h, k)$.

2. Исследование квадратичной формы $Q(h, k)$. Исходим из допущения, что коэффициенты формы a, b, c не равны одновременно нулю. Этого исключительного случая, когда $a=b=c=0$, мы не будем рассматривать; отметим лишь, что для его исследования надо взять формулу Тэйлора с остаточным членом более высокого порядка, чем R_2.

Относительно квадратичной формы Q возможны три различных случая.

1) При всех значениях h и k форма Q принимает значения одного лишь знака и обращается в нуль только при $h=k=0$. Такая форма называется *определенной* квадратичной формой, притом *положительно определенной*, если этот неизменный знак положительный, и *отрицательно определенной*, если он — отрицательный. Так, например, форма h^2+k^2, которая получается из Q при $a=c=1$ и $b=0$, является положительно определенной, а форма $-h^2+2hk-2k^2=$ $=-(h-k)^2-k^2$ есть отрицательно определенная.

2) Форма способна принимать значения, различные по знаку; например, форма $Q=2hk$ принимает значение 2 при $h=1$, $k=1$ и значение -2 при $h=1$, $k=-1$ или форма h^2-k^2 имеет значение -3 при $h=1$, $k=2$ и значение 8 при $h=3$, $k=1$. Такая квадратичная форма называется *неопределенной*.

3) Существует, наконец, третья возможность: форма способна обращаться в нуль и при значениях h и k, отличных от нуля, но при прочих значениях этих переменных она сохраняет неизменный знак. Такие формы называются *полуопределенными*. Примером может служить форма $(h+k)^2$, которая обращается в нуль всякий раз, как $h=-k$, а при всех других значениях переменных она положительна.

Квадратичная форма $Q=ah^2+2bhk+ck^2$ является определенной формой в том и только в том случае, если ее *дискриминант*

$$ac-b^2>0.$$

В этом случае коэффициенты a и c не равны нулю и имеют одинаковый знак; если этот общий знак положительный, то форма положительно определенная, если же — отрицательный, то форма отрицательно определенная.

Форма Q является неопределенной в том и только в том случае, если дискриминант

$$ac^2-b^2<0.$$

Наконец, полуопределенная форма вполне характеризуется равенством ее дискриминанта нулю:

$$ac-b^2=0.$$

Наметим путь вывода этих критериев. Если $a=c=0$, то обязательно $b\neq 0$ и форма $Q=2bhk$, очевидно, неопределенная; вместе с тем $ac-b^2=$ $=-b^2<0$.

Допустим теперь, что a и c не равны одновременно нулю. Так как эти коэффициенты играют в Q симметричную роль, то можно, без уменьшения общности, предположить, что $a \neq 0$. Тогда

$$ah^2 + 2bhk + ck^2 = a\left[\left(h + \frac{b}{a}k\right)^2 + \frac{ac - b^2}{a^2}k^2\right].$$

Теперь очевидно, что если $ac - b^2 > 0$, то форма определенная и принимает значения того же знака, что и коэффициент a. Если $ac - b^2 = 0$, то форма полуопределенная, ибо она тогда обращается в нуль при всех значениях h и k, связанных равенством $h = -\frac{b}{a}k$, а при всех других значениях переменных сохраняет постоянный знак. Наконец, если $ac - b^2 < 0$, то форма неопределенная, так как в этом случае она принимает значения различных знаков. Действительно, если $k = 0$, то значение формы будет того же знака, что a; если же $h + \frac{b}{a}k = 0$, то знак формы противоположен знаку a.

3. Достаточные условия максимума и минимума. Докажем теперь следующие достаточные критерии наличия в стационарной точке максимума и минимума.

Если квадратичная форма $Q(h, k)$ является положительно определенной, то стационарное значение, принимаемое функцией при $h = k = 0$, является ее *минимумом.* Если эта форма отрицательно определенная, то стационарное значение есть максимум функции. Если же форма неопределенная, то функция не имеет в стационарной точке ни максимума, ни минимума; такую точку мы назвали (гл. III, § 6, n° 2) *седловиной* или *точкой перевала.* Стало быть, определенность формы Q является достаточным условием экстремума, а неопределенность формы исключает возможность экстремума в стационарной точке. Полуопределенный случай приводит к сложным рассуждениям, и мы его не будем рассматривать.

Для доказательства рассмотрим отношение $\frac{Q(h, k)}{h^2 + k^2}$ и убедимся, что оно является функцией двух величин $u = \frac{h}{\sqrt{h^2 + k^2}}$ и $v = \frac{k}{\sqrt{h^2 + k^2}}$:

$$\frac{Q}{h^2 + k^2} = au^2 + 2buv + cv^2,$$

причем $u^2 + v^2 = 1$. Если форма Q положительно определенная, то мы воспользуемся тем, что выражение $\frac{Q}{h^2 + k^2}$ как непрерывная функция от u и v должно достигать на окружности $u^2 + v^2 = 1$ (в некоторой ее точке) своего наименьшего значения $2m$. Это число $2m$, очевидно, положительно; оно не равно нулю, так как на нашей окружности u и v нигде не обращаются в нуль одновременно. Поэтому в точках, лежащих на окружности $u^2 + v^2 = 1$,

$$Q \geqslant 2m(h^2 + k^2) = 2m\rho^2,$$

а следовательно, для этих точек

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{1}{2}Q(h, k) + \varepsilon\rho^2 \geqslant (m + \varepsilon)\rho^2.$$

Выберем теперь ρ столь малым, чтобы абсолютная величина числа ε была меньше чем $\frac{m}{2}$; тогда в точках, лежащих на окружности $u^2+v^2=1$,

$$f(x_0+h,\ y_0+k)-f(x_0,\ y_0)\geqslant\frac{m}{2}\rho^2.$$

Поэтому в выбранной нами окрестности точки $(x_0,\ y_0)$ значение функции f всюду больше, чем $f(x_0,\ y_0)$, за исключением, конечно, самой точки $(x_0,\ y_0)$, и $(x_0,\ y_0)$ есть точка минимума.

Аналогичным путем можно доказать, что если форма отрицательно определенная, то в точке $(x_0,\ y_0)$ функция имеет максимум.

Наконец, если форма Q неопределенная, то существует такая пара значений $(h_1,\ k_1)$, при которой Q отрицательна, и такая другая пара $(h_2,\ k_2)$, при которой Q положительна. Поэтому можно найти такое положительное число m, что

$$Q(h_1,\ k_1)<-2m\rho_1^2,\qquad Q(h_2,\ k_2)>2m\rho_2^2,$$

где $\rho_1^2=h_1^2+k_1^2$ и $\rho_2^2=h_2^2+k_2^2$.

Рассмотрим теперь переменную точку $(x_0+h,\ y_0+k)$, лежащую на прямой, соединяющей точки $(x_0,\ y_0)$ и $(x_0+h_1,\ y_0+k_1)$, и пусть $\rho^2=h^2+k^2$; тогда $h=th_1$ и $k=tk_1$ $(t\neq 0)$. Так как

$$Q(h,\ k)=t^2Q(h_1,\ k_1)\quad\text{и}\quad\rho^2=t^2\rho_1^2,$$

то имеем соответствующее неравенство

$$Q(h,\ k)<-2m\rho^2.$$

Поэтому путем выбора достаточно малого t (а следовательно, и соответствующего ρ) можно сделать приращение $f(x_0+h,\ y_0+k)-f(x_0,\ y_0)$ отрицательным. Для этого надо лишь выбрать t настолько малым, чтобы при $h=th_1$ и $k=tk_1$ абсолютная величина числа ε была меньше, чем $\frac{m}{2}$. Тогда будет

$$f(x_0+h,\ y_0+k)-f(x_0,\ y_0)<-\frac{m}{2}\rho^2,$$

т. е. значение функции в точке $(x_0+h,\ y_0+k)$ меньше, чем в стационарной точке $(x_0,\ y_0)$. Подобным же образом, рассматривая точки прямой, соединяющей $(x_0,\ y_0)$ с точкой $(x_0+h_2,\ y_0+k_2)$, т. е. полагая $h=th_2$ и $k=tk_2$, обнаружим, что в сколь угодно малой окрестности точки $(x_0,\ y_0)$ существуют точки, в которых значение функции f превосходит $f(x_0,\ y_0)$. Таким образом, если для стационарной точки форма Q неопределенна, то нет ни максимума, ни минимума; вместо экстремума там оказывается седловина.

Если $a=b=c=0$ в стационарной точке, т. е. если квадратичная форма тождественно равна нулю, а также если форма полуопределенная, то наше исследование отказывается служить. Формулировка достаточных критериев для этих исключительных случаев слишком сложна, да и вывод сложен.

В итоге мы имеем для решения вопроса о максимумах и минимумах следующее правило:

Если в точке (x_0, y_0), *удовлетворяющей уравнениям*

$$f_x(x, y)=0 \quad \textit{и} \quad f_y(x, y)=0,$$

выполняется неравенство

$$f_{xx}f_{yy}-f_{xy}^2>0,$$

то в этой точке функция $f(x, y)$ *имеет экстремум. При этом* $f_{xx}(x_0, y_0)$ *и* $f_{yy}(x_0, y_0)$ *имеют одинаковый знак; если их общий знак есть минус, то* $f(x_0, y_0)$ *является максимумом, если — плюс, то минимумом.*

Если, напротив, в стационарной точке

$$f_{xx}f_{yy}-f_{xy}^2<0,$$

то функция не имеет экстремума в этой точке.

Случай $f_{xx}f_{yy}-f_{xy}^2=0$ остается здесь невыясненным и требует дополнительного исследования.

Эти критерии допускают простое геометрическое истолкование. Необходимые условия $f_x=f_y=0$ утверждают, что касательная плоскость к поверхности $z=f(x, y)$ горизонтальна.

Если в точке (x_0, y_0) функция действительно имеет экстремум, то в окрестности этой точки касательная плоскость не пересекает поверхности. Если же (x_0, y_0) является точкой перевала (седловиной), то касательная плоскость пересекает поверхность по кривой, имеющей несколько ветвей, проходящих через исследуемую точку. Эта картина станет более ясной после изучения особых точек, в следующем параграфе.

4. Примеры. 1) Найдем экстремумы функции

$$f(x, y)=x^2+xy+y^2+ax+by.$$

Приравнивая нулю частные производные первого порядка, получаем систему двух уравнений

$$2x+y+a=0, \qquad x+2y+b=0;$$

решив эту систему, находим точку $x=\frac{1}{3}(b-2a)$, $y=\frac{1}{3}(a-2b)$.

Выражение $f_{xx}f_{yy}-f_{xy}^2=3>0$ всюду, следовательно и в найденной точке; стало быть, в этой точке функция имеет экстремум, а так как $f_{xx}=2>0$, то именно минимум.

2) Функция $f(x, y)=(y-x^2)^2+x^5$ имеет единственную стационарную точку — начало координат. В этой точке выражение $f_{xx}f_{yy}-f_{xy}^2$ обращается

в нуль, и наш критерий не дает ответа. Однако нетрудно убедиться, что данная функция не имеет в начале координат экстремума, ибо в любой окрестности начала она принимает как положительные, так и отрицательные значения.

3) Функция $f(x, y) = (x-y)^4 + (y-1)^4$ имеет единственную стационарную точку $x_0 = 1$, $y_0 = 1$. В этой точке дискриминант $f_{xx}f_{yy} - f_{xy}^2$ обращается в нуль. Однако, в отличие от предыдущего примера, эта функция имеет в своей стационарной точке (1; 1) минимум; действительно, в этой точке $f(1, 1) = 0$, а во всех других точках $f(x, y) > 0$.

Упражнение

Дана функция $\varphi(x)$, у которой $\varphi(a) = k \neq 0$ и $\varphi'(a) \neq 0$. Доказать, что функция $f(x) + f(y) + f(z)$ при дополнительном условии $\varphi(x)\varphi(y)\varphi(z) = k^3$ имеет при $x = y = z = a$ максимум, если только

$$f'(a)\left[\frac{\varphi''(a)}{\varphi'(a)} - \frac{\varphi'(a)}{\varphi(a)}\right] > f''(a).$$

§ 2. Особые точки плоских кривых

В гл. III, § 2, n° 2 мы видели, что кривая $f(x, y) = 0$ может иметь в точке (x_0, y_0), удовлетворяющей трем условиям

$$f(x_0, y_0) = 0, \qquad f_x(x_0, y_0) = 0, \qquad f_y(x_0, y_0) = 0,$$

особенность (особую точку). Для изучения этих особых точек предположим, что в окрестности рассматриваемой точки (x_0, y_0) функция $f(x, y)$ имеет непрерывные частные производные до второго порядка включительно и что в самой этой точке производные второго порядка не обращаются одновременно в нуль.

Пользуясь формулой Тэйлора с остаточным членом R_2, мы запишем уравнение кривой в следующем виде:

$$2f(x, y) = (x-x_0)^2 f_{xx}(x_0, y_0) + 2(x-x_0)(y-y_0) f_{xy}(x_0, y_0) + \\ + (y-y_0)^2 f_{yy}(x_0, y_0) + \varepsilon\rho^2 = 0,$$

где $\rho^2 = (x-x_0)^2 + (y-y_0)^2$, а $\varepsilon \to 0$ при $\rho \to 0$.

Любую прямую, проходящую через точку (x_0, y_0), можно задать параметрическими уравнениями вида

$$x - x_0 = at, \qquad y - y_0 = bt,$$

где t — переменный параметр, причем a и b — две произвольные постоянные, координаты направляющего вектора прямой; мы будем пользоваться *единичным* направляющим вектором, так что $a^2 + b^2 = 1$. Для определения точек пересечения этой прямой с кривой $f(x, y) = 0$ подставим выражения $x - x_0$ и $y - y_0$ через t в уравнение кривой, развернутое по формуле Тэйлора:

$$a^2t^2 f_{xx} + 2abt^2 f_{xy} + b^2t^2 f_{yy} + \varepsilon t^2 = 0;$$

из этого уравнения определятся значения параметра t для точек пересечения прямой с кривой. (Надо помнить, что в f_{xx}, f_{xy} и f_{yy} подставлены значения $x=x_0$, $y=y_0$; указание на это мы, для сокращения записи, опустили.)

Одно решение, $t=0$, видно сразу; оно дает точку пересечения (x_0, y_0), которая была известна заранее. Однако замечательно, что левая часть уравнения содержит множителем t^2, так что $t=0$ является *двойным* корнем уравнения. Поэтому точку (x_0, y_0), если она оказывается особой, называют двойной точкой кривой.

Опустив множитель t^2, мы имеем еще уравнение

$$a^2 f_{xx} + 2ab f_{xy} + b^2 f_{yy} + \varepsilon = 0$$

для определения остальных точек пересечения (неизвестное t содержится в скрытом виде в ε, которое зависит от x и y, а следовательно, и от t).

Поставим теперь вопрос: возможно ли, чтобы другая точка пересечения прямой с нашей кривой (или одна из других точек пересечения, если число их больше двух) стремилась к точке (x_0, y_0), когда прямая стремится к некоторому предельному положению, т. е. когда направляющий вектор $\{a, b\}$ секущей стремится к предельному вектору $\{a_0, b_0\}$. Такое предельное положение секущей мы, согласно известному определению, должны называть касательной. Чтобы разобраться в этом вопросе, заметим, что когда точка (x, y) стремится к (x_0, y_0), то и t стремится к нулю и ε стремится к нулю. Поэтому, так как написанное выше уравнение при указанном вращении секущей все время удовлетворяется, то выражение $a^2 f_{xx} + 2ab f_{xy} + b^2 f_{yy}$ должно тоже стремиться к нулю, когда секущая стремится к своему предельному положению. Стало быть, координаты направляющего вектора $\{a_0, b_0\}$ касательной должны удовлетворять уравнению

$$a_0^2 f_{xx} + 2a_0 b_0 f_{xy} + b_0^2 f_{yy} = 0.$$

Если пара чисел a_0, b_0 является решением этого квадратного уравнения, однородного относительно обеих неизвестных, то и пара λa_0, λb_0 является решением. Любую такую пару чисел можно принять за координаты касательного вектора. Присовокупив еще уравнение $a_0^2 + b_0^2 = 1$, получим единичный касательный вектор.

Возможны три различных случая.

1) Дискриминант квадратного уравнения имеет в точке (x_0, y_0) отрицательное значение, т. е.

$$f_{xx} f_{yy} - f_{xy}^2 < 0.$$

В этом случае существуют *две различные действительные касательные*. Кривая имеет *узловую точку*. Такую точку имеет, например, лемниската $(x^2 + y^2)^2 - (x^2 - y^2) = 0$ в начале координат и строфоида $(x^2 + y^2)(x - 2a) + a^2 x = 0$ в точке $x_0 = a$, $y_0 = 0$.

Так как уравнение каждой из двух касательных в точке (x_0, y_0) можно записать в виде $\frac{x-x_0}{a_0}=\frac{y-y_0}{b_0}$ (со своими значениями a_0, b_0), где теперь x и y — текущие координаты на касательной, то, подставив в уравнение для a_0, b_0 пропорциональные им разности $x-x_0$ и $y-y_0$, получим

$$(x-x_0)^2 f_{xx}+2(x-x_0)(y-y_0)f_{xy}+(y-y_0)^2 f_{yy}=0;$$

это уравнение выражает, очевидно, совокупность обеих касательных в узловой точке. Обратим внимание на то обстоятельство, что его можно получить, приравнивая нулю члены второго порядка в разложении функции $f(x, y)$ по формуле Тэйлора.

2) В точке (x_0, y_0) дискриминант положителен:

$$f_{xx}f_{yy}-f_{xy}^2>0.$$

В этом случае (*действительных*) *касательных вообще нет*. Точка (x_0, y_0) является *изолированной* точкой кривой. Изолированной или уединенной точкой кривой называется такая ее точка, в достаточно малой окрестности которой кривая не имеет других точек.

Например, кривая $(x^2-a^2)^2+(y^2-b^2)^2=a^4+b^4$ $(a\neq 0,\ b\neq 0)$, имеет уединенную точку в начале координат. Действительно, значения $x=0$, $y=0$ удовлетворяют уравнению кривой, но во всех других точках области $|x|<a\sqrt{2}$, $|y|<b\sqrt{2}$ левая часть уравнения меньше правой.

3) В точке (x_0, y_0) дискриминант обращается в нуль:

$$f_{xx}f_{yy}-f_{xy}^2=0.$$

Тогда кривая имеет в точке (x_0, y_0) две совпадающие касательные.

[Подробное исследование показывает, что здесь возможны различные ситуации. Точка (x_0, y_0) может быть *точкой заострения* (иначе называемой *точкой возврата*); такой точкой является начало координат у полукубической параболы $y^3=x^2$ и у кривой $(ay-x^2)^2=x^5$ $(a\neq 0)$. (Ср. упр. 15 на стр. 237.) У первой кривой двойной касательной является ось y, у второй — ось x. Возможна *изолированная* точка, как, например, начало координат у кривой $y^2=x^4(x-1)$.

Возможна также *точка самоприкосновения*, когда через точку (x_0, y_0) проходят две ветви кривой, причем каждая ветвь порознь не имеет никаких особенностей и все же для кривой в целом (x_0, y_0) является особой точкой, так как в ней обе ветви касаются друг друга. Например, кривая $y^2+x^2y-2x^4=0$ имеет точку самоприкосновения в начале координат. Эта кривая представляет собой совокупность двух парабол, $y=x^2$ и $y=-2x^2$, касающихся друг друга в начале координат.]

Мы опустили тот случай, когда в точке (x_0, y_0) все частные производные второго порядка обращаются в нуль. Такой случай требует более сложных и тонких исследований, и мы им заниматься не будем. Через такую точку может проходить несколько ветвей кривой; в такой точке могут быть и разнообразные другие особенности.

В заключение отметим вкратце связь между этими вопросами и теорией максимумов и минимумов. Так как в стационарной точке

(x_0, y_0) функции $z = f(x, y)$ частные производные первого порядка обращаются в нуль, то уравнение касательной плоскости к соответствующей поверхности упрощается и получает следующий вид:

$$z - f(x_0, y_0) = 0.$$

Поэтому уравнение

$$f(x, y) - f(x_0, y_0) = 0,$$

рассматриваемое на плоскости xy, дает проекцию на эту плоскость линии пересечения нашей поверхности со своей касательной плоскостью в точке (x_0, y_0); ясно, что точка (x_0, y_0) может оказаться особой точкой этой кривой. Если эта точка изолированная, то в некоторой ее окрестности касательная плоскость не имеет других общих точек с поверхностью, и функция $f(x, y)$ имеет в (x_0, y_0) максимум или минимум (ср. стр. 225). Если же точка (x_0, y_0) является узловой точкой проекции, то касательная плоскость пересекается с поверхностью по кривой с двумя ветвями и (x_0, y_0) является для функции точкой перевала (седловиной), в которой нет ни максимума, ни минимума. Эти замечания иллюстрируют достаточные условия экстремума, найденные в § 1.

Упражнения

1*. Найти кривизну в начале координат каждой из двух ветвей кривой

$$y(ax + by) = cx^3 + ex^2y + fxy^2 + gy^3.$$

2. Исследовать особые точки следующих кривых:

а) $F(x, y) \equiv ax^3 + by^3 - cxy = 0;$

б) $F(x, y) \equiv (y^2 - 2x^2)^2 - x^5 = 0;$

в) $F(x, y) \equiv (1 + e^{1/x})y - x = 0;$

г) $F(x, y) \equiv y^2(2a - x) - x^3 = 0;$

д) $F(x, y) \equiv (y - 2x)^2 - x^5 = 0.$

3. Пусть (x, y) — узловая точка кривой $F(x, y) = 0$. Вычислить угол φ между обеими касательными в точке (x, y), предполагая, что в этой точке не обращаются сразу в нуль все частные производные второго порядка от функции F.

Найти угол между касательными в узловой точке а) лемнискаты, б) декартова листа (см. стр. 139).

§ 3. Особые точки поверхностей

[Различение между обыкновенными и особыми точками поверхности вводится аналогичным образом. Точка поверхности $F(x, y, z) = 0$ называется *обыкновенной* или *регулярной*, если в окрестности этой точки можно одну из координат представить в виде непрерывно дифференцируемой функции двух других, например выразить z как функцию $z = f(x, y)$. Всякая точка поверхности, не обладающая этим свойством, называется *особой* точкой. В каждой своей обыкновенной точке поверхность имеет определенную касательную плоскость. Из теории неявных функций видно, что точка поверхности, в которой хотя бы одна из частных производных F_x, F_y, F_z не обращается в нуль, является обыкновенной точкой. Ясно, что особые точки дифференцируемой поверхности надо искать среди точек, удовлетворяющих системе уравнений

$$F=0, \quad F_x=0, \quad F_y=0, \quad F_z=0,$$

т. е. среди стационарных точек функции F, лежащих на поверхности. Пусть $P_0(x_0, y_0, z_0)$ есть такая точка; тогда уравнение касательной плоскости

$$(\xi-x_0)F_x+(\eta-y_0)F_y+(\zeta-z_0)F_z=0$$

обращается в точке P_0 в тождество $0=0$.

Если в точке P_0 не обращаются одновременно в нуль все частные производные второго порядка от функции $F(x, y, z)$, то геометрическим местом касательных ко всем кривым, лежащим на поверхности и проходящим через P_0, будет уже не плоскость, а конус второго порядка (действительный, мнимый, распадающийся на пару плоскостей или вырождающийся в прямую).

Для доказательства рассмотрим любую кривую, лежащую на поверхности $F(x, y, z)=0$ и проходящую через точку P_0; пусть она задана параметрическими уравнениями

$$x=x(t), \qquad y=y(t), \qquad z=z(t).$$

Эти функции удовлетворяют уравнению поверхности при всех значениях t. Представим себе, что они подставлены в уравнение $F(x,y,z)=0$, и продифференцируем полученное соотношение два раза по t. Второе дифференцирование дает

$$F_{xx}\left(\frac{dx}{dt}\right)^2+F_{yy}\left(\frac{dy}{dt}\right)^2+F_{zz}\left(\frac{dz}{dt}\right)^2+2F_{xy}\frac{dx}{dt}\frac{dy}{dt}+2F_{yz}\frac{dy}{dt}\frac{dz}{dt}+\\+2F_{xz}\frac{dx}{dt}\frac{dz}{dt}+F_x\frac{d^2x}{dt^2}+F_y\frac{d^2y}{dt^2}+F_z\frac{d^2z}{dt^2}=0.$$

Подставим сюда $x=x_0$, $y=y_0$, $z=z_0$ и значение t_0, соответствующее точке P_0. Тогда последние три члена обратятся в нуль, и уравнение примет следующий вид:

$$F_{xx}^0\left(\frac{dx}{dt}\right)_0^2+F_{yy}^0\left(\frac{dy}{dt}\right)_0^2+F_{zz}^0\left(\frac{dz}{dt}\right)_0^2+2F_{xy}^0\left(\frac{dx}{dt}\right)_0\left(\frac{dy}{dt}\right)_0+\\+2F_{yz}^0\left(\frac{dy}{dt}\right)_0\left(\frac{dz}{dt}\right)_0+2F_{xz}\left(\frac{dx}{dt}\right)_0\left(\frac{dz}{dt}\right)_0=0,$$

где значком 0 (наверху или внизу у скобки) отмечено, что соответствующая величина вычислена для точки P_0.

Для того чтобы получить уравнение геометрического места касательных, надо исключить $\left(\frac{dx}{dt}\right)_0$, $\left(\frac{dy}{dt}\right)_0$ и $\left(\frac{dz}{dt}\right)_0$ из последнего уравнения и из уравнений касательных

$$\frac{\xi-x_0}{\left(\frac{dx}{dt}\right)_0}=\frac{\eta-y_0}{\left(\frac{dy}{dt}\right)_0}=\frac{\zeta-z_0}{\left(\frac{dz}{dt}\right)_0}.$$

В результате получится уравнение

$$F_{xx}^0(\xi-x_0)^2+F_{yy}^0(\eta-y_0)^2+F_{zz}^0(\zeta-z_0)^2+2F_{xy}^0(\xi-x_0)(\eta-y_0)+\\+2F_{yz}^0(\eta-y_0)(\zeta-z_0)+2F_{xz}^0(\xi-x_0)(\zeta-z_0)=0.$$

Это уравнение представляет конус второго порядка, касающийся нашей поверхности в ее особой точке (x_0, y_0, z_0).

Если касательный конус действительный, то особая точка P_0 называется *конической точкой* поверхности. Если конус мнимый, то P_0 является изолированной точкой. Это кажется наглядно ясным, а строгое доказательство сложно, и мы его приводить не будем.

Примеры. 1) Эллипсоид $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$. Имеем $F_x=\frac{2x}{a^2}$, $F_y=\frac{2y}{b^2}$, $F_z=\frac{2z}{c^2}$. Эти выражения обращаются в нуль только при значениях $x=y=z=0$, которые не удовлетворяют уравнению эллипсоида. Стало быть, эллипсоид не имеет особых точек.

2) Найдем особые точки поверхности

$$F\equiv z^3+x^2+y^2-2y-3z-1=0.$$

Находим $F_x=2x$, $F_y=2y-2$, $F_z=3z^2-3$. Частные производные обращаются в нуль в двух точках: (0; 1; 1) и (0; 1; —1), но первая из них не лежит на поверхности. Остается точка P_0 (0; 1; —1). Вычислив значения частных производных второго порядка в этой точке, найдем касательный конус

$$\xi^2+(\eta-1)^2-3(\zeta+1)^2=0.$$

Это *действительный* конус с вершиной в точке P_0, которая является, стало быть, конической точкой поверхности.

3) Поверхность $F\equiv z^3-x^2-y^2=0$. Четырем уравнениям $F=0$, $F_x=0$, $F_y=0$, $F_z=0$ удовлетворяет единственная точка — начало координат. Касательный конус в этой точке имеет уравнение $\xi^2+\eta^2=0$. Конус выродился в прямую $\xi=0$, $\eta=0$. (Ср. пример 5 на экстремум на стр. 205.)

4) Конус второго порядка $z^2-x^2-y^2=0$. Единственной особой точкой оказывается его вершина — начало координат, а именно конической точкой. Касательный конус в вершине, как и следовало ожидать, совпадает с заданным конусом.

5) Поверхность $F\equiv z^3-x^4-y^4=0$ имеет единственную точку, в которой одновременно $F_x=F_y=F_z=0$, а именно, начало координат. Тем не менее начало O является обыкновенной точкой поверхности, ибо z выражается как явная непрерывно дифференцируемая функция (проверить!) от x и y: $z=(x^4+y^4)^{1/3}$. Переход к цилиндрическим координатам дает $z=(\cos^4\varphi+\sin^4\varphi)^{1/3}\rho^{4/3}$. Отсюда видно, что любое плоское сечение

поверхности, проходящее через ось z, есть кривая вида $z=ax^{4/3}$, напоминающая параболу по своему облику.

Аналогичный пример дает поверхность $(Ax+By+Cz+D)^2=0$, которая является плоскостью (считаемой два раза). Все ее точки, конечно, обыкновенные, несмотря на то, что во всех этих точках $F_x=F_y=F_z=0$.]

§ 4. Связь между уравнениями движения жидкости в форме Эйлера и в форме Лагранжа

Обозначим через (a, b, c) координаты частицы движущейся непрерывной среды (жидкости или газа) в момент времени $t=0$. Тогда движение можно описать с помощью трех функций

$$x=x(a, b, c, t), \quad y=y(a, b, c, t), \quad z=z(a, b, c, t) \qquad (L)$$

или, что то же самое, с помощью переменного радиус-вектора $\boldsymbol{r}=\boldsymbol{r}(a, b, c, t)=\boldsymbol{r}(\boldsymbol{r}_0, t)$, где постоянный вектор $\boldsymbol{r}_0=\{a, b, c\}$. [Эти уравнения позволяют проследить движение индивидуальной частицы, характеризуемой начальным положением $\boldsymbol{r}_0=\{a, b, c\}$.] Скорость и ускорение движущейся частицы определяются первой и второй производными по времени t. Стало быть, вектор-скорость есть $\dot{\boldsymbol{r}}=\{\dot{x}, \dot{y}, \dot{z}\}$, а вектор-ускорение есть $\ddot{\boldsymbol{r}}=\{\ddot{x}, \ddot{y}, \ddot{z}\}$, причем как эти два вектора, так и их координаты являются функциями начального положения (a, b, c) и параметра — времени t. Для каждого значения t мы имеем преобразование, которое переводит координаты (a, b, c), определяющие начальное положение различных точек движущейся среды, в их координаты (x, y, z) в момент времени t. Таковы уравнения движения жидкости в форме Лагранжа.

Эйлер ввел другой метод описания движения жидкости — уравнения движения в форме Эйлера. По этому методу движение жидкости задается тремя функциями

$$\dot{x}=u(x, y, z, t), \quad \dot{y}=v(x, y, z, t), \quad \dot{z}=w(x, y, z, t), \qquad (E)$$

определяющими координаты $\dot{x}, \dot{y}, \dot{z}$ вектор-скорости $\dot{\boldsymbol{r}}$ в каждой точке (x, y, z) движущейся жидкости в момент времени t. [Уравнения Эйлера дают возможность проследить в любой точке области пространства, занятой движущейся жидкостью, изменение скорости со временем; с другой стороны, в любой фиксированный момент времени t они описывают распределение скоростей в различных точках этой области.]

Как перейти от уравнений Лагранжа к уравнениям Эйлера? Для этого надо из уравнений (L) выразить a, b, c через x, y, z и t и найденные функции подставить в выражения $\dot{x}(a, b, c, t)$, $\dot{y}(a, b, c, t)$, $\dot{z}(a, b, c, t)$, полученные дифференцированием по t тех же уравнений:

$$u(x, y, z, t)=\dot{x}[a(x, y, z, t), b(x, y, z, t), c(x, y, z, t)], \text{ и т. д.}$$

Затем находим координаты вектор-ускорения из уравнений

$$\dot{x}(a, b, c, t)=u[x(a, b, c, t), y(a, b, c, t), z(a, b, c, t), t] \text{ и т. д.}$$

следующим образом:

$$\ddot{x} = u_x\dot{x} + u_y\dot{y} + u_z\dot{z} + u_t, \text{ и т. д.}$$

или

$$\begin{aligned}\ddot{x} &= u_x u + u_y v + u_z w + u_t,\\ \ddot{y} &= v_x u + v_y v + v_z w + v_t,\\ \ddot{z} &= w_x u + w_y v + w_z w + w_t.\end{aligned}$$

В динамике непрерывной среды играет фундаментальную роль следующее уравнение, связывающее представления движения по Эйлеру и Лагранжу:

$$\operatorname{div}\dot{r} \equiv u_x + v_y + w_z = \frac{\dot{D}}{D},$$

где

$$D(x, y, z, t) = \frac{\partial(x, y, z)}{\partial(a, b, c)}$$

есть характеризующий движение якобиан.

Предоставляем читателю доказать эту теорему, а также соответствующую теорему для двумерного случая, пользуясь различными правилами дифференцирования неявных функций.

§ 5. Представление замкнутой кривой с помощью семейства ее касательных

Дано семейство прямых, зависящее от параметра α:

$$x\cos\alpha + y\sin\alpha - p(\alpha) = 0, \tag{1}$$

где $p(\alpha)$ обозначает заданную периодическую дважды непрерывно дифференцируемую функцию с периодом 2π (так называемую *тангенциальную функцию*). Огибающая C этого семейства прямых является замкнутой кривой, определяемой уравнением (1) в совокупности с уравнением

$$-x\sin\alpha + y\cos\alpha - \dot{p}(\alpha) = 0.$$

Решив эту систему уравнений относительно x и y, получим параметрическое представление огибающей C:

$$\left.\begin{aligned}x &= p(\alpha)\cos\alpha - \dot{p}(\alpha)\sin\alpha,\\ y &= p(\alpha)\sin\alpha + \dot{p}(\alpha)\cos\alpha\end{aligned}\right\} \tag{2}$$

с тем же параметром α. Уравнение (1) определяет семейство касательных кривой C и называется *тангенциальным уравнением* этой кривой.

Так как $\dot{x} = -(p+\ddot{p})\sin\alpha$, $\dot{y} = (p+\ddot{p})\cos\alpha$, то сразу находим следующие выражения для длины L и площади F замкнутой кривой C:

$$L = \int_0^{2\pi} (p+\ddot{p})\,d\alpha = \int_0^{2\pi} p(\alpha)\,d\alpha,$$

$$F = \frac{1}{2}\int_0^{2\pi} (x\dot{y} - y\dot{x})\,d\alpha = \frac{1}{2}\int_0^{2\pi} (p+\ddot{p})p\,d\alpha = \frac{1}{2}\int_0^{2\pi} (p^2 - \dot{p}^2)\,d\alpha.$$

(Мы воспользовались тем, что $\dot{p}(\alpha)$, как и $p(\alpha)$, есть функция периода 2π.)

Так как тангенциальной функцией для параллельной кривой, отстоящей от линии C на расстоянии c, является сумма $p(\alpha)+c$, то из полученных формул нетрудно вывести выражения для длины и площади параллельной кривой. (Ср. т. I, стр. 334, упр. 21 и стр. 684.)

Из последних формул можно вывести изопериметрическое неравенство

$$L^2 \geqslant 4\pi F,$$

в котором знак равенства имеет место только для окружности. Это можно выразить словесно в виде следующей теоремы: *среди всех плоских замкнутых кривых данной длины наибольшую площадь охватывает окружность.*

Для доказательства разложим функцию $p(\alpha)$ в ряд Фурье (т. I, гл. IX):

$$p(\alpha) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\alpha + b_k \sin k\alpha);$$

тогда

$$\dot{p}(\alpha) = \sum_{k=1}^{\infty} k(b_k \cos k\alpha - a_k \sin k\alpha).$$

Подставив эти ряды вместо p и $\dot{p}$ в выведенные выше интегралы и воспользовавшись соотношениями ортогональности тригонометрических функций (т. I, стр. 247), получим

$$L = \pi a_0, \quad F = \frac{\pi}{2}\left[\frac{a_0^2}{2} - \sum_{k=2}^{\infty}(k^2-1)(a_k^2+b_k^2)\right].$$

Следовательно,

$$F \leqslant \frac{\pi a_0^2}{4} = \frac{L^2}{4\pi}.$$

Частный случай $F=\frac{L^2}{4\pi}$ наступит при том и только при том условии, если $a_k=b_k=0$ при $k\geqslant 2$, но тогда

$$p(\alpha)=\frac{a_0}{2}+a_1\cos\alpha+b_1\sin\alpha,$$

а при этой тангенциальной функции параметрические уравнения (2) определяют окружность.

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ III

1. Найти огибающую переменной окружности на плоскости, проходящей через заданную точку O, и центр которой описывает центральную кривую второго порядка с центром O.

2. Если Γ есть плоская кривая, а O — заданная точка в ее плоскости, то геометрическое место Γ' ортогональных проекций точки O на касательные кривой Γ называется *подэрой* этой кривой относительно точки O. Доказать, что, если точка M описывает кривую Γ, то подэра Γ' является огибающей переменной окружности, построенной на радиус-векторе $\overline{OM}$ как на диаметре.

3. Найти огибающую переменной сферы, построенной на радиус-векторе $\overline{OM}$ (ср. упр. 2) как на диаметре.

4. Какие будут огибающие у переменных окружностей и сфер в упр. 2 и 3, если Γ есть окружность, а точка O лежит на этой окружности?

5. Найти огибающую семейства окружностей, построенных на хордах эллипса, параллельных его малой оси, как на диаметрах.

6. Плоскость движется таким образом, что она касается двух парабол: $z=0$, $y^2=4x$ и $y=0$, $z^2=4x$. Показать, что огибающая этой переменной плоскости состоит из двух параболических цилиндров.

7 [1]. Обобщить исследование достаточных условий экстремума (Доп. к гл. III, § 1, n° 1) на функции n переменных и для этой цели доказать следующие предложения. Пусть $f(x_1,\ldots,x_n)$ — трижды непрерывно дифференцируемая функция в окрестности *стационарной* точки $P_0(x_1^0, x_2^0,\ldots, x_n^0)$, т. е. такой точки, в которой $f_{x_1}=f_{x_2}=\ldots=f_{x_n}=0$. Рассмотреть второй полный дифференциал функции f в точке P_0, $d^2f(P_0)=\sum\limits_{i,\,k=1}^{n} f^0_{x_ix_k}\,dx_i\,dx_k$. Он представляет собой квадратичную форму от n переменных $dx_1, dx_2,\ldots, dx_n$. Если эта квадратичная форма *не вырожденная*, т. е. если

$$D=\begin{vmatrix} f^0_{x_1x_1}\ldots f^0_{x_1x_n} \\ \vdots \qquad \vdots \\ f^0_{x_nx_1}\ldots f^0_{x_nx_n}\end{vmatrix}\neq 0,$$

то $d^2f(P_0)$ может быть 1) *положительно определенной* формой, 2) *отрицательно определенной* или 3) *неопределенной* формой. Доказать, что функция f имеет в точке P_0: в случае 1) минимум, в случае 2) максимум, а в случае 3) не имеет ни максимума, ни минимума.

8. Функция $f(x, y)=(y-x^2)(y-2x^2)$ имеет стационарное значение в начале O $(x=y=0)$. Доказать: 1) что вдоль любой прямой, проходящей через O, функция f имеет минимум в этой точке и 2) что f, рассматриваемая

---

[1]) В условиях задач 7, 9, 11, 12 предполагается, что читатель знаком с элементами теории квадратичных форм.

как функция свободной точки (x, y), не имеет в O ни максимума, ни минимума.

9. Дан плоский треугольник P_1, P_2, P_3, любой угол которого меньше 120°. Пользуясь критерием упр. 7, доказать, что во внутренней точке P этого треугольника, в которой $\angle P_2PP_3 = \angle P_3PP_1 = \angle P_1PP_2 = 120°$, сумма $PP_1 + PP_2 + PP_3$ действительно имеет минимум. (Ср. гл. III, § 6, n° 3, пример 4.)

10. В какой точке имеет минимум сумма $PP_1 + PP_2 + PP_3$, если в треугольнике из упр. 9 угол $P_2P_1P_3 \geqslant 120°$?

11. Для того чтобы исследовать стационарные точки функции $f = f(x_1, \ldots, x_n)$, где переменные подчинены m условиям

$$\varphi_1(x_1, \ldots, x_n) = 0, \ldots, \varphi_m(x_1, \ldots, x_n) = 0 \qquad (m<n), \tag{1}$$

построим вспомогательную функцию $F = f + \lambda_1\varphi_1 + \lambda_2\varphi_2 + \ldots + \lambda_m\varphi_m$. Предположим, что мы нашли численные значения переменных $x_1, \ldots, x_n$ и множителей λ_μ, удовлетворяющие условным уравнениям (1) и уравнениям

$$\frac{\partial F}{\partial x_1} = 0, \quad \frac{\partial F}{\partial x_2} = 0, \ldots, \frac{\partial F}{\partial x_n} = 0, \tag{2}$$

причем якобиан функций $\varphi_1, \ldots, \varphi_m$ по переменным $x_1, \ldots, x_m$ не равен нулю. Для применения критерия из упр. 7 можно действовать так. Рассматривая $x_{m+1}, \ldots, x_n$ как независимые переменные, дифференцируем m уравнений (1) и из них находим первые и вторые дифференциалы от $x_1, \ldots, x_m$ как функции от $x_{m+1}, \ldots, x_n$; затем вводим полученные для этих дифференциалов выражения в

$$d^2f = \sum_{i,k=1}^{n} f_{x_ix_k} dx_i\, dx_k + f_{x_1} d^2x_1 + \ldots + f_{x_m} d^2x_m. \tag{3}$$

Доказать нижеследующее второе правило, не требующее вычисления вторых дифференциалов $d^2x_1, \ldots, d^2x_m$. Трактуя все n переменных $x_1, \ldots, x_n$ как независимые, рассмотреть

$$d^2F = \sum F_{x_i x_k} d\,x_i\ dx_k = d^2f + \lambda_1 d^2\varphi_1 + \ldots + \lambda_m d^2\varphi_m;$$

вычислить $dx_1, \ldots, dx_m$ из уравнений

$$d\varphi_\mu = \varphi_{\mu x_1} dx_1 + \ldots + \varphi_{\mu x_n} dx_n = 0 \qquad (\mu = 1, \ldots, m)$$

и внести эти выражения в d^2F; в результате получится квадратичная форма δ^2F от переменных $dx_{m+1}, \ldots, dx_n$. Если эта квадратичная форма не вырожденная, то функция f имеет соответственно максимум либо минимум или вообще не имеет экстремума смотря по тому, является ли форма δ^2F отрицательно определенной, положительно определенной или неопределенной.

12. В задаче о нахождении максимума функции $f = x_1x_2 \ldots x_n$ при дополнительном условии $\varphi \equiv x_1 + x_2 + \ldots + x_n - a = 0$ $(a>0)$ правило неопределенных множителей дает стационарное значение функции f в точке $x_1 = x_2 = \ldots = x_n = \frac{a}{n}$. Показать с помощью правила из упр. 11, что функция f действительно имеет максимум в этой точке.

13. С помощью критерия из упр. 11 доказать, что среди всех треугольников с постоянным периметром наибольшую площадь имеет равносторонний треугольник. (Ср. гл. III, § 6, n° 7, пример 2.)

14. Кривая $x^3 + y^3 - 3axy = 0$ (декартов лист) имеет узловую точку в начале координат. Определить касательные к кривой в этой точке.

15. Построить график кривой $(y-x^2)^2-x^5=0$ и показать, что эта кривая имеет точку возврата в начале координат. Чем отличается эта точка возврата от такой же точки кривой $x^2-y^3=0$ и кривой $y^2-x^3=0$?

16. а) Доказать, что если все участвующие здесь буквы обозначают положительные величины, то стационарное значение функции $f=lx+my+$ $+nz$ при условии $x^p+y^p+z^p=c^p$ есть

$$c(l^q+m^q+n^q)^{1/q},$$

где $q=\frac{p}{p-1}$.

б) Показать, что это значение является максимумом, если $p>1$, и минимумом, если $p<1$.

17. Найти стационарные точки (x, y) функции $u=2x^3+(x-y)^2-6y$ и выяснить их характер.

18. Дана замкнутая выпуклая кривая; вокруг нее описан треугольник ABC, имеющий наименьшую площадь. Доказать, что точки касания кривой со сторонами треугольника лежат в серединах его сторон.

19. Показать, что каждая из кривых

$$(x\cos\alpha-y\sin\alpha-b)^3=c\,(x\sin\alpha+y\cos\alpha)^2,$$

где α — переменный параметр, имеет точку заострения и что все эти точки заострения лежат на окружности.

20. Пусть $C=f(a, b)$ есть собственный максимум или минимум функции $f(x, y)$ при дополнительном условии $\varphi(x, y)=C'$. Показать, что тогда вообще $C'=\varphi(a, b)$ является собственным максимумом или минимумом функции $\varphi(x, y)$ при условии $f(x, y)=C$.

21. Окружность радиуса a катится без скольжения по неподвижной прямой и несет на себе свою касательную, жестко связанную с ней (с окружностью). Примем за начало координат начальное положение точки касания, когда движущаяся касательная совпадает с неподвижной прямой, а за ось x — неподвижную прямую. Показать, что огибающая движущейся касательной выражается следующими параметрическими уравнениями:

$$x=a\,(\theta+\cos\theta\sin\theta-\sin\theta),\qquad y=a\,(\cos^2\theta-\cos\theta).$$

22. Координаты (x, y, z) точки шаровой поверхности даны уравнениями (гл. III, § 4, n° 1)

$$x=a\sin\theta\cos\varphi,\qquad y=a\sin\theta\sin\varphi,\qquad z=a\cos\theta$$

(θ — полярный угол, φ — географическая долгота). Через любую точку (θ_0, φ_0) шаровой поверхности проходит по одной кривой каждого из семейств $\theta+\varphi=\alpha$ и $\theta-\varphi=\beta$. Показать, что всякая такая пара кривых пересекается под углом $\arccos\frac{1-\sin^2\theta}{1+\sin^2\theta}$. (Ср. стр. 182.)

Показать, что радиус кривизны каждой из этих кривых равен $\frac{a(1+\sin^2\theta)^{3/2}}{(5+3\sin^2\theta)^{1/2}}$.

(Ср. упр. 10 на стр. 132.)

ГЛАВА IV

КРАТНЫЕ ИНТЕГРАЛЫ

Дифференцирование функции многих переменных и действия с частными производными получаются почти непосредственно путем приведения к соответствующим операциям для функции одной переменной. Много сложнее вопрос об интегрировании функции многих переменных и о его связи с дифференцированием, так как понятие интеграла допускает обобщение на функции многих переменных различными путями. В этой главе мы будем изучать наиболее естественное обобщение — интегралы по области или кратные интегралы, с которыми мы уже встречались в главе X первого тома. В дальнейшем, в главе V, мы займемся другими обобщениями понятия интеграла — криволинейными интегралами на плоскости, интегралами по поверхности, а также криволинейными интегралами в трехмерном пространстве. Однако всюду мы обнаружим, что в конечном итоге все вопросы интегрирования можно привести к первоначальному понятию интеграла от функции одной переменной.

§ 1. Обыкновенные интегралы как функции параметра

Прежде чем приступить к изучению новых свойств функций многих переменных, рассмотрим предварительно одно понятие, непосредственно примыкающее к вещам, нам уже знакомым.

1. Определения и примеры. Пусть в прямоугольной области $\alpha \leqslant x \leqslant \beta$, $a \leqslant y \leqslant b$ задана функция $f(x, y)$, непрерывная в этой области. Сперва представим себе величину x неизменной и, рассматривая $f(x, y)$ как функцию от одного лишь y, проинтегрируем ее по интервалу $a \leqslant y \leqslant b$. Получится

$$\int_a^b f(x, y)\, dy,$$

причем его значение еще зависит от выбора величины x, т. е. является функцией от x. В некотором смысле можно сказать, что рассматривается не один единственный интеграл, а сразу семейство таких интегралов, которые получаются при различных значениях x. Эта величина x, которую сохраняют неизменной в процессе интегрирования и которой можно придать любое значение из интервала ее изменения, называется *параметром*. Стало быть, наш обыкновенный *интеграл* является *функцией параметра* x.

Такие интегралы, являющиеся функциями параметра, часто встречаются в анализе и его приложениях.

Так, например,

$$\int_0^1 \frac{x\,dy}{\sqrt{1-x^2y^2}} = \arcsin x.$$

Этот интеграл легко вычисляется подстановкой $xy = t$.

Другой пример: при интегрировании степенной функции можно рассматривать показатель степени как параметр и в соответствии с этим писать

$$\int_0^1 y^x\,dy = \frac{1}{x+1},$$

предполагая, что $x > -1$.

На рис. 61 изображен прямоугольник, в котором определена функция $f(x, y)$; этот прямоугольник пересечен прямой AB, параллельной оси y и соответствующей выбранному значению параметра x.

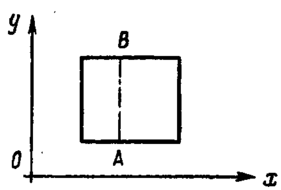


Рис. 61.

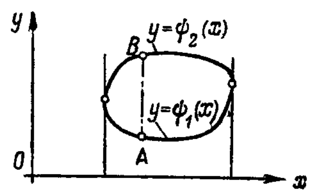


Рис. 62.

Функцию, подлежащую интегрированию по y, мы получим, рассматривая совокупность значений $f(x, y)$ как функцию одного лишь y, определенную на отрезке AB. Можно сказать, что мы интегрируем функцию $f(x, y)$ *вдоль отрезка* AB.

Это геометрическое рассмотрение подсказывает обобщение на произвольную область G, граница которой пересекается любой прямой, параллельной оси y, не более чем в двух точках (рис. 62). Если функция $f(x, y)$ определена в такой области G, то ее можно интегрировать, при фиксированном значении x, по переменной y, вдоль отрезка AB, по которому прямая, параллельная оси y, пересекает область G. В этом случае, при изменении параметра x, начальная и конечная точки промежутка интегрирования тоже будут изменяться. Другими словами, заслуживает изучения также интеграл типа

$$\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\,dy = F(x),$$

т. е. интеграл по y, содержащий параметр x, причем от x зависит не только подынтегральная функция, но и пределы интегрирования, Если, например, область определения функции $f(x, y)$ есть круг радиуса 1 с центром в начале, то рассматриваемый интеграл имеет следующий вид:

$$\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} f(x, y)\,dy.$$

2. Непрерывность и дифференцируемость интеграла как функции параметра. *Если функция $f(x, y)$ равномерно непрерывна в рассматриваемой области, то интеграл*

$$F(x)=\int_a^b f(x, y)\,dy$$

является непрерывной функцией параметра x.

В самом деле,

$$|F(x+h)-F(x)|=\left|\int_a^b (f(x+h, y)-f(x, y))\,dy\right| \leqslant$$

$$\leqslant \int_a^b |f(x+h, y)-f(x, y)|\,dy,$$

но, в силу равномерной непрерывности функции $f(x, y)$, подынтегральная функция в правой части становится сколь угодно малой при достаточно малом h, к тому же равномерно относительно y, а отсюда вытекает наше утверждение.

Из доказанной непрерывности функции $F(x)$ следует, что ее *можно интегрировать по параметру* x на отрезке $[\alpha, \beta]$:

$$\int_\alpha^\beta F(x)\,dx=\int_\alpha^\beta\left(\int_a^b f(x, y)\,dy\right)dx.$$

Стоящий справа интеграл называется *повторным интегралом*; его принято также писать в следующем более простом виде:

$$\int_\alpha^\beta\left(\int_a^b f(x, y)\,dy\right)dx=\int_\alpha^\beta\int_a^b f(x, y)\,dy\,dx.$$

Естественно возникает вопрос: а можно ли интеграл от функции, содержащей параметр, *дифференцировать* по этому параметру? Рассмотрим сначала тот случай, когда пределы интегрирования постоянны, и предположим, что $f(x, y)$ имеет непрерывную частную производную f_x во всем замкнутом прямоугольнике G. Тогда естественна мысль попытаться получить производную интеграла по параметру x следующим путем: вместо того, чтобы сначала интегрировать, а потом дифференцировать, выполнить эти операции в обратном порядке, т. е. сперва дифференцировать $f(x, y)$ по x, а затем полученную производную f_x интегрировать по y. В уточненном виде эта мысль выражается следующей теоремой:

Если подынтегральная функция $f(x, y)$ имеет в замкнутом прямоугольнике G: $\alpha \leqslant x \leqslant \beta$, $a \leqslant y \leqslant b$ непрерывную производную по параметру x, то можно дифференцировать по параметру под символом интеграла, т. е. при $\alpha \leqslant x \leqslant \beta$

$$\frac{d}{dx} F(x) = \frac{d}{dx} \int_a^b f(x, y)\, dy = \int_a^b f_x(x, y)\, dy.$$

Доказательство. Если оба значения, x и $x+h$, принадлежат интервалу $\alpha \leqslant x \leqslant \beta$, то приращение функции $F(x)$ запишется так:

$$F(x+h) - F(x) = \int_a^b f(x+h, y)\, dy - \int_a^b f(x, y)\, dy =$$
$$= \int_a^b \{f(x+h, y) - f(x, y)\}\, dy.$$

Так как $f(x, y)$ дифференцируема, то ее приращение, стоящее под знаком последнего интеграла, можно преобразовать по теореме о среднем значении (из дифференциального исчисления) в ее обычной форме

$$f(x+h, y) - f(x, y) = hf_x(x+\theta h, y) \qquad (0 < \theta < 1).$$

Участвующая здесь величина θ зависит от y и не исключено, что она не будет непрерывной функцией от y. Однако для дальнейшего это не имеет значения, ибо из равенства $f_x(x+\theta h, y) = \frac{f(x+h, y) - f(x, y)}{h}$ видно, что $f_x(x+\theta h, y)$ является непрерывной функцией от x и y, а стало быть, является интегрируемой функцией.

Наша цель — доказать, что

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = F'(x) = \int_a^b f_x(x, y)\, dy.$$

Для этого достаточно показать, что разность между отношением приращений $\frac{F(x+h) - F(x)}{h}$ и интегралом в правой части стремится к нулю при $h \to 0$. Но

$$\left| \frac{F(x+h) - F(x)}{h} - \int_a^b f_x(x, y)\, dy \right| = \left| \int_a^b \{f_x(x+\theta h, y) - f_x(x, y)\}\, dy \right| \leqslant$$
$$\leqslant \int_a^b |f_x(x+\theta h, y) - f_x(x, y)|\, dy.$$

Так как, по условию, производная f_x непрерывна в замкнутом прямоугольнике G, а стало быть, и равномерно непрерывна, то

$$|f_x(x+\theta h, y)-f_x(x, y)|<\varepsilon,$$

где ε — положительная величина, не зависящая от x и y и стремящаяся к нулю при $h\to 0$. Поэтому последний интеграл меньше, чем $\int\limits_a^b \varepsilon\, dy=\varepsilon(b-a)$, и, следовательно, стремится к нулю при $h\to 0$. Теорема доказана.

Попутное замечание. С помощью этого правила можно получить простое доказательство теоремы, которая уже была доказана в гл. II, § 3, стр. 71, о том, что при вычислении смешанной производной g_{xy} функции $g(x, y)$ порядок дифференцирования по x и по y безразличен, если только g_{yx} непрерывна. Для доказательства положим $f(x, y)=g_y(x, y)$. Тогда

$$g(x, y)=g(x, a)+\int\limits_a^y f(x, \eta)\, d\eta.$$

Так как $f(x, y)$ имеет в прямоугольнике $\alpha\leqslant x\leqslant\beta$, $a\leqslant y\leqslant b$ непрерывную производную по x, то

$$g_x=g_x(x, a)+\int\limits_a^y f_x(x, \eta)\, d\eta,$$

откуда

$$g_{xy}=f_x(x, y).$$

С другой стороны, $g_{yx}=\dfrac{\partial}{\partial x}g_y=f_x(x, y)$. Поэтому $g_{xy}=g_{yx}$.

Пример. Пусть $F(x)=\int\limits_0^1 \dfrac{x\, dy}{\sqrt{1-x^2y^2}}=\arcsin x$. Следовательно, $F'(x)=\dfrac{1}{\sqrt{1-x^2}}$. Но тот же результат получится, если сначала дифференцировать по параметру под знаком интеграла, а затем интегрировать

$$F'(x)=\int\limits_0^1 \frac{dy}{\sqrt{(1-x^2y^2)^3}}=\frac{1}{\sqrt{1-x^2}}.$$

Подобным же образом можно установить непрерывность интеграла, содержащего параметр, и правило дифференцирования по параметру под знаком интеграла в том более сложном случае, когда *пределы интегрирования* являются функциями параметра. Если, например, требуется дифференцировать по x функцию

$$F(x)=\int\limits_{\psi_1(x)}^{\psi_2(x)} f(x, y)\, dy,$$

то ее представляют как функцию трех параметров u, v, x так:

$$F(x)=\int_u^v f(x, y)\,dy=\Phi(u, v, x),$$

где $u=\psi_1(x)$ и $v=\psi_2(x)$. В данном случае рассматривается область G, ограниченная кривыми $y=\psi_1(x)$ и $y=\psi_2(x)$ и прямыми $x=\alpha$ и $x=\beta$. Предполагаем при этом, что $\psi_1(x)$ и $\psi_2(x)$ имеют непрерывные производные по x в интервале его изменения и что функция $f(x, y)$ непрерывно дифференцируема в области, охватывающей область G. По правилу цепочки имеем

$$F'(x)=\frac{\partial\Phi}{\partial x}+\frac{\partial\Phi}{\partial u}\frac{du}{dx}+\frac{\partial\Phi}{\partial v}\frac{dv}{dx}.$$

Пользуясь формулой Ньютона — Лейбница (т. I, гл. II, § 4, n° 4) приходим к следующему результату:

$$F'(x)=\int_{\psi_1(x)}^{\psi_2(x)} f_x(x, y)\,dy+\psi_2'(x)f[x, \psi_2(x)]-\psi_1'(x)f[x, \psi_1(x)].$$

Примеры. 1) Пусть $F(x)=\int_0^x \sin(xy)\,dy$; тогда

$$F'(x)=\int_0^x y\cos(xy)\,dy+\sin(x^2).$$

2) Рассмотрим совокупность интегралов

$$F_n(x)=\int_0^x \frac{(x-y)^n}{n!}f(y)\,dy, \qquad F_0(x)=\int_0^x f(y)\,dy,$$

где n — любое положительное целое число, а $f(y)$ есть функция одного только y, непрерывная в рассматриваемом интервале. Наше правило дает теперь

$$F_n'(x)=F_{n-1}(x);$$

[второй член исчезает, ибо подынтегральная функция обращается в нуль при $y=\psi_2(x)=x$, а третий член исчезает, ибо $\psi_1'(x)=0$]. Из полученной формулы в сочетании с тем, что $F_0'(x)=f(x)$, вытекает, что

$$F_n^{(n+1)}(x)=f(x).$$

Стало быть, $F_n(x)$ есть такая функция, $(n+1)$-я производная которой равна $f(x)$ и которая, вместе со своими n первыми производными, обращается в нуль при $x=0$; она получается из $F_{n-1}(x)$ интегрированием от 0 до x. Следовательно, функцию $F_n(x)$ можно получить, интегрируя $f(x)$ повторно $n+1$ раз между пределами 0 и x, т. е. повторное n-кратное интегрирование функции $f(x)$ от 0 до x можно заменить одним-единственным интегрированием функции $\frac{(x-y)^n}{n!}f(y)$ по y:

$$\underbrace{\int_0^x dx \int_0^x dx \ldots \int_0^x f(x)\,dx}_{n+1 \text{ раз}} = \frac{1}{n!}\int_0^x (x-y)^n f(y)\,dy.$$

Этот результат был получен в первом томе другим путем (см. т. I, гл. IV, § 4, n° 8).

Правило дифференцирования интеграла по параметру часто остается справедливым, хотя бы дифференцирование под знаком интеграла и привело к функции, не всюду непрерывной. Вместо того, чтобы пользоваться в таких случаях общими критериями с громоздкой формулировкой, целесообразнее, когда в этом возникает надобность, исследовать в каждом отдельном случае, допустимо ли дифференцировать под знаком интеграла.

В качестве примера рассмотрим эллиптический интеграл (т. I, стр. 283, 289)

$$F(k) = \int_{-1}^{+1} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \qquad (k^2 < 1).$$

Функция

$$f(k, x) = \frac{1}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

имеет разрыв при $x=1$ и $x=-1$, но интеграл сохраняет смысл как несобственный. Формальное дифференцирование по параметру k дает

$$F'(k) = \int_{-1}^{+1} \frac{kx^2\,dx}{\sqrt{(1-x^2)(1-k^2x^2)^3}}.$$

Для того чтобы исследовать, справедливо ли это равенство, надо повторить то рассуждение, которое привело нас раньше к общей формуле дифференцирования. Мы получаем

$$\frac{F(k+h)-F(k)}{h} = \int_{-1}^{+1} f_k(k+\theta h, x)\,dx = \int_{-1}^{+1} \frac{(k+\theta h)\,x^2\,dx}{\sqrt{(1-x^2)\,[1-(k+\theta h)^2 x^2]^3}}.$$

Это выражение отличается от интеграла, полученного путем формального дифференцирования, на

$$\Delta = \int_{-1}^{+1} \frac{x^2}{\sqrt{1-x^2}}\left(\frac{k+\theta h}{\sqrt{[1-(k+\theta h)^2 x^2]^3}} - \frac{k}{\sqrt{(1-k^2x^2)^3}}\right)dx.$$

Требуется доказать, что этот интеграл при $h \to 0$ стремится к нулю. Для этого мы окружим k интервалом $k_0 < k < k_1$, который не содержит значений $k=1$ и $k=-1$, и выберем h настолько малым, чтобы значение $k+\theta h$

непременно попало в этот интервал. Функция $\frac{k}{\sqrt{(1-k^2x^2)^3}}$ непрерывна в замкнутой области $-1 \leqslant x \leqslant 1$, $k_0 \leqslant k \leqslant k_1$, следовательно, она там равномерно непрерывна. Поэтому разность

$$\left| \frac{k+\theta h}{\sqrt{[1-(k+\theta h)^2 x^2]^3}} - \frac{k}{\sqrt{(1-k^2x^2)^3}} \right|$$

остается меньше некоторого числа ε, не зависящего от x и k и стремящегося к нулю, когда $h \to 0$. Вместе с тем интеграл Δ остается по абсолютному значению меньше чем

$$\int_{-1}^{+1} \frac{x^2\,dx}{\sqrt{1-x^2}}\,\varepsilon = M\varepsilon,$$

где M — определенное число, не зависящее от ε, т. е. интеграл Δ стремится к нулю, когда h стремится к нулю, что и требовалось доказать.

Следовательно, дифференцирование под знаком интеграла в данном случае законно. Подобные же рассуждения приводят к цели и в других случаях.

Несобственные интегралы (с бесконечным промежутком интегрирования), зависящие от параметра, будут рассмотрены в Дополнениях к этой главе, § 4.

Упражнения

1. Вычислить $F(y) = \int_0^1 x^{y-1}(y \ln x + 1)\,dx$.

2. Даны дважды непрерывно дифференцируемая функция $f(x, y)$ и функция $u(x, y, z)$, определенная следующим образом:

$$u(x, y, z) = \int_0^{2\pi} f(x + z\cos\varphi,\ y + z\sin\varphi)\,d\varphi.$$

Доказать, что

$$z(u_{xx} + u_{yy} - u_{zz}) - u_z = 0.$$

3*. Даны дважды непрерывно дифференцируемая функция $f(x)$ и интеграл

$$u(x, t) = \frac{1}{t^{p-2}} \int_{-t}^{+t} f(x+y)(t^2-y^2)^{\frac{p-3}{2}}\,dy \qquad (p > 1).$$

Доказать, что

$$u_{xx} = \frac{p-1}{t} u_t + u_{tt}.$$

4. Функцию Бесселя $J_0(x)$ можно определить следующим образом:

$$J_0(x) = \frac{1}{\pi} \int_{-1}^{+1} \frac{\cos xt}{\sqrt{1-t^2}}\,dt.$$

Доказать, что $J_0'' + \frac{1}{x} J_0' + J_0 = 0$.

5. При любом положительном целом индексе n функцию Бесселя $J_n(x)$ можно определить так:

$$J_n(x) = \frac{x^n}{1\cdot 3\cdot 5\,\ldots\,(2n-1)\pi}\int\limits_{-1}^{+1}\cos xt\,(1-t^2)^{n-\frac{1}{2}}\,dt.$$

Доказать, что

$$\text{а)}\quad J_n'' + \frac{1}{x}J_n' + \left(1-\frac{n^2}{x^2}\right)J_n = 0 \qquad (n\geqslant 0),$$

$$\text{б)}\quad J_{n+1} = J_{n-1} - 2J_n' \qquad (n\geqslant 1),$$

$$\text{в)}\quad J_1 = -J_0'.$$

§ 2. Интеграл от непрерывной функции по плоской или пространственной области

1. Интеграл по плоской области (двойной интеграл) как объем. Первое и самое важное обобщение понятия интеграла подсказано геометрической интуицией, как, впрочем, и сам обыкновенный интеграл. Рассмотрим замкнутую область G плоскости xy, ограниченную одной или несколькими криволинейными дугами, имеющими непрерывно вращающуюся касательную, и функцию $z=f(x, y)$, непрерывную в G. Предположим сначала, что функция f не принимает отрицательных значений, и представим себе ее геометрическое изображение в виде куска поверхности в пространстве xyz, расположенного над областью G. Через каждую точку границы области G проведем прямую, параллельную оси z; совокупность этих прямых образует цилиндрическую поверхность, перпендикулярную к плоскости xy. Построенная нами цилиндрическая поверхность, область G плоскости xy и заданная поверхность $z=f(x, y)$ отграничивают часть пространства или тело. Мы ставим себе целью найти (или, точнее, *определить*, так как такое определение еще не было дано) объем V описанного выше тела. Для прямоугольной области это было сделано подробно в т. I, гл. X, стр. 580; к тому же поставленная задача так похожа на задачу определения площади, приведшую к понятию обыкновенного интеграла, что читателю, надеемся, ясен естественный путь к определению требуемого объема.

Мы исходим из того, что знаем формулу для объема цилиндра, ограниченного двумя параллельными плоскостями; этот объем равен произведению площади основания на высоту; объем же тела, состоящего из нескольких частей, объемы которых известны, равен сумме объемов всех его частей. Трудность заключается в том, что наше тело, ограниченное сверху кривой поверхностью, невозможно разбить точно на такие цилиндры, но зато искомый объем можно определить как *предел* суммы цилиндрических объемов.

Разобьем область G на N частичных областей или ячеек $G_1, G_2, \ldots, G_N$ так, чтобы каждая из них имела кусочно гладкую границу [1]). Через граничную линию каждой ячейки проведем цилиндрическую поверхность с образующими, параллельными оси z. Эти поверхности разрежут наше тело на N столбиков. Объем каждого такого столбика (рис. 63) не легче выразить, чем объем V всего тела, но объем столбика можно заключить между двумя границами. В каждой ячейке G_i функция $f(x, y)$ принимает свое наименьшее значение m_i и наибольшее значение M_i; площадь ячейки G_i обозначим через ΔG_i. Тогда объем столбика, имеющего основание G_i, заключается между объемами двух цилиндров с тем же основанием G_i, но с разными высотами: m_i и M_i, т. е. между числами $m_i \Delta G_i$ и $M_i \Delta G_i$. Объем V всего тела равен сумме объемов всех N столбиков, а стало быть, удовлетворяет неравенству

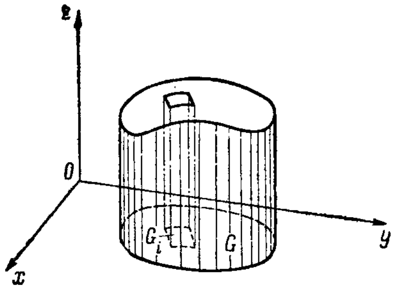


Рис. 63.

$$\sum_{i=1}^{N} m_i \Delta G_i \leqslant V \leqslant \sum_{i=1}^{N} M_i \Delta G_i.$$

Сумма $\sum\limits_{i=1}^{N} m_i \Delta G_i$ называется *нижней суммой* нашего разбиения, а сумма $\sum\limits_{i=1}^{N} M_i \Delta G_i$ — *верхней его суммой*.

Для сокращения речи назовем *поперечником* или *диаметром* области наибольшее расстояние между любыми двумя ее точками. Сделаем теперь наше разбиение (с помощью дробления ячеек) все более частым, заставляя число N безгранично возрастать, а наибольший из диаметров ячеек G_i стремиться к нулю. Тогда наглядно ясно (в Дополнениях к настоящей главе это будет строго доказано), что нижняя и верхняя суммы будут все более сближаться, так что *объем V можно рассматривать как общий предел верхней и нижней сумм при $N \to \infty$*.

Очевидно, тот же самый предел получится, если вместо m_i или M_i брать каждый раз любое число, заключенное между m_i и M_i, например, значение $f(\xi_i, \eta_i)$ функции f в некоторой точке (ξ_i, η_i) ячейки G_i.

2. Общее аналитическое определение двойного интеграла. Эти понятия, полученные путем геометрической интуиции, надлежит

[1]) Напомним, что кривая называется *кусочно гладкой*, если она состоит из конечного числа дуг, имеющих в некоторой системе координат уравнения вида $y = \varphi(x)$ или $x = \psi(y)$, где φ и ψ — непрерывные функции с непрерывными производными.

теперь с должной общностью аналитически формулировать и уточнить, не прибегая прямо к интуиции. Для этого действуем следующим образом. Рассматриваем функцию $f(x, y)$, определенную и непрерывную в замкнутой области G. Показанным выше способом разбиваем область G с помощью кусочно гладких дуг на N ячеек $G_1, G_2, \ldots, G_N$ с площадями $\Delta G_1, \Delta G_2, \ldots, \Delta G_N$. В каждой из ячеек G_i выбираем произвольно по точке (ξ_i, η_i) и строим сумму

$$V_N = \sum_{i=1}^{N} f_i \Delta G_i,$$

где $f_i = f(\xi_i, \eta_i)$ есть значение функции f в точке (ξ_i, η_i). Эта сумма называется *интегральной суммой.*

Теперь можно формулировать основную теорему:

Если число N неограниченно возрастает и при этом наибольший из диаметров ячеек G_i стремится к нулю, то интегральная сумма V_N стремится к пределу. Этот предел V не зависит ни от способа разбиения области G, ни от выбора точек (ξ_i, η_i) в ячейках G_i. Предел V называется (двойным) интегралом функции $f(x, y)$ по области G и обозначается символом

$$\iint_G f(x, y)\, dS.$$

Дополнение к теореме. Интегральная сумма V_N будет стремиться к тому же пределу V, если суммирование производить только по тем ячейкам G_i, которые лежат полностью *внутри* области G, т. е. не имеют общих точек с ее границей.

Формулированная выше теорема существования *двойного интеграла от непрерывной функции* должна быть доказана чисто аналитическим путем. Это доказательство, совершенно аналогичное доказательству теоремы существования обыкновенного интеграла, будет дано в Дополнениях к этой главе, § 1.

Примечание. Можно построить суммы несколько более общего вида, имеющие тот же предел. А именно, нет необходимости выбирать в качестве множителей f_i при ΔG_i такие значения, которые функция $f(x, y)$ действительно принимает в некоторых точках (ξ_i, η_i) ячеек G_i. Достаточно выбирать для f_i числа, отличающиеся от значений $f(\xi_i, \eta_i)$ на поправки, стремящиеся равномерно к нулю, когда количество N ячеек стремится к бесконечности, а наибольший из диаметров ячеек стремится к нулю. Другими словами, вместо значений $f(\xi_i, \eta_i)$ функции f можно брать числа $f_i = f(\xi_i, \eta_i) + \varepsilon_{i, N}$, причем $|\varepsilon_{i, N}| < \varepsilon_N$, а $\lim\limits_{N \to \infty} \varepsilon_N = 0$, и тогда

$$\lim \sum_{i=1}^{N} [f(\xi_i, \eta_i) + \varepsilon_{i, N}]\, \Delta G_i = \lim \sum_{i=1}^{N} f(\xi_i, \eta_i)\, \Delta G_i = \iint_G f(x, y)\, dS.$$

Эта теорема почти очевидна. В самом деле, абсолютная величина разности этих двух сумм есть

$$\left| \sum_{i=1}^{N} \varepsilon_{i,N} \Delta G_i \right| < \varepsilon_N \sum_{i=1}^{N} \Delta G_i = \varepsilon_N S \to 0$$

(S обозначает площадь всей области G). Например, если $f(x,y)=P(x,y)\,Q(x,y)$, то можно выбрать $f_i=P_iQ_i$, где P_i и Q_i суть наибольшие значения функций P и Q в ячейке G_i, которые, как правило, достигаются этими функциями в разных точках.

Поясним теперь понятие двойного интеграла путем рассмотрения различных способов разбиения. Простейшим является тот случай, когда область G есть *прямоугольник* $a\leqslant x\leqslant b$, $\alpha\leqslant y\leqslant\beta$; тогда естественно выбрать в качестве частичных областей (ячеек) G_i тоже прямоугольники со сторонами, параллельными осям координат. Для их построения разделим интервал на оси x на m равных частей длиною $h=\frac{b-a}{m}$, а интервал на оси y на n равных частей длиною $k=\frac{\beta-\alpha}{n}$. Точки деления обозначим соответственно буквами $x_0=a$, $x_1, x_2, \ldots, x_m=b$ и $y_0=\alpha, y_1, y_2, \ldots, y_n=\beta$. Через точки деления оси x проведем прямые, параллельные оси y, а через точки деления оси y — прямые, параллельные оси x. Ясно, что $N=mn$. Все ячейки являются равными прямоугольниками с площадью $\Delta G_i=hk=\Delta x\cdot\Delta y$, если положить $h=\Delta x$, $k=\Delta y$. В качестве точки $P_i(\xi_i, \eta_i)$ можно взять любую точку в каждом прямоугольнике и затем составить интегральную сумму

$$\sum_{i=1}^{nm} f(\xi_i, \eta_i)\,\Delta x\Delta y$$

по всем прямоугольникам нашего разбиения.

Заставим теперь числа m и n одновременно неограниченно возрастать; тогда наша сумма будет стремиться к пределу, и ее пределом будет двойной интеграл функции f по прямоугольнику G.

Каждый из прямоугольников-ячеек можно пометить двумя индексами μ и ν, соответствующими координатам левой нижней вершины прямоугольника: $x=a+\mu h$, $y=\alpha+\nu k$, причем ν пробегает целые значения от 0 до $(n-1)$, а μ — целые значения от 0 до $(m-1)$. При таком расположении прямоугольников по индексам μ и ν естественно и удобно записать нашу сумму в виде двойной суммы:

$$\sum_{\mu=0}^{m-1}\sum_{\nu=0}^{n-1} f(\xi_\mu, \eta_\nu)\,\Delta x\Delta y,$$

но для этого мы должны выбрать точки P_i так, чтобы они лежали по m точек на прямых, параллельных оси x, и по n точек на прямых, параллельных оси y.

Разбиение области G на ячейки проведением прямых, параллельных осям координат, оказывается часто целесообразным и в том случае, если эта область не является прямоугольной (рис. 64). Мы строим на плоскости xy прямоугольную сетку прямых, параллельных осям:

$$x=\mu h \quad (\mu=0, \pm 1, \pm 2, \ldots), \qquad y=\nu k \quad (\mu=0, \pm 1, \pm 2, \ldots),$$

причем h и k — произвольно выбранные числа, определяющие размеры каждой ячейки сети. Затем мы отбираем все те прямоугольники, которые полностью лежат внутри заданной области G. Эти-то прямоугольники и будут нашими ячейками G_i. Правда, они не заполняют всей области G, которая содержит кроме них еще и некоторое количество ячеек G_l, примыкающих к границе. Эти последние ячейки уже не являются прямоугольниками: граница каждой такой ячейки содержит, в качестве своей части, дугу граничной кривой области G. Однако на основании дополнения к теореме существования (начало этого номера) двойной интеграл от функции f по области G можно вычислить, беря предел суммы, распространенной только на *внутренние* ячейки области G.

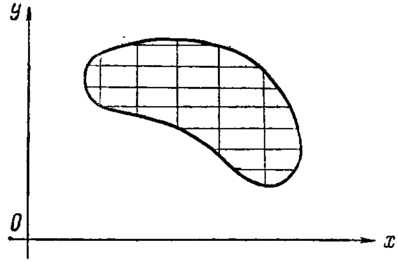


Рис. 64.

Часто применяется и другой способ разбиения — с помощью сетки *полярных координат* (рис. 65). Пусть полюс O системы полярных координат лежит внутри нашей области. Весь угол 2π, окружающий полюс, делим на n частей, величиной $\Delta\theta = \frac{2\pi}{n} = h$ каждая, и выбираем еще одну величину $\Delta r = k$. Теперь строим полупрямые $\theta = \nu h$ ($\nu = 0, 1, 2, \ldots, n-1$), выходящие из полюса, и концентрические окружности с центром в полюсе и радиусами $r_\mu = \mu k$ ($\mu = 1, 2, \ldots$). Эти два семейства линий разбивают плоскость на ячейки, изображенные на рис. 65. Те ячейки, которые лежат полностью внутри области G, обозначим через G_i, а их площади — через ΔG_i. Двойной интеграл функции $f(x, y) = F(r, \theta)$ по области G можно теперь рассматривать как предел суммы

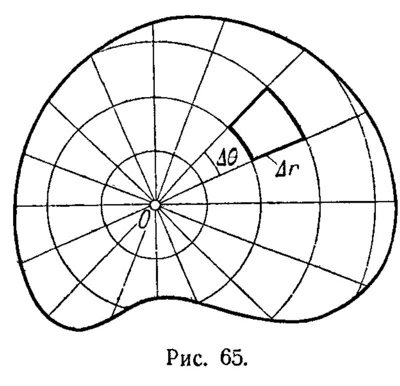


Рис. 65.

$$\sum F(r_i, \theta_i)\,\Delta G_i,$$

где $P_i(r_i, \theta_i)$ есть произвольно выбранная точка ячейки G_i. Сумма распространяется на все ячейки G_i, лежащие полностью *внутри* области G, а переход к пределу осуществляется путем одновременного стремления $h \to 0$ и $k \to 0$.

Площадь каждой ячейки G_i, лежащей в кольце между окружностями с радиусами $r_\mu = \mu k$ и $r_{\mu+1} = (\mu + 1)k$, выражается так:

$$\Delta G_i = \frac{1}{2}(r_{\mu+1}^2 - r_\mu^2)\Delta\theta = \frac{1}{2}(2\mu + 1)k^2h;$$

это вытекает из известной формулы элементарной геометрии для площади кругового сектора. Полученному выражению можно придать другой вид, введя в рассмотрение радиус промежуточной окружности $\rho_\mu = \frac{r_{\mu+1} + r_\mu}{2}$, а именно:

$$\Delta G_i = \rho_\mu \Delta r \Delta\theta = \rho_\mu kh.$$

3. Примеры. 1) Простейший пример дает функция $f(x, y) = 1$. Интегральная сумма для этой функции не зависит, очевидно, от способа разбиения области G и всегда равна ее площади. Поэтому двойной интеграл этой функции по области G как предел интегральной суммы тоже равен площади области G. Этого и следовало ожидать, так как искомый интеграл равен объему вертикального цилиндра высоты 1, для которого область G служит основанием.

2) Вычислим $\iint\limits_Q x\,dS$, где область Q — квадрат $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. Теперь $f(x, y) = x$. Этот интеграл дает, очевидно, объем прямой треугольной призмы, одно из боковых ребер которой совпадает с осью y. Объем призмы, а следовательно, и наш двойной интеграл равен $\frac{1}{2}$. Проверим этот результат, пользуясь аналитическим определением двойного интеграла. Для этого разобьем наш квадрат Q на квадратики со сторонами длиною $h = \frac{1}{n}$, параллельными осям. За точки (ξ_i, η_i) примем нижние левые вершины каждой квадратной ячейки. Тогда каждый из квадратиков вертикальной полосы, левый край которой имеет абсциссу μh, дает интегральной сумме член $\mu h \cdot h^2 = \mu h^3$, а все n квадратиков этой полосы дают сумму $n\mu h^3 = \mu h^2$. Это выражение надо просуммировать от $\mu = 0$ до $\mu = n - 1$:

$$\sum_{\mu=0}^{n-1} \mu h^2 = \frac{n(n-1)}{2} h^2 = \frac{1-h}{2}.$$

Это и есть наша интегральная сумма; ее предел при $h \to 0$ равен $\frac{1}{2}$, что и подтверждает результат, полученный геометрическим путем.

3) Функция $f(x, y) = \varphi(x)\psi(y)$, т. е. представляет собой произведение функции, зависящей только от x, на функцию, зависящую только от y. Область интегрирования есть прямоугольник R со сторонами, параллельными осям координат

$$a \leqslant x \leqslant b, \qquad \alpha \leqslant y \leqslant \beta.$$

Разбиение области интегрирования мы произведем тем же способом, что и на стр. 249, и воспользуемся той же двух-индексной нумерацией. Тогда искомый двойной интеграл будет пределом суммы

$$\sum_{\mu=0}^{m-1} \sum_{\nu=0}^{n-1} \varphi(\xi_\mu)\psi(\eta_\nu)\Delta x \Delta y,$$

которую можно представить в виде произведения двух сумм:

$$\left[\sum_{\mu=0}^{m-1}\varphi(\xi_\mu)\,\Delta x\right]\left[\sum_{\nu=0}^{n-1}\psi(\eta_\nu)\,\Delta y\right].$$

Но при $\Delta x\to 0$ и $\Delta y\to 0$ первая из этих сумм имеет своим пределом обыкновенный интеграл $\int\limits_a^b\varphi(x)\,dx$, а предел второй суммы есть $\int\limits_\alpha^\beta\psi(y)\,dy$. Таким образом, мы доказали следующее правило:

Если функция $f(x,y)$ может быть представлена как произведение двух функций $\varphi(x)$ и $\psi(y)$, то двойной интеграл функции f по прямоугольной области R: $a\leqslant x\leqslant b$, $\alpha\leqslant y\leqslant\beta$ равен произведению двух обыкновенных интегралов

$$\iint\limits_R\varphi(x)\,\psi(y)\,dS=\int\limits_a^b\varphi(x)\,dx\int\limits_\alpha^\beta\psi(y)\,dy.$$

4) В заключение решим задачу, в которой само собой напрашивается разбиение с помощью *сетки полярных координат.* Требуется вычислить двойной интеграл от функции $f(x,y)=\sqrt{1-x^2-y^2}$ по кругу радиуса 1 с центром в начале координат. Итак, область G определяется неравенством $x^2+y^2\leqslant 1$. Ясен и геометрический смысл задачи: требуется найти объем полушара радиуса 1.

Строим сетку полярных координат, как было показано выше. Выражаем нашу функцию в полярных координатах $f=\sqrt{1-r^2}$. Ячейка области G, лежащая между окружностями радиусов $r_\mu=\mu k$ и $r_{\mu+1}=(\mu+1)\,k$ и между полупрямыми $\theta=\nu h$ и $\theta=(\nu+1)h$ $\left(\text{вспомним, что } h=\frac{2\pi}{n}\right)$, дает интегральной сумме слагаемое

$$\frac{1}{2}\sqrt{1-\left(\frac{r_{\mu+1}+r_\mu}{2}\right)^2}\,(r_{\mu+1}^2-r_\mu^2)\,h=\sqrt{1-\rho_\mu^2}\,\rho_\mu kh,$$

причем в качестве значения функции в ячейке $G_{\mu\nu}$ мы выбрали то ее значение, которое она принимает на средней окружности с радиусом $\rho_\mu=\frac{r_{\mu+1}+r_\mu}{2}$. Все ячейки, лежащие в одном кольце, порождают одинаковые слагаемые, а так как число их $n=\frac{2\pi}{h}$, то общий вклад всего кольца равен $2\pi\sqrt{1-\rho_\mu^2}\,\rho_\mu k$, а вся интегральная сумма равна

$$\sum_{\mu=0}^{m-1}2\pi\sqrt{1-\rho_\mu^2}\,\rho_\mu k.$$

Эта же сумма, как известно, имеет своим пределом обыкновенный интеграл

$$2\pi\int\limits_0^1 r\sqrt{1-r^2}\,dr=-\frac{2\pi}{3}\sqrt{(1-r^2)^3}\,\Bigg|_0^1=\frac{2\pi}{3}.$$

Поэтому

$$\iint\limits_G\sqrt{1-x^2-y^2}\,dS=\frac{2\pi}{3}$$

в полном согласии с известной формулой для объема шара.

4. Обозначения, дополнения, основные правила. С разбиением области интегрирования на прямоугольники связано общепринятое обозначение двойного интеграла, ведущее свое начало от Лейбница. Исходя из записи интегральной суммы

$$\sum_{\nu=0}^{n-1}\sum_{\mu=0}^{m-1} f(\xi_\mu, \eta_\nu)\,\Delta x\,\Delta y,$$

распространенной на все прямоугольные ячейки, переход от суммы к ее пределу отмечают тем, что заменяют двойной символ суммирования двойным знаком интеграла, а *произведение* величин Δx и Δy *символом* $dx\,dy$. Таким образом, вместо записи

$$\iint_G f(x, y)\,dS,$$

в которой площадь ΔG_i заменена символом dS, двойной интеграл часто обозначают так:

$$\iint_G f(x, y)\,dx\,dy.$$

Подчеркнем еще раз, что символ $dx\,dy$ не следует рассматривать как произведение дифференциалов; он представляет собой только символическое указание на переход от суммы к ее пределу при $n\to\infty$ и $m\to\infty$ и одновременном стремлении Δx и Δy к нулю.

Ясно, что и у двойного интеграла, в полной аналогии с обыкновенным интегралом, обозначение *переменных интегрирования* не играет никакой роли, так что

$$\iint_G f(x, y)\,dx\,dy = \iint_G f(u, v)\,du\,dv = \iint f(\xi, \eta)\,d\xi\,d\eta.$$

Когда вводилось понятие двойного интеграла, мы видели, что для положительной функции $f(x, y)$ он дает объем вертикального цилиндра, построенного на области G как основании и усеченного сверху поверхностью $z = f(x, y)$. Однако при аналитическом определении двойного интеграла функция $f(x, y)$ отнюдь не должна быть непременно всюду положительной; она может быть отрицательной или же изменять свой знак в области интегрирования — в последнем случае поверхность $z = f(x, y)$ пересекает область G. Таким образом, двойной интеграл дает искомый объем с определенным знаком: со знаком плюс, если кусок поверхности $z = f(x, y)$, вырезанный цилиндром, лежит выше плоскости xy, и со знаком минус, если он лежит ниже этой плоскости. Если же часть, вырезанная цилиндром из поверхности $z = f(x, y)$, состоит из нескольких кусков, из которых одни лежат выше плоскости xy, а другие — ниже ее, то двойной интеграл по области G дает алгебраическую сумму соответствующих объемов, снабженных надлежащими знаками. Стало быть, в частности, двойной интеграл может обратиться в нуль, хотя бы подынтегральная функция и не была тождественно равна нулю.

Для двойных интегралов справедливы следующие основные правила, представляющие дословное повторение соответствующих правил для обыкновенных интегралов. Да и доказательства те же — на основании определения интеграла как предела интегральной суммы.

1) Если c — постоянная, то

$$\iint\limits_G cf(x, y)\, dS = c \iint\limits_G f(x, y)\, dS,$$

т. е. *постоянный множитель можно вынести из-под знака двойного интеграла.*

2)
$$\iint\limits_G [f(x, y) \pm g(x, y)]\, dS = \iint\limits_G f(x, y)\, dS \pm \iint\limits_G g(x, y)\, dS,$$

т. е. *интеграл от алгебраической суммы двух функций равен алгебраической сумме интегралов от слагаемых функций.* Ясно, что это правило распространяется на любое конечное число слагаемых функций.

Пользуясь этими двумя правилами и правилом вычисления двойного интеграла по прямоугольной области от функции вида $\varphi(x)\psi(y)$ (пример 3 предыдущего номера), мы теперь уже в состоянии интегрировать по такой области любой многочлен $P(x, y)$.

3) Если область интегрирования G является соединением двух областей G_1 и G_2, не имеющих общих внутренних точек, то

$$\iint\limits_G f(x, y)\, dS = \iint\limits_{G_1} f(x, y)\, dS + \iint\limits_{G_2} f(x, y)\, dS,$$

т. е. *при сложении (объединении) областей подлежат сложению интегралы по этим областям.*

5. Свойства двойного интеграла, его оценка и теорема о среднем значении. Ряд свойств обыкновенного определенного интеграла (см. т. I, гл. II, § 7) распространяется без изменения на двойные интегралы:

а) Если $f(x, y) \geqslant 0$ в области G, то

$$\iint\limits_G f(x, y)\, dS \geqslant 0,$$

если же $f(x, y) \leqslant 0$ в области G, то

$$\iint\limits_G f(x, y)\, dS \leqslant 0.$$

Это нетрудно доказать, исходя из определения двойного интеграла.

б) Если в области G всюду $f(x, y) \geqslant \varphi(x, y)$, то

$$\iint\limits_G f(x, y)\, dS \geqslant \iint\limits_G \varphi(x, y)\, dS.$$

Это свойство легко выводится из предыдущего свойства.

в)

$$\left|\iint\limits_G f(x, y)\,dS\right| \leqslant \iint\limits_G |f(x, y)|\,dS.$$

Доказательство. Из свойства б) вытекает, что

$$-\iint\limits_G |f(x, y)|\,dS \leqslant \iint\limits_G f(x, y)\,dS \leqslant \iint\limits_G |f(x, y)|\,dS,$$

а это и есть лишь другая запись доказываемого неравенства.

г) Так как функция $f(x, y)$ предполагается непрерывной в замкнутой области G, то она принимает в этой области свое наименьшее значение m и наибольшее значение M. Из свойства б) вытекает, что

$$m\iint\limits_G dS \leqslant \iint\limits_G f(x, y)\,dS \leqslant M\iint\limits_G dS$$

или

$$mS \leqslant \iint\limits_G f(x, y)\,dS \leqslant MS,$$

где S обозначает площадь области G. Наш интеграл равен поэтому произведению площади S на некоторое число μ, промежуточное между m и M:

$$\iint\limits_G f(x, y)\,dS = \mu S.$$

Это равенство называется *формулой оценки* двойного интеграла. Относительно значения промежуточного числа μ более точные общие указания сделать невозможно.

[Как и в случае непрерывной функции одной переменной, функция $f(x, y)$, непрерывная в замкнутой области G, непременно принимает значение μ в некоторой внутренней точке (ξ, η) в области G, так что $\mu = f(\xi, \eta)$, и формула оценки запишется так:

$$\iint\limits_G f(x, y)\,dS = Sf(\xi, \eta) \quad \text{или} \quad \frac{\iint\limits_G f(x, y)\,dS}{S} = f(\xi, \eta).$$

Эти формулы выражают *теорему о среднем значении* в теории двойных интегралов. Доказательство последней формулы см. в книге: Смирнов В. И., Курс высшей математики, т. II, «Наука», 1967, стр. 197—198. Смысл этого названия следующий. Отношение

$$\frac{\iint\limits_G f(x, y)\,dS}{S}$$

называется *средним значением* функции $f(x, y)$ в области G. Это — важное понятие, часто встречающееся в физике и технике. Теорема о среднем значении утверждает, стало быть, что функция, непрерывная в замкнутой области G, обязательно принимает свое среднее значение в этой области по крайней мере в одной ее внутренней точке.]

д) Формулу оценки двойного интеграла и теорему о среднем значении можно несколько обобщить. Пусть наряду с непрерывной функцией $f(x, y)$ задана еще произвольная функция $p(x, y)$, сохраняющая постоянный знак и непрерывная в G; тогда

$$\iint_G p(x, y) f(x, y)\, dS = \mu \iint_G p(x, y)\, dS,$$

где μ — некоторое число, лежащее между наименьшим и наибольшим значениями функции f в области G. [И далее:

$$\iint_G p(x, y) f(x, y)\, dS = f(\xi, \eta) \iint_G p(x, y)\, dS,$$

где (ξ, η) — некоторая внутренняя точка области G.]

е) На основании изложенного нетрудно доказать, что *двойной интеграл непрерывно зависит от подынтегральной функции*. Точнее, если $f(x, y)$ и $\varphi(x, y)$ — две функции, удовлетворяющие неравенству

$$|f(x, y) - \varphi(x, y)| < \varepsilon,$$

где ε есть положительное число, одинаковое для всей области G (с площадью S), то

$$\left| \iint_G f(x, y)\, dS - \iint_G \varphi(x, y)\, dS \right| < \varepsilon S,$$

т. е. меньше числа, которое стремится к нулю при $\varepsilon \to 0$.

ж) Тем же путем убеждаемся, что *двойной интеграл непрерывно зависит от области*. Это означает следующее. Пусть две области G_1 и G_2 получаются одна из другой добавлением или удалением частей, общая площадь которых не превышает ε, и пусть $f(x, y)$ есть функция, непрерывная в обеих областях, причем $|f(x, y)| < M$, где M — постоянное число. Тогда

$$\left| \iint_{G_2} f(x, y)\, dS - \iint_{G_1} f(x, y)\, dS \right| < M\varepsilon$$

и, стало быть, меньше числа, которое стремится к нулю, когда $\varepsilon \to 0$, т. е. когда разность площадей обеих областей стремится к нулю. Доказательство этого факта тоже понятно само собой на основании последней теоремы предыдущего номера.

Это свойство позволяет вычислить двойной интеграл по области G с любой точностью, заменив область G ее подобластью G', площадь которой отличается достаточно мало от площади G. Можно, например, вписать в область G прямолинейный многоугольник, площадь которого сколь угодно мало отличается от площади G. Можно также заменить область G многоугольником, стороны которого поочередно параллельны координатным осям; такой многоугольник составлен из прямоугольников со сторонами, параллельными осям.

6. Интегралы по трехмерным и многомерным областям (тройные и многократные интегралы). Все сказанное выше о двойном интеграле, т. е. об интеграле по области плоскости xy, можно распространить без каких-либо осложнений и не нуждаясь в новых идеях на области трех и большего числа измерений. Как ввести, например, понятие тройного интеграла, т. е. интеграла по трехмерной области G? Пусть дана функция $f(x, y, z)$, непрерывная в замкнутой области G. Разобьем область G с помощью конечного числа поверхностей на частичные области (трехмерные ячейки), которые мы обозначим через $G_1, G_2, \ldots, G_N$. В каждой ячейке G_i выберем по произвольной точке (ξ_i, η_i, ζ_i) и построим сумму (так называемую интегральную сумму)

$$\sum_{i=1}^{N} f(\xi_i, \eta_i, \zeta_i)\Delta G_i,$$

где ΔG_i теперь обозначает объем ячейки G_i. Суммирование производится по всем ячейкам G_i или, если это нам удобнее, лишь по тем ячейкам, которые не примыкают к границе области G. Если мы станем теперь увеличивать безгранично число N ячеек таким образом, чтобы наибольший из их диаметров стремился к нулю, то интегральная сумма будет стремиться к пределу, не зависящему ни от способа последовательного дробления ячеек, ни от выбора точек (ξ_i, η_i, ζ_i) во всех ячейках. Этот предел называется (*тройным*) *интегралом функции* $f(x, y, z)$ *по области* G и обозначается так:

$$\iiint\limits_G f(x, y, z)\,dV.$$

Если мы, в частности, разобьем область G на конгруэнтные ячейки, имеющие форму прямоугольных параллелепипедов с ребрами, параллельными осям координат, длиною соответственно Δx, Δy, Δz, то все внутренние (т. е. не примыкающие к границе) ячейки будут иметь одинаковый объем $\Delta x\,\Delta y\,\Delta z$. Совершенно так же, как у двойного интеграла, этот способ разбиения и переход от суммы к ее пределу отмечают введением для тройного интеграла, наряду с прежним, еще и следующего символа:

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz.$$

Все свойства, установленные выше для двойных интегралов, сохраняют силу и для тройных интегралов с учетом, конечно, необходимых изменений в обозначениях.

Понятие интеграла по многомерной области, т. е. по области пространства, число измерений которого больше трех, вводится таким же путем, надо только предварительно установить понятие объема многомерной области. Ограничимся пока прямоугольными областями. Прямоугольной областью называется множество точек $(x_1, x_2, \ldots, x_n)$, координаты которых удовлетворяют неравенствам следующего вида:

$$a_1 \leqslant x_1 \leqslant a_1 + h_1, \quad a_2 \leqslant x_2 \leqslant a_2 + h_2, \ \ldots, \ a_n \leqslant x_n \leqslant a_n + h_n.$$

Объем такого прямоугольника мы будем считать, в порядке определения, равным произведению $h_1 h_2 \ldots h_n$. Прямоугольные области мы будем разбивать на прямоугольные ячейки такого же типа. Тогда определение многомерного интеграла не нуждается в дальнейших разъяснениях. Такой интеграл по n-мерной области G обозначают так:

$$\iint\limits_{G}\ldots\int f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n.$$

Для непрямоугольных областей и для разбиений более общего вида придется опираться на абстрактное определение объема, которое будет дано в § 1 Дополнений к этой главе. Впрочем, если не считать § 3 Дополнений, мы удовольствуемся рассмотрением интегралов в пространстве с числом измерений, не бо́льшим трех.

7. Дифференцирование по области. Масса и плотность. Когда мы имеем обыкновенный интеграл от функции одной переменной, то подынтегральная функция получается обратно из интеграла процессом дифференцирования: берем интеграл по промежутку длины h, делим результат на h и находим предел отношения при $h \to 0$. Для функции одной переменной этот факт выражал основную связь между дифференциальным и интегральным исчислением и нашел свое наглядное истолкование с помощью физических понятий общей массы и плотности. Такая же связь существует и для кратных интегралов, т. е. для интегралов от функции многих переменных по многомерной области.

Рассмотрим интеграл

$$\iint\limits_{B} f(x, y)\, dS \quad \text{или} \quad \iiint\limits_{B} f(x, y, z)\, dV$$

от непрерывной функции f двух или трех переменных по области B (двумерной или трехмерной), содержащей внутри себя фиксированную точку $P(x_0, y_0)$ или $P(x_0, y_0, z_0)$, и пусть область B имеет площадь ΔS в первом случае и объем ΔV — во втором. Разделим значение интеграла на площадь ΔS (объем ΔV) области B; согласно теореме о среднем значении из n° 5, это отношение будет равно значению подынтегральной функции в некоторой внутренней точке Q области B. Заставим теперь диаметр $D(B)$ области B, окружающей точку P, стремиться к нулю, так что и площадь ΔS (объем ΔV) будет стремиться к нулю; тогда в силу непрерывности функции f, ее значение в точке Q будет стремиться к значению функции в точке P. Стало быть, получаем следующие предельные равенства:

$$\lim_{D(B)\to 0} \frac{1}{\Delta S} \iint_B f(x, y)\, dS = f(x_0, y_0),$$

$$\lim_{D(B)\to 0} \frac{1}{\Delta V} \iiint_B f(x, y, z)\, dV = f(x_0, y_0, z_0).$$

Этот предельный переход, соответствующий дифференцированию интеграла от функции одной переменной, называется *дифференцированием кратного интеграла по области*, а полученный предел — производной по области в точке P. Итак, *производная кратного интеграла по области интегрирования равна подынтегральной функции.*

Этот результат дает возможность истолковать роль подынтегральной функции по отношению к интегралу также и в случае многих переменных с помощью физических понятий *плотности* и *общей массы*. Представим себе, что масса некоторого вещества распределена непрерывно по двумерной или трехмерной области, т. е. в любой достаточно малой частичной области содержится сколь угодно малая масса. Для того чтобы определить плотность вещества в точке P, рассматриваем сначала некоторую окрестность B точки P, имеющую площадь ΔS (или объем ΔV), и делим массу, содержащуюся в этой окрестности, на площадь ΔS (объем ΔV). Полученное частное естественно назвать *средней плотностью* вещества в окрестности B точки P. Заставим теперь диаметр окрестности B стремиться к нулю; тогда предел средней плотности в окрестности B даст *плотность вещества в точке* P, предполагая, конечно, что такой предел, не зависящий от выбора последовательности окрестностей B_i, существует. Если обозначить эту плотность через $\mu(x, y)$ или $\mu(x, y, z)$, то сразу видно, что только что описанный процесс нахождения плотности представляет собой не что иное, как дифференцирование по области интеграла

$$\iint_G \mu(x, y)\, dS \quad \text{или} \quad \iiint_G \mu(x, y, z)\, dV \qquad (*)$$

в любой точке P области G. Естественно предположить, что этот интеграл, взятый по области G, дает *общую массу* вещества, содержащегося в G, какова бы ни была эта область. Этот вывод верен, но для его оправдания остается сделать еще один шаг.

В последних рассуждениях исходным пунктом было заранее заданное распределение массы (на плоскости, или в пространстве), так что любой указанной (в известных границах) области G соответствовала известная, заданная масса. Таким образом, масса была задана как *функция области*. С другой стороны, двойной и тройной интегралы тоже являются *функциями* своей *области интегрирования*. И вот мы только доказали, что первоначально заданная масса как функция области и кратный интеграл (*) имеют тождественные производные по области. Остается еще доказать, что исходная функция области определяется своей производной по области однозначно. Доказательство это нетрудно, но мы его приводить не будем. Оно напоминает доказательство теоремы Гейне — Бореля о покрытии.

С точки зрения физики понятие плотности распределения массы в точке P и представление массы с помощью выражений (*) означают, конечно, идеализацию. В том, что такая идеализация разумна, т. е. что она описывает действительное положение дел с достаточной точностью, заключается как раз уже *физическое* допущение.

Добавим еще следующее замечание. Построенные нами понятия сохраняют свой математический смысл и в том случае, когда функция точки $\mu(P)$ не является всюду положительной. Отрицательные плотности и массы встречаются и в физике, например в учении о распределении электрического заряда.

§ 3. Приведение кратного интеграла к повторному обыкновенному интегралу

Для создания общего метода вычисления кратных интегралов решающее значение имеет тот факт, что всякий кратный интеграл можно привести к повторному интегралу, т. е. к последовательному вычислению обыкновенных интегралов. Тем самым все ранее изученные методы неопределенного интегрирования ставятся на службу вычислению кратных интегралов.

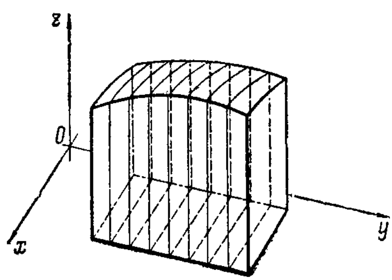


Рис. 66.

1. Двойной интеграл по прямоугольной области. В качестве области интегрирования возьмем сначала прямоугольник R: $a \leqslant x \leqslant b, \alpha \leqslant y \leqslant \beta$ и рассмотрим функцию $f(x,y)$, непрерывную в R. Приведение двойного интеграла $\iint\limits_R f(x, y)\,dS$ к последовательному вычислению двух простых интегралов мы сначала покажем путем наглядных соображений, которые будут затем аналитически уточнены. Наш двойной интеграл дает объем прямоугольного параллелепипеда, усеченного поверхностью $z=f(x, y)$. Вспомним, что этот объем мы сперва разбивали на столбики, основания которых — прямоугольники со сторонами $h=\frac{b-a}{m}$ и $k=\frac{\beta-\alpha}{n}$, а затем устремляли m и n к ∞ таким образом, что h и k стремились к нулю. Но эти столбики можно объединить в слои или пластинки (рис. 66) толщиною $k=\frac{\beta-\alpha}{n}$, ограниченные семейством параллельных плоскостей $y=\alpha+\nu k$ ($\nu=0, 1, 2, \ldots, n$). Таким образом, тело разбито на n пластинок, которые тем тоньше, чем больше n, и их суммарный объем равен значению двойного интеграла. Объем ν-й пластинки приближенно равен произведению толщины k на площадь ее левой боковой «стенки», т. е. выражению

$$k\int\limits_a^b f(x,\ \alpha+\nu k)\,dx,$$

— только приближенно, ибо площади левой и правой стенки несколько отличаются друг от друга.

Введем следующее обозначение:

$$\int\limits_a^b f(x,\ y)\,dx=\varphi(y);$$

тогда наш объем приближенно выразится суммой

$$\sum_{\nu=0}^{n-1}\varphi(\alpha+\nu k)\,k.$$

При $n\to\infty$ толщина пластинки $k=\frac{\beta-\alpha}{n}$ неограниченно убывает, и эта сумма будет иметь своим пределом $\int\limits_\alpha^\beta \varphi(y)\,dy$. Стало быть, естественно ожидать, что

$$\iint\limits_R f(x,\ y)\,dS=\int\limits_\alpha^\beta \varphi(y)\,dy.$$

Это равенство, которое еще нуждается в строгом доказательстве, дает искомое приведение двойного интеграла к последовательному вычислению двух простых интегралов:

Сначала надо интегрировать $f(x, y)$ при фиксированном y по x в промежутке от a до b. Этот интеграл $\int\limits_a^b f(x,\ y)\,dx=\varphi(y)$ является функцией параметра y, которую надлежит затем интегрировать от α до β. В результате получается

$$\iint\limits_R f(x,\ y)\,dS=\int\limits_\alpha^\beta dy\int\limits_a^b f(x,\ y)\,dx,$$

т. е. двойной интеграл равен повторному интегралу.

Для того чтобы доказать аналитически это утверждение, вернемся к определению двойного интеграла, данному в § 2, n° 2. Полагая $h=\frac{b-a}{m}$ и $k=\frac{\beta-\alpha}{n}$, имеем

$$\iint\limits_R f(x,\ y)\,dS=\lim_{\substack{m\to\infty\\ n\to\infty}}\sum_{\nu=1}^{n}\sum_{\mu=1}^{m} f(a+\mu h,\ \alpha+\nu k)\,hk,$$

предполагая, что существование двойного интеграла по области R, т. е. предела, записанного в правой части, доказано; притом самый этот предел надо понимать в том смысле, что абсолютная величина разности между значением двойного интеграла и двойной суммой правой стороны меньше любого сколь угодно малого положительного числа ε, коль скоро числа m и n оба превышают некоторое число N, зависящее только от ε.

Смысл следующего ниже рассуждения состоит просто в том, что двойной предел при одновременном возрастании m и n разбивают на два последовательных простых предельных перехода: сперва по $m \to \infty$ при фиксированном n, и лишь затем по $n \to \infty$ (ср. Дополнения к гл. II, § 2, n° 1).

Для внутренней суммы (по μ) введем обозначение

$$\Phi_\nu = \sum_{\mu=1}^{m} f(a+\mu h,\ \alpha+\nu k)\,h$$

при фиксированном ν. Тогда двойная сумма равна следующей сумме по ν:

$$\sum_{\nu=1}^{n} \Phi_\nu k.$$

Выберем теперь любое постоянное $n > N$; тогда, как мы уже знаем,

$$\left| \iint_R f(x,\ y)\,dS - k\sum_{\nu=1}^{n} \Phi_\nu \right| < \varepsilon$$

при любом значении числа m, лишь бы только оно тоже было больше, чем N. Сохраняя n неизменным, будем теперь неограниченно увеличивать m; при этом предельном переходе, на основании определения обычного интеграла, выражение Φ_ν имеет своим пределом интеграл

$$\int_a^b f(x,\ \alpha+\nu k)\,dx = \varphi(\alpha+\nu k),$$

а написанная выше абсолютная величина разности не должна превзойти ε:

$$\left| \iint_R f(x,\ y)\,dS - k\sum_{\nu=1}^{N} \varphi(\alpha+\nu k) \right| \leqslant \varepsilon.$$

Это неравенство, повторяем, сохраняет силу при произвольно малом ε для всех n, превышающих число N, зависящее только от ε. Если совершим теперь предельный переход $n \to \infty$ (так что $k \to 0$), то неравенство не нарушится, но при этом (на основании определения обычного интеграла и в силу непрерывности функции $\varphi(y) = \int_a^b f(x,y)\,dx$)

$$\lim_{n\to\infty} k\sum_{\nu=1}^{n} \varphi(\alpha+\nu k) = \int_\alpha^\beta \varphi(y)\,dy.$$

Следовательно, приходим к неравенству

$$\left|\iint_R f(x, y)\,dS-\int_\alpha^\beta \varphi(y)\,dy\right|\leqslant \varepsilon.$$

Так как левая часть есть постоянное число, а ε можно было выбрать сколь угодно малым, то левая часть может быть только нулем. Стало быть,

$$\iint_R f(x, y)\,dS=\int_\alpha^\beta \varphi(y)\,dy=\int_\alpha^\beta dy\int_a^b f(x, y)\,dx.$$

Тем самым доказано, что двойной интеграл по прямоугольной области может быть представлен как повторный обычный интеграл; другими словами, интегрирование по области R плоскости xy приводится к двум последовательным простым интегрированиям.

Так как переменные x и y можно в нашем рассуждении поменять ролями, то без дополнительного доказательства ясно, что повторный интеграл можно записать и в обратном порядке, т. е.

$$\iint_R f(x, y)\,dS=\int_a^b dx\int_\alpha^\beta f(x, y)\,dy.$$

2. Следствия. Изменение порядка интегрирования. Дифференцирование под знаком интеграла. Из последних двух формул вытекает соотношение

$$\int_\alpha^\beta dy\int_a^b f(x, y)\,dx=\int_a^b dx\int_\alpha^\beta f(x, y)\,dy,$$

т. е. *при повторном интегрировании непрерывной функции с постоянными промежутками интегрирования можно изменить порядок обоих интегрирований.*

Этой теореме можно дать еще и следующую формулировку:

Если функция $f(x, y)$ непрерывна в замкнутом прямоугольнике $R(a\leqslant x\leqslant b,\ \alpha\leqslant y\leqslant\beta)$, то вычисление интеграла по параметру y от выражения $\int_a^b f(x, y)\,dx$ можно выполнить, интегрируя по y под знаком интеграла, т. е. интегрируя сперва по y, а затем по x.

Эта теорема вполне соответствует правилу дифференцирования интеграла по параметру (§ 1).

Для того чтобы извлечь еще одно следствие, будем рассматривать один из верхних пределов, например b, как переменный параметр. Двойной интеграл можно тогда дифференцировать по этому параметру. По теореме о производной определенного интеграла по верхнему пределу имеем

$$\frac{\partial}{\partial b}\iint\limits_R f(x, y)\,dx\,dy = \int\limits_a^\beta f(b, y)\,dy.$$

Аналогично, рассматривая β как переменный параметр, получим

$$\frac{\partial}{\partial \beta}\iint\limits_R f(x, y)\,dx\,dy = \int\limits_a^b f(x, \beta)\,dx.$$

Из последних двух равенств, дифференцируя первое из них по β или второе по b, получим

$$\frac{\partial^2}{\partial b\,\partial\beta}\iint\limits_R f(x, y)\,dx\,dy = f(b, \beta).$$

Полученный результат выражается словесно так: *производная повторного интеграла по одному из верхних пределов равна обычному определенному интегралу по той стороне прямоугольника, которая соответствует этому верхнему пределу. Смешанная производная по обоим верхним пределам b и β равна значению подынтегральной функции в вершине (b, β) прямоугольника*[1].

Теорема об изменении порядка интегрирования имеет много приложений. Например, с помощью этой теоремы часто удается вычислить обычный определенный интеграл от функции, первообразная которой не выражается в конечном виде через элементарные функции.

В качестве примера вычислим несобственный интеграл

$$I = \int\limits_0^\infty \frac{e^{-ax} - e^{-bx}}{x}\,dx,$$

сходящийся при $a > 0$, $b > 0$. Величину I можно представить в виде повторного интеграла

$$I = \int\limits_0^\infty dx \int\limits_a^b e^{-xy}\,dy.$$

Но этот интеграл несобственный, и к нему недопустимо сразу применить теорему об изменении порядка интегрирования. Поэтому мы сначала напишем

$$I = \lim_{T\to\infty}\int\limits_0^T dx \int\limits_a^b e^{-xy}\,dy,$$

[1]) Обращаем внимание читателя на связь этой формулы с теоремой об изменении порядка дифференцирования (гл. II, § 3, n° 3) и рекомендуем продумать самостоятельно, в какой мере эти два факта можно считать эквивалентными.

и только теперь изменим порядок интегрирования

$$I=\lim_{T\to\infty}\int_a^b \frac{1-e^{-Ty}}{y}\,dy=\ln\frac{b}{a}-\lim_{T\to\infty}\int_a^b \frac{e^{-Ty}}{y}\,dy.$$

Интеграл в правой части

$$\int_a^b \frac{e^{-Ty}}{y}\,dy=\int_{Ta}^{Tb}\frac{e^{-y}}{y}\,dy$$

стремится к нулю при $T\to\infty$, а стало быть,

$$I=\int_0^\infty \frac{e^{-ax}-e^{-bx}}{x}\,dx=\ln\frac{b}{a}.$$

Тем же способом можно доказать следующую общую теорему. Если $f(t)$ есть кусочно гладкая функция при $t\geqslant 0$ и если $\int_1^\infty \frac{f(t)}{t}\,dt$ сходится, то

$$I=\int_0^\infty \frac{f(ax)-f(bx)}{x}\,dx=f(0)\ln\frac{b}{a}\qquad (a>0,\ b>0).$$

Действительно, и здесь можно представить искомый интеграл в виде повторного:

$$I=\int_0^\infty dx\int_b^a f'(xy)\,dy,$$

а потом действовать совершенно так же, как в последнем примере.

3. Распространение результата на двумерные области более общего вида. Перенесение полученного результата с прямоугольных областей на области более общего вида не вызывает никаких затруднений. Начнем с выпуклой области G (рис. 67), т. е. с такой области, граница которой пересекается любой прямой не более чем в двух точках, если только в состав границы не входит целый отрезок прямой (это допускается). Пусть область G лежит между опорными прямыми (т. I, стр. 314 и 316 и т. II, упр. 16 на стр. 121) $x=a$, $x=b$ и $y=\alpha$, $y=\beta$. Для точек области G абсцисса x изменяется в интервале $a\leqslant x\leqslant b$, а ордината y — в интервале $\alpha\leqslant y\leqslant\beta$. Рассмотрим интегралы

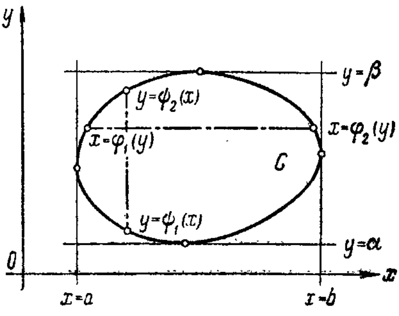


Рис. 67.

$$\int_{\varphi_1(y)}^{\varphi_2(y)} f(x,y)\,dx \quad\text{и}\quad \int_{\psi_1(x)}^{\psi_2(x)} f(x,y)\,dy;$$

они берутся вдоль отрезков, по которым прямые $y=\text{const}$ и прямые $x=\text{const}$ (соответственно) пересекают нашу область. Пределы интегрирования $\varphi_1(y)$ и $\varphi_2(y)$ обозначают абсциссы точек пересечения прямых $y=\text{const}$ с границей области, а $\psi_1(x)$ и $\psi_2(x)$ — ординаты точек пересечения прямых $x=\text{const}$ с границей области. Стало быть, интеграл $\int\limits_{\varphi_1(y)}^{\varphi_2(y)} f(x, y)\,dx$ является функцией параметра y, который содержится не только в подынтегральной функции, но и в каждом из пределов интегрирования. Аналогично $\int\limits_{\psi_1(x)}^{\psi_2(x)} f(x, y)\,dy$ является функцией параметра x. Искомое представление двойного интеграла по области G в виде повторного интеграла записывается теперь так:

$$\iint\limits_G f(x, y)\,dS=\int\limits_\alpha^\beta dy\int\limits_{\varphi_1(y)}^{\varphi_2(y)} f(x, y)\,dx=\int\limits_a^b dx\int\limits_{\psi_1(x)}^{\psi_2(x)} f(x, y)\,dy.$$

Для доказательства выберем на дуге $y=\psi_2(x)$ конечное число точек так, чтобы расстояние между любыми двумя соседними точками было меньше положительного числа δ. Каждую пару соседних точек мы соединим ломаной линией, лежащей в G и состоящей из одного горизонтального и одного вертикального отрезка прямой (рис. 68). Такую же операцию мы произведем с нижней границей $y=\psi_1(x)$. Таким образом

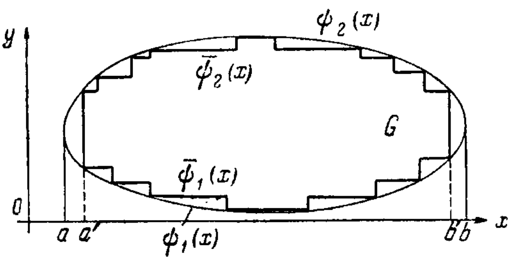


Рис. 68.

мы получили область $\bar G$, содержащуюся в G и разбивающуюся на конечное число прямоугольных областей. Нижняя часть границы области $\bar G$ имеет уравнение $y=\bar\psi_1(x)$, а верхняя часть — уравнение $y=\bar\psi_2(x)$, где $\bar\psi_1(x)$ и $\bar\psi_2(x)$ — кусочно непрерывные функции (рис. 68). На основании доказанной теоремы для прямоугольных областей имеем

$$\iint\limits_{\bar G} f(x, y)\,dS=\int\limits_{a'}^{b'} dx\int\limits_{\bar\psi_1(x)}^{\bar\psi_2(x)} f(x, y)\,dy.$$

Так как функции $\psi_1(x)$ и $\psi_2(x)$ равномерно непрерывны, то при безграничном увеличении числа точек деления и при $\delta\to 0$ функции $\bar\psi_1(x)$ и $\bar\psi_2(x)$ стремятся равномерно к $\psi_1(x)$ и $\psi_2(x)$, и в результате

$$\lim_{\delta\to 0}\int\limits_{\bar\psi_1(x)}^{\bar\psi_2(x)} f(x, y)\,dy=\int\limits_{\psi_1(x)}^{\psi_2(x)} f(x, y)\,dy$$

равномерно относительно x. Отсюда вытекает, что

$$\lim_{\delta\to 0}\int_{a'}^{b'} dx \int_{\bar\psi_1(x)}^{\bar\psi_2(x)} f(x, y)\, dy = \int_a^b dx \int_{\psi_1(x)}^{\psi_2(x)} f(x, y)\, dy.$$

С другой стороны, при $\delta \to 0$ область $\overline{G}$ стремится к области G. Поэтому

$$\lim_{\delta\to 0}\iint_{\overline{G}} f(x, y)\, dS = \iint_G f(x, y)\, dS.$$

Комбинируя три из полученных равенств, приходим к окончательному результату:

$$\iint_G f(x, y)\, dS = \int_a^b dx \int_{\psi_1(x)}^{\psi_2(x)} f(x, y)\, dy.$$

Доказательство второго утверждения можно провести тем же способом.

Сходное рассуждение применимо и к невыпуклым областям, имеющим, например, форму, изображенную на рис. 69. На этот раз мы предполагаем только, что граница области пересекается любой прямой, параллельной оси x или оси y в конечном числе точек (впрочем, и здесь допускается случай, когда такая прямая имеет с границей области общий отрезок). Под символом $\int f(x, y)\, dy$ мы будем теперь понимать *сумму* интегралов по y от функции $f(x, y)$ при постоянном x, распространенных на все отрезки, которые являются общими у рассматриваемой прямой $x = \text{const}$ и области G. У невыпуклой области число таких отрезков может быть больше единицы. Это число может при переходе через некоторую прямую $x = \xi$ измениться скачком (как на рис. 69), причем и интеграл $\int f(x, y)\, dy$ как функция параметра x может иметь конечный разрыв при $x = \xi$. Однако, двойной интеграл по области G можно и в этом случае представить в виде повторного интеграла

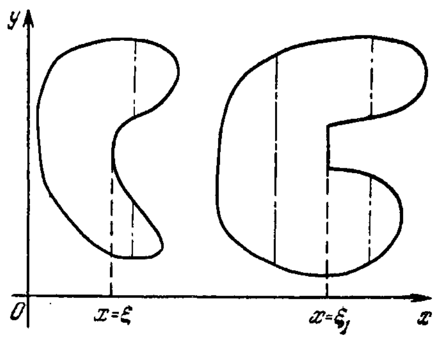


Рис. 69.

$$\iint_G f(x, y)\, dS = \int_a^b dx \int f(x, y)\, dy,$$

где интегрирование по x надо произвести по всему интервалу от a до b между крайними вертикальными опорными прямыми области G.

В доказательство этого предложения не придется вносить существенных изменений. Ясно, что возможно и приведение к другому повторному интегралу:

$$\iint_G f(x, y)\, dS = \int_\alpha^\beta dy \int f(x, y)\, dx.$$

Примеры. 1) Пусть область G есть круг $x^2+y^2 \leqslant 1$. Тогда преобразование двойного интеграла в повторный имеет следующий вид:

$$\iint_G f(x, y)\, dS = \int_{-1}^{+1} dx \int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} f(x, y)\, dy = \int_{-1}^{+1} dy \int_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}} f(x, y)\, dx.$$

2) Если область интегрирования — круговое кольцо между окружностями $x^2+y^2=1$ и $x^2+y^2=4$ (рис. 70), то

$$\iint_G f(x, y)\, dx\, dy = \int_{-2}^{-1} dx \int_{-\sqrt{4-x^2}}^{+\sqrt{4-x^2}} f(x, y)\, dy + \int_1^2 dx \int_{-\sqrt{4-x^2}}^{+\sqrt{4-x^2}} f(x, y)\, dy +$$
$$+ \int_{-1}^{+1} dx \int_{-\sqrt{4-x^2}}^{-\sqrt{1-x^2}} f(x, y)\, dy + \int_{-1}^{+1} dx \int_{+\sqrt{1-x^2}}^{+\sqrt{4-x^2}} f(x, y)\, dy.$$

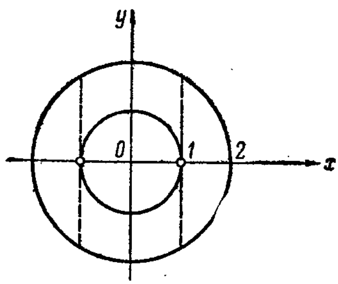


Рис. 70.

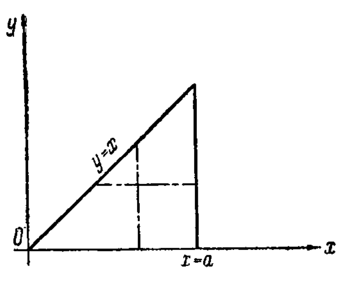


Рис. 71.

3) Область интегрирования — треугольник, ограниченный прямыми $x=y$, $y=0$ и $x=a$, $a>0$ (рис. 71). Если первое интегрирование произвести по x, то получим

$$\iint_G f(x, y)\, dS = \int_0^a dy \int_y^a f(x, y)\, dx,$$

а если в первую очередь интегрировать по y, то

$$\iint_G f(x, y)\, dS = \int_0^a dx \int_0^x f(x, y)\, dy.$$

Сравнение обоих результатов приводит к полезной *формуле Дирихле*:

$$\int_0^a dx \int_0^x f(x, y)\, dy = \int_0^a dy \int_y^a f(x, y)\, dx.$$

В частности, если $f(x, y) = f(y)$, т. е. зависит только от y, то формула Дирихле дает

$$\int_0^a dx \int_0^x f(y)\,dy = \int_0^a f(y)(a-y)\,dy.$$

Отсюда видно, что если интеграл $\int_0^x f(y)\,dy$ проинтегрировать еще раз (по x), то результат можно представить в виде простого интеграла (это частный случай более общего предложения, стр. 243, пример 2).

4. Приведение тройного интеграла к повторному. Если число измерений области интегрирования больше двух, то соответствующие теоремы настолько аналогичны, что можно ограничиться их формулировкой без доказательств. Начнем с интеграла по трехмерной *прямоугольной* области $R_3(x_0 \leqslant x \leqslant x_1,\ y_0 \leqslant y \leqslant y_1,\ z_0 \leqslant z \leqslant z_1)$ от функции $f(x, y, z)$, непрерывной в этой области. Такой тройной интеграл

$$V = \iiint\limits_{R_3} f(x, y, z)\,dV$$

можно привести различными способами к обыкновенным и двойным интегралам. Так, например,

$$\iiint\limits_{R_3} f(x, y, z)\,dV = \int_{z_0}^{z_1} dz \iint\limits_{R_2} f(x, y, z)\,dx\,dy,$$

где $\iint\limits_{R_2} f(x, y, z)\,dx\,dy$ есть двойной интеграл от функции $f(x, y, z)$ по двумерной прямоугольной области $R_2(x_0 \leqslant x \leqslant x_1,\ y_0 \leqslant y \leqslant y_1)$, причем z рассматривается при этом интегрировании как постоянный параметр, так что этот двойной интеграл является функцией параметра z. Таким же путем можно вместо z выделить в качестве параметра x или y.

Тройной интеграл V можно также представить в виде повторного интеграла

$$V = \int_{x_0}^{x_1} dx \int_{y_0}^{y_1} dy \int_{z_0}^{z_1} f(x, y, z)\,dz,$$

т. е. привести к последовательному вычислению трех обычных интегралов. При этом сначала вычисляют интеграл $\int_{z_0}^{z_1} f(x, y, z)\,dz$ при фиксированных значениях x и y, затем интеграл $\int_{y_0}^{y_1} dy \int_{z_0}^{z_1} f(x, y, z)\,dz$ при фиксированном x, и, наконец, интегрированием по x от x_0 до x_1 получают значение тройного интеграла V. Повторный интеграл можно с таким же успехом составить в любом другом порядке: например, сначала интегрировать по x при постоянных y и z, затем по

y при постоянном z и, наконец, — по z. Что порядок интегрирования можно изменить как угодно — вытекает из того факта, что повторный интеграл всегда равен тройному интегралу по области R_3. Итак, имеем следующую теорему:

Повторный интеграл от непрерывной функции $f(x, y, z)$ по замкнутой прямоугольной области R_3 не зависит от порядка интегрирования.

Едва ли нуждается в особом объяснении, как приводить к повторному тройной интеграл по непрямоугольной области G. Достаточно, пожалуй, показать такое приведение для шаровой области $x^2+y^2+z^2 \leqslant 1$:

$$\iiint_G f(x, y, z)\,dx\,dy\,dz = \int_{-1}^{+1} dx \int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} dy \int_{-\sqrt{1-x^2-y^2}}^{+\sqrt{1-x^2-y^2}} f(x, y, z)\,dz.$$

Упражнения

Вычислить интегралы в упр. 1—8:

1. $\iint x^2y^2\,dx\,dy$ по кругу $x^2+y^2 \leqslant 1$.

2. $\displaystyle\iint \frac{x^3+y^3-3xy(x^2+y^2)}{(x^2+y^2)^{3/2}}\,dx\,dy$ по кругу $x^2+y^2 \leqslant 1$.

3. $\iiint (x^2+y^2+z^2)\,xyz\,dx\,dy\,dz$ по шару $x^2+y^2+z^2 \leqslant a^2$.

4. $\iiint z\,dx\,dy\,dz$ по области, определенной неравенствами $x^2+y^2 \leqslant z^2$, $x^2+y^2+z^2 \leqslant 1$.

5. $\iiint (x+y+z)\,x^2y^2z^2\,dx\,dy\,dz$ по области $x+y+z \leqslant 1$, $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$.

6. $\displaystyle\iiint \frac{dx\,dy\,dz}{x^2+y^2+(z-2)^2}$ по шару $x^2+y^2+z^2 \leqslant 1$.

7. $\displaystyle\iiint \frac{dx\,dy\,dz}{x^2+y^2+\left(z-\frac{1}{2}\right)^2}$ по шару $x^2+y^2+z^2 \leqslant 1$.

8. $\displaystyle\iint \frac{dx\,dy}{\sqrt{x^2+y^2}}$ по квадрату $|x| \leqslant 1$, $|y| \leqslant 1$.

9. Доказать, что для непрерывной функции $f(x)$

$$\lim_{h \to +0} \int_{-1}^{+1} \frac{h}{h^2+x^2} f(x)\,dx = \pi f(0).$$

§ 4. Преобразование кратных интегралов

Мы уже знаем, что введение новой переменной является одним из главных методов преобразования и упрощения простых, т. е. одномерных интегралов. Введение новых переменных имеет большое значение и в теории кратных интегралов. Правда, несмотря на то, что кратный интеграл всегда приводится к последовательному вычислению простых интегралов, его явное вычисление, как правило, труднее и еще реже выполняется в элементарных функциях, чем у одномерного интеграла. Все же многие кратные интегралы удается вычислить после преобразования переменных интегрирования. Однако, и помимо вопроса об эффективном его вычислении, преобразование кратного интеграла к новым переменным интегрирования имеет само по себе важное значение, так как теория такого преобразования приводит к более полному овладению понятием интеграла.

1. Общая формула преобразования двойного интеграла к новым переменным. Практически самым важным частным случаем является преобразование двойного интеграла к полярным координатам, которое уже было выполнено в т. I, гл. X, стр. 590. Теперь же мы сразу перейдем к общей теории замены переменных под знаком двойного интеграла

$$\iint_G f(x, y)\, dS = \iint_G f(x, y)\, dx\, dy,$$

распространенного на область G плоскости xy. Пусть замкнутая область G отображается взаимно однозначно на замкнутую область G' плоскости uv с помощью преобразования

$$x = \varphi(u, v), \qquad y = \psi(u, v).$$

Предположим, что функции φ и ψ имеют в области G' непрерывные частные производные первого порядка и что их якобиан

$$D = \begin{vmatrix} \varphi_u & \varphi_v \\ \psi_u & \psi_v \end{vmatrix} = \varphi_u\psi_v - \psi_u\varphi_v$$

во всей замкнутой области G' нигде не обращается в нуль; для определенности допустим, что $D > 0$ всюду в G'. Как известно, при этих предположениях существует единственное обратное преобразование (гл. III, § 3, n° 6)

$$u = p(x, y), \qquad v = q(x, y).$$

Кроме того, два семейства кривых $u = \text{const}$ и $v = \text{const}$, т. е. $p(x, y) = \text{const}$ и $q(x, y) = \text{const}$, покрывают в виде сетки область G.

Нетрудно придумать наглядные соображения, которые подскажут, как преобразовать $\iint_G f(x, y)\, dx\, dy$ в интеграл по u и v. До сих пор мы для вычисления интеграла $\iint_G f(x, y)\, dS$ разбивали область G на ячейки сеткой прямых, параллельных осям x и y. Сама собой напрашивается мысль воспользоваться другим разбиением области G — сеткой кривых $u \equiv p(x, y) = \text{const}$ и $v \equiv q(x, y) = \text{const}$. С этой целью выберем какие-либо числа $h = \Delta u$ и $k = \Delta v$ и рассмотрим значения $u_\nu = \nu h$ и $v_\mu = \mu k$, где ν и μ пробегают все те целые числа, при которых прямые $u=\nu h$ и $v=\mu k$ пересекают область G', а стало быть, соответствующие им исходные кривые $p(x, y)=\nu h$ и $q(x, y)=\mu k$ пересекают область G. Эти последние кривые образуют на плоскости xy совокупность петель, имеющих вид криволинейных четырехугольников, напоминающих параллелограммы. Те из этих петель, которые лежат внутри области G, мы и примем за ячейки G_i (рис. 72 и 73). Надо теперь найти площадь такой ячейки.

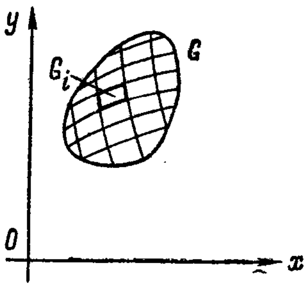


Рис. 72.

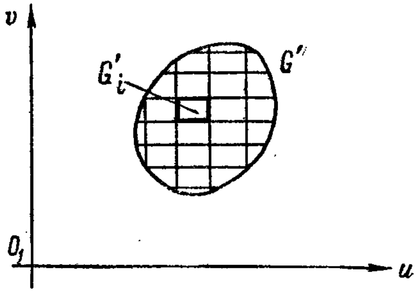


Рис. 73.

Если бы ячейка G_i была не криволинейным четырехугольником, а прямолинейным параллелограммом, то ее площадь была бы равна удвоенной площади треугольника с вершинами, соответствующими точкам (u_ν, v_μ), $(u_\nu+h, v_\mu)$ и $(u_\nu, v_\mu+k)$ плоскости uv, т. е. равнялась бы определителю

$$\begin{vmatrix} \varphi(u_\nu+h, v_\mu)-\varphi(u_\nu, v_\mu) & \varphi(u_\nu, v_\mu+k)-\varphi(u_\nu, v_\mu) \\ \psi(u_\nu+h, v_\mu)-\psi(u_\nu, v_\mu) & \psi(u_\nu, v_\mu+k)-\psi(u_\nu, v_\mu) \end{vmatrix},$$

который приближенно равен выражению

$$\begin{vmatrix} \varphi_u(u_\nu, v_\mu) & \varphi_v(u_\nu, v_\mu) \\ \psi_u(u_\nu, v_\mu) & \psi_v(u_\nu, v_\mu) \end{vmatrix} hk = hkD.$$

Помножим теперь площадь каждой ячейки G_i, вычисленную по этой формуле, на значение функции в какой-либо точке этой ячейки, и просуммируем полученные произведения по всем ячейкам, лежащим полностью внутри области G; совершив затем предельный переход $h\to 0$ и $k\to 0$, получим следующее выражение для преобразования двойного интеграла к новым переменным:

$$\iint_G f(x, y)\,dx\,dy=\iint_{G'} f\{\varphi(u, v), \psi(u, v)\}\,D\,du\,dv. \qquad (*)$$

Для того чтобы придать полную убедительность этому наводящему рассуждению, надо доказать, что замена ячеек параллелограммами и последующее приближение площади параллелограмма выражением $(\varphi_u\psi_v - \psi_u\varphi_v)hk$ дозволены, т. е. что возникающая в силу этого погрешность стремится к нулю при выполнении предельного перехода $h \to 0$ и $k \to 0$. Вместо того чтобы завершить это доказательство, выполнив оценку погрешности, мы предпочтем доказать формулу преобразования несколько отличным способом, который легко поддается обобщению на многомерные области.

Для этой цели мы воспользуемся результатами гл. III, § 3, n° 6 и выполним переход от переменных x, y к новым переменным u, v не сразу, а в две ступени. Сначала мы заменим переменные x, y новыми переменными x, v с помощью равенств

$$x = x, \qquad y = \Phi(x, v).$$

Мы предполагаем при этом, что частная производная Φ_v отлична от нуля во всей области G, например $\Phi_v > 0$ всюду в G, и что вся область G отображается взаимно однозначно на область B плоскости xv (рис. 74 и 75). После этого мы отобразим область B тоже

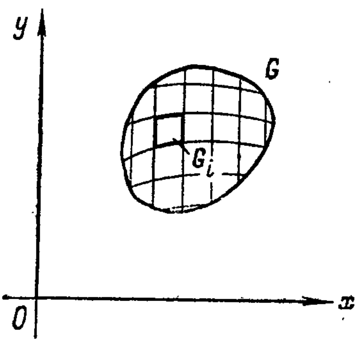


Рис. 74.

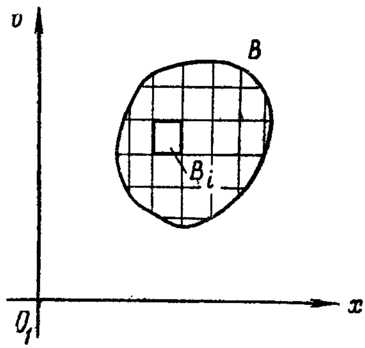


Рис. 75.

взаимно однозначно на область G' плоскости uv с помощью второго преобразования

$$x = \Psi(u, v), \qquad v = v;$$

при этом мы делаем второе предположение, что и производная $\Psi_u > 0$ во всей области B.

Преобразование нашего интеграла $\iint\limits_G f(x, y)\,dx\,dy$ мы выполним теперь тоже в два этапа. Начнем с того, что разобьем область B плоскости xv двумя семействами прямых $x = x_\nu = \mathrm{const}$ и $v = v_\mu = \mathrm{const}$ на прямоугольники со сторонами $\Delta x = h$ и $\Delta v = k$ (рис. 75). Этому разбиению области B соответствует разбиение области G плоскости xy на ячейки G_i (рис. 74), каждая из которых ограничена двумя параллельными прямыми $x = x_\nu$ и $x = x_\nu + h$ и дугами двух кривых $y=\Phi(x, v_\mu)$ и $y=\Phi(x, v_\mu+k)$. Площадь ячейки G_i (рис. 76) выражается так:

$$\Delta G_i = \int\limits_{x_\nu}^{x_\nu+h} \{\Phi(x, v_\mu+k)-\Phi(x, v_\mu)\}\,dx.$$

По теореме о среднем значении из интегрального исчисления это можно записать в следующем виде:

$$\Delta G_i = h\,\{\Phi(\bar{x}_\nu, v_\mu+k)-\Phi(\bar{x}_\nu, v_\mu)\},$$

где $\bar{x}_\nu$ есть некоторое промежуточное число между x_ν и $x_\nu+h$. Применив теперь теорему о среднем значении из дифференциального исчисления, получим

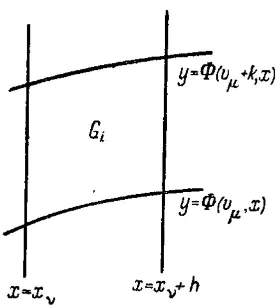


Рис. 76.

$$\Delta G_i = hk\,\Phi_v(\bar{x}_\nu, \bar{v}_\mu),$$

где $\bar{v}_\mu$ — некоторое промежуточное число между v_μ и $v_\mu+k$, так что $(\bar{x}_\nu, \bar{v}_\mu)$ являются координатами некоторой точки соответствующего прямоугольника B_i области B.

Следовательно, наш интеграл от функции $f(x, y)$ по области G есть предел интегральной суммы

$$\sum f_i\,\Delta G_i = \sum hkf\{\bar{x}_\nu, \Phi(\bar{x}_\nu, \bar{v}_\mu)\}\,\Phi_v(\bar{x}_\nu, \bar{v}_\mu)$$

при $h\to 0$, $k\to 0$. Сразу видно, что выражение в правой части имеет своим пределом интеграл

$$\iint\limits_B f(x, y)\,\Phi_v\,dx\,dv, \qquad \text{где} \qquad y=\Phi(x, v).$$

Поэтому

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_B f(x, y)\,\Phi_v\,dx\,dv.$$

К интегралу, стоящему в правой части, мы теперь применим те же рассуждения, которыми мы только что воспользовались для преобразования интеграла по области G в интеграл по области B, но теперь мы переведем область B в область G' с помощью преобразования $x=\Psi(u, v)$, $v=v$.

В результате интеграл по области B превратится в интеграл по области G' от подынтегральной функции $f\Phi_v\Psi_u$, и мы окончательно получим

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_{G'} f(x, y)\,\Phi_v\Psi_u\,du\,dv, \tag{$*$}$$

где переменные x и y должны быть выражены через новые переменные u и v с помощью введенных выше двух составляющих преобразований. Введя теперь результирующее преобразование

$$x = \varphi(u, v), \qquad y = \psi(u, v),$$

без труда находим

$$\frac{\partial(x, y)}{\partial(x, v)} = \Phi_v, \qquad \frac{\partial(x, v)}{\partial(u, v)} = \Psi_u$$

и, следовательно, якобиан результирующего преобразования

$$D = \frac{\partial(x, y)}{\partial(u, v)} = \Phi_v \Psi_u.$$

Поэтому формула преобразования двойного интеграла ($\overset{*}{*}$) приводится к тому виду (*), который был получен ранее с помощью наводящего рассуждения. Таким образом, эта формула доказана для всех тех преобразований вида $x = \varphi(u, v)$, $y = \psi(u, v)$, которые могут быть расщеплены на два последовательных примитивных преобразования[1]) $x = x$, $y = \Phi(x, v)$ и $x = \Psi(u, v)$, $v = v$.

Но в гл. III, § 3, стр. 170, мы видели, что замкнутую область G можно разбить на конечное число областей, в каждой из которых наше общее преобразование можно разложить на два примитивных. Правда, проведенные там рассуждения относились лишь ко всякой замкнутой области G_1, лежащей полностью внутри G. Однако область G_1 можно выбрать так, чтобы она занимала всю область G, за исключением части, имеющей сколь угодно малую площадь. Поэтому формула преобразования справедлива и для всей области G. Таким образом, приходим к следующему общему результату: *если преобразование $x = \varphi(u, v)$, $y = \psi(u, v)$ дает непрерывное взаимно однозначное отображение замкнутой области G плоскости xy на область G' плоскости uv, причем функции φ и ψ имеют непрерывные частные производные первого порядка, а якобиан $\frac{\partial(x, y)}{\partial(u, v)} = \varphi_u\psi_v - \psi_u\varphi_v$ всюду положителен, то*

$$\iint\limits_G f(x, y)\, dx\, dy = \iint\limits_{G'} f\{\varphi(u, v), \psi(u, v)\} \frac{\partial(x, y)}{\partial(u, v)}\, du\, dv.$$

Заметим еще, что формула преобразования двойного интеграла сохраняет силу и в том случае, когда якобиан $\frac{\partial(x, y)}{\partial(u, v)}$ обращается в нуль в конечном числе точек области, если только он при этом не

---

[1]) Правда, в ходе доказательства мы предполагали, что обе производные Φ_v и Ψ_u положительны, но легко убедиться, что это несерьезное ограничение. Действительно, из допущения $\frac{\partial(x, y)}{\partial(u, v)} > 0$ вытекает только, что эти две производные должны иметь одинаковые знаки. Если они обе отрицательны, то надо лишь заменить x на $-x$ и y на $-y$, что, согласно определению, не изменяет величины интеграла. Тогда оба примитивных преобразования будут иметь положительные якобианы.

изменяет своего знака. Для доказательства надо только вырезать эти точки из области G, окружив их малыми кругами радиуса ρ; наша формула справедлива для остающейся области. Заставим теперь ρ стремиться к нулю; в силу непрерывности всех участвующих функций, придем к выводу, что формула преобразования справедлива и для области G. Этим фактом пользуются всякий раз, когда вводится полярная система координат с полюсом, лежащим внутри области, ибо при этом якобиан преобразования равен r и, стало быть, обращается в полюсе в нуль.

В гл. V, § 4, стр. 400, мы еще вернемся к преобразованиям с отрицательным якобианом и убедимся, что рассуждения остаются в основном без изменения. Однако отметим уже здесь, что если якобиан D не обращается в нуль, то дополнительное условие $D > 0$ не приводит к ограничению общности, так как знак якобиана можно всегда изменить, поменяв ролями u и v. В гл. V, § 3, стр. 397, мы дадим другой способ доказательства формулы преобразования.

2. Преобразование n-кратного интеграла к новым переменным интегрирования. Тем же путем можно, конечно, вести рассуждения и для кратного интеграла по любой n-мерной области, например, для тройного интеграла. В результате получается следующая общая теорема:

Если взаимно однозначное преобразование отображает замкнутую область G пространства $xyz\ldots$ на область G^* пространства $uvw\ldots$ и если якобиан преобразования

$$\frac{\partial(x, y, z, \ldots)}{\partial(u, v, w, \ldots)}$$

всюду положителен, то n-кратный интеграл преобразуется по следующей формуле:

$$\iint\limits_{G}\ldots\int f(x, y, z, \ldots)\,dx\,dy\,dz\ldots = \iint\limits_{G^*}\ldots\int f(x, y, z, \ldots)\frac{\partial(x, y, z, \ldots)}{\partial(u, v, w, \ldots)}\,du\,dv\,dw\ldots$$

Заметим, что якобиан преобразования в пространстве n измерений представляет собой определитель n-го порядка, строение которого аналогично строению якобиана в двумерном случае.

Примером может служить преобразование прямоугольных декартовых координат в полярные. В плоском случае вместо u и v следует писать r и θ, и сразу получится для якобиана выражение $\frac{\partial(x, y)}{\partial(r, \theta)} = r$, что согласуется с результатом, полученным в гл. III, § 3, п° 4. Таким образом,

$$\iint\limits_{G} f(x, y)\,dx\,dy = \iint\limits_{G^*} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta.$$

В случае полярных координат в пространстве (сферических координат) формулы преобразования имеют следующий вид:

$$x = r\sin\theta\cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad z = r\cos\theta,$$

где $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \varphi < 2\pi$, $0 \leqslant r < +\infty$. Теперь мы должны положить $u = r$, $v = \theta$ и $w = \varphi$, и для якобиана D получится следующее выражение:

$$D = \frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} = \begin{vmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix} = r^2\sin\theta.$$

Поэтому преобразование тройного интеграла от прямоугольных к сферическим координатам производится по следующей формуле:

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz = \iiint\limits_{G^*} f(x, y, z)\,r^2\sin\theta\,dr\,d\theta\,d\varphi.$$

Как и в случае преобразования двойного интеграла к полярным координатам на плоскости, эту формулу можно получить и не прибегая к общей теории. Надо лишь воспользоваться разбиением пространства на ячейки с помощью шаровых поверхностей $r = \text{const}$, полуконусов $\theta = \text{const}$ и полуплоскостей $\varphi = \text{const}$. Выполнение этого элементарно-геометрического подсчета можно предоставить читателю.

При переходе к полярным координатам в пространстве якобиан D преобразования обращается в нуль, если $r = 0$ или $\theta = 0$, или $\theta = \pi$; таким образом, сделанные в общей теории допущения не выполняются во всех точках оси z. Тем не менее формула преобразования тройного интеграла сохраняет силу, в чем можно убедиться тем же способом, которым мы успешно воспользовались для плоской области.

Упражнения

1*. Вычислить $\iint e^{\frac{y-x}{y+x}}\,dx\,dy$ по области, ограниченной треугольником с вершинами $(0; 0)$, $(0; 1)$, $(1; 0)$.

2. Вычислить $\displaystyle\iint \frac{dx\,dy}{(1 + x^2 + y^2)^2}$ а) по области, ограниченной одной петлей лемнискаты $(x^2 + y^2)^2 - (x^2 - y^2) = 0$; б) по области, ограниченной треугольником с вершинами $(0; 0)$, $(2; 0)$, $(1; \sqrt{3})$.

3. Вычислить $\iiint xyz\,dx\,dy\,dz$ по объему части эллипсоида $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} \leqslant 1$, $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$.

4. Доказать, что $\displaystyle\iint\limits_G e^{-(x^2+y^2)}\,dx\,dy = ae^{-a^2}\int\limits_0^\infty \frac{e^{-u^2}}{a^2 + u^2}\,du$ (где область G есть полуплоскость $x \geqslant a > 0$), пользуясь преобразованием

$$x^2 + y^2 = u^2 + a^2, \qquad y = vx.$$

5. Доказать, что $\left|\iint (u_x^2 + u_y^2)\,dx\,dy\right|$ не изменяется при инверсии.

6. Вычислить тройной интеграл из упр. 4, стр. 270, пользуясь переходом к сферическим координатам.

7. Вычислить интеграл $I = \iiint \cos(x\xi + y\eta + z\zeta)\,d\xi\,d\eta\,d\zeta$ по объему шара $\xi^2 + \eta^2 + \zeta^2 \leqslant 1$.

8. Доказать, что $\displaystyle\iint \cos(x\xi + y\eta)\,d\xi\,d\eta = \frac{2\pi J_1(r)}{r}$, где $r = \sqrt{x^2 + y^2}$, если областью интегрирования служит круг $\xi^2 + \eta^2 \leqslant 1$, а J_1 есть символ бесселевой функции, определенной в упр. 5, стр. 246.

§ 5. Несобственные кратные интегралы

При изучении функций одной переменной мы встретились с необходимостью обобщить понятие определенного интеграла на такие функции, которые не являются непрерывными в замкнутом промежутке интегрирования. И мы действительно ввели в рассмотрение интегралы от функций, имеющих разрывы первого рода (конечные скачки), интегралы от функций, обращающихся в бесконечность в отдельных точках промежутка интегрирования, а также интегралы по бесконечным промежуткам. Теперь нам предстоит выполнить аналогичное обобщение кратного интеграла.

1. Интеграл от функции, имеющей конечные разрывы. Для функции, имеющей в области интегрирования G только конечные скачкообразные нарушения непрерывности, способ обобщения понятия кратного интеграла напрашивается сам собой. Достаточно будет пояснить этот способ на примере двойного интеграла по области G плоскости xy. Предполагаем, что область G можно разбить конечным числом гладких[1]) криволинейных дуг на конечное число таких частичных областей или районов G_1, G_2, ..., G_n, что подынтегральная функция $f(x, y)$ непрерывна внутри каждого района, а когда точка приближается изнутри каждого района к его границе, то значение функции в этой точке стремится к определенным предельным значениям (вообще говоря, различным для разных точек границы). Однако предельные значения, к которым стремится функция, когда мы приближаемся к какой-либо точке общей границы двух районов, могут быть различны, когда мы приближаемся к этой точке из одного или из другого района. Интеграл от функции $f(x, y)$ по каждому району G_k подходит под наше первоначальное определение двойного интеграла, если только мы дополним определение функции в G_k соответствующими ее значениями на границе до функции, непрерывной в замкнутом районе G_k. *Двойной интеграл от функции f по всей области G мы определим* теперь *как сумму интегралов от этой функции по всем частичным областям G_k.*

В качестве примера рассмотрим функцию $f(x, y)$, определенную на квадрате $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ следующим образом:

$$f(x, y) = \begin{cases} 1, & \text{если } y < x, \\ 2, & \text{если } y \geqslant x. \end{cases}$$

Для этой функции прямая $y = x$ является линией разрыва. Вычисляя, согласно сделанному определению, несобственный двойной интеграл по указанному выше квадрату, получаем

$$\iint\limits_{\square} f(x, y)\,dx\,dy = \frac{3}{2}.$$

[1]) Гладкой дугой называется дуга кривой с непрерывно вращающейся касательной.

2. Кратный интеграл от функции, обращающейся в бесконечность в изолированных точках. Если подынтегральная функция f обращается в бесконечность в единственной точке P области интегрирования G, то несобственный интеграл по области G мы определим процессом, аналогичным тому, которым пользовались в случае одной переменной. Вырезаем из области G вокруг точки разрыва P некоторую ее окрестность U_k, так что остающаяся замкнутая область $G_k = G - U_k$ уже не содержит точки P. Можно построить весьма разнообразные последовательности таких окрестностей U_k, диаметры которых стремятся к нулю с возрастанием k, например, последовательность кругов (в двумерной области) или шаров (в трехмерной области) с центром в точке P и с радиусом $c = \frac{1}{k}$. Если последовательность интегралов по остающейся области G_k стремится к определенному пределу (числу) I при $k \to \infty$, т. е. если

$$\lim_{k \to \infty} \iint_{G_k} f(x, y)\, dS = I$$

и если этот предел не зависит от выбора последовательности G_k, то *несобственный интеграл* от функции f по области G называется *сходящимся* и упомянутый предел I называется значением этого несобственного интеграла:

$$\iint_{G} f(x, y)\, dS = \lim_{k \to \infty} \iint_{G_k} f(x, y)\, dS = I.$$

Если предел I равен бесконечности или не существует, то несобственный интеграл по области G называется *расходящимся*.

Наше определение пригодно, очевидно, и для того случая, когда точка P является изолированной точкой неопределенности, как, например, начало координат для функции $\sin\frac{1}{x^2+y^2}$. Если в окрестности точки неопределенности P абсолютная величина функции остается меньше некоторого постоянного числа, то несобственный интеграл всегда сходится.

Общее условие сходимости несобственного интеграла можно, на основании сказанного выше, формулировать следующим образом. Пусть любому наперед заданному сколь угодно малому положительному числу ε соответствует число $\delta = \delta(\varepsilon)$, обладающее следующим свойством: если U и U' — какие угодно две различные (открытые) окрестности точки P (бесконечного разрыва или неопределенности), диаметры которых меньше δ, то разность интегралов от функции f, распространенных на замкнутые остающиеся области $G - U$ и $G - U'$, по абсолютной величине меньше ε. При выполнении этого условия несобственный интеграл по области G сходится.

Поясним эти идеи на примерах.

Функция $f(x, y) = \ln\sqrt{x^2+y^2}$ обращается в бесконечность в начале координат. Для того чтобы выяснить, сходится ли несобственный двойной интеграл от этой функции по области G, содержащей начало координат, например, по кругу $x^2+y^2 \leqslant 1$, надо вырезать из области G произвольную окрестность U_δ начала координат, диаметр которой меньше δ, а затем вычислить предел интеграла по остающейся области $G_\delta = G - U_\delta$, когда δ стремится к нулю. Любая окрестность начала координат, диаметр которой меньше δ, наверняка лежит внутри круга радиуса δ с центром в начале координат. Преобразуем интеграл по области G_δ к полярным координатам (§ 4, n° 2); тогда

$$\iint_{G_\delta} \ln\sqrt{x^2+y^2}\, dx\, dy = \iint_{G'_\delta} r \ln r\, dr\, d\theta,$$

причем интеграл в правой части берется по той области G'_δ плоскости $r\theta$, которая соответствует области G_δ. Следовательно, если область G есть круг $x^2+y^2 \leqslant 1$, то область G'_δ содержит прямоугольник $\delta \leqslant r \leqslant 1$, $0 \leqslant \theta \leqslant 2\pi$ плоскости $r\theta$, но не доходит до прямой $r=0$. Так как $\lim\limits_{r\to 0}(r\ln r) = 0$, то подынтегральная функция $r\ln r$ становится непрерывной при $r=0$, если приписать ей в этой точке значение нуль. Поэтому, заставляя δ стремиться к нулю, можно рассматривать преобразованный интеграл

$$\iint_{G'} r \ln r\, dr\, d\theta = \lim_{\delta\to 0} \iint_{G'_\delta} r \ln r\, dr\, d\theta$$

как обыкновенный (собственный) двойной интеграл. Это и доказывает сходимость предложенного интеграла.

В то же время этот пример показывает, что, как и в случае одной независимой переменной, подходящее преобразование координат может иногда превратить несобственный интеграл в собственный. Из этого факта ясно, какие недопустимые ограничения пришлось бы ввести, если бы мы отказались от рассмотрения несобственных интегралов.

В качестве следующего примера рассмотрим интеграл

$$I(G) = \iint_G \frac{dx\, dy}{(\sqrt{x^2+y^2})^\alpha},$$

взятый по области G. Пусть область G полностью лежит внутри круга K с центром в начале координат и радиусом R. Если сходится интеграл по кругу K, то сходится и интеграл по области G. Если, напротив, круг K постоянного радиуса R содержится в области G и интеграл по кругу K расходится, то, очевидно, расходится и наш интеграл по области G. И здесь мы сначала вырежем из круга K малый круг радиуса δ и возьмем интеграл по остающейся области K_δ. После перехода к полярным координатам получим

$$I(K_\delta) = \iint_{K'_\delta} \frac{1}{r^{\alpha-1}}\, dr\, d\theta$$

или, в виде повторного интеграла,

$$I(K_\delta)=\int_0^{2\pi} d\theta \int_\delta^R \frac{dr}{r^{\alpha-1}}=2\pi\int_\delta^R \frac{dr}{r^{\alpha-1}}.$$

Но мы уже знаем (т. I, стр. 286), что несобственный интеграл $\int_0^R \frac{dr}{r^{\alpha-1}}$ сходится в том и только в том случае, если $\alpha<2$. Отсюда вытекает, что и несобственный двойной интеграл $\iint_G \frac{dx\,dy}{(\sqrt{x^2+y^2})^\alpha}$ сходится в том и только в том случае, если $\alpha<2$. Ясно, что сходимость и этого интеграла не зависит от выбора окрестностей U_δ.

Пользуясь этим результатом, нетрудно вывести *достаточный* (но не необходимый) *признак сходимости* несобственного двойного интеграла:

Если функция $f(x, y)$ непрерывна всюду в замкнутой области G, за исключением одной точки P, в которой f обращается в бесконечность (эту точку P мы примем за начало координат), и если существует такое постоянное число M и такое положительное постоянное число $\alpha<2$, что

$$|f(x, y)|\leqslant\frac{M}{(\sqrt{x^2+y^2})^\alpha}$$

во всей области G, то несобственный двойной интеграл

$$\iint_G f(x, y)\,dx\,dy$$

сходится.

Доказательство получается на основании результата последнего примера из рассмотрения неравенства

$$\left|\iint_{G_\delta} f(x, y)\,dx\,dy\right|\leqslant\iint_{G_\delta}|f(x, y)|\,dx\,dy\leqslant M\iint_{G_\delta}\frac{dx\,dy}{(\sqrt{x^2+y^2})^\alpha},$$

где G_δ есть область, остающаяся после удаления из G круговой окрестности точки P радиуса δ.

Подобным же путем можно исследовать сходимость несобственного тройного интеграла

$$\iiint_G \frac{dx\,dy\,dz}{(\sqrt{x^2+y^2+z^2})^\alpha},$$

где трехмерная область G содержит в себе начало координат. Переходом к полярным координатам в пространстве этот интеграл приводится к виду

$$\iiint_{G'} \frac{1}{r^{\alpha-2}}\sin\theta\,dr\,d\theta\,d\varphi,$$

и рассуждение, аналогичное предшествующему, обнаруживает, что этот несобственный интеграл сходится в том и только в том случае, если $\alpha < 3$. Пользуясь этим результатом, получим следующий общий *достаточный* признак сходимости тройного интеграла.

Пусть функция $f(x, y, z)$ обращается в бесконечность в начале координат и непрерывна во всех остальных точках трехмерной области G, содержащей начало координат. Если существует такое постоянное число M и такое положительное постоянное число $\alpha < 3$, что

$$|f(x, y, z)| \leqslant \frac{M}{(\sqrt{x^2+y^2+z^2})^\alpha}$$

во всей области G, то несобственный интеграл

$$\iiint_G f(x, y, z)\,dx\,dy\,dz$$

сходится.

С помощью этих признаков сходимости нетрудно вывести следующие несколько более общие результаты: двойной интеграл вида

$$\iint_G \frac{g(x, y)\,dx\,dy}{(\sqrt{(x-a)^2+(y-b)^2})^\alpha} \qquad (\alpha < 2)$$

и тройной интеграл вида

$$\iiint_G \frac{g(x, y, z)\,dx\,dy\,dz}{(\sqrt{(x-a)^2+(y-b)^2+(z-c)^2})^\alpha} \qquad (\alpha < 3),$$

где (a, b) или (a, b, c) — постоянная точка, лежащая внутри области G, сходятся, если функция g непрерывна в замкнутой области G. Для того чтобы сделать возможным применение к этим интегралам наших признаков сходимости, надо лишь перенести начало координат в точку (a, b) или (a, b, c) с помощью параллельного переноса осей.

3. Интеграл от функции, обращающейся в бесконечность вдоль линии. Если функция $f(x, y)$ обращается в бесконечность не только в отдельной точке, но и вдоль некоторой линии C, лежащей в области G, то несобственному двойному интегралу от функции f по области G можно дать определение, следуя тому же пути. Линию разрыва C вырезаем из области G вместе с такой окрестностью U_ε этой линии, площадь которой меньше ε. Если при $\varepsilon \to 0$ интеграл от функции $f(x, y)$ по остающейся области $G - U_\varepsilon$ стремится к конечному пределу I, не зависящему от выбора изменяющейся окрестности U_ε, то несобственный *двойной интеграл от функции f по области G называется сходящимся* (говорят также, что он сходится), а этот предел I считается *значением несобственного интеграла.*

Простейший пример — тот, в котором линия C есть отрезок прямой, например отрезок оси y. Если функция f удовлетворяет во всей области G неравенству

$$|f(x, y)| < \frac{M}{x^\alpha},$$

где M и α — постоянные числа и $0 < \alpha < 1$, то $\iint\limits_G f(x, y)\,dx\,dy$ сходится. Доказательство проводится знакомым нам путем: можно, например, вырезать из области G полосу, содержащую ось y и ограниченную двумя прямыми, параллельными этой оси.

4. Интеграл по бесконечной области. Если область интегрирования G простирается в бесконечность, то ее «исчерпывают» последовательностью конечных частичных областей $G_1, G_2, \ldots, G_k, \ldots$, обладающих тем свойством, что *любая* ограниченная частичная область, содержащаяся в G, лежит внутри всех областей G_n, начиная с некоторого номера m, т. е. при всяком $n \geqslant m$. Если, например, областью G является вся плоскость, то за последовательность G_k можно принять совокупность кругов радиуса k ($k = 1, 2, \ldots$) с центром в начале координат. Если интеграл

$$\iint\limits_{G_k} f(x, y)\,dS$$

стремится к определенному пределу, числу I, при $k \to \infty$, и этот предел не зависит от выбора последовательности областей G_k, то говорят, что *несобственный интеграл*

$$\iint\limits_G f(x, y)\,dS$$

по бесконечной области G *сходится*, и число I называется *значением несобственного интеграла*

$$\iint\limits_G f(x, y)\,dS = \lim_{k \to \infty} \iint\limits_{G_k} f(x, y)\,dS = I.$$

Если же предел $\iint\limits_{G_k}$ бесконечен или вообще не существует, то говорят, что несобственный интеграл от функции f по бесконечной области G *расходится.*

Поясним это определение на примере интеграла

$$\iint\limits_G e^{-x^2-y^2}\,dx\,dy,$$

в котором областью интегрирования G является вся плоскость xy. Для того чтобы проверить, сходится ли этот несобственный интеграл, попытаемся сначала исчерпывать плоскость xy последовательностью кругов K_ν с радиусами ν: $x^2 + y^2 \leqslant \nu^2$, которые удовлетворяют, очевидно, поставленным выше требованиям. Стало быть, надо исследовать предел интеграла

$$\iint\limits_{K_\nu} e^{-x^2-y^2}\,dx\,dy$$

при $\nu \to \infty$. Но этот интеграл мы уже вычислили (т. I, стр. 592) и нашли, что он равен $\pi(1-e^{-\nu^2})$. Следовательно,

$$\lim_{\nu\to\infty} \iint_{K_\nu} e^{-x^2-y^2}\,dx\,dy = \lim_{\nu\to\infty} \pi(1-e^{-\nu^2}) = \pi.$$

Если нам удастся показать, что не только взятая нами последовательность кругов, но и всякая другая последовательность частичных областей плоскости G, обладающая требуемыми свойствами, приводит к тому же пределу π, то этим будет доказано, что наш несобственный интеграл сходится и равен π.

Итак, возьмем любую такую последовательность частичных областей G_1, G_2, ... Согласно условию, всякий круг K_m должен лежать внутри G_ν, если ν достаточно велико. С другой стороны, любая из областей G_ν ограничена, а потому сама содержится в круге K_M достаточно большого радиуса M. Так как подынтегральная функция $e^{-x^2-y^2}$ всюду положительна, то

$$\iint_{K_m} e^{-x^2-y^2}\,dx\,dy \leqslant \iint_{G_\nu} e^{-x^2-y^2}\,dx\,dy \leqslant \iint_{K_M} e^{-x^2-y^2}\,dx\,dy.$$

С возрастанием радиусов m и M интегралы по кругам K_m и K_M стремятся к общему пределу π; следовательно, и интеграл по области G_ν должен стремиться к тому же пределу. Стало быть, предложенный интеграл сходится и равен π.

Из этого факта можно извлечь особенно интересный результат, если в качестве областей G_ν взять последовательность квадратов $|x| \leqslant \nu$, $|y| \leqslant \nu$. Тогда интеграл

$$\iint_{G_\nu} e^{-x^2-y^2}\,dx\,dy$$

можно представить как произведение двух простых интегралов (см. § 2, n° 3, пример 3):

$$\iint_{G_\nu} e^{-x^2-y^2}\,dx\,dy = \int_{-\nu}^{\nu} e^{-x^2}\,dx \int_{-\nu}^{\nu} e^{-y^2}\,dy = \left(\int_{-\nu}^{\nu} e^{-x^2}dx\right)^2 = \left(2\int_0^{\nu} e^{-x^2}dx\right)^2.$$

При $\nu \to \infty$ должен получиться тот же предел π. Следовательно,

$$\left(2\int_0^{\infty} e^{-x^2}dx\right)^2 = \pi, \qquad \text{откуда} \qquad \int_0^{\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2},$$

в полном согласии с результатом, полученным в т. I, гл. VIII, стр. 482—484 и гл. X, стр. 592.

5. Заключительные замечания и некоторые дополнения. Полезно резюмировать с единой точки зрения введенные в этом параграфе понятия. Наше обобщение понятия кратного интеграла на случаи, когда определения, данные в § 2, непосредственно не приложимы, основано на том, что несобственный интеграл рассматривается как предел последовательности интегралов по областям G_ν, которые с возрастанием ν приближаются к заданной области интегрирования G, или, как принято говорить, исчерпывают эту область G. С этой целью область G рассматривают уже не как замкнутую, а как открытую область; при этом все точки разрыва подынтегральной функции относят к границе области, и эту границу не включают в область G. Тогда говорят, что *область G исчерпывается последовательностью замкнутых областей $G_1, G_2, \ldots, G_n, \ldots$, если все эти области G_n содержатся в G и если любая замкнутая частичная область, лежащая полностью внутри G, лежит также внутри G_n, коль скоро номер n достаточно велик.* Если эта последовательность частичных областей G_n выбрана так, что каждая из них содержит внутри себя предшествующую, то говорят, что последовательность G_n сходится к области G *монотонно*.

К каждой из областей G_n, лежащих внутри G, определение интеграла из § 2 непосредственно приложимо; так вот *несобственный интеграл от функции f по области G называется сходящимся, если интеграл по области G_n стремится к конечному пределу, не зависящему от выбора последовательности областей G_n.*

Полезно уяснить себе следующие факты, иллюстрацией которых могут служить приведенные выше примеры.

1) Если подынтегральная функция f *нигде не отрицательна* в области G, то достаточно доказать сходимость последовательности интегралов от функции f по *какой-либо одной монотонной последовательности областей* G_ν, и отсюда уже будет вытекать, что к тому же пределу сходится последовательность интегралов по *любой* последовательности областей G'_ν.

Доказательство. Так как замкнутая область G_ν лежит внутри G, то она содержится во всех областях G'_n, начиная с некоторого номера $n(\nu)$. На том же основании находим, что всякая область G'_n содержится в области G_m при достаточно большом m. Так как функция $f \geqslant 0$ всюду в G, то

$$\iint\limits_{G_\nu} f(x, y)\,dx\,dy \leqslant \iint\limits_{G'_n} f(x, y)\,dx\,dy \leqslant \iint\limits_{G_m} f(x, y)\,dx\,dy.$$

При $\nu \to \infty$ оба крайних интеграла стремятся к одному и тому же пределу; стало быть, и последовательность интегралов

$$\iint\limits_{G'_n} f(x, y)\,dx\,dy$$

должна сходиться к тому же пределу, что и требовалось доказать.

2) Более того, из этой теоремы вытекает, что если интегралы от неотрицательной функции f по монотонной последовательности областей G_k, сходящейся к области G, остаются ограниченными сверху некоторым числом M, то $\iint f\,dS$ по области G сходится. Действительно, интегралы по областям G_k образуют ограниченную сверху последовательность, члены которой монотонно возрастают (или по крайней мере не убывают). Стало быть, эта последовательность интегралов сходится.

Тот случай, когда функция $f \leqslant 0$ всюду в области G, легко привести к только что рассмотренному заменой f на $-f$.

3) Если функция f принимает в области G как положительные, так и отрицательные значения, то теоремы, доказанные в пунктах 1 и 2, можно применить к функции $|f|$. Если окажется, что интеграл от абсолютной величины функции f сходится, то непременно сходится и интеграл от самой функции f. Это легче всего доказать с помощью следующего приема. Строим две вспомогательные функции f_1 и f_2 следующим образом:

$$f_1 = \begin{cases} f \text{ в точках, где } f \geqslant 0, \\ 0 \text{ в точках, где } f < 0, \end{cases} \qquad f_2 = \begin{cases} -f \text{ в точках, где } f \leqslant 0, \\ 0 \text{ в точках, где } f > 0. \end{cases}$$

Тогда $f = f_1 - f_2$. Всюду в G $f_1 \geqslant 0$ и $f_2 \geqslant 0$; обе эти функции непрерывны там, где непрерывна функция f и, наконец, $f_1 \leqslant f$ и $f_2 \leqslant f$ во всей области G. Поэтому, если интеграл от $|f|$ остается ограниченным для монотонной последовательности областей G_k, сходящейся к G, то интегралы от f_1 и f_2 по области G сходятся, а вместе с ними сходится и интеграл по области G от их разности $f = f_1 - f_2$.

§ 6. Приложения к геометрии

1. Вычисление объема с помощью двойного интеграла. Примеры. Задача вычисления объема тела привела нас к определению двойного интеграла. Поэтому непосредственно ясно, как вычислять объемы с помощью двойных интегралов.

1) Для вычисления объема V *эллипсоида вращения*

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1$$

приводим уравнение к явному виду:

$$z = \pm \frac{b}{a} \sqrt{a^2 - x^2 - y^2}.$$

Объем половины эллипсоида, лежащей над плоскостью xy, равен двойному интегралу от функции z по области K круга $x^2 + y^2 \leqslant a^2$:

$$\frac{V}{2} = \frac{b}{a} \iint_K \sqrt{a^2 - x^2 - y^2}\, dx\, dy.$$

Переход к полярным координатам дает

$$\frac{V}{2} = \frac{b}{a} \iint_G \sqrt{a^2 - r^2}\, r\, dr\, d\theta =$$

$$= \frac{b}{a} \int_0^{2\pi} d\theta \int_0^a r \sqrt{a^2 - r^2}\, dr = 2\pi \frac{b}{a} \int_0^a r \sqrt{a^2 - r^2}\, dr,$$

откуда и получается искомый объем

$$V=\frac{4}{3}\pi a^2 b.$$

2) Для вычисления объема *трехосного эллипсоида*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

совершаем преобразование к новым, криволинейным, координатам с помощью формул

$$x=a\rho\cos\theta,\qquad y=b\rho\sin\theta$$

с якобианом

$$\frac{\partial(x, y)}{\partial(\rho, \theta)}=ab\rho.$$

Тогда половина объема эллипсоида выразится так:

$$\frac{V}{2}=c\iint\limits_{G}\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}\,dx\,dy=abc\iint\limits_{G'}\sqrt{1-\rho^2}\,\rho\,d\rho\,d\theta,$$

где G есть внутренняя область эллипса $\frac{x^2}{a^2}+\frac{y^2}{b^2}\leqslant 1$, а область G' определяется неравенствами $0\leqslant\rho\leqslant 1$, $0\leqslant\theta\leqslant 2\pi$. Поэтому

$$\frac{V}{2}=abc\int\limits_0^{2\pi}d\theta\int\limits_0^1\rho\sqrt{1-\rho^2}\,d\rho=\frac{2}{3}\pi abc\qquad\text{и}\qquad V=\frac{4}{3}\pi abc.$$

3) В заключение вычислим объем пирамиды, ограниченной плоскостями координат и плоскостью $ax+by+cz-1=0$, причем числа a, b, c будем считать положительными. Этот объем выражается следующим интегралом:

$$V=\frac{1}{c}\iint\limits_{T}(1-ax-by)\,dx\,dy,$$

в котором область интегрирования T есть треугольник

$$0\leqslant x\leqslant\frac{1}{a},\qquad 0\leqslant y\leqslant\frac{1}{b}(1-ax)$$

в плоскости xy. Следовательно,

$$V=\frac{1}{c}\int\limits_0^{\frac{1}{a}}dx\int\limits_0^{\frac{1}{b}(1-ax)}(1-ax-by)\,dy;$$

внутренний интеграл (по y) равен

$$\left[(1-ax)\,y-\frac{b}{2}y^2\right]_0^{\frac{1}{b}(1-ax)}=\frac{(1-ax)^2}{2b},$$

а последующее интегрирование по x дает

$$V=\frac{1}{2bc}\int\limits_0^{\frac{1}{a}}(1-ax)^2\,dx=-\frac{1}{6abc}(1-ax)^3\Bigg|_0^{\frac{1}{a}}=\frac{1}{6abc}.$$

Тот же результат дает, конечно, и теорема элементарной геометрии, по которой объем пирамиды равен одной трети произведения площади ее основания на высоту.

Для вычисления объема тела более сложной формы можно разбить это тело на такие части, объемы которых выражаются непосредственно двойными интегралами. Однако в дальнейшем (главным образом уже в следующей главе) мы выведем для объема, ограниченного замкнутой поверхностью, такие выражения, которые не требуют подобного разбиения.

2. Вычисление объема с помощью тройного интеграла. Объем в цилиндрических и сферических координатах. Подобно тому, как площадь плоской области G выражается двойным интегралом

$$S=\iint_G dS=\iint_G dx\,dy,$$

объем пространственной области G выражается тройным интегралом

$$V=\iiint_G dV=\iiint_G dx\,dy\,dz,$$

распространенным на эту трехмерную область. Это в точности соответствует нашему определению тройного интеграла и выражает тот геометрический факт, что объем тела можно вычислить следующим образом: рассекаем пространство на равные прямоугольные параллелепипеды, находим общий объем всех параллелепипедов, лежащих полностью в области G, а затем заставляем диаметры параллелепипедов стремиться к нулю.

Приведение тройного интеграла для V к виду $\int_{z_0}^{z_1} dz \iint dx\,dy$ выражает известный из элементарной геометрии *принцип Кавальери*, согласно которому объем тела вполне определяется заданием площадей всех его плоских поперечных сечений, перпендикулярных к некоторой прямой, например оси z.

Исходя из тройного интеграла, выражающего объем тела в декартовых координатах, можно получить разнообразные формулы для вычисления объема; для этого надо только преобразовать интеграл от переменных x, y, z к различным другим системам независимых переменных. Важнейшими из них являются цилиндрическая и сферическая системы координат.

В качестве примера применения *цилиндрических координат* приведем вычисление *объема тела вращения*, которое образуется при вращении кривой $x=\varphi(z)$ плоскости xz вокруг оси z. Предполагается при этом, что вращающаяся дуга кривой не пересекает оси z и что тело вращения ограничено снизу плоскостью $z=a$ и сверху — плоскостью $z=b$. Вводим вместо x, y, z цилиндрические координаты

z, ρ, θ с помощью формул $x=\rho\cos\theta$, $y=\rho\sin\theta$, $z=z$. Тело определяется, очевидно, неравенствами $0\leqslant\rho=\sqrt{x^2+y^2}\leqslant\varphi(z)$, $0\leqslant\theta\leqslant 2\pi$ и $a\leqslant z\leqslant b$. Для объема сразу получается выражение

$$V=\iiint\limits_G dx\,dy\,dz=\int\limits_a^b dz\int\limits_0^{2\pi} d\theta\int\limits_0^{\varphi(z)}\rho\,d\rho.$$

После выполнения двух простых интегрирований по ρ и по θ приходим к формуле (ср. т. I, гл. V, стр. 329)

$$V=\pi\int\limits_a^b[\varphi(z)]^2\,dz.$$

Эту формулу можно вывести и непосредственно, наглядным путем. Разрежем наше тело вращения плоскостями, перпендикулярными к оси z, на тонкие пластинки $z_k\leqslant z\leqslant z_{k+1}$ и обозначим через m_k наименьшее и через M_k наибольшее значения расстояния $\varphi(z)$ от точки поверхности до оси вращения в пластинке с номером k. Тогда объем этой пластинки заключается между объемами двух цилиндров с общей высотой $\Delta z=z_{k+1}-z_k$ и с радиусами m_k и M_k. Стало быть, объем V тела вращения удовлетворяет неравенствам

$$\sum\pi m_k^2\Delta z\leqslant V\leqslant\sum\pi M_k^2\Delta z.$$

Согласно определению обыкновенного определенного интеграла, отсюда вытекает, что

$$V=\pi\int\limits_a^b[\varphi(z)]^2\,dz.$$

Если поверхность тела выражается в сферических координатах r, θ, φ несложным уравнением вида

$$r=f(\theta,\varphi),$$

где функция $f(\theta,\varphi)$ однозначна, то для вычисления объема тела часто удобно перейти к этим сферическим координатам. Так как якобиан $\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)}=r^2\sin\theta$ (см. § 4, n° 2 этой главы), то

$$V=\iiint\limits_G dx\,dy\,dz=\iiint\limits_{G'}r^2\sin\theta\,dr\,d\theta\,d\varphi=\int\limits_0^{2\pi}d\varphi\int\limits_0^{\pi}\sin\theta\,d\theta\int\limits_0^{f(\theta,\varphi)}r^2\,dr.$$

Выполнив первое интегрирование (по r), получим

$$V=\frac{1}{3}\int\limits_0^{2\pi}d\varphi\int\limits_0^{\pi}[f(\theta,\varphi)]^3\sin\theta\,d\theta.$$

В частном случае шара, когда $f(\theta,\varphi)=R=\text{const}$, по этой формуле легко вычисляется объем шара $V=\frac{4}{3}\pi R^3$.

3. Площадь кривой поверхности. В свое время мы выразили длину дуги кривой с помощью обыкновенного определенного интеграла. Теперь мы найдем аналогичное выражение для площади кривой поверхности с помощью двойного интеграла. Длину дуги мы рассматривали как предел периметра вписанной ломаной при стремлении к нулю всех ее прямолинейных звеньев. Для определения площади кривой поверхности естественно напрашивается аналогичный путь: надо вписать в поверхность многогранник, состоящий из плоских треугольных граней, и найти площадь этого многогранника; затем надо изменять вписанную сеть треугольников таким образом, чтобы она становилась все более частой и чтобы наибольшее из ребер многогранника стремилось к нулю; остается наконец найти предел переменной площади многогранника. Этот-то предел и следовало бы рассматривать как определение площади кривой поверхности. Оказывается, однако, что такое определение площади было бы лишено точного смысла, ибо намеченный здесь общий процесс не дает единого предела. Объясняется это следующим обстоятельством. Ломаная, вписанная в гладкую дугу кривой, обладает тем свойством, что каждое ее звено дает сколь угодно точное приближение направления кривой в соответствующей вершине ломаной, если только все звенья достаточно малы. Это вытекает из теоремы дифференциального исчисления о среднем значении. Совершенно иначе обстоит дело с поверхностью. Грань вписанного в кривую поверхность многогранника может иметь сколь угодно крутой наклон относительно касательной плоскости в какой-либо соседней точке поверхности, как бы малы ни были стороны всех треугольных граней. Поэтому площадь поверхности такого вписанного многогранника отнюдь не может служить приближением к площади кривой поверхности. В последнем параграфе Дополнений к этой главе мы подробно разберем пример, иллюстрирующий эти соображения.

Однако, в основу определения длины гладкой кривой можно положить вместо вписанной ломаной *описанную*, т. е. такую ломаную, каждое звено которой касается кривой. Делается это следующим образом. Пусть дана кривая $y=f(x)$, причем функция f имеет непрерывную производную $f'(x)$; требуется определить длину дуги этой кривой между ее точками с абсциссами a и b. Разбиваем интервал от a до b точками деления $a=x_0, x_1, x_2, \ldots, x_n=b$ на n равных или неравных частей; выбираем в k-м частичном интервале произвольную точку ξ_k и в соответствующей точке кривой проводим касательную. Длина l_k того отрезка этой касательной, который лежит в полосе $x_k \leqslant x \leqslant x_{k+1}$, выражается так:

$$l_k=(x_{k+1}-x_k)\sqrt{1+[f'(\xi_k)]^2}.$$

Тогда длина дуги кривой между точками $x=a$ и $x=b$ определяется как предел суммы $\sum\limits_{k=1}^{n} l_k$, когда число точек деления n безгранично возрастает, а длина наибольшего из частичных интервалов стремится к нулю. Из определения интеграла сразу вытекает знакомая нам формула для длины дуги s:

$$s = \int_a^b \sqrt{1 + [f'(x)]^2}\,dx.$$

По аналогии с таким определением длины дуги как предела длины *описанной* ломаной действительно можно построить определение площади кривой поверхности.

Рассмотрим сначала поверхность $z = f(x, y)$, где функция $f(x, y)$ однозначна и имеет непрерывные частные производные. Пусть интересующий нас кусок этой поверхности проектируется на плоскость xy в виде замкнутой области G этой плоскости. Разобьем область G на n ячеек G_1, G_2, ..., G_n с площадями ΔG_1, ΔG_2, ..., ΔG_n и в каждой из этих ячеек выберем по точке: (ξ_1, η_1), (ξ_2, η_2), ..., (ξ_n, η_n). В точке поверхности с координатами ξ_k, η_k и $\zeta_k = f(\xi_k, \eta_k)$ мы построим касательную плоскость и найдем площадь $\Delta\sigma_k$ того куска касательной плоскости, который имеет своей проекцией на плоскость xy ячейку G_k. Обозначим через γ_k угол между касательной плоскостью

$$z - \zeta_k = f_x(\xi_k, \eta_k)(x - \xi_k) + f_y(\xi_k, \eta_k)(y - \eta_k)$$

и плоскостью xy. Согласно гл. III, § 2,

$$\cos\gamma_k = \frac{\pm 1}{\sqrt{1 + [f_x(\xi_k, \eta_k)]^2 + [f_y(\xi_k, \eta_k)]^2}}.$$

Так как ячейка G_k является проекцией куска $\Delta\sigma_k$ касательной плоскости на плоскость xy, то

$$\Delta G_k = \Delta\sigma_k \cdot |\cos\gamma_k|,$$

а следовательно, $\Delta\sigma_k = \sqrt{1 + f_x^2(\xi_k, \eta_k) + f_y^2(\xi_k, \eta_k)} \cdot \Delta G_k$. Составим теперь сумму всех этих площадей:

$$\sum_{k=1}^{n} \Delta\sigma_k,$$

а затем будем безгранично увеличивать число ячеек деления n, одновременно заставляя стремиться к нулю диаметр (а вместе с тем и площадь) наибольшей из ячеек G_k. На основании определения двойного интеграла эта сумма будет тогда стремиться к пределу

$$\sigma = \iint\limits_G \sqrt{1 + f_x^2 + f_y^2}\,dS.$$

Этот интеграл, значение которого не зависит ни от способа разбиения области G на ячейки G_k, ни от выбора точек (ξ_k, η_k), мы и *примем за определение площади заданного куска кривой поверхности.*

Если заданная поверхность есть плоскость, то это определение совпадает с данным ранее, что нетрудно установить. Проще всего это проверить, если $z=f(x, y)=0$; в этом случае получается

$$\sigma=\iint\limits_G dS=\iint\limits_G dx\,dy.$$

Подынтегральное выражение интеграла, выражающего площадь кривой поверхности,

$$d\sigma=\sqrt{1+f_x^2+f_y^2}\,dS=\sqrt{1+f_x^2+f_y^2}\,dx\,dy$$

принято называть *элементом площади* поверхности $z=f(x, y)$. Тогда интеграл, дающий площадь поверхности, можно записать символически в следующем виде:

$$\sigma=\iint\limits_G d\sigma.$$

В качестве примера вычислим площадь шаровой поверхности. Уравнение $z=\sqrt{R^2-x^2-y^2}$ представляет полусферу радиуса R. В этом случае

$$\frac{\partial z}{\partial x}=-\frac{x}{\sqrt{R^2-x^2-y^2}},\quad \frac{\partial z}{\partial y}=-\frac{y}{\sqrt{R^2-x^2-y^2}};$$

стало быть, площадь полусферы равна интегралу

$$\frac{\sigma}{2}=R\iint\limits_G\frac{dx\,dy}{\sqrt{R^2-x^2-y^2}},$$

в котором область интегрирования G есть круг $x^2+y^2\leqslant R^2$ в плоскости xy. Введем полярные координаты r, θ и представим двойной интеграл в виде повторного; тогда

$$\frac{\sigma}{2}=R\int\limits_0^{2\pi}d\theta\int\limits_0^R\frac{r\,dr}{\sqrt{R^2-r^2}}=2\pi R\int\limits_0^R\frac{r\,dr}{\sqrt{R^2-r^2}}.$$

Стоящий справа определенный интеграл легко вычисляется с помощью замены переменной $R^2-r^2=u$, откуда

$$\frac{\sigma}{2}=-\pi R\int\limits_{R^2}^0\frac{du}{\sqrt{u}}=\pi R\int\limits_0^{R^2}\frac{du}{\sqrt{u}}=2\pi R\sqrt{u}\,\Big|_0^{R^2}=2\pi R^2.$$

Следовательно, площадь шаровой поверхности $\sigma=4\pi R^2$, что совпадает с известной формулой элементарной геометрии.

Пусть теперь поверхность задана не уравнением вида $z=f(x, y)$, а неявным уравнением $\varphi(x, y, z)=0$. Предположим, что на рассматриваемом куске поверхности $\varphi_z\neq 0$, например $\varphi_z>0$; тогда

$$\frac{\partial z}{\partial x}=-\frac{\varphi_x}{\varphi_z},\quad \frac{\partial z}{\partial y}=-\frac{\varphi_y}{\varphi_z}$$

и для площади участка поверхности получится выражение

$$\sigma = \iint\limits_G \sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}\, \frac{1}{\varphi_z}\, dx\, dy;$$

напоминаем, что областью интегрирования G служит проекция куска поверхности на плоскость xy.

В обеих формулах для площади мы выделяли координату z. С тем же правом можно задать поверхность уравнением вида $x = x(y, z)$ или $y = y(x, z)$; тогда площадь куска поверхности выразится интегралами вида

$$\iint\limits_{G'} \sqrt{1 + x_y^2 + x_z^2}\, dy\, dz \quad \text{или} \quad \iint\limits_{G''} \sqrt{1 + y_x^2 + y_z^2}\, dz\, dx,$$

в которых область G' есть проекция куска поверхности на плоскость yz, а G'' — ее проекция на плоскость zx.

Если поверхность задана неявным уравнением $\varphi(x, y, z) = 0$, то формулу для площади можно также записать в трех разных видах: в виде интеграла, найденного выше, и в виде следующих двух интегралов:

$$\sigma = \iint\limits_{G'} \sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}\, \frac{1}{\varphi_x}\, dy\, dz \quad \text{или} \quad \sigma = \iint\limits_{G''} \sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}\, \frac{1}{\varphi_y}\, dz\, dx.$$

Что все эти выражения действительно определяют одну и ту же площадь, видно из самого процесса, послужившего определением площади кривой поверхности. Можно, однако, проверить равенство всех этих различных по форме выражений прямым преобразованием. Применим, например, к интегралу

$$\iint\limits_G \frac{\sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}}{\varphi_z}\, dx\, dy$$

преобразование $x = x(y, z)$, $y = y$, в котором мы $x = x(y, z)$ получили, разрешив уравнение $\varphi(x, y, z) = 0$ относительно x. Якобиан преобразования $\frac{\partial(x, y)}{\partial(y, z)} = -x_z = \frac{\varphi_z}{\varphi_x}$; следовательно,

$$\iint\limits_G \frac{\sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}}{\varphi_z}\, dx\, dy = \iint\limits_{G'} \frac{\sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}}{\varphi_x}\, dy\, dz.$$

Область G', по которой берется интеграл в правой части, есть проекция куска поверхности на плоскость yz.

Из трех интегральных формул для площади получаются следующие три выражения для элемента площади $d\sigma$ поверхности, заданной неявным уравнением $\varphi(x, y, z) = 0$:

$$d\sigma = \frac{\sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}}{\varphi_z}\, dx\, dy = \frac{\sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}}{\varphi_x}\, dy\, dz = \frac{\sqrt{\varphi_x^2 + \varphi_y^2 + \varphi_z^2}}{\varphi_y}\, dz\, dx.$$

4. Площадь поверхности, заданной параметрическими уравнениями. Если желательно получить для площади поверхности такую формулу, которая была бы свободна от всякого частного допущения о расположении поверхности относительно осей координат, то надо пользоваться параметрическим заданием поверхности:

$$x = \varphi(u, v), \qquad y = \psi(u, v), \qquad z = \chi(u, v). \tag{1}$$

Интересующему нас куску поверхности соответствует тогда определенная область G^* плоскости uv. Для того чтобы ввести в имеющиеся формулы параметры u и v, рассмотрим сначала такой кусок поверхности, на котором якобиан $\frac{\partial(x, y)}{\partial(u, v)} = D$ всюду положителен. Согласно гл. III, § 3, n°7 и n°4, на этом куске поверхности можно выразить u и v как функции от x и y, и частные производные от этих функций вычисляются по следующим формулам:

$$u_x = \frac{\psi_v}{D}, \qquad u_y = -\frac{\varphi_v}{D}, \qquad v_x = -\frac{\psi_u}{D}, \qquad v_y = \frac{\varphi_u}{D}.$$

В силу соотношений

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} u_x + \frac{\partial z}{\partial v} v_x \quad \text{и} \quad \frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} u_y + \frac{\partial z}{\partial v} v_y,$$

подынтегральная функция двойного интеграла для площади принимает следующий вид:

$$\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} =$$
$$= \frac{1}{D} \sqrt{(\varphi_u\psi_v - \psi_u\varphi_v)^2 + (\psi_u\chi_v - \chi_u\psi_v)^2 + (\chi_u\varphi_v - \varphi_u\chi_v)^2}.$$

Стало быть, по правилам преобразования двойного интеграла к новым переменным интегрирования u и v (§ 4), получаем следующую формулу для площади куска поверхности:

$$\sigma = \iint\limits_{G'} \sqrt{(\varphi_u\psi_v - \psi_u\varphi_v)^2 + (\psi_u\chi_v - \chi_u\psi_v)^2 + (\chi_u\varphi_v - \varphi_u\chi_v)^2}\, du\, dv. \tag{*}$$

В этом параметрическом выражении для площади уже ни одна из координат x, y, z не занимает особого положения. Ясно, что к этому же выражению преобразуется любое из непараметрических выражений площади, в котором одна из координат играет предпочтительную роль. Каждая из этих непараметрических формул пригодна на таком участке поверхности (1), на котором не обращается в нуль хотя бы один из якобианов

$$\frac{\partial(y, z)}{\partial(u, v)}, \qquad \frac{\partial(x, z)}{\partial(u, v)}, \qquad \frac{\partial(x, y)}{\partial(u, v)},$$

а следовательно, формула (*) дает площадь любого куска поверхности (1), в каждой точке которого хотя бы один из трех якобианов отличен от нуля.

Параметрическое выражение (*) для площади поверхности можно привести к другому, замечательному, виду. Для этого воспользуемся гауссовыми коэффициентами (гл. III, § 4, n° 2) квадрата линейного элемента поверхности (1):

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2,$$

т. е. выражениями

$$E = \varphi_u^2 + \psi_u^2 + \chi_u^2, \qquad F = \varphi_u\varphi_v + \psi_u\psi_v + \chi_u\chi_v, \qquad G = \varphi_v^2 + \psi_v^2 + \chi_v^2.$$

Несложное вычисление дает

$$EG - F^2 = (\varphi_u\psi_v - \psi_u\varphi_v)^2 + (\psi_u\chi_v - \chi_u\psi_v)^2 + (\chi_u\varphi_v - \varphi_u\chi_v)^2$$

(ср. стр. 181), а это — подынтегральное выражение в формуле (*). Следовательно, формула для площади куска поверхности (1) принимает следующий вид:

$$\sigma = \iint\limits_{G^*} \sqrt{EG - F^2}\,du\,dv, \tag{$*\atop*$}$$

а выражение элемента площади

$$d\sigma = \sqrt{EG - F^2}\,du\,dv.$$

В качестве примера вычислим снова площадь σ сферы радиуса R, но на этот раз по ее параметрическому заданию:

$$x = R\cos u \sin v, \qquad y = R\sin u \sin v, \qquad z = R\cos v,$$

где u и v изменяются в области $0 \leqslant u \leqslant 2\pi$, $0 \leqslant v \leqslant \pi$. Отсюда $d\sigma = R^2 \sin v\,du\,dv$. Несложное вычисление приводит к повторному интегралу $\sigma = R^2\int\limits_0^{2\pi} du \int\limits_0^{\pi} \sin v\,dv$ и дает для площади сферы тот же результат $\sigma = 4\pi R^2$.

Рассчитаем, какой вид примет формула ($*\atop*$) в приложении к поверхности вращения; пусть эта поверхность образуется при вращении кривой $z = f(x)$ плоскости xz вокруг оси z. В качестве параметров возьмем полярные координаты $r = u$, $\theta = v$ в плоскости xy. Наша поверхность вращения будет тогда представлена следующими параметрическими уравнениями:

$$x = u\cos v, \qquad y = u\sin v, \qquad z = f\left(\sqrt{x^2 + y^2}\right) = f(u).$$

Гауссовы коэффициенты имеют такие значения:

$$E = 1 + [f'(u)]^2, \qquad F = 0, \qquad G = u^2;$$

элемент площади

$$d\sigma = \sqrt{EG - F^2}\,du\,dv = u\sqrt{1 + [f'(u)]^2}\,du\,dv,$$

а площадь пояса поверхности вращения между плоскостями $z=z_1$ и $z=z_2$ (которым соответствуют значения параметра $u=u_1$ и $u=u_2$) выразится так:

$$\sigma=\int_0^{2\pi} dv \int_{u_1}^{u_2} u\sqrt{1+[f'(u)]^2}\,du=2\pi\int_{u_1}^{u_2} u\sqrt{1+[f'(u)]^2}\,du.$$

Если ввести в качестве параметра вместо u длину дуги s меридианной кривой $z=f(u)$, то формула для площади пояса *поверхности вращения* приведется к виду

$$\sigma=2\pi\int_{s_1}^{s_2} u\,ds,$$

где $u=u(s)$ есть расстояние от той точки вращающейся кривой, которая соответствует значению s, до оси вращения (правило Гульдина; ср. т. I, стр. 329).

В качестве примера вычислим площадь тора (ср. гл. III, § 4, конец п° 2), порождаемого вращением окружности $(y-a)^2+z^2=r^2$ вокруг оси z. Пусть $a>0$ и $0<r<a$. За параметр примем длину дуги s окружности, отсчитываемую от точки $y=a+r$, $z=0$ против часовой стрелки. Тогда $u=a+r\cos\frac{s}{r}$, и площадь всего тора равна

$$\sigma=2\pi\int_0^{2\pi r} u\,ds=2\pi\int_0^{2\pi r}\left(a+r\cos\frac{s}{r}\right)ds=2\pi a\cdot 2\pi r.$$

Стало быть, площадь тора равна произведению длины вращающейся окружности на длину пути, описываемого центром этой окружности.

Упражнения

1. Вычислить объем тела, ограниченного поверхностью

$$\frac{(\sqrt{x^2+y^2}-1)^2}{a^2}+\frac{z^2}{b^2}=1 \qquad (a<1).$$

2. Вычислить объем, отсекаемый от параболоида

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z$$

плоскостью $z=h$.

3. Найти объем, отсекаемый от эллипсоида

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

плоскостью $lx+my+nz=p$.

4. На поверхности, заданной в сферических координатах r, θ, φ уравнением

$$r^2=a^2\cos 2\theta,$$

начерчена какая-либо замкнутая кривая $\theta=f(\varphi)$.

а) Показать, что площадь участка поверхности, ограниченного этой кривой, равна площади центральной проекции этого участка на сферу $r=a$, причем центром проекций является полюс (начало координат).

б) Выразить эту площадь простым определенным интегралом.

в) Найти площадь всей поверхности.

5. Найти площадь поверхности сфероида, полученного вращением эллипса вокруг своей большой оси, и показать, что если можно пренебречь четвертой и более высокими степенями эксцентриситета ε, то эта площадь равна площади поверхности шара, объем которого равен объему сфероида.

6. Вычислить объем и площадь поверхности тела, полученного вращением треугольника ABC вокруг стороны AB.

7*. Трубчатая поверхность является огибающей семейства сфер радиуса 1, центры которых образуют замкнутую плоскую кривую L. Доказать, что площадь σ этой поверхности равна длине кривой L, умноженной на 2π.

8*. Из шара $x^2+y^2+z^2\leqslant a^2$ вырезаны отверстия двумя цилиндрическими поверхностями

$$x^2+y^2-ax=0 \quad \text{и} \quad x^2+y^2+ax=0.$$

а) Вычислить объем оставшейся части шара.

б) Вычислить площадь оставшейся части поверхности шара.

9. Вычислить площадь той части винтовой поверхности

$$y-x\operatorname{tg}\frac{z}{h}=0,$$

для которой $r^2\leqslant x^2+y^2\leqslant R^2$, $|z|\leqslant\frac{\pi}{2}h$.

10. Вычислить площадь поверхности $(x^2+y^2+z^2)^2=x^2-y^2$.

§ 7. Приложения к физике

В § 2, n° 7 было уже показано, каким образом понятие массы связано с понятием кратного интеграла. В этом параграфе мы займемся некоторыми другими понятиями механики. В первую очередь мы рассмотрим статический момент и момент инерции более подробно, чем это было возможно в т. I, гл. X, § 6, n° 7.

1. Статический момент и центр массы (центр тяжести). *Статическим моментом материальной точки массы m относительно плоскости xy называется произведение $T_z=mz$ ее массы m на аппликату z.* Соответственно этому статическим моментом относительно плоскости yz называется произведение $T_x=mx$, а относительно плоскости zx — выражение $T_y=my$. *Статические моменты нескольких материальных точек находят сложением,* т. е. статические моменты системы n масс $m_1, m_2, \ldots, m_n$, сосредоточенных в точках $(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)$ относительно плоскостей yz, zx и xy, определяются соответственно выражениями

$$T_x=\sum_{\nu=1}^{n} m_\nu x_\nu, \qquad T_y=\sum_{\nu=1}^{n} m_\nu y_\nu, \qquad T_z=\sum_{\nu=1}^{n} m_\nu z_\nu.$$

Если мы имеем дело не с конечным числом материальных точек, а с массой, распределенной непрерывно по пространственной области либо по поверхности, либо вдоль кривой с плотностью $\mu=\mu(x, y, z)$, то статические моменты такой распределенной массы определяют с помощью предельного перехода, как в т. I (гл. V, § 2, n° 9 и гл. X, § 6, n° 7), и таким образом статические моменты выражаются в виде интегралов. Если, например, масса распределена по пространственной области G, то эту область разбивают на n ячеек, массу каждой ячейки представляют себе сосредоточенной в какой-либо из ее точек и находят статические моменты этой системы n точечных масс, имеющие типичный вид интегральных сумм. При $n\to\infty$ с одновременным стремлением к нулю наибольшего из диаметров всех n ячеек полученные суммы будут иметь своими пределами интегралы

$$T_x=\iiint\limits_G \mu x\,dx\,dy\,dz,\quad T_y=\iiint\limits_G \mu y\,dx\,dy\,dz,\quad T_z=\iiint\limits_G \mu z\,dx\,dy\,dz,$$

которые и служат определением *статических моментов пространственно распределенной массы.*

Если масса распределена по поверхности Σ, заданной параметрическими уравнениями $x=\varphi(u, v)$, $y=\psi(u, v)$, $z=\chi(u, v)$ с поверхностной плотностью $\mu(u, v)$, *то статические моменты этой массы, распределенной по поверхности* Σ, определяются аналогичным образом как интегралы

$$T_x=\iint\limits_\Sigma \mu x\,d\sigma=\iint\limits_G \mu x\sqrt{EG-F^2}\,du\,dv,$$
$$T_y=\iint\limits_\Sigma \mu y\,d\sigma=\iint\limits_G \mu y\sqrt{EG-F^2}\,du\,dv,$$
$$T_z=\iint\limits_\Sigma \mu z\,d\sigma=\iint\limits_G \mu z\sqrt{EG-F^2}\,du\,dv.$$

Наконец, *статические моменты массы, распределенной вдоль пространственной кривой* $x=x(s)$, $y=y(s)$, $z=z(s)$ с линейной плотностью $\mu(s)$, определяются как обычные определенные интегралы

$$T_x=\int\limits_{s_0}^{s_1}\mu x\,ds,\quad T_y=\int\limits_{s_0}^{s_1}\mu y\,ds,\quad T_z=\int\limits_{s_0}^{s_1}\mu z\,ds,$$

где s обозначает длину дуги кривой.

Центром массы или *центром тяжести* массы M, распределенной по объему, по поверхности или вдоль линии, называется точка, имеющая следующие координаты:

$$\xi=\frac{T_x}{M},\quad \eta=\frac{T_y}{M},\quad \zeta=\frac{T_z}{M},$$

где M — общее количество распределенной массы. При этом, если масса распределена по пространственной области G, то

$$M=\iiint\limits_G \mu(x, y, z)\,dx\,dy\,dz;$$

если масса распределена по куску поверхности Σ, то

$$M=\iint_{\Sigma}\mu\,d\sigma=\iint_{G}\mu(u,v)\sqrt{EG-F^2}\,du\,dv;$$

если же масса распределена вдоль дуги кривой, то

$$M=\int_{s_0}^{s_1}\mu(s)\,ds.$$

Рассмотрим два примера. 1) Найдем центр массы однородного полушара H:

$$x^2+y^2+z^2\leqslant 1 \qquad (z\geqslant 0)$$

плотности $\mu=1$.

Статические моменты относительно плоскостей yz и zx равны нулю, т. е.

$$T_x=\iiint_H x\,dx\,dy\,dz=0 \quad\text{и}\quad T_y=\iiint_H y\,dx\,dy\,dz=0,$$

ибо уже интегрирование по x в первом случае и по y во втором дает в результате нуль. Для вычисления статического момента полушара относительно плоскости xy:

$$T_z=\iiint_H z\,dx\,dy\,dz$$

перейдем к цилиндрическим координатам r, θ, z с помощью преобразования

$$x=r\cos\theta,\quad y=r\sin\theta,\qquad z=z$$

и получим

$$T_z=\int_0^1 z\,dz\int_0^{\sqrt{1-z^2}} r\,dr\int_0^{2\pi}d\theta=2\pi\int_0^1\frac{1-z^2}{2}z\,dz=\pi\left(\frac{z^2}{2}-\frac{z^4}{4}\right)\Big|_0^1=\frac{\pi}{4}.$$

Так как масса нашего полушара $M=\frac{2\pi}{3}$, то центр массы полушара находится в точке $\left(0,\,0,\,\frac{3}{8}\right)$.

2) Масса распределена равномерно по половине шаровой поверхности радиуса $R=1$ с поверхностной плотностью $\mu=1$. Найти центр массы этой полусферы. Из параметрических уравнений шаровой поверхности

$$x=\cos u\sin v,\qquad y=\sin u\sin v,\qquad z=\cos v$$

вычисляем по формуле из § 6, n° 4 элемент площади

$$d\sigma=\sqrt{EG-F^2}\,du\,dv=\sin v\,du\,dv$$

и с его помощью вычисляем статические моменты

$$T_x=\int_0^{\frac{\pi}{2}}\sin^2 v\,dv\int_0^{2\pi}\cos u\,du=0,$$

$$T_y=\int_0^{\frac{\pi}{2}}\sin^2 v\,dv\int_0^{2\pi}\sin u\,du=0,$$

$$T_z=\int_0^{\frac{\pi}{2}}\sin v\cos v\,dv\int_0^{2\pi}du=2\pi\,\frac{\sin^2 v}{2}\Big|_0^{\frac{\pi}{2}}=\pi.$$

Так как, очевидно, полная масса $M=2\pi$, то центр массы однородной полусферы находится в точке $\left(0,\ 0,\ \frac{1}{2}\right)$.

2. Момент инерции. Понятие момента инерции тоже легко обобщается на случай бо́льшего числа измерений. Моментом инерции материальной точки относительно оси x называется произведение ее массы на $\rho^2=y^2+z^2$, т. е. на квадрат расстояния точки от оси x. Соответственно этому момент инерции массы, распределенной непрерывно по пространственной области G с объемной плотностью $\mu=\mu(x,\ y,\ z)$, относительно оси x определяется как тройной интеграл

$$I_x=\iiint\limits_G \mu(y^2+z^2)\,dx\,dy\,dz;$$

моменты инерции этой массы относительно других осей координат определяются аналогичными интегралами.

Порой вводят также определение момента инерции массы, распределенной по объемной области G, относительно *начала координат* [называемого *полярным моментом инерции*] как тройного интеграла

$$\iiint\limits_G \mu(x^2+y^2+z^2)\,dx\,dy\,dz;$$

легко понять, как надо видоизменить это определение для момента инерции относительно любой другой точки. Момент инерции того же распределения массы относительно *плоскости* yz определяется тройным интегралом

$$\iiint\limits_G \mu x^2 dx\,dy\,dz;$$

аналогичными интегралами выражаются моменты инерции относительно плоскостей zx и xy.

Момент инерции массы, распределенной по куску поверхности Σ, относительно оси x определяется как двойной интеграл

$$\iint\limits_\Sigma \mu(y^2+z^2)\,d\sigma,$$

где поверхностная плотность $\mu(u,\ v)$ задана как непрерывная функция параметров u и v на поверхности, а $d\sigma$ — элемент площади этой поверхности.

Момент инерции I_C тела переменной плотности $\mu(x,y,z)$, заполняющего область G пространства, относительно оси, параллельной оси x и проходящей через точку $C(\xi,\eta,\zeta)$, выражается интегралом

$$I_C=\iiint\limits_G \mu\left[(y-\eta)^2+(z-\zeta)^2\right]dx\,dy\,dz.$$

Это выражение можно представить в следующем виде:

$$I_C=\iiint_G \mu(y^2+z^2)\,dx\,dy\,dz+(\eta^2+\zeta^2)\iiint_G \mu\,dx\,dy\,dz-$$
$$-2\eta\iiint_G \mu y\,dx\,dy\,dz-2\zeta\iiint_G \mu z\,dx\,dy\,dz=$$
$$=I_x+(\eta^2+\zeta^2)M-2\eta T_y-2\zeta T_z$$

где M — масса тела, а T_y и T_z — статические моменты тела относительно плоскостей zx и xy. Если в качестве точки $C(\xi, \eta, \zeta)$ выбрать центр массы тела, то $T_y=M\eta$ и $T_z=M\zeta$; тогда последнее равенство приводит к соотношению

$$I_x=I_C+(\eta^2+\zeta^2)M.$$

Так как любая ось может быть принята за ось x, то полученное соотношение можно сформулировать в виде следующей теоремы, принадлежащей Штейнеру:

Момент инерции твердого тела относительно любой оси равен моменту инерции этого тела относительно параллельной оси, проходящей через его центр массы, сложенному с произведением массы тела на квадрат расстояния от центра массы до первой оси (квадрат расстояния между обеими осями).

Физическое значение момента инерции массы, распределенной по объему или по поверхности — совершенно такое же, как было отмечено в первом томе (гл. V, § 2, n° 11):

Кинетическая энергия тела, вращающегося равномерно вокруг оси, равна половине произведения момента инерции тела относительно оси вращения на квадрат его угловой скорости.

Решим два примера для иллюстрации понятия момента инерции и техники его вычисления в простых случаях.

1) Вычислим моменты инерции относительно осей координат однородного шара V радиуса $R=1$ с центром в начале координат, если объемная плотность $\mu=1$.

Из соображений симметрии ясно, что все три искомых момента инерции равны между собой; обозначим их общее значение через I:

$$I=\iiint_V (x^2+y^2)\,dx\,dy\,dz=\iiint_V (x^2+z^2)\,dx\,dy\,dz=\iiint_V (y^2+z^2)\,dx\,dy\,dz.$$

Сложив эти три интеграла, получим

$$3I=2\iiint_V (x^2+y^2+z^2)\,dx\,dy\,dz.$$

Этот интеграл легко вычисляется, если перейти к сферическим координатам r, θ, φ:

$$I=\frac{2}{3}\int_0^1 r^4\,dr\int_0^\pi \sin\theta\,d\theta\int_0^{2\pi} d\varphi=\frac{2}{3}\cdot\frac{1}{5}\cdot 2\cdot 2\pi=\frac{8\pi}{15}.$$

2) Балка имеет форму прямоугольного параллелепипеда, ребра которого параллельны осям x, y, z и равны соответственно a, b, c; начало координат поместим в центре массы балки. Объемная плотность $\mu = 1$. Найдем момент инерции балки относительно плоскости xy. Имеем

$$I_{xy} = \int\limits_{-\frac{a}{2}}^{\frac{a}{2}} dx \int\limits_{-\frac{b}{2}}^{\frac{b}{2}} dy \int\limits_{-\frac{c}{2}}^{\frac{c}{2}} z^2 dz = ab\frac{c^3}{12}.$$

3. Физический маятник. Введенные выше понятия механики находят приложение к математическому исследованию физического маятника, т. е. твердого тела, совершающего под действием силы тяжести вращательные колебания вокруг неподвижной оси.

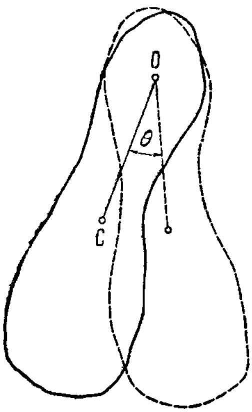


Рис. 77.

Проведем через центр массы C твердого тела плоскость, перпендикулярную к оси вращения; пусть эта плоскость пересекает ось в точке O (рис. 77). Движение тела будет, очевидно, полностью известно, если удастся определить угол $\theta = \theta(t)$, образуемый вектором $\overline{OC}$ с направленной вниз вертикалью, проходящей через точку O, как функцию времени t. Для того чтобы найти эту функцию $\theta(t)$, а затем и период колебания маятника, мы воспользуемся законом сохранения энергии, который утверждает, что во время движения тела сумма его кинетической и его потенциальной энергии остается постоянной (ср. гл. VI, § 1, n° 2). Потенциальная энергия нашего тела $U = Mgh$, где M — масса тела, g — ускорение силы тяжести, а h — высота центра массы над какой-либо фиксированной горизонтальной плоскостью, например над горизонтальной плоскостью, проходящей через положение центра массы в состоянии покоя. Обозначим расстояние OC центра массы от оси вращения через r; тогда $U = Mgr(1 - \cos\theta)$. Кинетическая энергия вращающегося тела (см. конец предыдущего номера) есть $T = \frac{1}{2} I\dot{\theta}^2$, где I — момент инерции тела относительно оси вращения, а $\dot{\theta} = \frac{d\theta}{dt}$. Следовательно, закон сохранения энергии приводит к уравнению

$$\frac{1}{2} I\dot{\theta}^2 - Mgr\cos\theta = \text{const} \quad \text{или} \quad \frac{1}{2}\dot{\theta}^2 - \frac{gMr}{I}\cos\theta = \text{const}. \qquad (1)$$

В т. I, гл. V, § 5, n° 3 было изучено колебание математического маятника, представляющего идеализацию физического маятника. Физическим осуществлением математического маятника может служить небольшое тело, подвешенное в поле силы тяжести на длинной нити и колеблющееся вокруг точки подвеса O, при условии, что можно пренебречь линейными размерами тела по сравнению с длиной l нити, а массой нити по сравнению с массой M тела. Составим уравнение энергии для математического маятника. Его кинетическая энергия $T_1 = \frac{1}{2} Mv^2 = \frac{1}{2} M (l\dot{\theta})^2$, а потенциальная энергия $U_1 = Mgl(1 - \cos\theta)$. Следовательно, уравнение энергии будет

$$\frac{1}{2} M(l\dot{\theta})^2 - Mgl\cos\theta = \text{const} \quad \text{или} \quad \frac{1}{2}\dot{\theta}^2 - \frac{g}{l}\cos\theta = \text{const}. \quad (2)$$

(Оно совпадает с уравнением $\frac{1}{2}\dot{s}^2 = -g(y - y_0)$, выведенным в т. I, § 5, n° 1 для несколько более общей задачи, если учесть, что для математического маятника $y = -l\cos\theta$.)

Сопоставляя уравнения энергии (1) и (2), видим, что они полностью совпадают, если ввести постоянную $l = \frac{I}{Mr}$, которая называется *приведенной длиной* физического маятника, — это длина такого математического маятника, который эквивалентен данному физическому маятнику.

Отсюда вытекает, что все формулы, выведенные в первом томе для математического маятника, можно непосредственно использовать для физического маятника. Так мы получим для периода колебания формулу

$$T = \sqrt{\frac{2l}{g}} \int_{-\alpha}^{\alpha} \frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}} = \sqrt{\frac{2I}{Mgr}} \int_{-\alpha}^{\alpha} \frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}},$$

где α — угол наибольшего отклонения, т. е. амлитуда колебания. Для малых колебаний получается приближенная формула

$$T = 2\pi\sqrt{\frac{l}{g}} = 2\pi\sqrt{\frac{I}{Mgr}}.$$

Естественно, что период математического маятника получается из этой формулы как частный случай. Действительно, если вся масса M сосредоточена в центре массы, то $I = Mr^2$, так что приведенная длина $l = \frac{I}{Mr} = r$.

Для продолжения исследования вспомним, что момент инерции I относительно оси вращения связан с моментом инерции I_O относительно параллельной оси, проходящей через центр массы соотношением

$$I = I_O + Mr^2$$

(см. в конце предыдущего номера). Стало быть,

$$l = \frac{I}{Mr} = r + \frac{I_O}{Mr}$$

или, если введем новую постояннную $a=\frac{I_0}{M}$,

$$l=r+\frac{a}{r}.$$

Осюда сразу видно, что у физического маятника приведенная длина $l>r$, а потому его период колебания всегда больше, чем период колебания математического маятника, который получится, если всю массу сосредоточить в центре массы. Ясно также, что период колебания вокруг любой из параллельных осей, находящихся на одинаковом расстоянии r от центра массы, всегда один и тот же. Это видно из того, что приведенная длина физического маятника зависит только от двух величин: r и $a=\frac{I_0}{M}$, и в силу этого остается неизменной, если сохраняется направление оси вращения и расстояние этой оси от центра массы.

Приведенная длина $l=r+\frac{a}{r}$ не изменяется и в том случае, если подставим вместо r величину $\frac{a}{r}$, т. е. если проведем ось вращения того же направления не на расстоянии r от центра массы, а на расстоянии $\frac{a}{r}$ от него. Это значит, что физический маятник имеет один и тот же период колебания вокруг всех параллельных осей, находящихся на расстоянии r или $\frac{a}{r}$ от центра массы.

Из формулы $T=2\pi\sqrt{\frac{r+\frac{a}{r}}{g}}$ видно, что период колебания T неограниченно возрастает, если r стремится к нулю или к бесконечности. Стало быть, он должен иметь минимум при некотором определенном значении r_0. Решая обычным путем задачу на экстремум, обнаружим, что период T имеет единственный экстремум, а следовательно, наименьшее значение при

$$r_0=\sqrt{a}=\sqrt{\frac{I_0}{M}}.$$

Маятник, ось вращения которого находится на расстоянии $r_0=\sqrt{\frac{I_0}{M}}$ от его центра массы, будет относительно нечувствителен к малым смещениям оси, ибо при $r=r_0$ производная $\frac{dT}{dr}$ обращается в нуль, так что изменениям расстояния r первого порядка будут соответствовать изменения периода T второго порядка. Этот факт используется в технике при конструировании точных часов (маятник Шулера).

4. **Потенциал поля тяготения.** В гл. II, § 7, n° 6 мы видели, что, согласно закону тяготения Ньютона, неподвижная частица Q массы m с координатами (ξ, η, ζ) притягивает другую точку P массы

1 с координатами (x, y, z) с силой

$$F = \operatorname{grad} \frac{m}{r},$$

где $r = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}$ есть расстояние между точками Q и P. Направление силы совпадает с направлением вектора $\overline{PQ}$, а ее величина обратно пропорциональна квадрату расстояния между точками Q и P.

[При вычислении градиента надо рассматривать ξ, η, ζ как постоянные параметры и дифференцировать по x, y, z. Автор здесь и ниже предполагает, что единицы участвующих в этой формуле величин выбраны таким образом, что постоянная тяготения равна единице. Если пользоваться любой заранее установленной системой единиц, то необходимо включить в правую часть множителем постоянную тяготения γ, так что $F = \gamma m \operatorname{grad} \frac{1}{r}$.]

Если источником поля тяготения является система n точек Q_1, Q_2, ..., Q_n, снабженных массами m_1, m_2, ..., m_n, то результирующая сила притяжения, действующая на материальную точку P массы 1, тоже можно представить в виде

$$F = \operatorname{grad}\left(\frac{m_1}{r_1} + \frac{m_2}{r_2} + \ldots + \frac{m_n}{r_n}\right),$$

где r_k обозначает расстояние между точками Q_k и P.

Если сила поля может быть представлена как градиент некоторой скалярной функции точки, то эту функцию точки обычно называют *потенциалом силового поля*; в соответствии с этим мы будем называть потенциалом поля тяготения системы частиц Q_1, Q_2, ..., Q_n в точке $P(x, y, z)$ скалярную функцию

$$U(x, y, z) = \sum_{k=1}^{n} \frac{m_k}{\sqrt{(x-\xi_k)^2 + (y-\eta_k)^2 + (z-\zeta_k)^2}}.$$

[В физике эту функцию $U(x, y, z)$ обычно называют силовой функцией, а потенциалом принято называть функцию $V = -U$, так что сила поля $F = \operatorname{grad} U = -\operatorname{grad} V$.]

Предположим теперь, что массы, порождающие поле тяготения, не сосредоточены в конечном числе точек, а распределены по области G пространства либо по участку Σ поверхности, либо по дуге L кривой с непрерывной плотностью $\mu(\xi, \eta, \zeta)$. [В первом случае это будет объемная плотность, во втором — поверхностная, в третьем — линейная плотность.] Тогда в качестве определения потенциала поля в точке P, лежащей вне области G или участка Σ поверхности, или дуги L, естественно, принимают: в первом случае тройной интеграл

$$\iiint\limits_{G} \frac{\mu(\xi, \eta, \zeta)}{r}\, d\xi\, d\eta\, d\zeta,$$

во втором случае — интеграл по поверхности[1])

$$\iint_{\Sigma} \frac{\mu}{r}\, d\sigma,$$

в третьем случае — обычный определенный интеграл

$$\int_{s_0}^{s_1} \frac{\mu(s)}{r}\, ds,$$

где переменной интегрирования служит длина дуги s кривой L. Во всех трех интегралах r обозначает расстояние между точкой $Q(\xi, \eta, \zeta)$ области интегрирования и точкой наблюдения P поля тяготения.

Так, например, потенциал в точке $P(x, y, z)$ сплошного шара K радиуса 1, с центром в начале, при постоянной плотности $\mu = 1$, приводится к следующему виду:

$$U = \iiint_K \frac{d\xi\, d\eta\, d\zeta}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}} =$$

$$= \int_{-1}^{+1} d\xi \int_{-\sqrt{1-\xi^2}}^{+\sqrt{1-\xi^2}} d\eta \int_{-\sqrt{1-\xi^2-\eta^2}}^{+\sqrt{1-\xi^2-\eta^2}} \frac{1}{r}\, d\zeta.$$

Во всех этих выражениях координаты точки P участвуют *не* в качестве переменных интегрирования, а в качестве параметров, так что потенциал является всякий раз функцией этих параметров, т. е. функцией точки P.

Для того чтобы получить из потенциала проекции силы на оси координат, надо дифференцировать соответствующий интеграл по параметрам x, y, z. Правила дифференцирования по параметру распространяются непосредственно на кратные интегралы; поэтому на основании § 1, n° 2 можно выполнить дифференцирование под знаком интеграла, если только точка P не принадлежит области интегрирования, т. е. если расстояние r не обращается в нуль ни для какой точки области интегрирования. Например, для координат силы притяжения телом, занимающим область G пространства, единичной массы, помещенной в точку $P(x, y, z)$, получаются следующие выражения:

$$F_1 = -\iiint_G \frac{x-\xi}{r^3}\, d\xi\, d\eta\, d\zeta, \qquad F_2 = -\iiint_G \frac{y-\eta}{r^3}\, d\xi\, d\eta\, d\zeta,$$

$$F_3 = -\iiint_G \frac{z-\zeta}{r^3}\, d\xi\, d\eta\, d\zeta.$$

[1]) В этом месте автор рассматривает интеграл по поверхности как понятие интуитивно очевидное. Точное изложение вопроса и способ вычисления такого интеграла даются в Дополнениях к этой главе, § 3, n° 1 и гл. V, § 4. (*Прим. перев.*)

В заключение подчеркнем, что выражения для потенциала и его первых производных по x, y, z сохраняют все же смысл и в том случае, если точка P лежит внутри области интегрирования, заполненной притягивающими массами. В этом случае интегралы являются несобственными, но их сходимость нетрудно установить с помощью признаков сходимости из § 5, n° 2.

В качестве примера вычислим потенциал, порожденный массами, распределенными по шаровой поверхности Σ радиуса a с плотностью $\mu = 1$, как во внешней, так и во внутренней точке шара.

Из соображений симметрии ясно, что потенциал должен зависеть только от расстояния точки поля P от центра шара. Примем центр шаровой поверхности за начало координат, а ось x проведем через внешнюю или внутреннюю точку P, считая положительным направление от центра к точке P. Тогда точка P будет иметь координаты $(x, 0, 0)$ $(x > 0)$, и потенциал запишется так:

$$U = \oiint_{\Sigma} \frac{d\sigma}{\sqrt{(x-\xi)^2 + \eta^2 + \zeta^2}},$$

где колечко, надетое на символ поверхностного интеграла, показывает, что он распространяется на замкнутую поверхность. Введем на шаровой поверхности Σ сферические координаты с осью абсцисс в качестве полярной оси:

$$\xi = a\cos\theta, \qquad \eta = a\sin\theta\cos\varphi, \qquad \zeta = a\sin\theta\sin\varphi.$$

Тогда

$$U = \int_0^{2\pi} d\varphi \int_0^{\pi} \frac{a^2\sin\theta}{\sqrt{(x-a\cos\theta)^2 + a^2\sin^2\theta}}\, d\theta =$$

$$= 2\pi \int_0^{\pi} \frac{a^2\sin\theta}{\sqrt{x^2 + a^2 - 2ax\cos\theta}}\, d\theta = \frac{2\pi a}{x}\sqrt{x^2 + a^2 - 2ax\cos\theta}\,\Bigg|_{\theta=0}^{\theta=\pi} =$$

$$= \frac{2\pi a}{x}\left(\sqrt{x^2 + a^2 + 2ax} - \sqrt{x^2 + a^2 - 2ax}\right) = \frac{2\pi a}{x}(x + a - |x - a|).$$

Стало быть, если $x > a$, то

$$U = \frac{4\pi a^2}{x},$$

если же $0 < x < a$, то

$$U = 4\pi a.$$

При $x = 0$ требуется отдельное вычисление, которое приводит к тому же результату $U = 4\pi a$. Так как любая полупрямая, выходящая из центра сферы, может быть принята за положительную полуось абсцисс, то x означает здесь просто расстояние от центра до любой точки поля.

Следовательно, потенциал поля вне сферы — такой, как будто бы вся масса сосредоточена в ее центре, внутри же сферы потенциал имеет постоянное значение. На самой поверхности потенциал остается непрерывной функцией; выражение для U сохраняет на ней смысл как сходящийся несобственный интеграл и имеет значение $4\pi a$. Сама же сила притяжения (точнее, напряженность поля) $\boldsymbol{F}$ имеет на поверхности шара конечный разрыв. Чтобы это было яснее, заменим букву x, означающую расстояние точки P от центра шара, более привычной для этого расстояния буквой r. Тогда потенциал

$$U=\begin{cases} \dfrac{4\pi a^2}{r} & \text{при } r\geqslant a, \\ 4\pi a & \text{при } r\leqslant a. \end{cases}$$

Напряженность поля

$$\boldsymbol{F}=\operatorname{grad} U=\begin{cases} -\dfrac{4\pi a^2}{r^2}\boldsymbol{r}^\circ & \text{при } r>a, \\ 0 & \text{при } r<a, \end{cases}$$

где $\boldsymbol{r}^\circ$ — единичный вектор направления радиус-вектора. Стало быть, при переходе через поверхность вектор $\boldsymbol{F}$ совершает скачок от значения 0 до значения $-4\pi\boldsymbol{r}^\circ$.

Потенциал U_1 шара радиуса R, заполненного веществом с объемной плотностью $\mu=1$, в любой точке P, лежащей вне шара, легко вычислить так: помножить потенциал шаровой поверхности радиуса a во внешней точке на da и интегрировать по a от 0 до R:

$$U_1=\int_0^R \frac{4\pi a^2}{r}\,da=\frac{4\pi R^3}{3r} \quad \text{при} \quad r>R.$$

Отсюда видно, что потенциал однородного шарообразного тела в любой точке, лежащей вне его, таков, как будто вся масса $\frac{4}{3}\pi R^3$ сосредоточена в центре.

Упражнения

1. Для следующих ниже тел найти объем, центр массы и моменты инерции относительно каждой из осей x, y и z (объемная плотность $\mu=1$):

а) параллелепипед $0\leqslant x\leqslant a$, $0\leqslant y\leqslant b$, $0\leqslant z\leqslant c$;

б) полушар $0\leqslant z\leqslant\sqrt{a^2-x^2-y^2}$;

в) треугольная призма с вершинами $(0, 0, 0)$, $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$.

2. Определить центр массы боковой поверхности прямого кругового конуса (поверхностная плотность постоянна).

3. Найти координаты центра массы сегмента параболоида $y^2+z^2=px$, отсекаемого плоскостью $x=a$.

4*. Трубчатая поверхность задана как огибающая семейства сфер радиуса 1 с центрами на некоторой кривой плоскости xy. Обозначим через Σ площадь куска этой поверхности, лежащего над плоскостью xy, и через S — площадь проекции этого куска поверхности на плоскость xy. Доказать, что аппликата ζ центра массы куска Σ трубчатой поверхности равна S/Σ.

5. Вычислить момент инерции тела, ограниченного двумя цилиндрическими поверхностями $x^2+y^2=R^2$ и $x^2+y^2=R'^2$ $(R>R')$ и двумя плоскостями $z=h$ и $z=-h$ относительно а) оси z; б) оси x (объемная плотность $\mu=1$).

6. Пусть A, B, C — моменты инерции произвольного тела положительной плотности относительно осей x, y и z. Доказать, что эти три величины удовлетворяют «неравенствам треугольника»

$$A+B>C,\ A+C>B,\ B+C>A.$$

7. Вычислить момент инерции эллипсоида $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ (объемная плотность $\mu=1$) относительно а) оси z; б) произвольной прямой, проходящей через начало координат

$$\frac{x}{\alpha}=\frac{y}{\beta}=\frac{z}{\gamma} \qquad (\alpha^2+\beta^2+\gamma^2=1).$$

8*. Найти огибающую поверхность семейства плоскостей, относительно которых эллипсоид

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

имеет один и тот же момент инерции h. Объемная плотность эллипсоида равна единице.

9. Дано произвольное тело, и пусть O — произвольная точка. На всяком луче, выходящем из точки O, отложим отрезок $OL=1/\sqrt{I_l}$, где I_l — момент инерции тела относительно прямой, совпадающей с этим лучом. Доказать, что геометрическое место концов L этих отрезков есть эллипсоид (так называемый *эллипсоид инерции* тела для точки O).

10. Найти эллипсоид инерции эллипсоида $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ для точки (ξ, η, ζ). Объемная плотность $\mu=1$.

11. Найти координаты центра массы шаровой поверхности $x^2+y^2+z^2=1$, если поверхностная плотность $\mu=\frac{1}{\sqrt{(x-1)^2+y^2+z^2}}$.

12. Найти абсциссу центра массы октанта (восьмой части) эллипсоида:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \qquad x\geqslant 0, \qquad y\geqslant 0, \qquad z\geqslant 0.$$

13. Тело A состоит из двух частей A_1 и A_2. Обозначим через I, I_1, I_2 моменты инерции тела A и его частей A_1 и A_2 относительно трех параллельных осей, проходящих через соответствующие центры массы этих трех тел. Доказать, что

$$I=I_1+I_2+\frac{m_1m_2}{m_1+m_2}d^2,$$

где m_1 и m_2 — массы частей A_1 и A_2, а d — расстояние между осями, проходящими через их центры массы.

14. Вычислить потенциал эллипсоида вращения $\frac{x^2+y^2}{a^2}+\frac{z^2}{b^2}=1$ в его центре $(b>a)$. Объемная плотность $\mu=1$.

15. Вычислить потенциал тела вращения

$$r=\sqrt{x^2+y^2}\leqslant f(z), \qquad a\leqslant z\leqslant b$$

в начале координат. Объемная плотность $\mu=1$.

ДОПОЛНЕНИЯ К ГЛАВЕ IV

§ 1. Существование кратного интеграла

1. Понятие меры плоской и пространственной области. Прежде чем приступить к аналитическому доказательству существования кратного интеграла, необходимо сначала рассмотреть понятие *меры области* (площади плоской, объема пространственной области).

В т. I, гл. V, § 2, n° 2 мы видели, каким образом можно при некоторых общих условиях выразить площадь плоской области с помощью интеграла. Мы не будем здесь ссылаться на этот факт, не станем также и на ту точку зрения, что существование площади обеспечивается геометрической интуицией, но дадим общее определение *меры* области и заодно исследуем, при каких предположениях процесс построения этого понятия имеет смысл.

Начнем с понятия меры *плоской* области. Сначала дадим определение меры (площади) прямоугольника как произведения его основания на высоту. Если разбить прямоугольник на меньшие прямоугольники проведением некоторого числа прямых, параллельных его сторонам, то из принятого определения вытекает, что мера данного прямоугольника равна сумме мер всех составляющих его прямоугольников.

Теперь можно определить меру области, состоящей из конечного числа прямоугольников, стороны которых параллельны осям координат, как сумму мер всех этих прямоугольников.

В дальнейшем тексте этого параграфа слово прямоугольник будет всегда означать прямоугольник, стороны которого параллельны осям.

Определенная таким образом мера области, составленной из прямоугольников (совпадающая, очевидно, с обычным определением ее площади), не зависит от способа разбиения этой области на прямоугольники. Действительно, если имеются два различных разбиения, то можно построить такое новое разбиение, которое получается дроблением каждого из двух исходных разбиений. Для этого достаточно продолжить все прямые, параллельные осям, фигурирующие в каждом из двух разбиений, через всю область, рассекая тем самым каждое из этих разбиений на еще более мелкие прямоугольники. Сумма мер всех этих мелких прямоугольников равна сумме мер всех прямоугольников как первого, так и второго разбиения.

Для того чтобы дать определение меры произвольной ограниченной области B, построим нижнее и верхнее приближение к измеряемой области, т. е. построим две области, состоящие из прямоугольников: внутреннюю область B_i (i — первая буква латинского слова internus, внутренний) и охватывающую область B_e (латинское слово externus означает внешний), таким образом, что область B_i лежит полностью внутри B, а область B_e содержит внутри себя данную область B.

Это можно выполнить, например, так. Сначала строим большой квадрат, содержащий область B. Этот квадрат разбиваем прямыми, параллельными осям, на малые прямоугольники. Совокупность тех прямоугольников, которые лежат целиком внутри B, образует область B_i, заключенную внутри B, а совокупность всех прямоугольников, имеющих общие точки с областью B, составляет область B_e, которая включает внутри себя область B.

Понятию меры $\mathcal{M}(B)$ области B мы должны теперь дать такое определение, чтобы она удовлетворяла неравенству

$$\mathcal{M}(B_i) \leqslant \mathcal{M}(B) \leqslant \mathcal{M}(B_e)$$

при любом выборе областей B_i и B_e.

Если производить последовательное дробление наших разбиений таким образом, что диагонали всех прямоугольников будут стремиться к нулю, то меры областей B_i образуют монотонно возрастающую последовательность, а меры областей B_e — монотонно убывающую последовательность, так как при этой операции области B_i могут только приобретать новые прямоугольники, а области B_e — только терять прямоугольники. Следовательно, как мера B_i, так и мера B_e стремятся к определенным пределам. *Если пределы B_i и B_e равны, то этот общий предел называется мерой области B.*

При каких же обстоятельствах предел меры области B_i равен пределу меры области B_e? Конечно, в том случае, когда разность

$$\mathcal{M}(B_e) - \mathcal{M}(B_i) = \mathcal{M}(B_e - B_i)$$

стремится к нулю при неограниченном дроблении обоих разбиений, когда диагонали всех прямоугольников будут стремиться к нулю. Но область $B_e - B_i$ состоит из тех прямоугольников, которые имеют общие точки с границей области B, — она образует как бы кайму из прямоугольников, которая содержит границу области B. Стало быть, если площадь этой каймы $B_e - B_i$ стремится к нулю, то это значит, что граница области B может быть заключена в кайму из прямоугольников, площадь которой сколь угодно мала. Обратно, если граница области B может быть заключена в кайму K сколь угодно малой площади, состоящую из прямоугольников, то при достаточно мелком разбиении все прямоугольники пригpаничной области $B_e - B_i$ окажутся внутри каймы K, так что мера этой области станет меньше меры каймы K и в силу этого будет стремиться к нулю.

Резюмируем полученный результат: *предел меры B_i равен пределу меры B_e в том и только в том случае, если граница измеряемой нами области B может быть заключена в кайму сколь угодно малой меры, состоящую из прямоугольников. В этом случае наше построение действительно дает возможность приписать области B определенную меру.*

Наше определение меры имеет с геометрической точки зрения тот недостаток, что оно исходит из некоторой, тем самым выделенной системы координат. Однако не представляет существенных затруднений доказать, что мера области не зависит от выбора системы координат не только в случае двух, но и в общем случае n измерений. Мы здесь опускаем это доказательство по двум причинам. Во-первых, оно не используется для нашей конкретной цели — доказательства существования кратного интеграла. Во-вторых, независимость меры области от выбора системы координат выяснится впоследствии сама собой, когда мы получим выражение меры с помощью кратного интеграла; что значение этого интеграла не изменяется при переходе к новым прямоугольным координатам, будет тогда видно из самих формул преобразования.

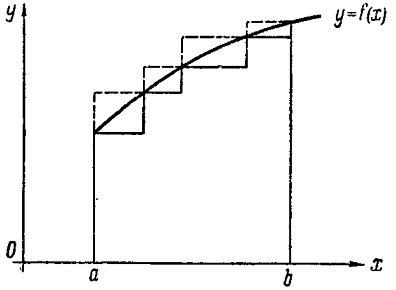


Рис. 78.

Представляется наглядно правдоподобным, что *всякая кусочно гладкая кривая*, т. е. непрерывная кривая, состоящая из конечного числа дуг с непрерывно изменяющейся касательной, *может быть заключена внутри области, имеющей вид цепочки прямоугольников, площадь которой сколь угодно мала.* В следующем номере этот факт будет доказан аналитически. Из этого будет вытекать, что для области, состоящей из конечного числа частей, имеющих кусочно гладкую границу, наше условие существования меры непременно выполняется. Стало быть, такие области имеют однозначно определенную меру, а на практике только такие области и встречаются.

В следующем номере будет также доказано, что если область B разбивается кусочно гладкими кривыми на конечное число частей, то сумма мер этих частичных областей равна мере всей области B. Здесь же мы только еще покажем, что наше определение меры плоской области находится в согласии с прежними определениями площади при помощи интегралов.

Рассмотрим сначала область B (рис. 78), ограниченную осью x, прямыми $x=a$ и $x=b$ и дугой кривой

$$y=f(x).$$

Области B_i и B_e выбираем, как показано на рис. 78: область B_i ограничена сверху сплошной ломаной, область B_e — пунктирной ломаной. Согласно определению обычного интеграла (т. I, гл. II, § 1, n°1), мера (площадь) B_i представляет собой нижнюю сумму $\underline{F}_n$, а мера (площадь) B_e — верхнюю сумму $\overline{F}_n$ для интеграла $\int\limits_a^b f(x)dx$. Поэтому

$$\mathcal{M}(B_i)\leqslant\int\limits_a^b f(x)\,dx\leqslant\mathcal{M}(B_e).$$

Сопоставляя эту формулу с неравенством

$$\mathscr{M}(B_i) \leqslant \mathscr{M}(B) \leqslant \mathscr{M}(B_e)$$

и учитывая, что $\lim \mathscr{M}(B_i) = \lim \mathscr{M}(B_e)$, приходим к заключению, что $\mathscr{M}(B) = \int\limits_a^b f(x)\,dx$, т. е. определение меры криволинейной трапеции совпадает с прежним определением ее площади.

Таким же путем нетрудно показать и для общего случая произвольной плоской области B, пользуясь разбиением области прямыми, параллельными осям, что наше определение меры эквивалентно выражению площади с помощью двойного интеграла

$$\iint\limits_B dx\,dy.$$

Наше определение меры области допускает непосредственное обобщение на *пространственные* и даже n-мерные области. Меру или объем прямоугольного параллелепипеда определяют как произведение длин трех его ребер, выходящих из одной вершины. Мера (объем) области, состоящей из конечного числа таких параллелепипедов, определяется как сумма мер всех составляющих параллелепипедов. Для произвольной области B строят области B_i, составленные из параллелепипедов, ребра которых параллельны осям, и лежащие внутри B, и такого же типа области B_e, охватывающие данную область B. Тогда определение меры области B как общего предела объемов областей B_i и B_e имеет смысл при том условии, что граница области B может быть заключена внутри окаймляющей ее оболочки, составленной из параллелепипедов и имеющей сколь угодно малый объем. В следующем номере будет показано, что это условие выполняется для всякой области, граница которой состоит из конечного числа кусков поверхностей, имеющих непрерывно изменяющиеся касательные плоскости. Впредь, как и раньше, мы будем рассматривать только такие области; слово область будет всегда означать ограниченную замкнутую область, граница которой состоит из конечного числа кусков поверхностей, выражающихся непрерывно дифференцируемыми функциями.

Мера цилиндра, образующие которого параллельны оси z, а основание лежит в плоскости xy, равна произведению площади основания на высоту. Это очевидно, если основание состоит из прямоугольников, стороны которых параллельны осям x и y. В общем случае цилиндр можно заключить между двумя цилиндрами, основания которых состоят из прямоугольников и меры (объемы) которых разнятся сколь угодно мало; стало быть, наше правило вычисления меры годится и для цилиндра с произвольным основанием. Отсюда, как и выше, нетрудно вывести, что двойной интеграл

$$\iint\limits_B f(x, y)\,dx\,dy$$

выражает меру тела, ограниченного областью B плоскости xy, поверхностью $z=f(x, y)$ и цилиндрической поверхностью, образующие которой параллельны оси z, а направляющей служит граница области B.

Далее, тем же путем убеждаемся, что и для любой пространственной области G наше определение меры приводит к знакомому тройному интегралу

$$\iiint\limits_G dx\,dy\,dz,$$

т. е. определение меры пространственной области совпадает с прежним определением объема.

2. Теоремы о кусочно гладкой дуге плоской кривой и о кусочно гладком куске поверхности. В нашем построении понятия меры (площади) плоской области мы пользовались теоремой о том, что кусочно гладкая дуга плоской кривой может быть включена в такую область-кайму, которая составлена из прямоугольников со сторонами, параллельными осям координат, и имеет сколь угодно малую площадь. Достаточно, очевидно, доказать эту теорему для любой гладкой части этой дуги, т. е. для непрерывной дуги, вдоль которой и касательная изменяется непрерывно. Такая дуга может быть представлена параметрическими уравнениями

$$x=\varphi(s), \quad y=\psi(s), \quad a \leqslant s \leqslant b,$$

где параметром служит длина дуги s, а $\varphi(s)$ и $\psi(s)$ — непрерывно дифференцируемые функции. Тогда

$$|\varphi'(s)| \leqslant 1 \quad \text{и} \quad |\psi'(s)| \leqslant 1.$$

Пусть s и s_1 — два значения длины дуги из интервала $a \leqslant s \leqslant b$, и пусть $x_1=\varphi(s_1)$, $y_1=\psi(s_1)$. На основании теоремы о среднем значении из дифференциального исчисления имеем

$$|x-x_1|=|\varphi(s)-\varphi(s_1)| \leqslant |s-s_1|,$$
$$|y-y_1|=|\psi(s)-\psi(s_1)| \leqslant |s-s_1|.$$

Разделим нашу дугу $a \leqslant s \leqslant b$ на n равных частей длиною $\varepsilon=\frac{b-a}{n}$ каждая и обозначим начальную точку k-й части через (x_k, y_k), а произвольную точку этой части — через (x, y). Тогда

$$|x-x_k| \leqslant \varepsilon \quad \text{или} \quad x_k-\varepsilon \leqslant x \leqslant x_k+\varepsilon,$$
$$|y-y_k| \leqslant \varepsilon \quad \text{или} \quad y_k-\varepsilon \leqslant y \leqslant y_k+\varepsilon.$$

Стало быть, все точки k-й части дуги лежат в квадрате с центром в точке (x_k, y_k), стороной 2ε и площадью $4\varepsilon^2$. Вся дуга содержится в области, объединяющей эти n квадратов, которые частично покрывают друг друга. Поэтому площадь всей этой области не превышает величины

$$4\varepsilon^2 n = 4\varepsilon (b - a),$$

которая, очевидно, стремится к нулю при $\varepsilon \to 0$.

Из этой теоремы вытекает важное следствие. Если плоская область G, ограниченная замкнутой кусочно гладкой кривой, разбита на две части G' и G'', отделенные друг от друга кусочно гладкими дугами, то мера площади области G равна сумме мер площадей обеих ее частей:

$$\mathcal{M}(G) = \mathcal{M}(G') + \mathcal{M}(G'').$$

Для доказательства разобьем область G сеткой прямых, параллельных осям координат, со столь малым расстоянием между этими прямыми, что кайма, состоящая из прямоугольников, имеющих общие точки с границей области G или с линиями, отделяющими друг от друга области G' и G'', имеет сколь угодно малую площадь. Как и в n° 1, обозначим через G_i область, состоящую из всех прямоугольников, лежащих целиком внутри G, а через G_e — область, объединяющую все прямоугольники, имеющие общие точки с G. Соответствующие аппроксимирующие области для G' и G'' обозначим через G'_i, G'_e, G''_i, G''_e. Области G'_i и G''_i содержатся в G_i, а друг от друга они полностью отделены; поэтому

$$\mathcal{M}(G) \geqslant \mathcal{M}(G_i) \geqslant \mathcal{M}(G'_i) + \mathcal{M}(G''_i).$$

С другой стороны, области G'_e и G''_e вместе покрывают G_e, но некоторые из прямоугольников сосчитаны при этом два раза; поэтому

$$\mathcal{M}(G'_e) + \mathcal{M}(G''_e) \geqslant \mathcal{M}(G_e) \geqslant \mathcal{M}(G).$$

Когда диагонали прямоугольников сетки стремятся к нулю, то $\mathcal{M}(G'_i)$ и $\mathcal{M}(G'_e)$ стремятся к общему пределу $\mathcal{M}(G')$; $\mathcal{M}(G''_i)$ и $\mathcal{M}(G''_e)$ стремятся к $\mathcal{M}(G'')$, а $\mathcal{M}(G_i)$ и $\mathcal{M}(G_e)$ — к $\mathcal{M}(G)$. Следовательно, из выведенных неравенств вытекают два предельных неравенства

$$\mathcal{M}(G') + \mathcal{M}(G'') \leqslant \mathcal{M}(G) \quad \text{и} \quad \mathcal{M}(G') + \mathcal{M}(G'') \geqslant \mathcal{M}(G).$$

Стало быть,

$$\mathcal{M}(G') + \mathcal{M}(G'') = \mathcal{M}(G).$$

Ясно, что эта теорема сложения справедлива и в том случае, если область G разбивается на любое конечное число n частей $G_1, G_2, \ldots, G_n$.

Теореме о кусочно гладкой дуге плоской кривой соответствует в пространстве теорема о куске поверхности, заданном параметрическими уравнениями

$$x = \varphi(u, v), \quad y = \psi(u, v), \quad z = \chi(u, v),$$
$$a \leqslant u \leqslant b, \quad \alpha \leqslant v \leqslant \beta,$$

где функции φ, ψ, χ имеют кусочно непрерывные частные производные. При этом предполагается еще, что одну из переменных x, y, z можно выразить как кусочно непрерывно дифференцируемую функцию двух других. Такой кусок поверхности всегда можно включить в окаймляющую область сколь угодно малого объема, состоящую из прямоугольных параллелепипедов, ребра которых параллельны осям координат.

Теореме сложения для плоских областей соответствует в пространстве трех и большего числа измерений своя аналогичная теорема сложения, которую читатель без труда сформулирует сам.

Доказательства обеих теорем совершенно аналогичны доказательствам соответствующих теорем для плоскости.

3. Доказательство существования двойного интеграла от непрерывной функции. Пусть функция $f(x, y)$ непрерывна в замкнутой области G. Разобьем область G на N ячеек, каждая из которых имеет кусочно гладкую границу, как это описано в гл. IV, § 2, пп° 1 и 2. Докажем, что когда наибольший диаметр ячеек G_i стремится к нулю, то нижние суммы $\sum m_i \Delta G_i$ и верхние суммы $\sum M_i \Delta G_i$ (введенные там же) стремятся к общему пределу, не зависящему от способа разбиения. Доказательство по существу такое же, как у соответствующей теоремы для функции одной переменной (т. I, Доп. к гл. II, § 1), и мы поэтому позволим себе наметить его весьма кратко.

Сначала предположим, что разбиение области G на ячейки G_i выполнено при помощи ломаных. Наибольший диаметр δ ячеек G_i мы выберем столь малым, что в любых двух точках одной ячейки значения функции будут отличаться меньше, чем на произвольно заданное число ε, так что

$$M_i - m_i < \varepsilon.$$

Тогда разность между верхней суммой и нижней будет удовлетворять неравенству

$$0 < \sum M_i \Delta G_i - \sum m_i \Delta G_i < \varepsilon \sum \Delta G_i = \varepsilon \mathcal{M}(G).$$

Всякое разбиение, которое получается из данного разбиения путем его дробления, имеет, очевидно, такую нижнюю сумму, которая лежит между верхней и нижней суммами исходного разбиения.

Для доказательства теоремы достаточно проверить, что у любых двух разбиений области G на ячейки, диаметры которых меньше δ, верхние и нижние суммы одного разбиения сколь угодно мало отличаются от одноименных сумм другого разбиения, если только число δ выбрано достаточно малым.

Итак, пусть наряду с разбиением области G на ячейки G_i имеется еще одно ее разбиение на ячейки G'_k, диаметры которых тоже меньше δ; тогда и для этого разбиения можно написать соответствующее неравенство

$$0 < \sum M'_k \Delta G'_k - \sum m'_k \Delta G'_k < \varepsilon \mathcal{M}(G).$$

Оба эти разбиения в совокупности определяют новое разбиение, которое может быть получено дроблением любого из двух исходных разбиений. Для получения этого нового разбиения надо только выделить общую часть каждой пары ячеек G_i и G'_k (если такая общая часть существует) в качестве новой ячейки G''_{ik}. Согласно сделанному выше замечанию, нижняя сумма нового разбиения не меньше, чем нижняя сумма каждого из двух исходных разбиений, и отличается от каждой из этих сумм меньше чем на $\varepsilon\mathcal{M}(G)$. Следовательно, сумма $\sum m_i \Delta G_i$ отличается от суммы $\sum m'_k \Delta G'_k$ меньше чем на $2\varepsilon\mathcal{M}(G)$. Так как ε можно выбрать сколь угодно малым, то отсюда (на основании критерия сходимости Коши) вытекает, что нижние суммы стремятся к пределу, не зависящему от способа разбиения области G. Выше мы видели, что верхние суммы отличаются сколь угодно мало от нижних сумм; стало быть, и верхние суммы стремятся к тому же пределу. Таким образом, существование двойного интеграла $\iint\limits_G f(x, y)\, dS$ от непрерывной функции $f(x, y)$ доказано, но пока только для разбиений с помощью ломаных.

Это ограничение мы наложили на способ разбиения с целью обеспечить уверенность, что получающееся новое разбиение действительно состоит из *конечного* числа ячеек G''_{ik}. Предположим, напротив, что разрешается иметь криволинейные границы ячеек, и пусть, например, часть границы одного разбиения есть отрезок прямой $x = 0$, а часть границы другого — кусок кривой $y = x^2 \sin \frac{1}{x}$; тогда результирующее разбиение будет иметь в окрестности точки $x = 0$ бесконечное множество ячеек! Однако, нетрудно теперь освободиться от стеснительного запрета пользоваться криволинейной сеткой. Действительно, согласно теореме из предыдущего номера, всякое криволинейное разбиение можно заменить таким разбиением с помощью ломаных, что итоговая разность площадей, а стало быть, и разность соответствующих нижних сумм будет сколь угодно мала. Тем самым общий случай произвольного разбиения области на ячейки приводится к уже рассмотренному частному случаю.

Доказательство не зависит, очевидно, от числа измерений.

Дополнения к теореме существования двойного интеграла (гл. IV, § 2, n°2) вытекают непосредственно из приведенных в гл. IV, § 2, n°5 формул для оценки интегралов и не нуждаются в дальнейшем обосновании.

§ 2. Обобщенные формулы Гульдина. Полярный планиметр

1. Об одном преобразовании двойного и тройного интеграла. Если область G плоскости xy покрыта семейством кривых $\varphi(x, y) = \text{const}$ таким образом, что через каждую точку области проходит одна и только одна кривая семейства, то можно принять величину

$\varphi(x, y)=\xi$ в качестве новой независимой переменной, т. е. можно принять семейство линий $\varphi(x, y)=\text{const}$ за одну из двух систем координатных линий. В качестве второй независимой переменной можно оставить величину $\eta=y$, конечно, при том условии, что точки области G определяются линиями $\varphi(x, y)=\text{const}$ и $y=\text{const}$ однозначно.

Если ввести эти новые переменные в двойной интеграл, то он преобразуется по следующей формуле:

$$\iint\limits_{G} f(x, y)\,dx\,dy=\iint\limits_{G'} \frac{f(x, y)}{\varphi_x}\,d\xi\,d\eta.$$

В правой части будем сперва интегрировать по η при постоянном ξ; тогда внутренний интеграл можно записать в следующем виде:

$$\int \frac{f(x, y)}{\sqrt{\varphi_x^2+\varphi_y^2}}\,\frac{\sqrt{\varphi_x^2+\varphi_y^2}}{\varphi_x}\,d\eta.$$

Обозначим длину дуги кривой $\varphi(x, y)=\xi=\text{const}$, отсчитываемую от какой-либо ее точки, через s; тогда $\frac{ds}{d\eta}=\frac{\sqrt{\varphi_x^2+\varphi_y^2}}{\varphi_x}$, и внутренний интеграл можно рассматривать как интеграл по s вдоль кривой $\varphi(x, y)=\xi$, а двойной интеграл представится в виде следующего повторного:

$$\iint f(x, y)\,dx\,dy=\int d\xi \int \frac{f(x, y)}{\sqrt{\varphi_x^2+\varphi_y^2}}\,ds=\int d\xi \int \frac{f(x, y)}{|\operatorname{grad}\varphi|}\,ds. \qquad (1)$$

Наглядный смысл этого представления легче всего понять, если допустить, что существует другое семейство кривых, ортогональное данному семейству $\varphi(x, y)=\text{const}$; это значит, что кривые второго семейства пересекают каждую кривую данного семейства под прямым углом, т. е. по направлению вектора $\operatorname{grad}\varphi$. Каждую из ортогональных кривых представим себе заданной в параметрическом виде, причем параметром служит длина дуги σ этой же кривой. Тогда

$$\frac{dx}{d\sigma}=\frac{\varphi_x}{\sqrt{\varphi_x^2+\varphi_y^2}}, \qquad \frac{dy}{d\sigma}=\frac{\varphi_y}{\sqrt{\varphi_x^2+\varphi_y^2}},$$

а так как

$$\frac{d\xi}{d\sigma}=\varphi_x\frac{dx}{d\sigma}+\varphi_y\frac{dy}{d\sigma},$$

то

$$\frac{d\xi}{d\sigma}=\sqrt{\varphi_x^2+\varphi_y^2}=|\operatorname{grad}\varphi|.$$

Теперь наш двойной интеграл преобразуется в новый двойной интеграл

$$\iint f(x, y)\,dx\,dy=\iint f(x, y)\,d\sigma\,ds.$$

Отсюда видно, что для вычисления двойного интеграла можно вместо разбиения области G на прямоугольники со сторонами, параллельными осям x и y, пользоваться разбиением на ячейки с помощью кривых $\varphi(x, y) = \mathrm{const}$ и их ортогональных траекторий. Эти ячейки имеют вид «криволинейных прямоугольников» со сторонами Δs_i и $\Delta\sigma_j$.

Аналогичную формулу можно вывести и для тройного интеграла. Если область G пространства покрыта семейством поверхностей $\varphi(x, y, z) = \mathrm{const}$ таким образом, что через каждую точку области проходит одна и только одна поверхность семейства, то за одну из переменных интегрирования можно принять величину $\xi = \varphi(x, y, z)$. В итоге тройной интеграл

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz = \int d\xi \iint \frac{f(x, y, z)}{\sqrt{\varphi_x^2+\varphi_y^2+\varphi_z^2}} \frac{\sqrt{\varphi_x^2+\varphi_y^2+\varphi_z^2}}{\varphi_x}\,dy\,dz$$

разлагается на два последовательных интегрирования: сначала интегрирование по поверхности $\varphi(x, y, z) = \xi = \mathrm{const}$:

$$\iint \frac{f(x, y, z)}{\sqrt{\varphi_x^2+\varphi_y^2+\varphi_z^2}}\,d\sigma$$

(см. формулу для $d\sigma$ в гл. IV, § 6, в конце n° 3), а затем интегрирование по ξ, так что

$$\iiint f(x, y, z)\,dx\,dy\,dz = \int d\xi \iint \frac{f(x, y, z)}{\sqrt{\varphi_x^2+\varphi_y^2+\varphi_z^2}}\,d\sigma = \\ = \int d\xi \iint \frac{f(x, y, z)}{|\operatorname{grad}\varphi|}\,d\sigma. \qquad (2)$$

2. Обобщенная формула Гульдина для плоскости и для пространства. Полярный планиметр. Формулы преобразования двойного и тройного интеграла, выведенные в предыдущем номере, позволяют дать простое доказательство следующих теорем:

I. *Если прямолинейный отрезок L постоянной или переменной длины $l = l(t)$, зависящей от времени t, двигается на плоскости, то площадь F, описанная движущимся отрезком, выражается формулой*

$$F = \int_{t_0}^{t_1} l(t)\,v_n(t)\,dt,$$

где t_0 и t_1 — моменты времени, соответствующие начальному и конечному положениям отрезка, а $v_n(t)$ — проекция скорости середины отрезка L на направление перпендикуляра к этому отрезку.

II. *Объем тела, описываемого движением в пространстве куска плоскости Π площади $F = F(t)$, выражается формулой*

$$V = \int_{t_0}^{t_1} F(t)\,v_n(t)\,dt,$$

где $v_n(t)$ *есть проекция скорости центра массы куска плоскости* П *на нормаль к этой плоскости.*

В обеих этих формулах и при их доказательстве мы сначала будем предполагать, что движущийся отрезок L или кусок плоскости П проходит один и только один раз через каждую точку описываемой области (рис. 79).

Сперва докажем теорему I. Уравнение прямой, на которой лежит движущийся отрезок, возьмем в следующем виде:

$$\alpha(t)x+\beta(t)y+\gamma(t)=0, \qquad \text{(a)}$$
$$\alpha^2+\beta^2=1,$$

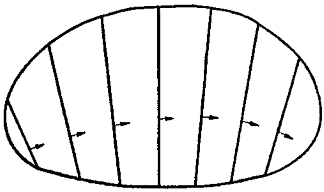


Рис. 79.

но его можно разрешить относительно t и. записать так:

$$t=\varphi(x, y).$$

Площадь фигуры, описанной отрезком L, выражается двойным интегралом $F=\iint dx\,dy$, который мы преобразуем по формуле (1) предыдущего номера, причем теперь $f(x, y)=1$ и $\xi=t$, так что

$$F=\int_{t_0}^{t_1} dt \int_L \frac{ds}{|\operatorname{grad}\varphi|},$$

где интегрирование по s производится вдоль отрезка L при $t=\text{const}$. Подставим в уравнение (a) $t=\varphi(x, y)$ и продифференцируем по x и по y; без труда найдем, что

$$\frac{1}{|\operatorname{grad}\varphi|}=|\alpha'x+\beta'y+\gamma'|,$$

причем штрихами обозначено дифференцирование по t. Стало быть, площадь выражается так:

$$F=\pm\int_{t_0}^{t_1} dt \int_L (\alpha'x+\beta'y+\gamma')\,ds.$$

Внутренний интеграл (по s при постоянном t) равен

$$l(t)(\alpha'X+\beta'Y+\gamma'),$$

где X и Y — координаты центра массы отрезка L, т. е. его середины (линейная плотность $=1$). Но X и Y удовлетворяют уравнению $\alpha X+\beta Y+\gamma=0$. Дифференцируя это уравнение по t, имеем

$$\alpha'X+\beta'Y+\gamma'+\alpha X'+\beta Y'=0.$$

Таким образом,

$$-(\alpha'X+\beta'Y+\gamma')=\alpha X'+\beta Y'=\boldsymbol{n}^\circ\boldsymbol{v}=v_n(t),$$

где $\boldsymbol{n}^\circ=\{\alpha, \beta\}$ — единичный нормальный вектор отрезка L, вектор

$v=\{X', Y'\}$ есть скорость середины отрезка в момент времени t, а $v_n(t)$ обозначает проекцию этой скорости на нормаль к отрезку. Тем самым требуемая формула доказана, но при этом мы из двух знаков выбрали один, т. е. в сущности опустили знак абсолютной величины, а это значит, что наша формула приписывает площади фигуры, описанной отрезком L, определенный знак. Этот знак зависит от того, в каком из двух направлений нормали выбран вектор n°. Площадь оказывается положительной, если середина отрезка двигается в ту сторону, куда указывает нормальный вектор n°; в противном случае площадь отрицательна.

Таким же путем доказывается и соответствующая теорема II для пространства с помощью формулы (2) предшествующего номера. Полагаем, что нет надобности излагать здесь это доказательство. Частный случай этой теоремы, когда плоская фигура II вращается вокруг оси, сохраняя неизменными свою форму и размеры, известна в механике под названием правила Гульдина для объема тела вращения. [Правило Гульдина формулируется так: объем тела, описанного вращением плоской фигуры около оси, лежащей в ее плоскости и не пересекающей самой фигуры, равен произведению площади этой фигуры на длину окружности, описанной ее центром массы. Читатель легко выведет это правило из теоремы II.]

Объем тела по теореме II тоже получается с определенным знаком, положительным или отрицательным, смотря по тому, двигается ли центр массы фигуры II в сторону, указываемую нормальным вектором n° плоскости этой фигуры или в противоположную. Ясно, что при другом выборе направления n° (из двух возможных) изменится и знак объема.

Эти замечания о знаке площади и объема позволяют распространить теоремы I и II на те случаи, когда отрезок L или плоская фигура II двигается не всегда в одном направлении, либо покрывает часть плоскости (или пространства) более одного раза. В этих случаях выведенные выше интегралы выражают алгебраическую сумму площадей (или объемов) частей описанной плоской фигуры (или описанного тела), причем каждая такая часть берется с надлежащим знаком. Предоставляем читателю разобраться самостоятельно, как это выполнить на практике.

В качестве примера рассмотрим отрезок L постоянной длины l, движущийся на плоскости таким образом, что его конечные точки постоянно находятся на двух фиксированных замкнутых кривых C и C' этой плоскости (рис. 80). По стрелкам, показывающим выбранное направление нормального вектора n° и направление движения концов отрезка L по кривым C и C', можно определить знаки, с которыми две части площади входят в окончательное значение интеграла. Нетрудно убедиться, что интеграл теоремы I дает разность площадей, ограниченных кривыми C и C'. Если кривая C' заключает площадь, равную нулю, когда она, скажем, вырождается в дугу кривой, пробегаемую несколько раз, то интеграл дает площадь, ограниченную кривой C.

Этот принцип используется в конструкции известного полярного планиметра Амслера. Это — механический прибор для измерения площади плоских фигур. Он состоит из стержня AB (играющего роль отрезка L), конец B которого связан шарниром с другим стержнем BD. Середина M стержня AB снабжена мерным колесиком, которое способно катиться по плоскости чертежа, вращаясь в плоскости, перпендикулярной к стержню AB. На конце D стержня BD укреплена острая игла, которая втыкается неподвижно в какой-либо точке O (полюсе) чертежа, а в начале A стержня AB приделан штифт, которым обводят контур C измеряемой площади. При этом точка B описывает, очевидно, дугу окружности (возможно, многократно), т. е. замкнутую линию C' площади нуль, а колесико в M отчасти скользит, отчасти катится по плоскости, причем на величину его угла поворота влияет только сумма бесконечно малых перемещений точки M по нормали к стержню AB. Согласно сказанному выше, эта сумма нормальных перемещений равна произведению площади, ограниченной кривой C, на длину l стержня AB. Но, с другой стороны, эта сумма нормальных перемещений пропорциональна углу поворота колесика (стало быть, и числу его оборотов).

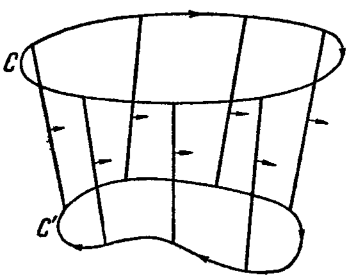


Рис. 80.

В обычной конструкции прибора колесико не помещено точно в середине стержня AB, но это изменяет лишь коэффициент пропорциональности, который для каждого экземпляра определяют раз навсегда его калибрированием.

[Подробнее о планиметре и работе с ним см. Крылов А. Н., Лекции о приближенных вычислениях, М. — Л., Гостехиздат, 1950, гл. IV.]

Упражнение

Трубчатая поверхность S (ср. стр. 197 и 199, упр. 1) является огибающей семейства единичных окружностей, центры которых лежат на замкнутой кривой C плоскости xy, а их плоскости нормальны к кривой C. Доказать, что объем, ограниченный поверхностью S, равен длине кривой C, помноженной на π.

§ 3. Объем и площадь в пространстве любого числа измерений

1. Площадь поверхности и интегрирование по поверхности в пространстве, число измерений которого больше трех. В n-мерном пространстве, т. е. во множестве, точками которого являются системы n чисел-координат, поверхность $n-1$ измерений (гиперповерхность) определяется уравнением

$$\varphi(x_1, x_2, \ldots, x_n) = \text{const}.$$

Какой-либо кусок этой гиперповерхности соответствует определенной области B переменных $x_1, x_2, \ldots, x_{n-1}$; при этом x_n подлежит вычислению из уравнения $\varphi(x_1, x_2, \ldots, x_n) = \text{const}$.

Площадью этого куска поверхности называют абсолютную величину интеграла

$$\sigma = \int\int_B \ldots \int \frac{\sqrt{\varphi_{x_1}^2 + \varphi_{x_2}^2 + \ldots + \varphi_{x_n}^2}}{\varphi_{x_n}} dx_1 dx_2 \ldots dx_{n-1}.$$

Это определение является только формальным обобщением формулы для площади поверхности в трехмерном пространстве, полученной из наглядных соображений. Оно находит, однако, известное оправдание в том факте, что величина σ не зависит от выбора выделенной координаты x_n. Это можно доказать тем же методом, что и в трехмерном пространстве (ср. гл. IV, § 6, n° 3).

Для того чтобы полностью оправдать введенное для σ название «площадь куска поверхности», надо будет доказать, что величина σ не изменяется при любом преобразовании переменных $x_1, x_2, \ldots, x_n$ к новым переменным $\xi_1, \xi_2, \ldots, \xi_n$ с помощью уравнений вида

$$x_k = f_k(\xi_1, \xi_2, \ldots, \xi_n) \qquad (k = 1, 2, \ldots, n).$$

Это значит, что надо раньше развить общую теорию преобразования n-кратных интегралов, как в частном случае $n = 2$ (ср. гл. IV, § 4), и уж затем доказать независимость величины σ от таких преобразований. В принципе это не труднее, чем в двумерном случае, и мы здесь этого излагать не станем.

Теперь мы дадим определение интеграла от функции $f(x_1, x_2, \ldots, x_n)$ по данному выше куску поверхности $n-1$ измерений следующим образом:

$$\int\int \ldots \int f(x_1, x_2, \ldots, x_n)\, d\sigma =$$
$$= \int\int_B \ldots \int f(x_1, x_2, \ldots, x_n) \frac{\sqrt{\varphi_{x_1}^2 + \ldots + \varphi_{x_n}^2}}{\varphi_{x_n}} dx_1 dx_2 \ldots dx_{n-1};$$

и в этом интеграле имеется в виду, что вместо x_n подставлено его выражение через $x_1, x_2, \ldots, x_{n-1}$ из уравнения $\varphi(x_1, \ldots, x_n) = \text{const}$. Здесь, как и выше, можно доказать, что результат не зависит от выбора выделенной переменной x_n.

Как и в случае двух или трех измерений (ср. § 2, n° 1), кратный интеграл по n-мерной области G можно представить как интеграл по поверхности Σ $n-1$ измерений, который потом интегрируется по переменной ξ, характеризующей непрерывное изменение поверхности Σ. Предположим, что область G покрыта семейством поверхностей

$$\varphi(x_1, x_2, \ldots, x_n) = \text{const}$$

таким образом, что через каждую точку $(x_1, x_2, \ldots, x_n)$ области G проходит одна и только одна поверхность семейства. Пользуемся по-прежнему независимыми переменными $x_1, x_2, \ldots, x_{n-1}$, но вместо x_n вводим новую независимую переменную

$$\xi = \varphi(x_1, x_2, \ldots, x_n);$$

тогда наш n-кратный интеграл приводится к следующему виду:

$$\int d\xi \int \ldots \int \frac{f(x_1, \ldots, x_n)}{\sqrt{\varphi_{x_1}^2 + \ldots + \varphi_{x_n}^2}} \frac{\sqrt{\varphi_{x_1}^2 + \ldots + \varphi_{x_n}^2}}{\varphi_{x_n}} dx_1 \ldots dx_{n-1} =$$
$$= \int d\xi \int \ldots \int \frac{f(x_1, \ldots, x_n)}{\sqrt{\varphi_{x_1}^2 + \ldots + \varphi_{x_n}^2}} d\sigma. \quad (*)$$

2. Площадь поверхности и объем единичного шара в n-мерном пространстве. В качестве примера вычислим площадь шаровой поверхности в n-мерном пространстве, т. е. площадь поверхности $n-1$ измерений, выражаемой уравнением

$$x_1^2 + \ldots + x_n^2 = R^2,$$

и объем ограниченного ею n-мерного тела, определенного неравенством $x_1^2 + x_2^2 + \ldots + x_n^2 \leqslant R^2$.

Допустим, что внутри шара задана функция $f(r)$, непрерывно зависящая от величины $r = \sqrt{x_1^2 + \ldots + x_n^2}$. Вычислим n-кратный интеграл этой функции $\int\int \ldots \int f(r)\, dx_1\, dx_2 \ldots dx_n$ по объему шара $x_1^2 + \ldots + x_n^2 \leqslant R^2$. Для этого введем новую переменную r равенством

$$r^2 = \varphi(x_1, \ldots, x_n) = x_1^2 + \ldots + x_n^2.$$

Так как $\sqrt{\varphi_1^2 + \ldots + \varphi_{x_n}^2} = 2r$ и $d(r^2) = 2r\, dr$, то по формуле (*) предыдущего номера получаем

$$\int\int \ldots \int f(r)\, dx_1 \ldots dx_n = \int_0^R f(r)\, dr \int \ldots \int d\sigma = \int_0^R f(r)\, \Omega_n(r)\, dr,$$

где $\Omega_n(r)$ есть площадь шаровой поверхности $x_1^2 + \ldots + x_n^2 = r^2$.

Согласно нашему общему определению, площадь половины шаровой поверхности радиуса r дается интегралом

$$\frac{1}{2}\Omega_n(r) = r \int \ldots \int \frac{dx_1 \ldots dx_{n-1}}{x_n},$$

причем интеграл берется по внутренней области $(n-1)$-мерного шара

$$x_1^2+\ldots+x_{n-1}^2\leqslant r^2.$$

Введем теперь вместо переменных x_k величины $\xi_k=\frac{x_k}{r}$; тогда $\sum\limits_{k=1}^{n}\xi_k^2=1$, и мы получаем

$$\Omega_n(r)=2r^{n-1}\int\int\ldots\int\frac{d\xi_1\ldots d\xi_{n-1}}{\xi_n}=r^{n-1}\,\omega_n,$$

где $\omega_n=2\int\int\ldots\int\frac{d\xi_1\ldots d\xi_{n-1}}{\xi_n}$ обозначает площадь единичной шаровой поверхности $\xi_1^2+\ldots+\xi_n^2=1$. Отсюда вытекает, что

$$\int\int\ldots\int f(r)\,dx_1\ldots dx_n=\omega_n\int\limits_0^R f(r)\,r^{n-1}\,dr. \qquad (A)$$

Из этой формулы можно вычислить ω_n следующим образом: в левой части мы распространим интегрирование на все пространство $x_1\ x_2\ \ldots\ x_n$ (т. е. заставим R безгранично возрастать), а в качестве $f(r)$ выберем такую функцию, для которой можно вычислить в явном виде как n-кратный интеграл слева, так и одномерный интеграл справа. Для такой роли годится функция

$$f(r)=e^{-(x_1^2+x_2^2+\ldots+x_n^2)}=e^{-r^2}.$$

При таком выборе функции $f(r)$ имеем

$$\left(\int\limits_{-\infty}^{\infty}e^{-x^2}\,dx\right)^n=\omega_n\int\limits_0^{\infty}e^{-r^2}\,r^{n-1}\,dr.$$

Мы уже знаем, что $\int\limits_{-\infty}^{\infty}e^{-x^2}\,dx=\sqrt{\pi}$ (см. гл. IV, § 5, n° 4), а интеграл справа вычисляем подстановкой $r^2=t$:

$$\int\limits_0^{\infty}e^{-r^2}\,r^{n-1}\,dr=\frac{1}{2}\int\limits_0^{\infty}e^{-t}\,t^{\frac{n}{2}-1}\,dt=\frac{1}{2}\Gamma\left(\frac{n}{2}\right).$$

[Отсюда, в частности, получается при $n=1$, что $\int\limits_0^{\infty}e^{-r^2}\,dr=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$, так что $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.] Поэтому

$$\omega_n=\frac{2(\sqrt{\pi})^n}{\Gamma\left(\frac{n}{2}\right)}.$$

(Если n четно, то

$$\Gamma\left(\frac{n}{2}\right)=\left(\frac{n-2}{2}\right)!,$$

если же n нечетно, то

$$\Gamma\left(\frac{n}{2}\right)=\frac{(n-2)(n-4)\dots 1}{2^{\frac{n-1}{2}}}\sqrt{\pi}.$$

По поводу общего определения гамма-функции и начальных сведений о ней см. т. I, стр. 291, и ниже § 6 настоящих Дополнений к гл. IV.)

Для того чтобы вычислить объем V_n n-мерного единичного шара, положим в формуле (A) $f(r)=1$, и мы получим

$$V_n=\int\int\dots\int dx_1\,dx_2\dots dx_n=\omega_n\int_0^1 r^{n-1}\,dr=\frac{\omega_n}{n},$$

откуда

$$V_n=\frac{(\sqrt{\pi})^n}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)}=\frac{(\sqrt{\pi})^n}{\Gamma\left(\frac{n+2}{2}\right)}.$$

3. Обобщения. Параметрические представления. В пространстве n измерений можно рассматривать любое r-мерное многообразие при всяком $r\leqslant n$ и дать определение его *меры*. Для этой цели удобнее всего пользоваться параметрическим представлением. Пусть наше r-мерное многообразие задано уравнениями

$$\begin{gathered} x_1=\varphi_1(u_1,\dots,u_r),\\ \dots\dots\dots\dots\\ x_n=\varphi_n(u_1,\dots,u_r),\end{gathered}$$

причем функции φ_k имеют непрерывные производные в некоторой области B переменных $u_1,\dots,u_r$. Когда переменные $u_1,\dots,u_r$ пробегают эту область, то точка $(x_1, x_2, \dots, x_n)$ описывает кусок r-мерной поверхности.

Из прямоугольной матрицы (схемы, таблицы)

$$\begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_2}{\partial u_1} & \dots & \frac{\partial x_n}{\partial u_1}\\ \frac{\partial x_1}{\partial u_2} & \frac{\partial x_2}{\partial u_2} & \dots & \frac{\partial x_n}{\partial u_2}\\ \dots & \dots & \dots & \dots\\ \frac{\partial x_1}{\partial u_r} & \frac{\partial x_2}{\partial u_r} & \dots & \frac{\partial x_n}{\partial u_r}\end{bmatrix}$$

составим теперь все возможные определители порядка r; мы получим $\binom{n}{r}$ таких определителей, которые мы обозначим через D_k, где

$k=1, 2, \ldots, \binom{n}{r}$. Первым из них будет, например,

$$D_1=\begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_2}{\partial u_1} & \cdots & \frac{\partial x_r}{\partial u_1} \\ \frac{\partial x_1}{\partial u_2} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_r}{\partial u_2} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial x_1}{\partial u_r} & \frac{\partial x_2}{\partial u_r} & \cdots & \frac{\partial x_r}{\partial u_r} \end{vmatrix}.$$

В качестве определения *меры* куска r-мерной поверхности мы примем r-кратный интеграл

$$\int \ldots \int \sqrt{D_1^2+D_2^2+\ldots+D_{\binom{n}{r}}^2}\, du_1 \ldots du_r.$$

Опираясь на теорему преобразования кратных интегралов (гл. IV, § 4, n° 2), с помощью несложных вычислений с определителями (которые мы здесь опустим) можно доказать, что принятое для меры выражение остается неизменным, если ввести вместо параметров $u_1, u_2, \ldots, u_r$ другие параметры. Нетрудно убедиться, что в случае $r=1$ мера многообразия сводится к длине дуги, а в случае $r=2$ в трехмерном пространстве получается обычная формула для площади куска поверхности.

Мы дадим доказательство для случая $r=n-1$ при произвольном n, т. е. докажем следующую теорему:

Если кусок поверхности $n-1$ измерений $\varphi(x_1, \ldots, x_n)=0$ в n-мерном пространстве может быть представлен в параметрическом виде уравнениями

$$x_i=\psi_i(u_1, \ldots, u_{n-1}) \qquad (i=1, \ldots, n),$$

то его площадь (n° 1) *дается также интегралом*

$$\sigma=\int\ldots\int\sqrt{D_1^2+\ldots+D_n^2}\, du_1 \ldots du_{n-1},$$

где D_i есть якобиан $(n-1)$-го порядка:

$$D_i=\frac{\partial(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)}{\partial(u_1, \ldots, u_{n-1})}=\frac{1}{\dfrac{\partial(u_1, \ldots, u_{n-1})}{\partial(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)}}.$$

Здесь, как обычно, предполагается существование и непрерывность всех встречающихся производных.

Без потери общности можно допустить, что $\varphi_{x_n}\neq 0$. Так как, согласно определению, данному на стр. 323,

$$\sigma=\int \ldots \int \frac{|\operatorname{grad}\varphi|}{\varphi_{x_n}}\, dx_1 \ldots dx_{n-1},$$

то достаточно показать, что

$$\frac{1}{\varphi_{x_n}}|\operatorname{grad}\varphi|\,dx_1\ldots dx_{n-1}=\sqrt{\sum_i D_i^2}\,du_1\ldots du_{n-1},$$

или

$$|\operatorname{grad}\varphi|^2=\varphi_{x_n}^2\left(\sum_i D_i^2\right)\left[\frac{\partial(u_1,\ldots,u_{n-1})}{\partial(x_1,\ldots,x_{n-1})}\right]^2=\frac{\varphi_{x_n}^2}{D_n^2}\sum_i D_i^2.$$

Но из свойств якобиана вытекает, что

$$\frac{D_\nu}{D_n}=\frac{\dfrac{\partial(x_1,\ldots,x_{\nu-1},x_{\nu+1},\ldots,x_n)}{\partial(u_1,\ldots,\partial u_{n-1})}}{\dfrac{\partial(x_1,\ldots,x_{n-1})}{\partial(u_1,\ldots,u_{n-1})}}=\frac{\partial(x_1,\ldots,x_{\nu-1},x_{\nu+1},\ldots,x_n)}{\partial(x_1,\ldots,x_{n-1})}.$$

Последний якобиан соответствует введению вместо $(x_1,\ldots,x_{n-1})$ новых независимых переменных $(x_1,\ldots,x_{\nu-1},x_{\nu+1},\ldots,x_n)$, а так как частные производные $\frac{\partial x_n}{\partial x_i}$ получаются из уравнений

$$\varphi_{x_n}\frac{\partial x_n}{\partial x_i}+\varphi_{x_i}=0\qquad(i=1,\ldots,n-1),$$

то $\frac{D_\nu}{D_n}=\pm\frac{\varphi_{x_\nu}}{\varphi_{x_n}}$. Поэтому

$$\frac{D_\nu^2}{D_n^2}=\frac{\varphi_{x_\nu}^2}{\varphi_{x_n}^2},$$

что и доказывает формулу для σ.

Отметим, что выражение $\sum_{i=1}^{n} D_i^2$ можно представить в виде определителя $(n-1)$-го порядка

$$\sum_{i=1}^{n} D_i^2=|x_{u_i}x_{u_k}|=\begin{vmatrix} x_{u_1}^2 & x_{u_1}x_{u_2} & \ldots & x_{u_1}x_{u_{n-1}}\\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\\ x_{u_{n-1}}x_{u_1} & x_{u_{n-1}}x_{u_2} & \ldots & x_{u_{n-1}}^2\end{vmatrix}=G,$$

элементы которого равны скалярным произведениям векторов

$$x_{u_i}=\left\{\frac{\partial x_1}{\partial u_i},\ldots,\frac{\partial x_n}{\partial u_i}\right\}\quad\text{и}\quad x_{u_k}=\left\{\frac{\partial x_1}{\partial u_k},\ldots,\frac{\partial x_n}{\partial u_k}\right\},$$

т. е. выражения

$$x_{u_i}x_{u_k}=\sum_{\nu=1}^{n}\frac{\partial\psi_\nu}{\partial u_i}\frac{\partial\psi_\nu}{\partial u_k}.$$

Таким образом,

$$\sigma=\int\ldots\int\sqrt{G}\,du_1\,du_2\ldots du_{n-1}.$$

Упражнения

1. Вычислить объем n-мерного эллипсоида

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\ldots+\frac{x_n^2}{a_n^2}=1.$$

2. Привести интеграл I от функции $f(x_1)$, зависящей только от x_1, по поверхности единичной сферы $x_1^2+x_2^2+\ldots+x_n^2=1$ в n-мерном пространстве к виду обычного определенного интеграла.

§ 4. Несобственные интегралы как функции параметра

1. Равномерная сходимость. Непрерывная зависимость интеграла от параметра. Несобственные интегралы тоже часто являются функциями параметра; так, например, интеграл от степенной функции

$$\int_0^1 y^x\,dy=\frac{1}{x+1}$$

оказывается сходящимся несобственным интегралом при $-1<x<0$ и является функцией параметра x в этом интервале.

Мы уже знаем, что интеграл по конечному промежутку интегрирования является непрерывной функцией параметра, если подынтегральная функция (как функция двух переменных) непрерывна. Однако в случае бесконечного промежутка интегрирования дело обстоит далеко не так просто. Рассмотрим, например, интеграл

$$F(x)=\int_0^\infty \frac{\sin xy}{y}\,dy.$$

Пользуясь преобразованием $xy=z$, $x\,dy=dz$, получаем

$$F(x)=\int_0^\infty \frac{\sin z}{z}\,dz \quad \text{при} \quad x>0$$

и

$$F(x)=\int_0^{-\infty} \frac{\sin z}{z}\,dz=-\int_0^\infty \frac{\sin z}{z}\,dz \quad \text{при} \quad x<0.$$

Интеграл $\int_0^\infty \frac{\sin z}{z}\,dz$ сходится, что было показано в т. I (стр. 292 и 485). Более того, мы знаем, что этот интеграл равен $\frac{\pi}{2}$ (т. I, стр. 527; см. также ниже, n° 4, пример 3). Стало быть, несмотря на то, что функция $\frac{\sin xy}{y}$ (рассматриваемая как функция от x и y)

непрерывна всюду, а ее интеграл сходится при всяком значении x, функция $F(x)$ имеет разрыв при $x=0$, ибо

$$F(x)=\begin{cases} -\frac{\pi}{2} & \text{при} \quad x<0, \\ 0 & \text{при} \quad x=0, \\ \frac{\pi}{2} & \text{при} \quad x>0. \end{cases}$$

Сам по себе этот факт не содержит ничего удивительного; с аналогичной ситуацией мы уже встречались при изучении бесконечных рядов (т. I, гл. VIII, стр. 447 и 449), а ведь процесс интегрирования можно рассматривать как обобщенное суммирование! Для того чтобы сходящийся ряд непрерывных функций непременно имел непрерывную сумму, пришлось в свое время потребовать, чтобы сходимость ряда была *равномерной*. Теперь, при изучении сходящихся несобственных интегралов, зависящих от параметра, тоже понадобится ввести понятие равномерной сходимости.

Определение. *Сходящийся интеграл*

$$F(x)=\int_B^{\infty} f(x, y)\,dy \tag{1}$$

называется равномерно сходящимся (относительно x) в интервале $\alpha \leqslant x \leqslant \beta$, если «остаток» интеграла может быть сделан сколь угодно малым одновременно для всех значений x из рассматриваемого интервала.

Точнее, интеграл (1) называется равномерно сходящимся на $[\alpha, \beta]$, если для всякого наперед заданного положительного числа ε существует такое положительное число $A=A(\varepsilon)$, *не* зависящее от x, что для любого $B \geqslant A$ будет

$$\left|\int_B^{\infty} f(x, y)\,dy\right| < \varepsilon.$$

Приведем следующий полезный для практики признак равномерной сходимости интеграла (1). *Если при всяком $y \geqslant y_0$ во всем интервале $\alpha \leqslant x \leqslant \beta$ выполняется неравенство*

$$|f(x, y)| < \frac{M}{y^a},$$

где M — положительная постоянная и $a>1$, то интеграл (1) *сходится равномерно (и абсолютно) при $\alpha \leqslant x \leqslant \beta$.*

В самом деле, при этом условии

$$\left|\int_B^{\infty} f(x, y)\,dy\right| < M\int_B^{\infty} \frac{dy}{y^a} = \frac{M}{(a-1)B^{a-1}} \leqslant \frac{M}{(a-1)A^{a-1}},$$

но правую часть, не зависящую от x, можно сделать сколь угодно малой, выбрав A достаточно большим. Этот признак представляет собой естественную аналогию соответствующему критерию для функциональных рядов (т. I, стр. 453).

Теперь нетрудно убедиться, что *равномерно сходящийся интеграл* (1) *от непрерывной функции параметра является в свою очередь непрерывной функцией параметра.* В самом деле, если выбрать число A так, чтобы для всех значений x из рассматриваемого интервала выполнялось неравенство

$$\left|\int_A^\infty f(x, y)\,dy\right| < \varepsilon,$$

то отсюда будет вытекать, что

$$|F(x+h) - F(x)| < \left|\int_0^A \{f(x+h, y) - f(x, y)\}\,dy\right| + 2\varepsilon.$$

В силу непрерывности функции $f(x, y)$, можно выбрать приращение h столь малым, что абсолютная величина интеграла в правой части будет меньше чем ε, а это и доказывает непрерывность интеграла (1).

Аналогично обстоит дело у интеграла с конечным промежутком интегрирования, когда подынтегральная функция имеет в этом промежутке точку бесконечного разрыва. Пусть, например, $f(x, y) \to \infty$ при $y \to a + 0$. Тогда *сходящийся несобственный интеграл*

$$F(x) = \int_a^b f(x, y)\,dy \qquad (b > a) \tag{2}$$

называется равномерно сходящимся в интервале $\alpha \leqslant x \leqslant \beta$, *если для любого положительного числа* ε *можно найти такое число* $k > 0$, *не зависящее от* x, *что*

$$\left|\int_a^{a+h} f(x, y)\,dy\right| < \varepsilon, \qquad \textit{коль скоро} \quad 0 < h \leqslant k.$$

И здесь существует аналогичный признак равномерной сходимости. *Если вблизи точки* $y = a$ *выполняется для всего интервала* $\alpha \leqslant x \leqslant \beta$ *неравенство*

$$|f(x, y)| < \frac{M}{(y-a)^\nu},$$

где M — положительная постоянная, а $\nu < 1$, *то несобственный интеграл* (2) *сходится равномерно при* $\alpha \leqslant x \leqslant \beta$.

Теперь можно доказать (тем же путем, что и выше), что *в случае равномерной сходимости несобственного интеграла* (2) *он является непрерывной функцией параметра* x.

Если несобственные интегралы (1) и (2) сходятся равномерно, скажем, в промежутке $\alpha \leqslant x \leqslant \beta$, то они представляют собой функции параметра x, непрерывные в этом промежутке. Поэтому их можно интегрировать по этому промежутку и тем самым составить соответствующие несобственные повторные интегралы

$$\int_{\alpha}^{\beta} dx \int_{0}^{\infty} f(x, y)\, dy \quad \text{и} \quad \int_{\alpha}^{\beta} dx \int_{a}^{b} f(x, y)\, dy.$$

Наряду с этими интегралами с конечным промежутком интегрирования по x ($\alpha \leqslant x \leqslant \beta$) заслуживают, конечно, изучения и несобственные повторные интегралы с бесконечным промежутком интегрирования по x.

2. Интегрирование несобственных интегралов по параметру. При интегрировании и дифференцировании несобственного интеграла по параметру далеко не всегда дозволено производить эти операции под знаком интеграла, т. е. не всегда допустимо менять порядок этих операций с первоначальной операцией интегрирования [ср. пример 4 в n° 4].

а) Для того чтобы решить вопрос о правомерности изменения порядка интегрирования в каком-либо заданном несобственном повторном интеграле, можно воспользоваться излагаемым ниже критерием, либо произвести специальное исследование по образцу доказательства этого критерия.

Если несобственный интеграл

$$F(x) = \int_{0}^{\infty} f(x, y)\, dy$$

сходится равномерно относительно x в интервале $\alpha \leqslant x \leqslant \beta$, то

$$\int_{\alpha}^{\beta} dx \int_{0}^{\infty} f(x, y)\, dy = \int_{0}^{\infty} dy \int_{\alpha}^{\beta} f(x, y)\, dx.$$

Для доказательства положим

$$\int_{0}^{\infty} f(x, y)\, dy = \int_{0}^{A} f(x, y)\, dy + R_A(x),$$

где $|R_A(x)| < \varepsilon(A)$, причем, по условию, $\varepsilon(A)$ есть число, зависящее только от A, но не от x и стремящееся к нулю при $A \to \infty$. В силу элементарной теоремы об изменении порядка интегрирования в собственном повторном интеграле имеем

$$\int_{\alpha}^{\beta} dx \int_{0}^{\infty} f(x, y)\, dy = \int_{\alpha}^{\beta} dx \int_{0}^{A} f(x, y)\, dy + \int_{\alpha}^{\beta} R_A(x)\, dx =$$

$$= \int_{0}^{A} dy \int_{\alpha}^{\beta} f(x, y)\, dx + \int_{\alpha}^{\beta} R_A(x)\, dx;$$

стало быть, по теореме о среднем значении из интегрального исчисления,

$$\left|\int_\alpha^\beta dx \int_0^\infty f(x, y)\,dy - \int_0^A dy \int_\alpha^\beta f(x, y)\,dx\right| \leqslant \varepsilon(A)\,|\beta-\alpha|.$$

При $A \to \infty$ отсюда вытекает формула, подлежавшая доказательству.

Если интегрирование по параметру тоже производится по бесконечному промежутку, то изменение порядка не всегда допустимо даже в том случае, если сходимость равномерна. Однако если соответствующий *несобственный интеграл по двумерной области* сходится (гл. IV, § 5, n° 5), то перемена порядка интегрирования допустима. Так, например, если двойной интеграл $\iint f(x, y)\,dx\,dy$, распространенный на всю первую четверть, сходится, то

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

Доказательство этой теоремы вытекает из того факта, что, в силу своего определения, сходящийся несобственный двойной интеграл не зависит от способа исчерпывания области интегрирования. С одной стороны, это исчерпывание можно произвести с помощью бесконечных полос, параллельных оси x, а с другой стороны — с помощью полос, параллельных оси y.

б) Аналогичный результат получается и в том случае, если промежуток интегрирования исходного интеграла конечен, но подынтегральная функция имеет в области интегрирования конечное число линий разрыва (например, прямых $y = \text{const}$). Соответствующая теорема формулируется, например, так:

Пусть в полосе $\alpha \leqslant x \leqslant \beta$ функция $f(x, y)$ имеет разрывы только на конечном числе прямых $y = a_1$, $y = a_2$, ..., $y = a_r$, и пусть интеграл $\int_a^b f(x,y)\,dy$ сходится равномерно относительно x; тогда этот интеграл представляет в промежутке $\alpha \leqslant x \leqslant \beta$ непрерывную функцию от x, причем

$$\int_\alpha^\beta dx \int_a^b f(x, y)\,dy = \int_a^b dy \int_\alpha^\beta f(x, y)\,dx,$$

т. е. при указанных условиях порядок интегрирования безразличен. Доказывается эта теорема тем же способом, что и предыдущая.

3. Дифференцирование несобственных интегралов по параметру. Столь же легко распространить на несобственные интегралы и правило дифференцирования по параметру. Это правило выражается в виде следующей теоремы.

Если в интервале $\alpha \leqslant x \leqslant \beta$ функция $f(x, y)$ имеет кусочно непрерывную производную по x и оба интеграла

$$F(x)=\int_0^\infty f(x, y)\,dy \quad \text{и} \quad \int_0^\infty f_x(x, y)\,dy$$

сходятся равномерно, то производная

$$F'(x)=\int_0^\infty f_x(x, y)\,dy,$$

т. е. при указанных выше предположениях можно переставить порядок действий интегрирования и дифференцирования по параметру. Для доказательства положим

$$G(x)=\int_0^\infty f_x(x, y)\,dy.$$

Тогда, по доказанной выше теореме о перестановке двух действий интегрирования, имеем

$$\int_\alpha^\xi G(x)\,dx=\int_\alpha^\xi dx\int_0^\infty f_x(x, y)\,dy=\int_0^\infty dy\int_\alpha^\xi f_x(x, y)\,dx.$$

Так как внутренний интеграл в правой части

$$\int_\alpha^\xi f_x(x, y)\,dx=f(\xi, y)-f(\alpha, y),$$

то

$$\int_\alpha^\xi G(x)\,dx=\int_0^\infty dy\int_\alpha^\xi f_x(x, y)\,dx=$$
$$=\int_0^\infty f(\xi, y)\,dy-\int_0^\infty f(\alpha, y)\,dy=F(\xi)-F(\alpha).$$

Дифференцируя правую и левую части по ξ и заменив затем ξ через x, получим требуемый результат:

$$F'(x)=G(x)=\int_0^\infty f_x(x, y)\,dy.$$

Правило дифференцирования по параметру можно распространить и на такой интеграл, у которого один из пределов интегрирования тоже зависит от x. Для этого достаточно разбить промежуток интегрирования на две части каким-либо постоянным числом a, лежащим внутри его:

$$\int_{\varphi(x)}^\infty f(x, y)\,dy=\int_{\varphi(x)}^a f(x, y)\,dy+\int_a^\infty f(x, y)\,dy,$$

а затем применить доказанные ранее правила к каждому интегралу в отдельности.

Правило дифференцирования по параметру справедливо и для несобственного интеграла по конечному промежутку, но с разрывной подынтегральной функцией.

4. Примеры. 1) Рассмотрим при $x > 0$ интеграл

$$\int_0^\infty e^{-xy}\,dy = \frac{1}{x}.$$

При $x \geqslant 1$ этот интеграл сходится равномерно, так как для любого положительного числа A

$$\int_A^\infty e^{-xy}\,dy \leqslant \int_A^\infty e^{-y}\,dy = e^{-A},$$

причем правая часть уже не зависит от x и может быть сделана сколь угодно малой путем выбора достаточно большого значения A. Это утверждение справедливо и для интегралов от частных производных подынтегральной функции по x. Поэтому путем последовательного дифференцирования по x находим

$$\int_0^\infty ye^{-xy}\,dy = \frac{1}{x^2}, \quad \int_0^\infty y^2e^{-xy}\,dy = \frac{2}{x^3}, \;\ldots,\; \int_0^\infty y^ne^{-xy}\,dy = \frac{n!}{x^{n+1}}.$$

Подставив в последнее равенство $x = 1$, получим

$$\Gamma(n+1) = \int_0^\infty y^n e^{-y}\,dy = n!.$$

Эту формулу мы уже вывели другим путем в первом томе, гл. IV, стр. 292.

2) Рассмотрим интеграл $\displaystyle\int_0^\infty \frac{dy}{x^2+y^2} = \frac{\pi}{2}\cdot\frac{1}{x}$.

И здесь легко убедиться, что при $x \geqslant a$, где a — любое положительное число, все условия, обеспечивающие право дифференцирования под знаком интеграла, выполнены. Поэтому повторное дифференцирование дает последовательность формул:

$$\int_0^\infty \frac{dy}{(x^2+y^2)^2} = \frac{\pi}{2}\cdot\frac{1}{2}\cdot\frac{1}{x^3},$$

$$\int_0^\infty \frac{dy}{(x^2+y^2)^3} = \frac{\pi}{2}\cdot\frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{x^5},$$

. . . .

$$\int_0^\infty \frac{dy}{(x^2+y^2)^n} = \frac{\pi}{2}\cdot\frac{1\cdot 3\ldots(2n-3)}{2\cdot 4\ldots(2n-2)}\cdot\frac{1}{x^{2n-1}}.$$

Кстати, из этих формул можно извлечь другое доказательство формулы Валлиса, дающей разложение числа π в бесконечное произведение (ср. т. I, гл. IV, § 4, n° 7). Для этого, подставляя $x=\sqrt{n}$, получаем

$$\int_0^\infty \frac{dy}{\left(1+\frac{y^2}{n}\right)^n}=\frac{\pi}{2}\cdot\frac{1\cdot 3\ldots(2n-3)}{2\cdot 4\ldots(2n-2)}\sqrt{n}.$$

Интеграл в левой части этого равенства при $n\to\infty$ имеет своим пределом интеграл $\int\limits_0^\infty e^{-y^2}\,dy=\frac{1}{2}\sqrt{\pi}$. Для доказательства заметим, что абсолютная величина разности обоих интегралов удовлетворяет неравенству

$$\left|\int_0^\infty e^{-y^2}\,dy-\int_0^\infty \frac{dy}{\left(1+\frac{y^2}{n}\right)^n}\right|\leqslant$$

$$\leqslant\int_0^T\left|e^{-y^2}-\frac{1}{\left(1+\frac{y^2}{n}\right)^n}\right|dy+\int_T^\infty e^{-y^2}\,dy+\int_T^\infty\frac{dy}{\left(1+\frac{y^2}{n}\right)^n},$$

а так как $\left(1+\frac{y^2}{n}\right)^n>y^2$, то

$$\left|\int_0^\infty e^{-y^2}\,dy-\int_0^\infty \frac{dy}{\left(1+\frac{y^2}{n}\right)^n}\right|\leqslant\int_0^T\left|e^{-y^2}-\frac{1}{\left(1+\frac{y^2}{n}\right)^n}\right|dy+\int_T^\infty e^{-y^2}\,dy+\frac{1}{T}.$$

Выберем теперь T столь большим, что $\int\limits_T^\infty e^{-y^2}\,dy+\frac{1}{T}<\frac{\varepsilon}{2}$, а затем при выбранном T возьмем настолько большое n, что также и

$$\int_0^T\left|e^{-y^2}-\frac{1}{\left(1+\frac{y^2}{n}\right)^n}\right|dy<\frac{\varepsilon}{2};$$

это выполнимо, так как $\lim\limits_{n\to\infty}\left(1+\frac{y^2}{n}\right)^{-n}=e^{-y^2}$. Тогда

$$\left|\int_0^\infty\left(e^{-y^2}-\frac{1}{\left(1+\frac{y^2}{n}\right)^n}\right)dy\right|<\varepsilon,$$

а отсюда вытекает соотношение

$$\lim_{n\to\infty}\frac{1\cdot 3\ldots(2n-3)}{2\cdot 4\ldots(2n-2)}\sqrt{n}=\frac{1}{\sqrt{\pi}},$$

которое эквивалентно формуле Валлиса, выведенной на стр. 265 первого тома.

3) Рассмотрим несобственный интеграл

$$F(x)=\int_0^\infty e^{-xy}\frac{\sin y}{y}\,dy,$$

зависящий от параметра x. Мы имеем в виду с его помощью вычислить $\int_0^\infty \frac{\sin y}{y}\,dy$. Интеграл $F(x)$ сходится равномерно при $x\geqslant 0$, интеграл же

$$\int_0^\infty e^{-xy}\sin y\,dy,$$

полученный дифференцированием по x подынтегральной функции интеграла $F(x)$, сходится равномерно при $x\geqslant\delta>0$, где δ — сколь угодно малое положительное число. Оба эти утверждения будут доказаны ниже. Поэтому функция $F(x)$ непрерывна при $x\geqslant 0$, а при $x\geqslant\delta$ ее производная

$$F'(x)=-\int_0^\infty e^{-xy}\sin y\,dy.$$

Этот интеграл легко вычислить с помощью формулы (т. I, стр. 253):

$$\int e^{ax}\sin bx\,dx=\frac{e^{ax}}{a^2+b^2}(a\sin bx-b\cos bx)+C.$$

В итоге получим

$$F'(x)=-\frac{1}{1+x^2},$$

откуда

$$F(x)=\operatorname{arcctg} x+C,$$

где C — некоторая постоянная. Для определения C заметим, что

$$|F(x)|=\left|\int_0^\infty e^{-xy}\frac{\sin y}{y}\,dy\right|\leqslant\int_0^\infty e^{-xy}\,dy=\frac{e^{-xy}}{x}\Bigg|_\infty^0=\frac{1}{x},$$

так что $\lim\limits_{x\to\infty}F(x)=0$; с другой стороны, и $\lim\limits_{x\to\infty}\operatorname{arcctg} x=0$. Следовательно, $C=0$ и

$$F(x)=\operatorname{arcctg} x.$$

Выполним в этом равенстве предельный переход $x\to+0$. Так как

$$\lim_{x\to+0}\operatorname{arcctg} x=\frac{\pi}{2}\quad\text{и}\quad\lim_{x\to+0}F(x)=\int_0^\infty\frac{\sin y}{y}\,dy,$$

то

$$\int_0^\infty \frac{\sin y}{y}\,dy = \frac{\pi}{2}.$$

Это и есть искомая формула (ср. т. I, стр. 527).

Остается теперь доказать те два утверждения, на которые мы опирались.

а) Докажем, что $\int_0^\infty e^{-xy}\frac{\sin y}{y}\,dy$ сходится равномерно при $x \geqslant 0$. Пусть A — произвольное положительное число, и пусть $k\pi$ — наименьшее целое кратное от π, превосходящее число A. Тогда «остаток» нашего интеграла можно записать в следующем виде:

$$\int_A^\infty e^{-xy}\frac{\sin y}{y}\,dy = \int_A^{k\pi} e^{-xy}\frac{\sin y}{y}\,dy + \sum_{\nu=k}^{\infty}\int_{\nu\pi}^{(\nu+1)\pi} e^{-xy}\frac{\sin y}{y}\,dy.$$

Ряд в правой части — знакочередующийся, абсолютные величины его членов монотонно убывают и стремятся к нулю; следовательно, этот ряд сходится по признаку Лейбница (т. I, стр. 431), и его сумма по абсолютной величине меньше его первого члена. Поэтому из последней формулы вытекает неравенство

$$\left|\int_A^\infty e^{-xy}\frac{\sin y}{y}\,dy\right| < \int_A^{(k+1)\pi} e^{-xy}\frac{|\sin y|}{y}\,dy < \int_A^{(k+1)\pi}\frac{1}{A}\,dy < \frac{2\pi}{A},$$

правая часть которого не зависит от x и может быть сделана сколь угодно малой. Это и доказывает, что интеграл сходится равномерно.

б) Остается доказать, что $\int_0^\infty e^{-xy}\sin y\,dy$ сходится равномерно при $x \geqslant \delta > 0$. Но это сразу вытекает из следующего неравенства:

$$\left|\int_A^\infty e^{-xy}\sin y\,dy\right| \leqslant \int_A^\infty |e^{-xy}\sin y|\,dy \leqslant \int_A^\infty e^{-xy}\,dy = \frac{e^{-Ax}}{x} \leqslant \frac{e^{-A\delta}}{\delta}.$$

4) В n° 2 мы выяснили, что для интеграла $\int_\alpha^\beta dx\int_0^\infty f(x, y)\,dy$ достаточным условием допустимости изменения порядка интегрирования является равномерная сходимость интеграла по бесконечному промежутку. Из нижеследующего примера видно, что простой сходимости может оказаться недостаточно.

Положим $f(x, y) = (2 - xy)\,xye^{-xy}$. Тогда

$$f(x, y) = \frac{\partial}{\partial y}(xy^2e^{-xy}),$$

и интеграл $\int_0^\infty f(x, y)\,dy$ сходится при всяком x из интервала $0 \leqslant x \leqslant 1$, причем для всех этих значений x

$$\int_0^\infty f(x, y)\,dy = \left[xy^2e^{-xy}\right]_0^\infty = 0,$$

Поэтому

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

С другой стороны, $f(x, y) = \frac{\partial}{\partial x}(x^2 y e^{-xy})$, так что

$$\int_0^1 f(x, y)\, dx = ye^{-y} \quad \text{при} \quad y \geqslant 0$$

и

$$\int_0^\infty dy \int_0^1 f(x, y)\, dx = \int_0^\infty ye^{-y}\, dy = 1.$$

Стало быть,

$$\int_0^1 dx \int_0^\infty f(x, y)\, dy \neq \int_0^\infty dy \int_0^1 f(x, y)\, dx.$$

5. Вычисление интегралов Френеля. Встречающиеся в теории света интегралы

$$F_1 = \int_{-\infty}^{+\infty} \sin(\tau^2)\, d\tau, \qquad F_2 = \int_{-\infty}^{+\infty} \cos(\tau^2)\, d\tau$$

называются интегралами Френеля. Их сходимость была доказана в конце гл. IV первого тома. Для вычисления этих интегралов введем вместо τ новую переменную $t = \tau^2$. Тогда

$$F_1 = \int_0^\infty \frac{\sin t}{\sqrt{t}}\, dt, \qquad F_2 = \int_0^\infty \frac{\cos t}{\sqrt{t}}\, dt.$$

Пользуясь тем, что $\frac{1}{\sqrt{t}} = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-x^2 t}\, dx$ (это можно проверить, вычисляя интеграл в правой части с помощью замены переменной $x = \frac{y}{\sqrt{t}}$, $dx = \frac{dy}{\sqrt{t}}$), приводим наши интегралы к следующему виду:

$$F_1 = \frac{2}{\sqrt{\pi}} \int_0^\infty dt \int_0^\infty e^{-x^2 t} \sin t\, dx, \qquad F_2 = \frac{2}{\sqrt{\pi}} \int_0^\infty dt \int_0^\infty e^{-x^2 t} \cos t\, dx.$$

Нетрудно убедиться, что внутренние интегралы сходятся равномерно. Стало быть, можно поменять порядок интегрирования:

$$F_1 = \frac{2}{\sqrt{\pi}} \int_0^\infty dx \int_0^\infty e^{-x^2 t} \sin t\, dt, \qquad F_2 = \frac{2}{\sqrt{\pi}} \int_0^\infty dx \int_0^\infty e^{-x^2 t} \cos t\, dt.$$

Внутренние интегралы можно теперь вычислить по формулам т. I, гл. IV, § 4, n° 4 (своеобразный случай интегрирования произведения) и интегралы Френеля приводятся к сравнительно простым интегралам от рациональных функций:

$$F_1=\frac{2}{\sqrt{\pi}}\int\limits_0^\infty \frac{1}{1+x^4}\,dx, \qquad F_2=\frac{2}{\sqrt{\pi}}\int\limits_0^\infty \frac{x^2}{1+x^4}\,dx.$$

С помощью подстановки $x=\frac{1}{\xi}$ убеждаемся, что второй интеграл равен первому, а первообразная функция для F_1 вычислена в первом томе (стр. 274). В результате получается

$$F_1=F_2=\sqrt{\frac{\pi}{2}}.$$

Упражнения

1. Вычислить $\int\limits_0^\infty x^n e^{-x^2}\,dx.$

2. При каких значениях a, b, c

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-(ax^2+2bxy+cy^2)}\,dx\,dy=1?$$

3. Вычислить несобственные интегралы

а) $$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} e^{-(ax^2+2bxy+cy^2)}(Ax^2+2Bxy+Cy^2)\,dx\,dy;$$

б) $$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} e^{-(ax^2+2bxy+cy^2)}(ax^2+2bxy+cy^2)\,dx\,dy$$

$$(a>0,\ ac-b^2>0).$$

4. Вычислить следующие интегралы:

а) $$K(a)=\int\limits_0^\infty e^{-ax^2}\cos x\,dx;$$

б) $$\int\limits_0^\infty \frac{e^{-bx}-e^{-ax}}{x}\cos x\,dx \qquad (a>0,\ b>0);$$

в) $$I(a)=\int\limits_0^\infty e^{-x^2-a^2/x^2}\,dx;$$

г) $$\int\limits_0^\infty \frac{\sin(ax)\,J_0(bx)}{x}\,dx,$$

где J_0 обозначает функцию Бесселя нулевого индекса, определенную в упр. 4, стр. 245.

5*. Доказать, что $\int\limits_0^{n\pi} \frac{\sin^2 ax}{x}\,dx$ имеет тот же порядок роста, что и $\ln n$ при $n \to \infty$ и что

$$\int\limits_0^{\infty} \frac{\sin^2 ax - \sin^2 bx}{x}\,dx = \frac{1}{2}\ln\frac{a}{b}.$$

6. Заменить утверждение «интеграл $\int\limits_0^{\infty} f(x, y)\,dy$ не является равномерно сходящимся» эквивалентным утверждением, не содержащим ни в какой форме слов «равномерно сходящийся». (Ср. т. I, стр. 67, упр. 1.)

§ 5. Интеграл Фурье

1. Введение. Теория, изложенная в § 4, n° 2, иллюстрируется важной теоремой, известной под названием *интегральной теоремы Фурье* или, короче, *интеграла Фурье*. Вспомним, что ряд Фурье дает представление кусочно гладкой, но в остальном произвольной, периодической функции с помощью тригонометрических функций. Интеграл Фурье дает соответствующее тригонометрическое представление функции $f(x)$, определенной на всей оси $-\infty < x < +\infty$ и не подчиненной никакому условию периодичности.

Относительно функции $f(x)$ мы сделаем следующие допущения:

1) $f(x)$ есть кусочно гладкая функция; это значит, что $f(x)$ и ее первая производная в любом конечном интервале непрерывны или имеют конечное число разрывов первого рода, т. е. конечных скачков.

2) Интеграл $\int\limits_{-\infty}^{\infty} |f(x)|\,dx = C$ является сходящимся.

3) Во всякой точке разрыва x нашей функции значение $f(x)$ принимается равным среднему арифметическому предельных значений функции слева и справа, т. е.

$$f(x) = \frac{1}{2}[f(x-0) + f(x+0)].$$

При этих допущениях интегральная теорема Фурье выражается так:

$$f(x) = \frac{1}{\pi}\int\limits_0^{\infty} d\tau \int\limits_{-\infty}^{\infty} f(t)\cos\tau(t-x)\,dt \qquad (1)$$

или, в комплексной записи,

$$f(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} d\tau \int\limits_{-\infty}^{\infty} f(t)\,e^{-i\tau(t-x)}\,dt. \qquad (2)$$

Интегральную теорему Фурье можно изложить и в следующем виде:

Если

$$
\left.\begin{aligned}
g(\tau)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-it\tau}\,dt,\\
&\text{то}\\
f(t)&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(\tau)e^{i\tau t}\,d\tau.
\end{aligned}\right\}\qquad(3)
$$

В последней записи теоремы обе формулы являются взаимно обратными, т. е. если рассматривать первую формулу как уравнение для неизвестной $f(t)$, то вторая формула дает его решение, а если во второй формуле видеть уравнение для $g(\tau)$, то решением будет первая формула.

Введем новую переменную $u=\frac{\tau}{2\pi}$, а в окончательном результате вновь заменим букву u буквой τ. Тогда интегральная теорема Фурье выразится с помощью следующих двух взаимно обратных формул:

$$
\left.\begin{aligned}
h(\tau)=\int_{-\infty}^{\infty} f(t)\,e^{-2\pi it\tau}\,dt,\qquad f(t)=\int_{-\infty}^{\infty} h(\tau)\,e^{2\pi it\tau}\,d\tau,\\
\text{где}\qquad h(\tau)=\sqrt{2\pi}\,g(2\pi\tau).
\end{aligned}\right\}\qquad(4)
$$

Возвращаясь к записи (1) интегральной теоремы Фурье, заметим, что, если $f(x)$ есть четная функция, т. е. $f(-x)=f(x)$, то формула (1) приводится к более простому виду:

$$
f(x)=\frac{2}{\pi}\int_0^{\infty}\cos(\tau x)\,d\tau\int_0^{\infty} f(t)\cos(\tau t)\,dt.
$$

Если же $f(x)$ — функция нечетная, т. е. если $f(-x)=-f(x)$, то

$$
f(x)=\frac{2}{\pi}\int_0^{\infty}\sin(\tau x)\,d\tau\int_0^{\infty} f(t)\sin(\tau t)\,dt.
$$

Прежде чем перейти к доказательству теоремы, приведем несколько примеров для ее иллюстрации.

а) Пусть $f(x)=1$, когда $|x|<1$, $f(x)=0$, когда $|x|>1$. Перед нами четная функция, и

$$
f(x)=\frac{2}{\pi}\int_0^{\infty}\cos(\tau x)\,d\tau\int_0^{1}\cos(t\tau)\,dt=\frac{2}{\pi}\int_0^{\infty}\frac{\sin\tau\cos(\tau x)}{\tau}\,d\tau=\begin{cases}0, & |x|>1,\\ \frac{1}{2}, & |x|=1,\\ 1, & |x|<1.\end{cases}
$$

Последний интеграл часто встречается в математической литературе под названием *разрывного множителя Дирихле*.

б) Пусть $f(x)=e^{-kx}$ $(k>0)$ при $x>0$, и пусть $f(-x)=f(x)$. Тогда

$$f(x)=\frac{2}{\pi}\int_0^\infty \cos(\tau x)\,d\tau\int_0^\infty e^{-kt}\cos(t\tau)\,dt=\frac{2}{\pi}\int_0^\infty \frac{k\cos(\tau x)}{k^2+\tau^2}\,d\tau=e^{-k|x|}.$$

Если же положить $f(-x)=-f(x)$, то получим

$$f(x)=\frac{2}{\pi}\int_0^\infty \sin(\tau x)\,d\tau\int_0^\infty e^{-kt}\sin(t\tau)\,dt=$$

$$=\frac{2}{\pi}\int_0^\infty \frac{\tau\sin(\tau x)}{k^2+\tau^2}\,d\tau=\begin{cases} e^{-kx} & \text{при} \quad x>0,\\ 0 & \text{при} \quad x=0,\\ -e^{kx} & \text{при} \quad x<0.\end{cases}$$

Отсюда вытекают при $k>0$ следующие две формулы:

$$\int_0^\infty \frac{\cos(\tau x)}{k^2+\tau^2}\,d\tau=\frac{\pi}{2}\,\frac{e^{-k|x|}}{k},\qquad \int_0^\infty \frac{\tau\sin\tau x}{k^2+\tau^2}\,d\tau=\begin{cases} \frac{\pi}{2}e^{-kx} & \text{при} \quad x>0,\\ 0 & \text{при} \quad x=0,\\ -\frac{\pi}{2}e^{kx} & \text{при} \quad x<0.\end{cases}$$

в) Функция $f(x)=e^{-\frac{x^2}{2}}$ дает интересную иллюстрацию взаимно обратных формул. Из решения упр. 4а, стр. 340 и 634, при $a=\frac{1}{2\tau^2}$ находим

$$\sqrt{\frac{2}{\pi}}\int_0^\infty e^{-t^2/2}\cos\tau t\,dt=e^{-\tau^2/2}.$$

Отсюда вытекает, что для функции $f(x)=e^{-\frac{x^2}{2}}$ обе взаимно обратные формулы для $g(\tau)$ и $f(t)$ совпадают.

2. Доказательство интегральной теоремы Фурье. В доказательстве этой теоремы мы будем опираться на предельную формулу Дирихле

$$\pi f(x)=\lim_{\lambda\to\infty}\int_{-a}^{a} f(x+t)\frac{\sin(\lambda t)}{t}\,dt,$$

справедливую при любом положительном значении a. Мы докажем эту формулу с помощью следующей леммы, доказанной в первом томе, гл. IX, стр. 524:

Если функция $s(t)$ кусочно непрерывна в интервале $\alpha\leqslant t\leqslant\beta$, то

$$\lim_{\lambda\to\infty}\int_\alpha^\beta s(t)\sin\lambda t\,dt=0,$$

Рассмотрим сперва интервал $0 \leqslant t \leqslant a$ и введем вспомогательную функцию

$$s(t) = \frac{f(x+t) - f(x+0)}{t},$$

которая, очевидно, является кусочно непрерывной в этом интервале. В силу требований, наложенных на функцию $f(x)$, ясно, что $\lim\limits_{t \to 0} s(t) = f'(x+0)$. Имеем

$$\int_0^a f(x+t) \frac{\sin(\lambda t)}{t} dt = \int_0^a f(x+0) \frac{\sin(\lambda t)}{t} dt + \int_0^a s(t) \sin(\lambda t)\, dt.$$

Согласно цитированной выше лемме, последний интеграл этой формулы стремится к нулю при $\lambda \to \infty$; что касается первого интеграла в правой части, то его предел

$$\lim_{\lambda \to \infty} \int_0^a f(x+0) \frac{\sin \lambda t}{t} dt = \lim_{\lambda \to \infty} f(x+0) \int_0^{a\lambda} \frac{\sin y}{y} dy =$$
$$= f(x+0) \int_0^{\infty} \frac{\sin y}{y} dy = \frac{\pi}{2} f(x+0)$$

(см. стр. 338). Аналогичное рассуждение покажет, что

$$\lim_{\lambda \to \infty} \int_{-a}^{0} f(x+t) \frac{\sin \lambda t}{t} dt = \frac{\pi}{2} f(x-0),$$

а так как $\frac{1}{2}[f(x+0) + f(x-0)] = f(x)$, то формула Дирихле доказана.

Следующей ступенью в доказательстве интегральной формулы Фурье является подстановка в формулу Дирихле выражения

$$\frac{\sin \lambda t}{t} = \int_0^{\lambda} \cos t\tau \, d\tau.$$

Интеграл, участвующий в формуле Дирихле, является функцией от a и от λ, и мы его обозначим через $F(\lambda, a)$. Тогда

$$F(\lambda, a) = \int_{-a}^{a} f(x+t) \frac{\sin(\lambda t)}{t} dt = \int_{-a}^{a} f(x+t)\, dt \int_0^{\lambda} \cos(t\tau)\, d\tau =$$
$$= \int_0^{\lambda} d\tau \int_{-a}^{a} f(x+t) \cos(t\tau)\, dt.$$

Формула Дирихле утверждает, что

$$\pi f(x) = \lim_{\lambda \to \infty} F(\lambda, a).$$

Так как этот предел не зависит от a, то можно писать и так:

$$\pi f(x) = \lim_{a\to\infty} \lim_{\lambda\to\infty} F(\lambda, a).$$

Если быть уверенным, что порядок предельных переходов в этой формуле можно переставить, т. е. что позволено выполнить предельный переход $a\to\infty$ над символом интеграла по t, то получится формула

$$\pi f(x) = \lim_{\lambda\to\infty} \int_0^\lambda d\tau \int_{-\infty}^{\infty} f(x+t)\cos t\tau\, dt = \int_0^\infty d\tau \int_{-\infty}^{\infty} f(x+t)\cos t\tau\, dt.$$

Сделаем теперь в правой части преобразование $x+t=u$, $dt=du$, а после этого заменим букву u буквой t, и формула Фурье будет доказана. Стало быть, доказательство будет завершено, когда мы докажем правомерность перестановки предельных переходов:

$$\lim_{a\to\infty} \lim_{\lambda\to\infty} F(\lambda, a) = \lim_{\lambda\to\infty} \lim_{a\to\infty} F(\lambda, a).$$

Согласно полученным ранее результатам (стр. 333), для этого достаточно показать, что предел

$$\lim_{a\to\infty} F(\lambda, a) = \int_{-\infty}^{\infty} f(x+t)\frac{\sin\lambda t}{t}\, dt$$

является интегралом, сходящимся равномерно относительно λ.

Для этого требуется установить, что по любому наперед заданному $\varepsilon>0$ можно найти такое число A, не зависящее от λ, что $|F(\lambda, a)-F(\lambda, b)|<\varepsilon$, лишь только оба числа a и b превзойдут A. Но ведь при $b>a>A$

$$|F(\lambda, a)-F(\lambda, b)| \leqslant \int_a^b |f(x+t)|\frac{|\sin\lambda t|}{t}\, dt + \int_{-b}^{-a} |f(x+t)|\frac{|\sin\lambda t|}{|t|}\, dt \leqslant \frac{2}{A}\int_{-\infty}^{\infty} |f(t)|\, dt.$$

Так как, согласно допущению 2), последний интеграл меньше C, то

$$|F(\lambda, a)-F(\lambda, b)| < \frac{2C}{A},$$

так что в качестве числа A, соответствующего числу ε, можно взять $A=\frac{2C}{\varepsilon}$. Тем самым завершено доказательство интегральной теоремы Фурье.

§ 6. Интегралы Эйлера (гамма-функция и бета-функция)

Одной из важнейших функций, определенных несобственным интегралом, содержащим параметр, является *гамма-функция* $\Gamma(x)$. Мы дадим здесь довольно подробный очерк свойств этой функции, введенной Эйлером. Изложение, во многом близкое к нашему, см. в книге: Artin E., Einführung in die Theorie der Γ-Funktion, Leipzig, 1931 [русск. перев.: Артин Е., Введение в теорию гамма-функции, М. — Л., ГТТИ, 1934].

1. Определение и функциональное уравнение гамма-функции. Функция $\Gamma(x)$ определяется для всех значений $x>0$ как несобственный интеграл

$$\Gamma(x)=\int_0^\infty e^{-t}t^{x-1}\,dt.$$

Этот интеграл мы уже изучали в т. I, гл. IV, стр. 291, при целых положительных значениях n. Там же мы обнаружили, что этот интеграл сходится при всяком $x>0$, и нетрудно показать, что он сходится *равномерно* во всяком замкнутом интервале положительной части оси x, не содержащем точки $x=0$. Стало быть, *функция* $\Gamma(x)$ *непрерывна при* $x>0$.

Интеграл для $\Gamma(x)$ можно привести с помощью простых преобразований к другим видам, которыми часто пользуются. Мы укажем здесь только подстановку $t=u^2$, которая приводит гамма-функцию к виду

$$\Gamma(x)=2\int_0^\infty e^{-u^2}u^{2x-1}\,du.$$

Стало быть, и обратно, часто встречающийся интеграл $\int_0^\infty e^{-u^2}u^\alpha\,du$ $(\alpha>-1)$, выражается через гамма-функцию:

$$\int_0^\infty e^{-u^2}u^\alpha\,du=\frac{1}{2}\Gamma\left(\frac{\alpha+1}{2}\right)\qquad(\alpha>-1)$$

(ср. стр. 325 и стр. 340, упр. 1).

Методом интегрирования произведения обнаруживаем (как и в т. I, стр. 292), что соотношение

$$\Gamma(x+1)=x\Gamma(x)$$

справедливо при любом $x>0$. Это равенство называется *функциональным уравнением гамма-функции.*

Конечно, требование, чтобы $\Gamma(x)$ было решением функционального уравнения $\varphi(x+1)=x\varphi(x)$, не определяет этой функции однозначно.

Действительно, произведение $\Gamma(x)$ на любую периодическую функцию $p(x)$ с периодом 1 тоже будет решением этого уравнения. С другой стороны, множество функций

$$\varphi(x)=\Gamma(x)p(x) \quad \text{при условии} \quad p(x+1)=p(x)$$

представляет совокупность всех решений нашего функционального уравнения. В самом деле, пусть $\varphi(x)$ есть любое решение; тогда отношение $f(x)=\frac{\varphi(x)}{\Gamma(x)}$, которое всегда возможно составить, так как $\Gamma(x)\neq 0$, удовлетворяет уравнению

$$f(x+1)=f(x).$$

Вместо функции $\Gamma(x)$ часто удобнее рассматривать функцию $u(x)=\ln\Gamma(x)$. Так как $\Gamma(x)$ определена при $x>0$ и $\Gamma(x)>0$, то функция $u(x)=\ln\Gamma(x)$ тоже определена при $x>0$. Эта функция удовлетворяет функциональному уравнению (на сей раз разностному уравнению)

$$u(x+1)-u(x)=\ln x. \tag{1}$$

Очевидно, сумма функции $\ln\Gamma(x)$ с произвольной периодической функцией периода 1 тоже будет решением этого уравнения. Для того чтобы функция $\ln\Gamma(x)$ определялась как решение функционального уравнения (1) однозначно, необходимо поэтому дополнить это уравнение еще какими-то условиями. Одно такое простое условие дает так называемая теорема Бора (H. Bohr), но для ее формулировки необходимо предварительно ввести понятие выпуклой функции.

2. Выпуклые функции и их свойства. [Рассмотрим дугу $\widehat{A_1A_2}$ кривой $y=f(x)$, ограниченную точками $A_1(x_1, f(x_1))$ и $A_2(x_2, f(x_2))$. Если α и β — любые два положительных числа, удовлетворяющих условию $\alpha+\beta=1$ (см. стр. 26, где $\lambda_0=\alpha$, $\lambda_1=\beta$), то точка $x=\alpha x_1+\beta x_2$, $y=\alpha f(x_1)+\beta f(x_2)$ пробегает хорду A_1A_2, так что абсциссе $\alpha x_1+\beta x_2$ соответствует точка хорды с ординатой $\alpha f(x_1)+\beta f(x_2)$ и точка дуги $\widehat{A_1A_2}$ с ординатой $f(\alpha x_1+\beta x_2)$. Теперь будет понятно следующее определение.]

Функция $f(x)$ называется *выпуклой* на отрезке $a\leqslant x\leqslant b$, если для любых двух точек x_1 и x_2 этого отрезка и любых двух положительных чисел α и β, связанных соотношением $\alpha+\beta=1$, выражение

$$\alpha f(x_1)+\beta f(x_2)-f(\alpha x_1+\beta x_2)$$

ни разу не меняет знака; или, выражаясь наглядно, если хорда, соединяющая любые две точки рассматриваемой дуги кривой $y=f(x)$, нигде не лежит ниже либо нигде не лежит выше дуги, стягиваемой этой хордой (рис. 81). (Ср. стр. 121.)

Прежде всего установим несколько свойств выпуклых функций. При этом можно ограничиться рассмотрением таких функций $f(x)$, для которых

$$\alpha f(x_1)+\beta f(x_2)-f(\alpha x_1+\beta x_2)\geqslant 0 \qquad (2)$$

и которые можно называть *выпуклыми книзу*, ибо функции, *выпуклые кверху*, можно всегда превратить умножением на -1 в функции, выпуклые книзу.

1) Если функция $f(x)$ дважды непрерывно дифференцируема, то

$$\alpha f(x_1)+\beta f(x_2)-f(\alpha x_1+\beta x_2)=$$
$$=\beta(x_2-x_1)^2\int_0^1 dt\int_{\beta t}^{t} f''[x_1+(x_2-x_1)\tau]\,d\tau,$$

что нетрудно проверить. Отсюда вытекает, что если $f''(x)$ сохраняет постоянный знак на отрезке $a\leqslant x\leqslant b$, то $f(x)$ есть выпуклая функция на этом отрезке, причем она выпукла книзу, если $f''(x)\geqslant 0$, и выпукла кверху, если $f''(x)\leqslant 0$. С другой стороны, вычислив предел

$$\lim_{x_2\to x_1}\frac{\alpha f(x_1)+\beta f(x_2)-f(\alpha x_1+\beta x_2)}{(x_2-x_1)^2}=\frac{\alpha\beta f''(x_1)}{2},$$

обнаружим, что постоянство знака второй производной является также и необходимым условием выпуклости функции; стало быть, оно представляет собой характеристическое свойство дважды непрерывно дифференцируемой выпуклой функции.

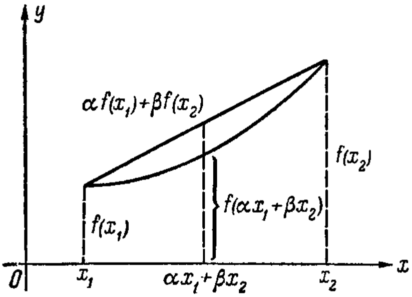


Рис. 81.

2) Замечательным и полезным для приложений свойством является тот факт, что нет надобности особо требовать непрерывности какой-либо выпуклой функции $f(x)$, ибо из определения выпуклой функции уже вытекает ее непрерывность. Более того, можно даже при этом заменить определяющее неравенство (2) значительно более слабым по виду, но в действительности равносильным неравенством, как это выражено в следующей теореме:

Если ограниченная функция $f(x)$ удовлетворяет неравенству

$$f(x+h)+f(x-h)-2f(x)\geqslant 0 \qquad (3)$$

при всех значениях x и h, для которых аргументы $x\pm h$ еще лежат в интервале определения функции, т. е. если середина любой хорды кривой $y=f(x)$ лежит не ниже дуги кривой, то функция $f(x)$ выпукла (книзу) (ср. т. I, стр. 164, упр. 60).

Доказательство этой теоремы мы проведем в два этапа, а) Сначала докажем, что *всякая ограниченная функция* $f(x)$, *удовлетворяющая неравенству* (3), *непрерывна.*

Для доказательства перепишем неравенство (3) в несколько ином виде:

$$f(x)-f(x-h)\leqslant f(x+h)-f(x).$$

Из этой записи легко вывести следующие неравенства:

$$f(x-kh)-f(x-(k+1)h)\leqslant f(x+h)-f(x)\leqslant$$
$$\leqslant f(x+(k+1)h)-f(x+kh),$$

справедливые при любом целом $k\geqslant 0$. Просуммировавши эти неравенства по k от $k=0$ до $k=n-1$, получим оценку

$$\frac{f(x)-f(x-nh)}{n}\leqslant f(x+h)-f(x)\leqslant\frac{f(x+nh)-f(x)}{n};$$

так как, в силу ограниченности функции $f(x)$, $|f(x)|\leqslant C$, то

$$|f(x+h)-f(x)|\leqslant\frac{2C}{n}.$$

Правда, при этом n может быть любым положительным целым числом, но только подчиненным условию, что аргументы $x\pm nh$ еще лежат в области определения функции $f(x)$. Однако, если мы заставим h стремиться к нулю, то наибольшее возможное значение n будет расти неограниченно, и, значит, выражение $f(x+h)-f(x)$ будет стремиться к нулю, чем и доказывается непрерывность функции $f(x)$.

б) Используя непрерывность функции $f(x)$, теперь нетрудно доказать ее выпуклость, т. е. вывести для нее неравенство

$$\alpha f(x_1)+\beta f(x_2)-f(\alpha x_1+\beta x_2)\geqslant 0.$$

Для этого из доказанного ранее неравенства

$$f(x)-f(x-nh)\leqslant n\,[f(x+h)-f(x)],$$

полагая $\xi=x-nh$, выводим соотношение

$$\frac{f(\xi+nh)-f(\xi)}{n}\leqslant\frac{f(\xi+(n+1)h)-f(\xi)}{n+1},$$

а из него более общее неравенство

$$\frac{f(\xi+mh)-f(\xi)}{m}\leqslant\frac{f(\xi+nh)-f(\xi)}{n}\qquad(0<m\leqslant n).$$

Полагая в нем $\xi+nh=\xi_1$, получим после некоторых преобразований

$$\left(1-\frac{m}{n}\right)f(\xi)+\frac{m}{n}f(\xi_1)\geqslant f\left(\left(1-\frac{m}{n}\right)\xi+\frac{m}{n}\xi_1\right),$$

а это и есть неравенство, подлежавшее доказательству, но для рациональных значений чисел α и β. Справедливость же этого неравенства при любых α и β вытекает теперь из непрерывности функции $f(x)$.

3) Из определения выпуклости вытекает двойное неравенство

$$\frac{f(x+a)-f(x)}{a} \leqslant \frac{f(x+b)-f(x)}{b} \leqslant \frac{f(x+c)-f(x)}{c}$$

при любых отличных от нуля значениях a, b, c, для которых $a \leqslant b \leqslant c$. Отсюда вытекает, что отношение приращений $\frac{f(x+h)-f(x)}{h}$ ограничено и изменяется монотонно, когда h стремится к нулю по положительным, либо по отрицательным значениям. Следовательно, оно имеет как правый, так и левый предел, и стало быть, выпуклая функция имеет в каждой точке правую и левую производную.

4) Наконец, приведем еще для выпуклых (книзу) функций следующее неравенство:

$$f(x+h)+f(x-h)-[f(x+\delta)+f(x-\delta)] \geqslant 0,$$

коль скоро $h \geqslant \delta > 0$. Оно становится очевидным при одном взгляде на график, а доказательство неравенства получается при сложении двух соотношений:

$$\frac{1}{2}\left(1+\frac{\delta}{h}\right)f(x-h)+\frac{1}{2}\left(1-\frac{\delta}{h}\right)f(x+h)-f(x-\delta) \geqslant 0,$$

$$\frac{1}{2}\left(1-\frac{\delta}{h}\right)f(x-h)+\frac{1}{2}\left(1+\frac{\delta}{h}\right)f(x+h)-f(x+\delta) \geqslant 0,$$

вытекающих из определяющего неравенства (2).

3. Теорема Бора. Теперь мы можем формулировать и доказать теорему Бора, упомянутую в конце n° 1:

Всякое выпуклое решение разностного уравнения

$$u(x+1)-u(x)=\ln x$$

в интервале $0 < x < \infty$ *может отличаться от функции* $\ln \Gamma(x)$ *только на аддитивную константу* (постоянное слагаемое).

Прежде всего убедимся, что $\ln \Gamma(x)$ есть выпуклая функция. Для доказательства мы воспользуемся неравенством Шварца для интегралов

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \leqslant \int_a^b f^2(x)\,dx \int_a^b g^2(x)\,dx,$$

доказательство которого будет дано в примечании 1 в конце этого номера. Запишем интеграл, определяющий $\Gamma(x)$, в следующем виде:

$$\Gamma(x)=\int_0^\infty \underbrace{e^{-\frac{t}{2}}t^{\frac{x-1+h}{2}}}_{f(x)}\underbrace{e^{-\frac{t}{2}}t^{\frac{x-1-h}{2}}}_{g(x)}\,dt,$$

где h — любое положительное число, а x — любое число, большее чем h, и к этому интегралу применим неравенство Шварца. Тогда сразу получится

$$[\Gamma(x)]^2 \leqslant \Gamma(x+h)\,\Gamma(x-h),$$

откуда

$$\ln\Gamma(x+h)+\ln\Gamma(x-h)-2\ln\Gamma(x)\geqslant 0, \tag{4}$$

чем и доказывается, что функция $\ln\Gamma(x)$ выпукла книзу.

Пусть теперь функции $f(x)$ и $g(x)$ — два ограниченных в любом конечном интервале выпуклых решения функционального уравнения

$$u(x+1)-u(x)=\ln x;$$

тогда их разность

$$\varphi(x)=f(x)-g(x)$$

является непрерывной периодической функцией периода 1. Кроме того, из неравенств

$$f(x+1)-f(x)=\ln x \quad \text{и} \quad f(x)-f(x-1)=\ln(x-1)$$

вытекает для $f(x)$ соотношение

$$f(x+1)+f(x-1)-2f(x)=\ln\frac{x}{x-1}.$$

Отсюда в сочетании со свойством 4 выпуклой функции (n° 2) получается неравенство

$$f(x+h)+f(x-h)-2f(x)\leqslant\ln\frac{x}{x-1},$$

справедливое при любом h из интервала $0<h\leqslant 1$. Аналогичное неравенство получается, конечно, и для $g(x)$:

$$g(x+h)+g(x-h)-2g(x)\leqslant\ln\frac{x}{x-1}.$$

Следовательно,

$$|\varphi(x+h)+\varphi(x-h)-2\varphi(x)|\leqslant 2\ln\frac{x}{x-1}.$$

Заставим теперь x безгранично возрастать; тогда $\ln\frac{x}{x-1}$ будет стремиться к нулю, а следовательно и функция $\varphi(x+h)+\varphi(x-h)-2\varphi(x)$ будет стремиться к нулю. Так как эта функция является периодической, то неизбежно приходим к выводу, что

$$\varphi(x+h)+\varphi(x-h)-2\varphi(x)=0$$

при всяком $x>0$.

Непрерывная периодическая функция $\varphi(x)$, удовлетворяющая этому соотношению при любом h из интервала $0<h\leqslant 1$ и любом $x>h$ может быть только постоянной.

Действительно, так как $\varphi(x)$ периодична с периодом 1, то $\varphi(1)=\varphi(2)$. Обозначим это общее значение через a; тогда $\varphi\left(\frac{3}{2}\right)=\frac{1}{2}[\varphi(1)+\varphi(2)]$, а следовательно, и $\varphi\left(\frac{3}{2}\right)=a$. Далее оказывается, что $\varphi(x_k)=a$ во всех точках x_k отрезка $1 \leqslant x \leqslant 2$, получающихся при последовательном делении пополам всех частичных отрезков. Так как эти точки x_k расположены всюду плотно, то из непрерывности функции $\varphi(x)$ следует, что $\varphi(x) = a$ на всем отрезке $1 \leqslant x \leqslant 2$, а в силу периодичности этой функции также и при любом $x > 0$.

Таким образом доказано, что всякое выпуклое решение уравнения $u(x+1) - u(x) = \ln x$, ограниченное в любом конечном интервале $a \leqslant x \leqslant b$, где $a > 0$, может отличаться от функции $\ln \Gamma(x)$ только постоянным слагаемым.

Примечание 1. *Неравенство Шварца для интегралов*

$$\left[\int_a^b f(x)\, g(x)\, dx\right]^2 \leqslant \int_a^b [f(x)]^2\, dx \cdot \int_a^b [g(x)]^2\, dx$$

можно вывести из алгебраического неравенства Шварца

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leqslant \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right),$$

доказанного в т. I, стр. 27—28, рассматривая интегралы как пределы сумм.

Другое доказательство. Из основных свойств двойного интеграла вытекает соотношение

$$0 \leqslant \int_a^b\!\!\int_a^b [f(x)\, g(y) - f(y)\, g(x)]^2\, dx\, dy = \int_a^b f^2(x)\, dx \int_a^b g^2(y)\, dy +$$

$$+ \int_a^b f^2(y)\, dy \int_a^b g^2(x)\, dx - 2\int_a^b f(x)\, g(x)\, dx \int_a^b f(y)\, g(y)\, dy.$$

Так как $\int_a^b \varphi(x)\, dx = \int_a^b \varphi(y)\, dy$, то правая часть равна

$$2\int_a^b f^2(x)\, dx \int_a^b g^2(x)\, dx - 2\left(\int_a^b f(x)\, g(x)\, dx\right)^2,$$

что и доказывает неравенство Шварца.

Еще одно доказательство получается из неравенства

$$0 \leqslant \int_a^b [pf(x) - qg(x)]^2\, dx =$$

$$= p^2 \int_a^b f^2(x)\, dx + q^2 \int_a^b g^2(x)\, dx - 2pq \int_a^b f(x)\, g(x)\, dx,$$

справедливого при любых постоянных p и q. Надо лишь положить

$$\frac{1}{p^2} = \int_a^b f^2(x)\, dx \quad \text{и} \quad \frac{1}{q^2} = \int_a^b g^2(x)\, dx.$$

[Интегральное неравенство, называемое обычно неравенством Шварца, было впервые установлено академиком В. Я. Буняковским в 1859 г.]

Примечание 2 (к неравенству (4)). Этот факт является частным случаем более общей теоремы. Если функции $f_\nu(x)$ ($\nu = 1, 2, \ldots, n$) удовлетворяют условиям

$$f_\nu(x) \geqslant 0 \quad \text{и} \quad f_\nu^2(x) \leqslant f_\nu(x-h)\,f_\nu(x+h),$$

т. е. если функции $\ln f_\nu(x)$ выпуклы, то и сумма $\sum\limits_1^n f_\nu(x)$ удовлетворяет этим условиям.

В самом деле, запишем $\Sigma f_\nu(x)$ в следующем виде:

$$\sum_1^n f_\nu(x) = \sum_1^n \frac{f_\nu(x)}{\sqrt{f_\nu(x-h)}\sqrt{f_\nu(x+h)}} \sqrt{f_\nu(x-h)}\sqrt{f_\nu(x+h)}.$$

Так как $\dfrac{f_\nu(x)}{\sqrt{f_\nu(x-h)}\sqrt{f_\nu(x+h)}} \leqslant 1$, то

$$\left(\sum_1^n f_\nu(x)\right)^2 \leqslant \left(\sum_1^n \sqrt{f_\nu(x-h)}\sqrt{f_\nu(x+h)}\right)^2.$$

Применив к правой части алгебраическое неравенство Шварца, получим

$$\left(\sum_1^n f_\nu(x)\right)^2 \leqslant \sum_1^n f_\nu(x-h) \cdot \sum_1^n f_\nu(x+h).$$

Аналогичную теорему можно доказать для интегралов вида

$$\int_a^b f(x, t)\,dt,$$

где функции $f(x, t)$ удовлетворяют при всех значениях параметра t условиям

$$f(x, t) \geqslant 0 \quad \text{и} \quad [f(x, t)]^2 \leqslant f(x-h, t)\, f(x+h, t).$$

Гамма-функция является интегралом именно такого типа.

4. Представление гамма-функции в виде бесконечного произведения. Дадим, следуя Гауссу и Вейерштрассу, разложение гамма-функции в бесконечное произведение. Сначала докажем, что

$$\Gamma(x) = \lim_{n\to\infty} G_n(x),$$

где

$$G_n(x) = \frac{1 \cdot 2 \ldots (n-1)}{x(x+1)\ldots(x+n-1)}\, n^x.$$

Это утверждение кажется правдоподобным, потому что при целых значениях $x = k$

$$G_n(k) = (k-1)!\,\frac{n}{n}\,\frac{n}{n+1}\cdots\frac{n}{n+k-1}$$

и при $n \to \infty$ стремится, очевидно, к пределу $(k-1)! = \Gamma(k)$.

Нашей целью является доказать, что последовательность функций $G_n(x)$ сходится при всех значениях x, за исключением значений

$x=0, -1, -2, \ldots$, и что предельная функция $G(x)$ совпадает при $x>0$ с функцией $\Gamma(x)$.

Для того чтобы доказать сходимость последовательности $G_n(x)$ при $x \neq 0, -1, -2, \ldots$, введем для числа n выражение

$$n=(1+1)\left(1+\frac{1}{2}\right)\left(1+\frac{1}{3}\right)\ldots\left(1+\frac{1}{n-1}\right)$$

и соответственно этому напишем

$$G_n(x)=\frac{1}{x}\prod_{\nu=1}^{n-1}\frac{\left(1+\frac{1}{\nu}\right)^x}{1+\frac{x}{\nu}}. \qquad (1)$$

Согласно признаку сходимости, доказанному в т. I, стр. 488, бесконечное произведение, стоящее в правой части, сходится абсолютно и равномерно, если ряд

$$\sum_{\nu=1}^{\infty}\left|\frac{\left(1+\frac{1}{\nu}\right)^x}{1+\frac{x}{\nu}}-1\right| \qquad (2)$$

сходится равномерно. Разлагая $\left(1+\frac{1}{\nu}\right)^x$ по степеням $\alpha=\frac{1}{\nu}$ по формуле Тэйлора с остаточным членом второго порядка, приведем общий член ряда (2) к виду

$$\frac{\left(1+\frac{1}{\nu}\right)^x-\left(1+\frac{x}{\nu}\right)}{1+\frac{x}{\nu}}=\frac{1}{2\nu^2}\,\frac{x(x-1)\left(1+\frac{\theta}{\nu}\right)^{x-2}}{1+\frac{x}{\nu}},$$

где θ — некоторое число между 0 и 1. Отсюда вытекает, что

$$\left|\frac{\left(1+\frac{1}{\nu}\right)^x-\left(1+\frac{x}{\nu}\right)}{1+\frac{x}{\nu}}\right|\leqslant\frac{C}{\nu^2},$$

где C не зависит от номера ν. Во всякой замкнутой области, не содержащей точек $x=-1, -2, \ldots$, можно принять за C постоянную, не зависящую не только от ν, но и от x. Отсюда вытекает, что во всякой такой области сходится равномерно ряд (2), а следовательно и бесконечное произведение (1).

Предельная функция

$$G(x)=\lim_{n\to\infty}\frac{1\cdot 2\ldots(n-1)}{x(x+1)\ldots(x+n-1)}\,n^x$$

непрерывна при всяком $x\neq 0, -1, -2, \ldots$; нетрудно убедиться, что она удовлетворяет функциональному уравнению

$$G(x+1)=xG(x),$$

а ее логарифм, $u = \ln G(x)$, удовлетворяет уравнению

$$u(x+1) - u(x) = \ln x.$$

Для того чтобы доказать, что при $x > 0$ функция $G(x)$ совпадает с функцией $\Gamma(x)$, по теореме Бора достаточно показать, что $\ln G(x)$ является выпуклой функцией и что в какой-либо одной точке x функции $\ln G(x)$ и $\ln \Gamma(x)$ принимают равные значения. Замечаем, что $\ln G(x)$ есть предельная функция последовательности

$$\ln G_n(x) = \ln (n-1)! + x \ln n - \sum_{\nu=0}^{n-1} \ln (x+\nu).$$

При любом значении $h > 0$ и любом значении $x > h$ функции $\ln G_n(x)$ удовлетворяют условию выпуклости

$$\ln G_n(x+h) + \ln G_n(x-h) - 2 \ln G_n(x) =$$
$$= \sum_{\nu=0}^{n-1} [2 \ln (x+\nu) - \ln (x+h+\nu) - \ln (x-h+\nu)] \geqslant 0,$$

а стало быть, и предельная функция $\ln G(x)$ выпукла. Так как к тому же

$$\ln G(1) = \ln \Gamma(1) = 0,$$

то по теореме Бора $G(x) \equiv \Gamma(x)$. Стало быть, мы получили для функции $\Gamma(x)$ *разложение Гаусса в бесконечное произведение*

$$\Gamma(x) = \lim_{n\to\infty} \frac{1\cdot 2 \ldots (n-1)}{x(x+1)\ldots(x+n-1)} n^x = \frac{1}{x} \prod_{\nu=1}^{\infty} \frac{\left(1+\frac{1}{\nu}\right)^x}{1+\frac{x}{\nu}}.$$

Это представление гамма-функции имеет принципиальное значение, так как его можно рассматривать как определение $\Gamma(x)$ не только для всех положительных значений x, но также и для всех нецелых отрицательных значений x.

Полученное бесконечное произведение можно привести к несколько иному виду. В выражение $n^x = e^{x \ln n}$ подставим вместо $\ln n$ выражение

$$\ln n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - \gamma + \varepsilon_n,$$

где γ — постоянная Эйлера (т. I, стр. 444), а $\varepsilon_n \to 0$ при $n \to \infty$, Тогда для $\frac{1}{\Gamma(x)}$ получится следующее выражение:

$$\frac{1}{\Gamma(x)} = x \lim_{n\to\infty} (1+x)\left(1+\frac{x}{2}\right)\ldots\left(1+\frac{x}{n-1}\right) \cdot e^{-x-\frac{x}{2}-\ldots-\frac{x}{n}+\gamma x - \varepsilon_n x} =$$
$$= x e^{\gamma x} \lim_{n\to\infty} e^{-\varepsilon_n x - \frac{x}{n}} \prod_{\nu=1}^{n-1} \left(1+\frac{x}{\nu}\right) e^{-\frac{x}{\nu}}.$$

Так как при $n \to \infty$ множитель $e^{-\varepsilon_n x - \frac{x}{n}} \to 1$, то для функции $\frac{1}{\Gamma(x)}$ получается сходящееся *бесконечное произведение Вейерштрасса*

$$\frac{1}{\Gamma(x)} = xe^{\gamma x} \prod_{1}^{\infty} \left(1 + \frac{x}{\nu}\right) e^{-\frac{x}{\nu}},$$

из которого сразу видно, что $\frac{1}{\Gamma(x)}$ имеет нули (нулевые точки) кратности 1 при $x = 0, -1, -2, \dots$

5. Функция $\ln \Gamma(x)$ и ее производные. Логарифмируя бесконечное произведение Вейерштрасса (конец предыдущего номера), получим для функции $\ln \Gamma(x)$ следующее выражение:

$$\ln \Gamma(x) = -\ln x - \gamma x - \sum_{\nu=1}^{\infty} \left[\ln\left(1 + \frac{x}{\nu}\right) - \frac{x}{\nu}\right].$$

Докажем, что ряд в правой части сходится абсолютно и равномерно в любом замкнутом интервале положительной оси x. Действительно, его общий член

$$\ln\left(1 + \frac{x}{\nu}\right) - \frac{x}{\nu} = -\frac{1}{\nu} \int_0^x \frac{t\,dt}{\nu + t},$$

так что

$$\left|\ln\left(1 + \frac{x}{\nu}\right) - \frac{x}{\nu}\right| \leqslant \frac{1}{\nu^2} \int_0^x t\,dt = \frac{x^2}{2\nu^2}.$$

Так как ряд $\sum \frac{1}{\nu^2}$ сходится, то наше утверждение доказано.

Особый интерес представляют производные от функции $\ln \Gamma(x)$, ибо они дают возможность выразить в явном виде значения рядов $\sum_{\nu=0}^{\infty} \left(\frac{1}{x+\nu}\right)^m$.

Дифференцируя выражение для $\ln \Gamma(x)$ почленно по x, получим

$$\frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)} = -\frac{1}{x} - \gamma - \sum_{\nu=1}^{\infty} \left(\frac{1}{x+\nu} - \frac{1}{\nu}\right).$$

[Почленное дифференцирование законно, ибо общий член полученного ряда

$$\frac{1}{x+\nu} - \frac{1}{\nu} = -\frac{x}{\nu(x+\nu)},$$

откуда видно, что полученный дифференцированием ряд сходится абсолютно и равномерно в любом замкнутом интервале положительной оси x.] Дифференцируя почленно еще раз, получим

$$\frac{d^2}{dx^2}\ln\Gamma(x)=\sum_{\nu=0}^{\infty}\frac{1}{(x+\nu)^2}$$

и, наконец, после m-кратного дифференцирования.

$$\sum_{\nu=0}^{\infty}\frac{1}{(x+\nu)^m}=\frac{(-1)^m}{(m-1)!}\frac{d^m}{dx^m}\ln\Gamma(x)\qquad(m\geqslant 2).$$

6. Формула дополнения. Выведем теперь формулу, которая дает возможность легко вычислять значения функции $\Gamma(x)$ при отрицательных x по ее значениям при положительных x. Для этого составим произведение

$$\Gamma(x)\Gamma(-x)=$$
$$=\lim_{n\to\infty}\frac{1\cdot 2\ldots(n-1)}{x(x+1)\ldots(x+n-1)}n^x\lim_{n\to\infty}\frac{1\cdot 2\ldots(n-1)}{-x(1-x)(2-x)\ldots(n-1-x)}n^{-x}$$

и объединим оба предельных перехода в один:

$$\Gamma(x)\Gamma(-x)=-\frac{1}{x^2}\lim_{n\to\infty}\frac{1}{\left(1-\frac{x^2}{1^2}\right)\left(1-\frac{x^2}{2^2}\right)\ldots\left(1-\frac{x^2}{(n-1)^2}\right)}$$

Вспомнив разложение синуса в бесконечное произведение

$$\frac{\sin\pi x}{\pi x}=\lim_{n\to\infty}\left(1-\frac{x^2}{1^2}\right)\left(1-\frac{x^2}{2^2}\right)\ldots\left(1-\frac{x^2}{(n-1)^2}\right),$$

выведенное в т. I, стр. 518, получим окончательно

$$\Gamma(x)\Gamma(-x)=-\frac{\pi}{x\sin\pi x}.$$

Стало быть,

$$\Gamma(-x)=-\frac{\pi}{x\sin\pi x}\cdot\frac{1}{\Gamma(x)}.$$

Отсюда можно вывести выражение для произведения $\Gamma(x)\Gamma(1-x)$ и тем самым придать полученной формуле несколько иной вид. Так как

$$\Gamma(1-x)=-x\Gamma(-x),$$

то

$$\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin\pi x}.$$

Эта формула называется *формулой дополнения* для гамма-функции.

Подставив в эту формулу $x=\frac{1}{2}$, получим $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$. Так как $\Gamma\left(\frac{1}{2}\right)=$ $=2\int\limits_0^{\infty}e^{-u^2}du$, то перед нами новое доказательство формулы $\int\limits_0^{\infty}e^{-u^2}du=$

$=\frac{\sqrt{\pi}}{2}$. Далее, зная $\Gamma\left(\frac{1}{2}\right)$, можно вычислить значение гамма-функции для любого полуцелого положительного значения $x=n+\frac{1}{2}$, где $n=1, 2, 3, \ldots$:

$$\Gamma\left(n+\frac{1}{2}\right)=\left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right)\ldots\frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{(2n-1)(2n-3)\ldots3\cdot1}{2^n}\sqrt{\pi},$$

а затем и для любого отрицательного полуцелого числа $x=-n+\frac{1}{2}$, где $n=1, 2, 3, \ldots$:

$$\Gamma\left(-n+\frac{1}{2}\right)=-\frac{\pi}{\sin\left(n-\frac{1}{2}\right)\pi}\,\frac{1}{\Gamma\left(n+\frac{1}{2}\right)}=\frac{(-1)^n 2^n\sqrt{\pi}}{(2n-1)(2n-3)\ldots3\cdot1}.$$

В частности, при $n=1$ получаем $\Gamma\left(-\frac{1}{2}\right)=-2\sqrt{\pi}$.

7. Бета-функция и ее функциональное уравнение. Эйлер ввел в научный обиход еще одну функцию, определяемую несобственным интегралом, содержащим *два* параметра x и y. Эта функция, носящая название бета-функции, определена интегралом

$$B(x, y)=\int_0^1 t^{x-1}(1-t)^{y-1}\,dt. \tag{1}$$

Если одновременно $x\geqslant1$ и $y\geqslant1$, то это обычный, собственный интеграл. Если же x или y меньше единицы, то интеграл становится несобственным; однако, если ограничиваться значениями $x\geqslant\varepsilon$, $y\geqslant\eta$, где ε и η — любые сколь угодно малые положительные числа, то согласно нашему критерию из § 4, n° 1 (стр. 331), интеграл сходится равномерно относительно x и y. Стало быть, при всех положительных значениях x и y функция $B(x, y)$ является непрерывной функцией. [Напротив, если хотя бы одна из переменных x, y принимает значение $\leqslant0$, то интеграл расходится.

Почти сразу убеждаемся (с помощью преобразования $t=1-u$), что

$$B(y, x)=B(x, y).$$

Это свойство словесно выражают так: $B(x, y)$ является симметрической функцией своих аргументов.]

С помощью замены переменной $t=\tau+\frac{1}{2}$ получаем для бета-функции несколько иное выражение:

$$B(x, y)=\int_{-\frac{1}{2}}^{\frac{1}{2}}\left(\frac{1}{2}+\tau\right)^{x-1}\left(\frac{1}{2}-\tau\right)^{y-1}d\tau, \tag{2}$$

а этот интеграл преобразуется подстановкой $\tau=\frac{t}{2s}$, где постоянная $s>0$, к виду:

$$(2s)^{x+y-1}\mathrm{B}(x, y)=\int_{-s}^{s}(s+t)^{x-1}(s-t)^{y-1}\,dt. \tag{3}$$

Наконец, преобразование $t=\sin^2\varphi$ приводит исходную формулу (1) к четвертому виду:

$$\mathrm{B}(x, y)=2\int_{0}^{\frac{\pi}{2}}\sin^{2x-1}\varphi\cos^{2y-1}\varphi\,d\varphi. \tag{4}$$

[Выведем *функциональное уравнение* для бета-функции. Для этого напишем

$$\mathrm{B}(x+1, y)=\int_{0}^{1}t^{x}(1-t)^{y-1}\,dt$$

и преобразуем интеграл в правой части по правилу интегрирования произведения, причем сперва интегрируем второй множитель $(1-t)^{y-1}$; если $x>0$, то

$$\mathrm{B}(x+1, y)=\left[-t^{x}\frac{(1-t)^{y}}{y}\right]_0^1+\frac{x}{y}\int_{0}^{1}t^{x-1}(1-t)^{y}\,dt=$$

$$=\frac{x}{y}\int_{0}^{1}t^{x-1}(1-t)^{y-1}\,dt-\frac{x}{y}\int_{0}^{1}t^{x}(1-t)^{y-1}\,dt=$$

$$=\frac{x}{y}\mathrm{B}(x, y)-\frac{x}{y}\mathrm{B}(x+1, y),$$

откуда и получается функциональное уравнение

$$\mathrm{B}(x+1, y)=\frac{x}{x+y}\mathrm{B}(x, y).$$

Пользуясь соотношением $\mathrm{B}(x, y+1)=\mathrm{B}(y+1, x)$, функциональному уравнению можно придать другой вид:

$$\mathrm{B}(x, y+1)=\frac{y}{x+y}\mathrm{B}(x, y).]$$

8. Связь между бета-функцией и гамма-функцией. Теперь мы покажем, что бета-функцию можно выразить через гамма-функцию. Для этого нам придется сделать некоторые преобразования, которые могут показаться на первый взгляд немного странными.

Помножим обе стороны формулы (3), выражающей бета-функцию, на e^{-2s} и проинтегрируем по s от 0 до A:

$$\mathrm{B}(x, y)\int_{0}^{A}e^{-2s}(2s)^{x+y-1}\,ds=\int_{0}^{A}e^{-2s}\,ds\int_{-s}^{s}(s+t)^{x-1}(s-t)^{y-1}\,dt. \quad (*)$$

Повторный интеграл в правой части можно рассматривать как двойной интеграл функции $e^{-2s}(s+t)^{x-1}(s-t)^{y-1}$ по области, имеющей вид равнобедренного треугольника, стороны которого лежат на прямых $s-t=0$, $s+t=0$ и $s=A$.

Этот интеграл мы теперь преобразуем с помощью отображения

$$u=s-t, \quad v=s+t, \quad \text{откуда} \quad s=\frac{u+v}{2}, \quad t=\frac{v-u}{2},$$

в другой интеграл

$$\frac{1}{2}\iint\limits_{G} e^{-(u+v)}\, v^{x-1} u^{y-1}\, du\, dv. \qquad (\overset{*}{*})$$

Областью интегрирования G является теперь треугольник плоскости uv, ограниченный прямыми $u=0$, $v=0$ и $u+v=2A$.

Заставим теперь A неограниченно возрастать. Тогда левая часть равенства (*) имеет своим пределом функцию

$$\frac{1}{2}\,\mathrm{B}(x, y)\,\Gamma(x+y).$$

Поэтому должна сходиться и правая часть, равная интегралу ($\overset{*}{*}$), и ее пределом будет двойной интеграл по всей первой четверти плоскости uv, причем этот бесконечный квадрант исчерпывается с помощью последовательности равнобедренных прямоугольных треугольников, катеты которых лежат на осях u и v. Так как подынтегральная функция положительна, а интеграл ($\overset{*}{*}$) сходится для монотонной последовательности областей, то, согласно гл. IV, § 5, стр. 285, получающийся в пределе несобственный двойной интеграл не зависит от способа исчерпывания (или аппроксимирования) результирующего квадранта. Стало быть, можно воспользоваться для этой цели квадратами, стороны которых, длины A, лежат на осях u и v, и соответственно этому писать

$$\mathrm{B}(x, y)\,\Gamma(x+y)=\lim_{A\to\infty}\int_0^A\int_0^A e^{-(u+v)}\, v^{x-1} u^{y-1}\, du\, dv=$$

$$=\int_0^\infty e^{-v} v^{x-1}\, dv \int_0^\infty e^{-u} u^{y-1}\, du.$$

Таким образом, мы получили важное соотношение

$$\mathrm{B}(x, y)=\frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}.$$

Эту формулу, выражающую бета-функцию через гамма-функцию, можно также получить с помощью теоремы Бора. Сначала убеждаемся с помощью функционального уравнения бета-функции (n° 7)

$$\mathrm{B}(x+1, y)=\frac{x}{x+y}\,\mathrm{B}(x, y),$$

что функция $u(x, y) = \Gamma(x+y)\,\mathrm{B}(x, y)$, рассматриваемая как функция от x, удовлетворяет функциональному уравнению гамма-функции

$$u(x+1) = xu(x).$$

Так как из теоремы примечания 2 в конце n° 3 вытекает, что $\ln u(x, y)$ является выпуклой функцией от x, то

$$\Gamma(x+y)\,\mathrm{B}(x, y) = \Gamma(x)\cdot a(y),$$

и, наконец, полагая $x=1$, находим $a(y) = \Gamma(y)$.

Мы знаем, что функция $\Gamma(x)$ является распространением факториала $\Gamma(n) = (n-1)!$ на все положительные числа. Подобно этому, из формулы, выражающей $\mathrm{B}(x, y)$ через гамма-функцию, можно вывести выражение *биномиальных коэффициентов* $\binom{m+n}{n} = \frac{(m+n)!}{m!\,n!}$ через бета-функцию $\mathrm{B}(m+1,\ n+1)$. Действительно,

$$\mathrm{B}(m+1,\ n+1) = \frac{\Gamma(m+1)\,\Gamma(n+1)}{\Gamma(m+n+2)} = \frac{m!\,n!}{(m+n+1)!},$$

откуда $\binom{m+n}{m} = \binom{m+n}{n} = \frac{(m+n)!}{m!\,n!} = \frac{1}{(m+n+1)\,\mathrm{B}(m+1,\ n+1)}$.

Следовательно, функция $\frac{1}{(x+y+1)\,\mathrm{B}(x+1,\ y+1)}$ при целых $x=m$, $y=n$ принимает значения $\binom{m+n}{n}$.

Отметим еще, что из формулы (4) предыдущего номера видно, что

$$\int_0^{\frac{\pi}{2}} \sin^\alpha t\,dt = \int_0^{\frac{\pi}{2}} \cos^\alpha t\,dt =$$

$$= \frac{1}{2}\,\mathrm{B}\left(\frac{\alpha+1}{2},\ \frac{1}{2}\right) = \frac{1}{2}\,\mathrm{B}\left(\frac{1}{2},\ \frac{\alpha+1}{2}\right) = \frac{\sqrt{\pi}}{\alpha}\,\frac{\Gamma\left(\frac{\alpha+1}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}.$$

Упражнения

1. Доказать, что объем тела, лежащего в положительном октанте и ограниченного плоскостями $x=0$, $y=0$, $z=h$ и поверхностью

$$\frac{x^m}{a^m} + \frac{y^m}{b^m} = \frac{z}{c} \qquad (m > 0),$$

равен

$$abh\left(\frac{h}{c}\right)^{\frac{2}{m}} \frac{\Gamma\left(1+\frac{1}{m}\right)^2}{\Gamma\left(2+\frac{2}{m}\right)}.$$

2. Доказать, что интеграл

$$\iiint f\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) x^{p-1} y^{q-1} z^{r-1}\,dx\,dy\,dz,$$

распространенный на положительный октант эллипсоида $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\leqslant 1$, равен

$$\frac{a^p b^q c^r}{8}\,\frac{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{r}{2}\right)}{\Gamma\left(\frac{p+q+r}{2}\right)}\int_0^1 f(\xi)\,\xi^{\frac{p+q+r}{2}-1}\,d\xi.$$

Указание. Ввести новые переменные ξ, η, ζ с помощью преобразования

$$\left.\begin{aligned}\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}&=\xi,\\ \frac{y^2}{b^2}+\frac{z^2}{c^2}&=\xi\eta,\\ \frac{z^2}{c^2}&=\xi\eta\zeta,\end{aligned}\right\}\quad\text{или}\quad\left\{\begin{aligned}x&=a\sqrt{\xi(1-\eta)},\\ y&=b\sqrt{\xi\eta(1-\zeta)},\\ z&=c\sqrt{\xi\eta\zeta}\end{aligned}\right.$$

и выполнить интегрирование по η и по ζ.

3. Найти абсциссу ξ центра массы тела

$$\left(\frac{x}{a}\right)^{\frac{1}{n}}+\left(\frac{y}{b}\right)^{\frac{1}{n}}+\left(\frac{z}{c}\right)^{\frac{1}{n}}\leqslant 1,\qquad x\geqslant 0,\quad y\geqslant 0,\quad z\geqslant 0.$$

4. Найти момент инерции площади, ограниченной астроидой $x^{\frac{2}{3}}+y^{\frac{2}{3}}=R^{\frac{2}{3}}$ относительно оси x.

5. Доказать, что

$$2^{2x}\frac{\Gamma(x)\,\Gamma\left(x+\frac{1}{2}\right)}{\Gamma(2x)}=2\sqrt{\pi}.$$

§ 7. Дифференцирование и интегрирование нецелого порядка. Интегральное уравнение Абеля.

Мы воспользуемся знакомством с гамма-функцией для того, чтобы построить простое обобщение понятий интегрирования и дифференцирования. На стр. 244 мы видели, что формула

$$F(x)=\int_0^x\frac{(x-t)^{n-1}}{(n-1)!}f(t)\,dt=\frac{1}{\Gamma(n)}\int_0^x(x-t)^{n-1}f(t)\,dt$$

дает n-кратный (повторный) интеграл от функции $f(x)$ по интервалу от 0 до x. (См. также т. I, стр. 266.) Обозначим символом D оператор дифференцирования и символом D^{-1} оператор $\int_0^x\ldots\,dx$, который представляет действие, обратное дифференцированию. Тогда можно записать символически

$$F(x)=D^{-n}f(x).$$

Математическое содержание этой формулы таково: функция $F(x)$ и ее первые $(n-1)$ производных обращаются в нуль при $x=0$, и к тому же n-я производная от $F(x)$ равна $f(x)$. Теперь представляется возможным построить весьма естественным путем такое определение оператора $D^{-\lambda}$, которое сохраняет смысл, когда положительное число λ не является целым. *Интеграл порядка* λ *от функции* $f(x)$ *по интервалу от* 0 *до* x *определяется выражением*

$$D^{-\lambda} f(x)=\frac{1}{\Gamma(\lambda)}\int_0^x (x-t)^{\lambda-1} f(t)\,dt.$$

С помощью этого определения можно теперь обобщить и действие дифференцирования n-го порядка, символизируемое оператором $D^n=\frac{d^n}{dx^n}$, и построить оператор дифференцирования μ-го порядка D^μ, где μ — любое неотрицательное число. Пусть m — наименьшее целое число, превосходящее μ, так что $\mu=m-\rho$, где $0<\rho\leqslant 1$. Тогда оператор D^μ определяется следующим выражением:

$$D^\mu f(x)=D^m D^{-\rho} f(x)=\frac{d^m}{dx^m}\frac{1}{\Gamma(\rho)}\int_0^x (x-t)^{\rho-1} f(t)\,dt.$$

Этому определению можно придать другой вид, поменявши порядок обоих действий (дифференцирования и интегрирования):

$$D^\mu f(x)=D^{-\rho} D^m f(x)=\frac{1}{\Gamma(\rho)}\int_0^x (x-t)^{\rho-1} f^{(m)}(t)\,dt.$$

Предоставим читателю доказать, с помощью формул для гамма-функции, что

$$D^\alpha D^\beta f(x)=D^\beta D^\alpha f(x),$$

каковы бы ни были действительные числа α и β. Надлежит показать, что это соотношение, а также и обобщенная операция дифференцирования имеют смысл, если функция $f(x)$ имеет обычные производные достаточно высокого порядка. В частности, $D^\mu f(x)$ существует, если $f(x)$ имеет непрерывные производные до m-го порядка включительно.

В тесной связи с этими понятиями находится *интегральное уравнение* Абеля, имеющее важные приложения. Так как $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$, то интеграл от функции $f(x)$ порядка $\frac{1}{2}$ выражается так:

$$D^{-\frac{1}{2}} f(x)=\frac{1}{\sqrt{\pi}}\int_0^x \frac{f(t)}{\sqrt{x-t}}\,dt=\psi(x).$$

Если $\psi(x)$ есть заданная функция, а функцию $f(x)$ требуется найти, то последнее соотношение есть интегральное уравнение Абеля. Пусть $\psi(x)$ — непрерывно дифференцируемая функция, удовлетворяющая условию $\psi(0)=0$. Тогда решение уравнения Абеля будет

$$f(x)=D^{\frac{1}{2}}D^{-\frac{1}{2}}f(x)=D^{\frac{1}{2}}\psi(x)$$

или

$$f(x)=\frac{1}{\sqrt{\pi}}\frac{d}{dx}\int\limits_0^x \frac{\psi(t)}{\sqrt{x-t}}\,dt.$$

§ 8. Замечание по поводу определения площади кривой поверхности

Определение площади кривой поверхности, данное в § 6 гл. IV, стр. 291, заметно отличается от определения длины дуги кривой, данного в т. I, гл. V, стр. 320. Определение длины дуги строилось на процессе вписывания ломаных, а в определении площади кривой поверхности использовались касательные плоскости вместо вписанных многогранников. [В этом параграфе мы для краткости вместо термина «многогранная поверхность» будем пользоваться словом «многогранник».]

Мы сейчас покажем, что определение площади поверхности невозможно строить на процессе вписывания многогранников. Для этой цели рассмотрим в пространстве xyz кусок цилиндрической поверхности $x^2+y^2=1$, ограниченный плоскостями $z=0$ и $z=1$. Боковая поверхность этого цилиндра равна 2π. В эту боковую поверхность мы теперь впишем многогранник, все грани которого — равные треугольники, следующим образом. Построим на поверхности цилиндра m равноотстоящих горизонтальных окружностей, лежащих в плоскостях $z=0$ (окружность основания), $z=h$, $z=2h, \ldots, z=(m-1)h$, где $h=1/m$. Каждую из этих m окружностей разделим на n равных дуг таким образом, чтобы точки деления каждой окружности (кроме окружности основания $z=0$) лежали на образующих цилиндра, проходящих через середины дуг соседней снизу окружности. Рассмотрим теперь вписанный в цилиндр многогранник, ребрами которого служат хорды наших окружностей и отрезки прямых, соединяющие ближайшие между собою точки соседних окружностей. Все грани этого многогранника — равные между собой равнобедренные треугольники, и если выбрать m и n достаточно большими, то многогранник будет лежать сколь угодно близко к боковой поверхности цилиндра. Дадим теперь числу n какое-либо постоянное значение; тогда можно число m выбрать столь большим, что каждая треугольная грань будет сколь угодно близка к плоскости, параллельной основанию цилиндра, а следовательно, образует с его боковой поверхностью угол, как угодно близкий к прямому. Из этого видно, что нельзя ожидать, чтобы площадь нашего вписанного многогранника давала приближение к боковой поверхности цилиндра. Нетрудно показать, что предел этой площади зависит от способа стремления чисел m и n к бесконечности.

Действительно, основание каждого равнобедренного треугольника $a=2\sin\frac{\pi}{n}$, высоту же его h можно вычислить по теореме Пифагора:

$$h=\sqrt{\frac{1}{m^2}+\left(1-\cos\frac{\pi}{n}\right)^2} = \sqrt{\frac{1}{m^2}+4\sin^4\frac{\pi}{2n}}.$$

Так как число треугольных граней равно, очевидно, $2mn$, то площадь многогранника

$$F_{n,\,m}=2mn\sin\frac{\pi}{n}\cdot\sqrt{\frac{1}{m^2}+4\sin^4\frac{\pi}{2n}}=2n\sin\frac{\pi}{n}\cdot\sqrt{1+4m^2\sin^4\frac{\pi}{2n}}.$$

Ясно, что предел этого выражения зависит от того, каким образом мы заставим стремиться к бесконечности числа m и n. Если, например, фиксировать значение n и выполнить предельный переход $m\to\infty$, то площадь $F_{n,m}\to\infty$. Если m и n одновременно стремятся к бесконечности, притом так, что $m=n$, то $F_{n,m}\to 2\pi$. Если положить $m=n^2$, то $\lim\limits_{n\to\infty} F_{n,\,m}=2\pi\sqrt{1+\frac{\pi^4}{4}}$ и т. д. Однако, из того, что $F_{n,m}\geqslant 2n\sin\frac{\pi}{n}$, а $\lim\limits_{n\to\infty} 2n\sin\frac{\pi}{n}=2\pi$, видно, что нижний предел (нижняя точка сгущения) числового множества $F_{n,m}$ равен 2π.

В заключение сообщим без доказательства следующий факт, касающийся определения площади поверхности. Этот факт представляет принципиальный интерес и приведенный выше пример можно рассматривать как его иллюстрацию. Если имеется последовательность многогранников, вписанных в данный кусок поверхности и неограниченно к нему приближающихся, то, как мы видели, площадь многогранника может и не сходиться к площади этого куска поверхности. Однако, предел последовательности площадей многогранников (если он существует) или, общée, всякая точка сгущения этой последовательности площадей всегда больше или по крайней мере равна площади куска кривой поверхности. Если для всех таких последовательностей многогранников найдем нижние пределы их площадей, то эти нижние пределы составляют определенное числовое множество, связанное с изучаемым куском кривой поверхности. Так вот *площадь куска поверхности можно определить как нижний предел* (нижнюю точку сгущения) *этого множества.*

Это замечательное свойство площади поверхности называется *полунепрерывностью*, точнее, *полунепрерывностью снизу.*

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ IV

1. Пусть $f(x)=\int\limits_0^{\frac{\pi}{2}}\ln(1-x^2\cos^2\theta)\,d\theta$. Доказать, что а) если $x^2\leqslant 1$, то $f(x)$ конечна; б) если $x^2<1$, то

$$f'(x)=\int\limits_0^{\frac{\pi}{2}}\frac{d}{dx}\ln(1-x^2\cos^2\theta)\,d\theta.$$

Пользуясь этим, вычислить интеграл.

2. Показать, что площадь $\sum$ поверхности прямого геликоида

$$x=r\cos\theta,\qquad y=r\sin\theta,\qquad z=f(\theta),$$

заключенной между двумя плоскостями, проходящими через ось z, и цилиндрической поверхностью, уравнение которой в цилиндрических координатах есть $r=f'(\theta)$, относится к своей ортогональной проекции на плоскость xy как

$$[\sqrt{2}+\ln(1+\sqrt{2})]:1.$$

3. Допустим, что земля есть шар радиуса R, объемная плотность которого зависит только от расстояния r от центра и выражается функцией вида

$$\mu=A-Br^2,$$

причем у поверхности плотность равна 2,5 $г/см^3$, а средняя плотность земли равна 5,5 $г/см^3$. Показать, что при этих допущениях напряженность силы тяжести в любой внутренней точке на расстоянии r от центра есть

$$g_1=\frac{g}{11}\,\frac{r^3}{R^3}\left(20-g\frac{r^2}{R^2}\right),$$

где g есть напряженность силы тяжести на поверхности.

4. Вершины треугольника площади S находятся в точках (x_1, y_1), (x_2, y_2), (x_3, y_3), причем порядок индексов соответствует положительному направлению обхода. Доказать, что момент инерции площади треугольника относительно оси x равен

$$\frac{S}{6}(y_1^2+y_2^2+y_3^2+y_1y_2+y_2y_3+y_3y_1).$$

5. Полушар радиуса a и постоянной объемной плотности ρ расположен так, что его центр находится в начале координат, а круг, его ограничивающий, лежит в плоскости xy. Показать, что потенциал в точке $(0, 0, z)$ равен

$$\frac{2\pi\rho}{3z}\left[(a^2+z^2)^{\frac{3}{2}}-a^3+\frac{3}{2}a^2z\right]-\frac{4}{3}\pi\rho z^2\qquad \text{если } 0<z<a$$

и

$$\frac{2\pi\rho}{z}\left[(a^2+z^2)^{\frac{3}{2}}+a^3-\frac{3}{2}a^2z\right]-\frac{2}{3}\pi\rho z^2,\qquad \text{если } z>a.$$

6. Построить кривую

$$x=\frac{1-\lambda^2}{1+\lambda^2},\qquad y=\frac{\lambda-\lambda^3}{1+\lambda^2}\qquad(-1\leqslant\lambda\leqslant+1)$$

и вычислить ограниченную ею площадь.

7. Сфероид с полуосями a, a, c покрыт веществом с постоянной поверхностной плотностью μ. Показать, что сила притяжения на каждом полюсе равна

$$2\pi\mu \int_0^{2c} r\,(1-\cos\theta)\,dr,$$

где

$$r = \frac{2a^2 c\cos\theta}{a^2\cos^2\theta + c^2\sin^2\theta}.$$

8. В интеграле

$$I = \int_2^4 dx \int_{4/x}^{\frac{20-4x}{8-x}} (y-4)\,dy$$

изменить порядок интегрирования, а затем вычислить интеграл.

9. а) Вычислить интеграл

$$K = \int_0^{a\sin\beta} dy \int_{y\operatorname{ctg}\beta}^{\sqrt{a^2-y^2}} \ln(x^2+y^2)\,dx \qquad \left(0<\beta<\frac{\pi}{2}\right)$$

преобразованием к полярным координатам.

б) В исходном интеграле изменить порядок интегрирования.

10. Найти объем V тела, вырезаемого из прямого конуса

$$x^2+y^2 = (h-z)^2\operatorname{tg}^2\alpha \qquad (0\leqslant z\leqslant h)$$

вертикальной цилиндрической поверхностью, направляющая которой в плоскости xy имеет в полярных координатах r, θ уравнение

$$(h\operatorname{tg}\alpha - r)^3 = h^3\operatorname{tg}^3\alpha\,\sin^4\theta\cos^2\theta.$$

Имеется в виду объем, расположенный вне цилиндра и внутри конуса.

11. Показать, что интеграл

$$I = \iint \frac{z_{xx}z_{yy} - z_{xy}^2}{(1+z_x^2+z_y^2)^2}\,d\sigma,$$

взятый по куску поверхности гиперболического параболоида $z = xy$, ограниченному двумя прямолинейными образующими, проходящими через начало координат, и двумя, проходящими через его точку (ξ, η, ζ), равен

$$-\operatorname{arctg}\frac{\xi^2\eta^2}{\sqrt{\eta^2\zeta^2+\zeta^2\xi^2+\xi^2\eta^2}}.$$

12. Доказать, что если $-1<a<1$, то

$$K(a) = \int_0^{\pi} \frac{\ln(1+a\cos x)}{\cos x}\,dx = \pi\arcsin a.$$

13. Показать, что площадь, ограниченная кривыми

$x^3 = a^2 y$, $x^3 = b^2 y$, $y^3 = \alpha x$ и $y^3 = \beta x$, где $b>a>0$ и $\beta>\alpha>0$, равна

$$\frac{5}{12}\left(b^{\frac{6}{5}} - a^{\frac{6}{5}}\right)\left(\beta^{\frac{4}{5}} - \alpha^{\frac{4}{5}}\right).$$

ГЛАВА V

КРИВОЛИНЕЙНЫЕ ИНТЕГРАЛЫ. ИНТЕГРАЛЫ ПО ПОВЕРХНОСТИ

Кратные интегралы, которыми мы занимались в предшествующей главе, представляют собой не единственное возможное обобщение понятия интеграла на функции многих переменных. Напротив того, существуют еще и другие обобщения, в соответствии с тем фактом, что в многомерные области можно вложить многообразия меньшего числа измерений, а затем рассматривать интегралы по таким многообразиям. У функции двух независимых переменных мы будем рассматривать, наряду с интегралом по двумерной области (двойным интегралом), интеграл вдоль кривой, т. е. по одномерному многообразию. Функции трех переменных мы будем интегрировать не только по трехмерным областям (тройные интегралы) и вдоль кривых (криволинейные интегралы), но и по поверхностям, т. е. по двумерным многообразиям, вложенным в трехмерное пространство. Эти понятия: криволинейного интеграла, поверхностного интеграла и т. д. (все они непосредственно связаны с практикой) — мы введем в этой главе и исследуем их взаимные связи.

§ 1. Криволинейные интегралы

Определение простого интеграла мы связали с наглядным представлением *площади*, а естественное обобщение на функции двух и трех переменных привело нас к двойному и тройному интегралу. С другой стороны, и физическое понятие *работы* нас тоже привело к простому, одномерному интегралу. Если мы пожелаем дать математическое определение работы в произвольном пространственном силовом поле, то в качестве нового обобщения исходного понятия интеграла от функции одной переменной получим *криволинейный интеграл.*

1. Определение криволинейного интеграла. Обозначения. Начнем с математического определения криволинейного интеграла для трехмерного пространства xyz. Пусть в этом пространстве задана кусочно гладкая[1]) кривая своими параметрическими уравнениями

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

---

[1]) Напомним (ср. стр. 56), что кривая называется *кусочно гладкой*, если она состоит из конечного числа дуг, каждая из которых имеет касательную, изменяющуюся непрерывно вдоль всей дуги, включая ее конечные точки.

где $x(t)$, $y(t)$, $z(t)$ — непрерывные функции с кусочно непрерывными первыми производными. Рассмотрим дугу $\overset{\frown}{P_0P}$ этой кривой, ограниченную точками $P_0(x_0, y_0, z_0)$ и $P(\bar{x}, \bar{y}, \bar{z})$. Пусть точки этой дуги, которую мы обозначим через C, соответствуют значениям параметра t из интервала $\alpha \leqslant t \leqslant \beta$. Если в какой-либо области, содержащей эту дугу, определена непрерывная функция точки $f(x, y, z)$, то вдоль дуги C кривой эта функция обращается в функцию $f[x(t), y(t), z(t)]$, зависящую только от параметра t. Для того чтобы, по аналогии с обычным интегралом, определить криволинейный интеграл этой функции вдоль дуги C, разобьем эту дугу с помощью точек P_0, P_1, $P_2, \ldots, P_n$ $(P_n = P)$ на мелкие дужки и обозначим разность абсцисс точек P_{i+1} и P_i через $\Delta x_i = x_{i+1} - x_i$. Построим теперь сумму

$$\sum_{i=0}^{n-1} f[x(t_i), y(t_i), z(t_i)]\, \Delta x_i,$$

где через t_i обозначено любое число, выбранное в интервале изменения параметра t, соответствующем дужке $\overset{\frown}{P_iP_{i+1}}$, а затем заставим число n точек деления безгранично возрастать таким образом, чтобы длина наибольшей из дужек $\overset{\frown}{P_iP_{i+1}}$ стремилась к нулю. Тогда мы вправе ожидать, что построенная нами сумма будет стремиться к определенному пределу. Этот предел мы обозначим символом

$$\int_C f(x, y, z)\, dx$$

и будем называть *криволинейным интегралом функции* $f(x, y, z)$ *вдоль дуги C нашей кривой*. Дуга C называется *путем интегрирования*. Что этот предел действительно существует и притом не зависит от выбора промежуточных точек деления, можно доказать тем же методом, которым было доказано существование обычного интеграла. Но это предложение доказывается совсем просто, если переписать нашу сумму в следующем виде:

$$\sum_{i=0}^{n-1} f[x(t_i), y(t_i), z(t_i)]\, \frac{\Delta x_i}{\Delta t_i}\, \Delta t_i,$$

где Δt_i обозначает приращение параметра t при переходе от точки P_i к точке P_{i+1}. Теперь можно опираться на теорему существования обычного определенного интеграла: при нашем предельном переходе эта сумма имеет своим пределом определенный интеграл

$$\int_\alpha^\beta f[x(t), y(t), z(t)]\, \frac{dx}{dt}\, dt,$$

где α и β — значения параметра t, соответствующие начальной и конечной точкам дуги C.

Заодно мы получили формулу

$$\int_C f(x, y, z)\,dx = \int_\alpha^\beta f(x, y, z)\frac{dx}{dt}\,dt,$$

которая приводит криволинейный интеграл к обычному определенному интегралу от функции параметра t по этому параметру.

Ясно, что обычный интеграл есть частный случай криволинейного интеграла, когда в качестве пути интегрирования берется отрезок оси x.

Следуя выполненному выше образцу, нетрудно теперь определить и следующие криволинейные интегралы:

$$\int_C f(x, y, z)\,dy = \int_\alpha^\beta f(x, y, z)\frac{dy}{dt}\,dt$$

и

$$\int_C f(x, y, z)\,dz = \int_\alpha^\beta f(x, y, z)\frac{dz}{dt}\,dt.$$

Три полученные формулы, выражающие криволинейные интегралы эффективно через определенные интегралы, позволяют легко доказать, что *криволинейные интегралы зависят только от выбора пути интегрирования C и совершенно не зависят от способа его задания, т. е. от выбора параметра t*. Действительно, если ввести новый параметр $\tau = \varphi(t)$, где $\varphi(t)$ — непрерывно дифференцируемая функция, причем $\varphi'(t) > 0$ в промежутке интегрирования $\alpha \leqslant t \leqslant \beta$, то этот последний интервал преобразуется взаимно однозначно на интервал $\alpha_1 \leqslant \tau \leqslant \beta_1$ и

$$\int_\alpha^\beta f(x, y, z)\frac{dx}{dt}\,dt = \int_{\alpha_1}^{\beta_1} f(x, y, z)\frac{dx}{dt}\frac{dt}{d\tau}\,d\tau = \int_{\alpha_1}^{\beta_1} f(x, y, z)\frac{dx}{d\tau}\,d\tau.$$

2. Векторная запись криволинейного интеграла. В приложениях криволинейные интегралы обычно возникают в следующей стандартной комбинации:

$$\int_C F_1(x, y, z)\,dx + \int_C F_2(x, y, z)\,dy + \int_C F_3(x, y, z)\,dz,$$

где $F_1(x, y, z)$, $F_2(x, y, z)$, $F_3(x, y, z)$ — три функции точки, непрерывные в некоторой области, содержащей дугу C. Эту комбинацию можно привести к следующему виду:

$$\int_C (F_1\,dx + F_2\,dy + F_3\,dz) = \int_\alpha^\beta (F_1\dot{x} + F_2\dot{y} + F_3\dot{z})\,dt,$$

где, как обычно, $\dot{x}=\frac{dx}{dt}$ и т. д. Обозначим через $\boldsymbol{r}$ переменный радиус-вектор (вектор положения) точки области, $\boldsymbol{r}=\{x, y, z\}$. Функции F_1, F_2, F_3 можно рассматривать как прямоугольные координаты вектора $\boldsymbol{F}(x, y, z)$, точнее, вектор-функции точки

$$\boldsymbol{F}=\boldsymbol{F}(\boldsymbol{r})=\{F_1, F_2, F_3\}.$$

Величины $\dot{x}, \dot{y}, \dot{z}$ являются координатами вектора $\dot{\boldsymbol{r}}=\frac{d\boldsymbol{r}}{dt}$ (производной от радиус-вектора переменной точки дуги), и подынтегральное выражение можно записать как скалярное произведение $\boldsymbol{F}\,d\boldsymbol{r}=\boldsymbol{F}\dot{\boldsymbol{r}}\,dt$. Таким образом, комбинированный криволинейный интеграл приводится к следующему виду:

$$\int_C (F_1\,dx+F_2\,dy+F_3\,dz)=\int_C \boldsymbol{F}\,d\boldsymbol{r}=\int_\alpha^\beta \boldsymbol{F}\dot{\boldsymbol{r}}\,dt.$$

Такой интеграл мы будем называть криволинейным интегралом переменного вектора $\boldsymbol{F}$ вдоль дуги C.

[С этой точки зрения, если криволинейный интеграл содержит лишь один член:

$$\int_C f(x, y, z)\,dx, \quad \int_C f(x, y, z)\,dy \quad \text{или} \quad \int_C f(x, y, z)\,dz,$$

то это означает, что подлежащая интегрированию вектор-функция имеет постоянное направление, параллельное одной из осей координат.]

Криволинейные интегралы на плоскости

$$\int_C f(x, y)\,dx, \quad \int_C f(x, y)\,dy, \quad \int_C \{F_1(x, y)\,dx+F_2(x, y)\,dy\}$$

можно ввести как частный случай криволинейных интегралов в пространстве.

Далее, понятие криволинейного интеграла можно распространить и на функции от n переменных. В этом общем случае мы предполагаем, что заданы n непрерывных функций параметра t:

$$x_1=x_1(t), \quad x_2=x_2(t), \ldots, x_n=x_n(t),$$

имеющих кусочно непрерывные производные в интервале $\alpha \leqslant t \leqslant \beta$. Эти n уравнений определяют криволинейную дугу в n-мерном пространстве. *Криволинейный интеграл мы определяем тогда выражением*

$$\int_C f(x_1, x_2, \ldots, x_n)\,dx_i=\int_\alpha^\beta f(x_1(t), x_2(t), \ldots, x_n(t))\frac{dx_i}{dt}\,dt$$
$$(i=1, 2, \ldots, n).$$

Так, в n-мерном пространстве проще всего определить криволинейный интеграл сразу как некоторый обычный определенный интеграл.

И здесь можно рассматривать n функций $F_1, F_2, \ldots, F_n$ от n переменных $x_1, x_2, \ldots, x_n$ и составить криволинейный интеграл общего вида (в координатной и в векторной записи):

$$\int\limits_C (F_1\,dx_1 + F_2\,dx_2 + \ldots + F_n\,dx_n) = \int\limits_C \boldsymbol{F}\,d\boldsymbol{x} = \int\limits_\alpha^\beta \boldsymbol{F}\dot{\boldsymbol{x}}\,dt,$$

где радиус-вектор (вектор положения) теперь удобно обозначить полужирной буквой $\boldsymbol{x}$, $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$, а $\boldsymbol{F} = \{F_1, F_2, \ldots, F_n\}$ есть вектор-функция точки.

С примером криволинейного интеграла (на плоскости) мы уже фактически встречались при вычислении площади фигуры, ограниченной замкнутой кривой C (т. I, гл. V, стр. 317). Если замкнутая кусочно гладкая кривая C плоскости xy, сама себя не пересекающая, задана параметрическими уравнениями $x = x(t)$, $y = y(t)$, то площадь S, ею ограниченная, выражается так:

$$S = -\int\limits_\alpha^\beta y\dot{x}\,dt = \int\limits_\alpha^\beta x\dot{y}\,dt = \frac{1}{2}\int\limits_\alpha^\beta (x\dot{y} - y\dot{x})\,dt.$$

В нашей нынешней терминологии это просто криволинейные интегралы

$$S = -\oint\limits_C y\,dx = \oint\limits_C x\,dy = \frac{1}{2}\oint\limits_C (-y\,dx + x\,dy),$$

взятые вдоль замкнутой кривой C в направлении возрастания параметра t. Колечком, надетым на символ интеграла, отмечено, что интеграл берется вдоль *замкнутой* кривой, а стрелкой указано направление ее обхода.

3. Основные свойства. Нетрудно вывести некоторые свойства криволинейного интеграла, пользуясь его выражением через определенный интеграл.

1) *Значение криволинейного интеграла зависит от направления, в котором описывается кривая C, и умножается на -1, если эта же кривая пробегается в противоположном направлении*, т. е. от P к P_0 вместо прежнего направления от P_0 к P. Доказательство этого свойства очевидно. В силу этого свойства целесообразно всегда считать, что символ C обозначает кривую, снабженную определенным направлением или, как принято говорить, *ориентированную* кривую (ср. т. I, гл. V, § 2, стр. 311—312). Если определенная ориентированная кривая обозначена через C, то эту же кривую, но описываемую в противоположном направлении, мы будем обозначать символом $-C$.

2) Если путь интегрирования C образован путем соединения двух дуг C_1 и C_2, пробегаемых последовательно одна за другой (это записывается символически: $C = C_1 + C_2$), то соответствующие криволинейные интегралы связаны соотношением

$$\int\limits_C = \int\limits_{C_1} + \int\limits_{C_2}.$$

3) Особенно важна следующая *теорема разбиения*, относящаяся к интегралу по *замкнутой* кривой. Возьмем сначала случай двух переменных x и y и рассмотрим криволинейный интеграл

$$\oint_C (F_1\,dx + F_2\,dy),$$

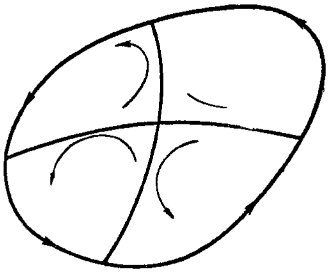


Рис. 82.

взятый по ориентированной замкнутой кривой C (рис. 82), причем предполагаем, что векторное поле $\boldsymbol{F} = \{F_1, F_2\}$ определено и непрерывно в некоторой области, содержащей кривую C. Разобьем область G, ограниченную кривой C, дугами кривых на n частичных областей $G_1, G_2, \ldots, G_n$, ограниченных замкнутыми кривыми $C_1, C_2, \ldots$ $\ldots, C_n$ с тем же направлением обхода, что и у кривой C. Тогда

$$\oint_C (F_1\,dx + F_2\,dy) = \oint_{C_1} (F_1\,dx + F_2\,dy) + \\ + \oint_{C_2} (F_1\,dx + F_2\,dy) + \ldots + \oint_{C_n} (F_1\,dx + F_2\,dy)$$

при любом разбиении описанного выше типа. Еще раз подчеркиваем, что как кривая C, так и границы всех частичных областей следует обходить в одном и том же направлении.

Для доказательства заметим, что все те дуги кривых $C_1, C_2, \ldots, C_n$, которые лежат внутри данной кривой C, служат общими границами двух соседних частичных областей. Поэтому при сложении интегралов в правой части доказываемого равенства каждая из этих дуг обходится два раза в двух противоположных направлениях, так что соответствующие этим дугам интегралы взаимно сокращаются. Таким образом, остаются лишь интегралы по тем дугам кривых $C_1, C_2, \ldots, C_n$, которые лежат и на кривой C, но эти интегралы складываются в должном направлении в интеграл по замкнутой кривой C. Теорема доказана.

Эта теорема легко обобщается на пространство трех (и большего числа) измерений следующим образом. Представим себе, что на пространственную замкнутую кривую C, не имеющую самопересечений, «натянута» какая-нибудь поверхность. Разобьем кусок поверхности, ограниченный кривой C, на n частей, ограниченных замкнутыми кривыми $C_1, C_2, \ldots, C_n$. Все эти кривые снабдим одинаковой ориентацией. Тогда для соответствующих криволинейных интегралов справедлива формула

$$\oint_C \boldsymbol{F}\,d\boldsymbol{x} = \oint_{C_1} \boldsymbol{F}\,d\boldsymbol{x} + \oint_{C_2} \boldsymbol{F}\,d\boldsymbol{x} + \ldots + \oint_{C_n} \boldsymbol{F}\,d\boldsymbol{x}.$$

4) Несколько иное применение этого принципа разбиения представляет следующая теорема. На две одинаково ориентированные замкнутые кривые C и C' (рис. 83) нанесены соответственно точки $A_1, A_2, \ldots, A_n$ и $B_1, B_2, \ldots, B_n$, занумерованные в порядке ориентации кривых. Каждая пара точек с одинаковым номером соединена какой-нибудь линией. Обозначим через C_i замкнутую ориентированную кривую $A_iA_{i+1}B_{i+1}B_i$; тогда

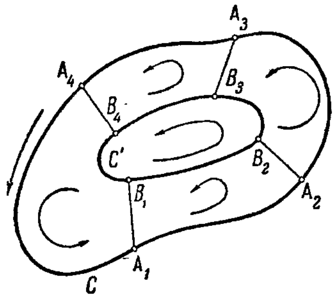


Рис. 83.

$$\sum_{i=1}^{n} \oint_{C_i} F\,dx = \oint_{C} F\,dx - \oint_{C'} F\,dx.$$

Доказывается эта теорема, как это ясно из рис. 83, тем же методом, что и предыдущая. Кстати, для ее справедливости не обязательно, чтобы кривые C и C' не пересекали сами себя или друг друга.

5) В заключение приведем *формулу оценки криволинейного интеграла*

$$\left|\int_C \boldsymbol{F}\,d\boldsymbol{r}\right| = \left|\int_C (F_1\,dx + F_2\,dy + F_3\,dz)\right| \leqslant ML,$$

где M — верхняя граница модуля вектора $\boldsymbol{F}$ на пути интегрирования C:

$$|\boldsymbol{F}| = \sqrt{F_1^2 + F_2^2 + F_3^2} \leqslant M,$$

а L — длина дуги C.

Доказательство легко получается из очевидного неравенства

$$\left|\boldsymbol{F}\frac{d\boldsymbol{r}}{dt}\right| \leqslant |\boldsymbol{F}|\left|\frac{d\boldsymbol{r}}{dt}\right| = |\boldsymbol{F}|\frac{ds}{dt},$$

где $\frac{ds}{dt}$ — производная от длины дуги линии C по параметру.

4. Механическое истолкование криволинейного интеграла. Выше уже было сказано, что понятие криволинейного интеграла тесно связано с физическим понятием *работы*. Если материальная точка движется под действием силового поля $\boldsymbol{F} = \{F_1, F_2, F_3\}$, изменяющегося, вообще говоря, от точки к точке, и описывает при этом кривую C, то криволинейный интеграл $\int_C \boldsymbol{F}\,d\boldsymbol{r}$ выражает работу, затраченную полем при этом движении. Действительно, если сила постоянна по величине и направлению, а движение прямолинейно, то работа, по определению, равна скалярному произведению вектора-силы на вектор-перемещение. Для того чтобы естественным образом обобщить это определение, заменим путь C вписанной в него ломаной линией с вершинами $P_0, P_1, P_2, \ldots, P_n = P$, а действующую фактически переменную силу $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{r})$ — воображаемой силой, которая сохраняет постоянную величину и постоянное направление на каждом звене $\overline{P_iP_{i+1}}$ ломаной и равна вектору-силе, действующей в начальной точке P_i. Работа этой воображаемой силы на перемещении $\overline{P_iP_{i+1}}$ равна

$$\boldsymbol{F}(\boldsymbol{r}_i)\Delta\boldsymbol{r}_i = f_1(x_i, y_i, z_i)\Delta x_i + f_2(x_i, y_i, z_i)\Delta y_i + f_3(x_i, y_i, z_i)\Delta z_i.$$

Суммируя эти выражения по всем звеньям ломаной и переходя к пределу $n \to \infty$, при условии, что длина наибольшего звена ломаной стремится к нулю, мы и получим криволинейный интеграл $\int\limits_C \boldsymbol{F}\,d\boldsymbol{r}$, который, стало быть, действительно дает работу силового поля при указанном движении материальной точки.

Впоследствии (§ 3, стр. 393) будут даны еще и другие физические истолкования криволинейного интеграла.

5. Криволинейный интеграл в поле градиента. Интегрирование полного дифференциала. Особенно важен тот случай, когда вектор поля $\boldsymbol{F} = \{F_1, F_2, F_3\}$ является градиентом некоторой скалярной функции точки $U(x, y, z)$, т. е. существует такая функция $U(x, y, z)$, что

$$\boldsymbol{F} = \operatorname{grad} U \quad \text{или} \quad F_1 = U_x, \quad F_2 = U_y, \quad F_3 = U_z.$$

Если $\boldsymbol{F} = \operatorname{grad} U$, то скалярную функцию U часто называют *потенциальной функцией* или, короче, *потенциалом* векторного поля $\boldsymbol{F}$. [Физики называют обычно функцию U силовой функцией, а потенциалом V они называют силовую функцию с обратным знаком: $V = -U$.]

Значение криволинейного интеграла в векторном поле зависит вообще не только от положения начальной и конечной точек пути интегрирования C, но и от всего хода кривой C. Однако если вектор поля является градиентом, $\boldsymbol{F} = \operatorname{grad} U$, то имеет место следующая т е о р е м а:

Криволинейный интеграл в поле градиента равен приращению потенциальной функции при переходе от начальной до конечной точки пути интегрирования C и не зависит от выбора кривой C, соединяющей эти точки. Это значит, что мы получим одно и то же значение криволинейного интеграла при любом выборе пути интегрирования C, соединяющего заданные начальную и конечную точки, пока кривая C остается полностью в области определения потенциальной функции $U(x, y, z)$.

Действительно, если $F = \operatorname{grad} U$, то

$$\int\limits_C \boldsymbol{F}\,d\boldsymbol{r} = \int\limits_C \operatorname{grad} U\,d\boldsymbol{r} = \int\limits_C (U_x\,dx + U_y\,dy + U_z\,dz) =$$
$$= \int\limits_\alpha^\beta (U_x\dot{x} + U_y\dot{y} + U_z\dot{z})\,dt = \int\limits_\alpha^\beta \frac{dU}{dt}\,dt.$$

Стало быть, можно выполнить интегрирование в явном виде, и в итоге получится:

$$\int\limits_C \boldsymbol{F}\,d\boldsymbol{r} = U(x(\beta), y(\beta), z(\beta)) - U(x(\alpha), y(\alpha), z(\alpha)),$$

что и требовалось доказать.

Это относится, например, к полю тяготения, порождаемому одной единичной точечной массой; мы уже знаем (гл, II, § 7, стр. 112), что напряженность этого поля является градиентом потенциала $1/r$. Стало быть, работа силы тяготения при перемещении материальной точки из данного начального в данное конечное положение не зависит от пути перемещения.

Подынтегральное выражение криволинейного интеграла является в этом случае *полным дифференциалом* функции $U(x, y, z)$ (гл. II, § 4, n° 4):

$$\boldsymbol{F}\,d\boldsymbol{r} = F_1\,dx + F_2\,dy + F_3\,dz = U_x\,dx + U_y\,dy + U_z\,dz = dU.$$

Поэтому результат интегрирования можно записать в следующем виде:

$$\int\limits_C dU = U(x(\beta), y(\beta), z(\beta)) - U(x(\alpha), y(\alpha), z(\alpha)),$$

и естественно толковать его как интегрирование полного дифференциала $U_x\,dx + U_y\,dy + U_z\,dz$.

Крайне важно понять и усвоить, что утверждение «*криволинейный интеграл не зависит от пути*» равносильно утверждению «*криволинейный интеграл по любой замкнутой кривой равен нулю*». Действительно, если разбить замкнутую кривую (рис. 84) какими-либо двумя ее точками A и B на две дуги AmB и AnB, то из равенства нулю криволинейного интеграла по замкнутой кривой $AmBnA$ следует равенство криволинейных интегралов от точки A до точки B по двум различным путям AmB и AnB. Обратно, если интегралы от A до B по двум различным кривым AmB и AnB равны между собой, то отсюда вытекает обращение в нуль суммы $\int\limits_{AmB} + \int\limits_{BnA}$, а эта сумма и есть интеграл, взятый по всей замкнутой кривой $AmBnA$.

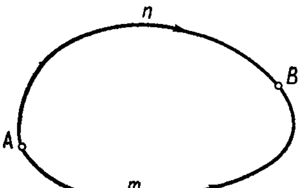


Рис. 84.

6. Условие независимости криволинейного интеграла от пути интегрирования. Как мы уже подчеркивали, независимость криволинейного интеграла от пути интегрирования или равносильное этому обращение в нуль интеграла по замкнутой кривой представляет собой исключительное свойство, характеризующее специальный класс векторных полей. Если, например, ориентированная замкнутая кривая C ограничивает на плоскости xy положительную площадь, то, согласно n°2, криволинейный интеграл $\oint x\,dy$ или $\oint (x\,dy - y\,dx)$ по этой кривой C не равен нулю. Основная задача теории криволинейных интегралов и состоит в том, чтобы показать, что данное в предыдущем номере *достаточное* условие независимости от пути является также и *необходимым*, а затем привести это необходимое и достаточное условие к виду, удобному для практического применения.

Этот вопрос о независимости криволинейного интеграла от пути интегрирования мы сначала исследуем для плоского векторного поля, но уже сейчас заметим, что в случае трех и большего числа измерений результаты получаются совершенно аналогичные.

При этом мы будем предполагать, что в плоском векторном поле $\boldsymbol{F} = \{F_1(x, y), F_2(x, y)\}$ функции $F_1(x, y)$ и $F_2(x, y)$ непрерывны и имеют непрерывные частные производные $\frac{\partial F_1}{\partial y}$ и $\frac{\partial F_2}{\partial x}$ в некоторой области G плоскости xy. Тогда справедлива следующая теорема:

Криволинейный интеграл $\int\limits_C \boldsymbol{F}\,d\boldsymbol{r} = \int\limits_C (F_1\,dx + F_2\,dy)$, *взятый вдоль линии C, лежащей в области G, не зависит от выбора этого пути интегрирования C и вполне определяется положением начальной и конечной точек этого пути в том и только в том случае, если выражение $F_1\,dx + F_2\,dy$ является полным дифференциалом некоторой функции $U(x, y)$, т. е. если в G существует такая функция $U(x, y)$, что*

$$\frac{\partial U}{\partial x} = F_1(x, y), \quad \frac{\partial U}{\partial y} = F_2(x, y) \quad \text{или} \quad \boldsymbol{F} = \operatorname{grad} U \tag{*}$$

всюду в области G.

В предшествующем номере мы уже доказали достаточность этого условия, т. е. доказали, что при его выполнении интеграл действительно не зависит от пути интегрирования.

Остается доказать *необходимость* условия (*). Для этого заметим, что если интеграл не зависит от выбора кривой C, то, при фиксированной начальной точке P_0 этой кривой, интеграл является (однозначной) функцией координат ее конечной точки $P(\xi, \eta)$. Обозначим эту функцию точки P через $U(\xi, \eta)$. Функция $U(\xi, \eta)$, очевидно, дифференцируема по ξ и η, причем во всякой внутренней точке области G

$$U_\xi(\xi, \eta) = \lim_{h\to 0} \frac{1}{h}\{U(\xi + h, \eta) - U(\xi, \eta)\} =$$
$$= \lim_{h\to 0} \frac{1}{h}\left\{\int\limits_{C+C_h} \boldsymbol{F}\,d\boldsymbol{r} - \int\limits_C \boldsymbol{F}\,d\boldsymbol{r}\right\} = \lim_{h\to 0} \frac{1}{h}\int\limits_{C_h} \boldsymbol{F}\,d\boldsymbol{r} =$$
$$= \lim_{h\to 0} \frac{1}{h}\int\limits_{C_h} (F_1\,dx + F_2\,dy);$$

здесь C обозначает *любую* кусочно гладкую кривую, идущую в области G от точки P_0 до точки $P(\xi, \eta)$, а C_h — любую кусочно гладкую кривую области G, идущую от точки P до точки $P_1(\xi+h, \eta)$. Так как при достаточно малом h прямолинейный отрезок PP_1 лежит в G, то в качестве пути интегрирования C_h можно выбрать этот отрезок PP_1, выражаемый параметрическими уравнениями

$$x=t, \quad y=\eta=\text{const}, \quad \xi \leqslant t \leqslant \xi+h.$$

Следовательно,

$$U_\xi(\xi, \eta)=\lim_{h \to 0} \frac{1}{h} \int_{\xi}^{\xi+h} F_1(t, \eta)\, dt=F_1(\xi, \eta).$$

Таким же способом получается

$$U_\eta(\xi, \eta)=\lim_{h \to 0} \frac{1}{h} \int_{\eta}^{\eta+h} F_2(\xi, t)\, dt=F_2(\xi, \eta).$$

Действительно, стало быть, $U_x(x, y)=F_1$ и $U_y(x, y)=F_2$ и необходимость нашего условия доказана. Хотя этот результат доказан пока только для внутренних точек области G, но, в силу непрерывности всех участвующих в этом вычислении функций, он справедлив и для всех точек границы области G.

7. Условие, при котором вектор поля является градиентом — условие интегрируемости выражения $F_1\,dx+F_2\,dy$. Нельзя, однако, считать, что только что доказанная теорема решает проблему до конца, так как мы еще не располагаем общим способом, позволяющим распознать, является ли вектор $\boldsymbol{F}$ поля градиентом или нет. Мы установим теперь критерий, дающий возможность решить этот вопрос по виду функций $F_1(x, y)$ и $F_2(x, y)$.

Если область G односвязна, то выполнение соотношения

$$\frac{\partial F_1}{\partial y}=\frac{\partial F_2}{\partial x} \tag{1}$$

во всех точках области G является необходимым и достаточным условием того, чтобы вектор поля $\boldsymbol{F}=\{F_1, F_2\}$ был градиентом некоторой скалярной функции $U(x, y)$ и чтобы, следовательно, выражение

$$F_1\,dx+F_2\,dy$$

было полным дифференциалом этой функции. Напомним, что мы заранее предположили, что $\frac{\partial F_1}{\partial y}$ и $\frac{\partial F_2}{\partial x}$ непрерывны во всей области G.

Соотношение (1) принято называть *условием* или *критерием интегрируемости* дифференциального выражения

$$\boldsymbol{F}\,d\boldsymbol{r}=F_1\,dx+F_2\,dy.$$

Необходимость условия (1) вытекает из следующих соображений. Если $F = \operatorname{grad} U$, то $U_x = F_1$ и $U_y = F_2$. Так как $U_{xy} = \frac{\partial F_1}{\partial y}$ и $U_{yx} = \frac{\partial F_2}{\partial x}$ непрерывны, то порядок дифференцирования безразличен (гл. II, § 3, n° 3) и $U_{xy} = U_{yx}$, так что $\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$.

Достаточность условия (1) будет доказана тем, что, пользуясь этим условием, мы фактически построим в области G потенциальную функцию $U(x, y)$, для которой

$$U_x = F_1(x, y) \quad \text{и} \quad U_y = F_2(x, y).$$

Рассмотрим сначала простой случай, когда наша область есть прямоугольник $R(\alpha \leqslant x \leqslant \beta,\ \gamma \leqslant y \leqslant \delta)$ со сторонами, параллельными осям координат. Проведем в области R ломаную $P_0P'P$ (рис. 85) от постоянной точки $P_0(\xi_0, \eta_0)$ до переменной точки $P(\xi, \eta)$ так, что отрезок P_0P' ломаной параллелен оси y, а $P'P$ параллелен оси x; промежуточная вершина ломаной есть $P'(\xi_0, \eta)$. Отрезок P_0P' выражается параметрически уравнениями $x = \xi_0 = \text{const}$, $y = t$, причем $\eta_0 \leqslant t \leqslant \eta$, так что $dx = 0$, $dy = dt$, а отрезок $P'P$ — уравнениями $x = t$, $y = \eta = \text{const}$, причем $\xi_0 \leqslant t \leqslant \xi$, так что $dx = dt$, $dy = 0$. Поэтому криволинейный интеграл вдоль ломаной $P_0P'P$ выразится так:

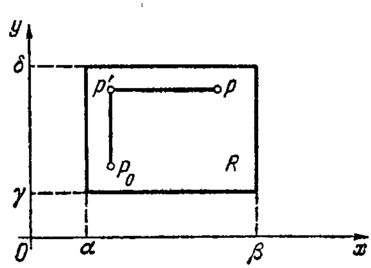


Рис. 85.

$$\int\limits_{P_0P'P} F\,dr = \int\limits_{P_0P'P} (F_1\,dx + F_2\,dy) = \int\limits_{\eta_0}^{\eta} F_2(\xi_0, t)\,dt + \int\limits_{\xi_0}^{\xi} F_1(t, \eta)\,dt.$$

Определенная таким образом функция

$$U(\xi, \eta) = \int\limits_{\eta_0}^{\eta} F_2(\xi_0, t)\,dt + \int\limits_{\xi_0}^{\xi} F_1(t, \eta)\,dt$$

и является искомой функцией. Действительно, дифференцируя по ξ, сразу получаем

$$U_\xi(\xi, \eta) = F_1(\xi, \eta).$$

Далее,

$$U_\eta(\xi, \eta) = F_2(\xi_0, \eta) + \frac{\partial}{\partial \eta} \int\limits_{\xi_0}^{\xi} F_1(t, \eta)\,dt.$$

Так как частная производная $\frac{\partial F_1(t,\eta)}{\partial\eta}$, по условию, непрерывна, то можно дифференцировать под знаком интеграла в правой части:

$$U_\eta(\xi,\eta)=F_2(\xi_0,\eta)+\int_{\xi_0}^{\xi}\frac{\partial F_1(t,\eta)}{\partial\eta}dt=F_2(\xi_0,\eta)+\int_{\xi_0}^{\xi}\frac{\partial F_2(t,\eta)}{\partial\xi}dt$$

(последнее — в силу условия $\frac{\partial F_1}{\partial y}=\frac{\partial F_2}{\partial x}$). Следовательно,

$$U_\eta(\xi,\eta)=F_2(\xi_0,\eta)+F_2(\xi,\eta)-F_2(\xi_0,\eta)=F_2(\xi,\eta).$$

Таким образом, действительно вектор поля $\boldsymbol{F}=\operatorname{grad} U$. На основании предыдущего номера криволинейный интеграл не зависит от пути; поэтому можно положить вообще

$$U(\xi,\eta)=\int_C(F_1\,dx+F_2\,dy),$$

где C — произвольная кусочно гладкая линия, лежащая в данном прямоугольнике R и идущая от постоянной точки P_0 до переменной точки P. Стало быть, теорема доказана для случая прямоугольной области.

Для того чтобы обобщить полученный результат на любую односвязную область G, требуется лишь продолжить построение функции U на такую более общую область. Заметим, что двумерная область называется односвязной, если всякая замкнутая ломаная, полностью лежащая внутри области, может быть стянута в одну точку путем непрерывной деформации, не выводящей за пределы области. Эту наглядную картину стягивания в точку можно уточнить следующим образом. Пусть вершины замкнутой ломаной Π находятся в точках $P_0(x_0,y_0)$, $P_1(x_1,y_1)$, ..., $P_n(x_n,y_n)$. Представим себе, что эти вершины движутся непрерывно со временем, выходя из точек P_0, P_1, P_2, ..., P_n в момент $t=0$ и что все они сходятся в одной и той же точке (ξ,η) области G в момент времени $t=1$. Другими словами, мы допускаем существование переменных точек $P_0(t)$, $P_1(t)$, ..., $P_n(t)$, координаты которых $(x_0(t),y_0(t))$, $(x_1(t),y_1(t))$, ..., $(x_n(t),y_n(t))$ являются непрерывными функциями от t при $0\leqslant t\leqslant 1$ и к тому же

$$P_0(0)=(x_0(0),y_0(0))=P_0,\ \ldots,\ P_n(0)=(x_n(0),y_n(0))=P_n$$

и

$$P_0(1)=(x_0(1),y_0(1))=(\xi,\eta),\ \ldots,\ P_n(1)=(x_n(1),y_n(1))=(\xi,\eta).$$

Конечно, всякая замкнутая ломаная может быть стянута в точку, если никак не ограничивать пути движения ее вершин, но существенной чертой нашего определения односвязной области является требование, чтобы всякая замкнутая ломаная $\Pi(t)$ *этой области* могла быть стянута в точку, да еще *при условии, чтобы многоугольник* $\Pi(t)$ *с вершинами* $P_0(t)$, $P_1(t)$, ..., $P_n(t)$ *оставался в области* G

в продолжение всего процесса стягивания, т. е. при всех значениях t из интервала $0 \leqslant t \leqslant 1$.

Наглядно ясно, что это определение находится в согласии с определением односвязной области, данным в гл. II, § 1, n° 1. Действительно, если область G многосвязна в смысле прежнего определения, то в ней есть «дырка» и многоугольник, содержащий эту «дырку», не может превратиться в точку, не пересекая этой «дырки», т. е. не оставляя области G. Обратно, если область G не имеет «дырок», т. е. не имеет внутренних границ, то всякий многоугольник, лежащий в этой области, может быть стянут в точку. Мы не станем приводить аналитического доказательства эквивалентности обоих определений, ибо оно потребовало бы обширного отступления, а для нашей цели достаточно нынешнего определения.

Итак, для того чтобы обобщить критерий интегрируемости на любую односвязную область G, надо построить в этой области, аналогично прежнему, такую функцию $U(x, y)$, для которой

$$U_x = F_1(x, y) \quad \text{и} \quad U_y = F_2(x, y).$$

Исходя из произвольной, но фиксированной точки P_0 области G, мы определим функцию $U(x, y)$ как криволинейный интеграл

$$U(x, y) = \int_{P_0}^{P} (F_1\,dx + F_2\,dy),$$

причем путем интегрирования служит любая ломаная области G, соединяющая постоянную точку P_0 с переменной точкой $P(x, y)$. Если удастся показать, что определенное таким образом значение $U(x, y)$ не зависит от выбора ломаной, соединяющей точку P_0 с P, то наше определение действительно дает построение функции $U(x, y)$, удовлетворяющей условиям $U_x = F_1$ и $U_y = F_2$.

Стало быть, надо лишь доказать, что наш интеграл не зависит от пути или, вместо этого, что интеграл $\int (F_1\,dx + F_2\,dy)$ вдоль любой замкнутой ломаной Π, лежащей в области G, обращается в нуль. С этой целью стянем многоугольник Π в какую-либо точку области G. Это значит, что мы строим в G замкнутую ломаную $\Pi(t)$ с вершинами $Q_0(t)$, $Q_1(t)$, ..., $Q_n(t)$, которая при $t = 0$ совпадает с Π, а при $t = 1$ обращается в точку. Когда путь интегрирования обращается в точку (кривую длины нуль), то криволинейный интеграл, очевидно, равен нулю. Поэтому достаточно показать, что криволинейный интеграл вдоль $\Pi(t)$ остается постоянным, когда t изменяется от 0 до 1, и тогда станет ясно, что интеграл вдоль замкнутой ломаной $\Pi(t)$ равен нулю при любом значении t $(0 \leqslant t \leqslant 1)$, и в частности при $t = 0$, т. е. интеграл вдоль замкнутой ломаной Π равен нулю.

Рассмотрим теперь любое значение τ параметра t. Так как замкнутая ломаная $\Pi(\tau)$ лежит в G, то можно на ней выбрать последовательность точек (не обязательно вершин) B_0, B_1, B_2, ..., $B_m = B_0$, настолько близких одна к другой, что каждый кусок B_iB_{i+1} ломаной лежит внутри некоторого прямоугольника R_i, содержащегося в G. Если t есть любое другое значение параметра, достаточно близкое к τ, то замкнутая ломаная $\Pi(t)$ настолько близка к $\Pi(\tau)$, что на $\Pi(t)$ можно выбрать точки $A_0, A_1, \ldots, A_m = A_0$, для которых отрезки B_iA_i и $B_{i+1}A_{i+1}$, а также весь кусок A_iA_{i+1} ломаной лежат внутри прямоугольника R_i. Тогда из доказанного критерия интегрируемости для прямоугольной области вытекает, что криволинейный интеграл вдоль замкнутой ломаной $A_iA_{i+1}B_{i+1}B_iA_i$ равен нулю. Обозначим эту замкнутую ломаную через C_i; тогда (см. конец n° 3)

$$\oint_{\Pi(t)} (F_1\,dx + F_2\,dy) - \oint_{\Pi(\tau)} (F_1\,dx + F_2\,dy) = \sum_{i=0}^{m-1} \oint_{C_i} (F_1\,dx + F_2\,dy) = 0.$$

Поэтому для всех значений t, достаточно близких к τ, интеграл вдоль $\Pi(t)$ равен интегралу вдоль $\Pi(\tau)$. Но интеграл вдоль $\Pi(t)$ можно рассматривать как функцию $\varphi(t)$ параметра t; стало быть, эта функция $\varphi(t) = \mathrm{const}$, т. е. интеграл вдоль $\Pi(t)$ имеет одинаковое значение при всех значениях t. Тем самым критерий интегрируемости доказан для любой односвязной области G.

Отметим в заключение, что в пространствах трех и большего числа измерений действует аналогичный критерий, который и доказывается аналогичным методом. Мы удовольствуемся тем, что сформулируем соответствующую теорему для трехмерной области.

Если в трехмерной области G, внутри которой любая замкнутая ломаная может быть стянута путем непрерывной деформации в точку, задано поле непрерывной векторной функции

$$F = \{F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)\},$$

имеющей непрерывные частные производные

$$\frac{\partial F_1}{\partial y}, \quad \frac{\partial F_1}{\partial z}, \quad \frac{\partial F_2}{\partial z}, \quad \frac{\partial F_2}{\partial x}, \quad \frac{\partial F_3}{\partial x}, \quad \frac{\partial F_3}{\partial y},$$

то необходимым и достаточным условием независимости криволинейного интеграла

$$\int_C F\,d\mathbf{r} = \int_C (F_1\,dx + F_2\,dy + F_3\,dz)$$

от пути C, лежащего полностью в области G, является выполнение соотношений

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}, \qquad \frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}, \qquad \frac{\partial F_3}{\partial x} = \frac{\partial F_1}{\partial z}$$

или (в векторной записи) критерия

$$\operatorname{rot} F = 0.$$

Если это условие выполнено, то криволинейный интеграл представляет собой при фиксированной начальной точке P_0 пути C функцию $U(P)=U(x, y, z)$ его конечной точки P, точнее,

$$\int_{P_0}^{P}(F_1\,dx+F_2\,dy+F_3\,dz)=U(x, y, z)-U(x_0, y_0, z_0)$$

или (в векторной записи)

$$\int_{P_0}^{P}\boldsymbol{F}\,d\boldsymbol{r}=U(P)-U(P_0).$$

При этом вектор поля $\boldsymbol{F}=\operatorname{grad}U$.

8. Важность условия односвязности. Для всего исследования, проведенного в n° 7, существенно, чтобы рассматриваемая область была односвязной. Если область G не односвязна, то не может быть уверенности в том, что функцию U возможно определить во всей области G однозначно путем интегрирования по ломаным путям. И действительно, мы покажем сейчас на примере, что в многосвязной области наш критерий интегрируемости не обеспечивает независимости криволинейного интеграла от пути интегрирования.

Функции

$$F_1(x, y)=-\frac{y}{x^2+y^2},\qquad F_2(x, y)=\frac{x}{x^2+y^2}$$

определены и непрерывны всюду, кроме начала координат. Их частные производные

$$\frac{\partial F_1}{\partial y}=-\frac{1}{x^2+y^2}+\frac{2y^2}{(x^2+y^2)^2},\qquad \frac{\partial F_2}{\partial x}=\frac{1}{x^2+y^2}-\frac{2x^2}{(x^2+y^2)^2}$$

тоже непрерывны всюду, кроме начала координат, и во всех этих точках удовлетворяют условию

$$\frac{\partial F_1}{\partial y}=\frac{\partial F_2}{\partial x}=\frac{y^2-x^2}{(x^2+y^2)^2}.$$

Возьмем теперь интеграл $\oint\limits_C(F_1\,dx+F_2\,dy)$ вдоль положительно ориентированной окружности C: $x=\cos t$, $y=\sin t$ с центром в начале координат. Имеем

$$\oint\limits_C(F_1\,dx+F_2\,dy)=\int_0^{2\pi}\{-\sin t\,(-\sin t)+\cos t\cos t\}\,dt=\int_0^{2\pi}dt=2\pi;$$

таким образом, интеграл вдоль замкнутой кривой C отличен от нуля и, стало быть, наш криволинейный интеграл зависит от пути. Это объясняется тем, что окружность C невозможно включить в такую *односвязную* область, в которой выполнялись бы все условия теоремы; для того чтобы эти условия были выполнены, придется уже взять кольцевидную область, содержащую окружность C, но не содержащую точки (0, 0).

Заметим кстати, что интеграл $\oint(F_1dx+F_2dy)$ имеет вдоль любой замкнутой кривой, окружающей начало координат, одно и то же значение, а именно 2π. Это можно вывести из общей теоремы о разбиении (n° 3, свойство 4, рис. 83) следующим образом. Разбиваем кольцевую область между двумя такими кривыми C и C' с помощью поперечных линий A_iB_i на некоторое число односвязных областей C_i и учитываем тот факт, что вдоль контура каждой из этих областей криволинейный интеграл равен нулю.

Упражнения

1. Вычислить интеграл $\int\limits_C (e^x \sin y\, dx + e^x \cos y\, dy)$, где путь интегрирования C есть кривая, идущая от точки $(0, 0)$ до точки (ξ, η).

2*. Вычислить интеграл

$$\int\limits_C \left(\frac{e^x}{x^2+y^2}(x\cos y + y \sin y)\, dy + \frac{e^x}{x^2+y^2}(x \sin y - y \cos y)\, dx\right),$$

где путь C есть замкнутая кривая, окружающая начало координат и не пересекающая сама себя.

§ 2. Связь между криволинейным и двойным интегралом на плоскости — интегральные теоремы для плоских векторных полей

1. Интегральная теорема Гаусса [теорема Остроградского для плоскости [1])]. Для функций одной независимой переменной мы в свое время вывели формулу Ньютона — Лейбница

$$\int\limits_{x_0}^{x_1} f'(x)\, dx = f(x_1) - f(x_0),$$

устанавливающую связь между дифференцированием и интегрированием. Аналогичную формулу можно вывести и для функций от двух переменных — интегральную теорему Гаусса. В этой формуле дифференцирование тоже погашается интегрированием в том смысле, что двойные интегралы вида

$$\iint\limits_G f_x\, dx\, dy \qquad \text{или} \qquad \iint\limits_G g_y\, dx\, dy$$

преобразовываются в (одномерные) криволинейные интегралы вдоль граничной кривой C области G. При этом мы рассматриваем границу C области G как ориентированную кривую и направление обхода отмечаем знаком плюс или минус.

Теорема Гаусса формулируется следующим образом:

Если функции $f(x, y)$ и $g(x, y)$ непрерывны и имеют непрерывные частные производные в замкнутой области G, ограниченной кусочно гладкой кривой C, то

$$\iint\limits_G \{f_x(x, y) + g_y(x, y)\}\, dx\, dy = \oint\limits_{+C} \{f(x, y)\, dy - g(x, y)\, dx\}, \qquad (1)$$

---

[1]) См. сноску на стр. 408 (§ 5, n° 1). (*Прим. перев.*)

причем интеграл в правой части есть криволинейный интеграл вдоль границы области G, описываемой в положительном направлении, т. е. так, что внутренние точки области G остаются слева.

Для доказательства рассмотрим сначала тот случай, когда любая прямая, параллельная оси x или оси y, пересекает контур C не более чем в двух точках; к тому же предположим, что $g(x, y) \equiv 0$ в G. Тогда $\iint_G f_x(x, y)\,dx\,dy$ можно записать как повторный интеграл

$$\iint_G f_x(x, y)\,dx\,dy = \int dy \int f_x(x, y)\,dx,$$

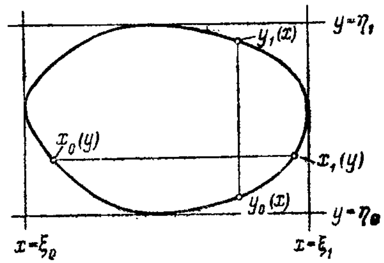


Рис. 86.

где внутренний интеграл (по x) берется вдоль тех отрезков прямых $y = \text{const}$, которые лежат в области G, а внешний интеграл берется по тому промежутку значений y, которому соответствуют точки области G (рис. 86). Обозначим абсциссу точки входа в область G прямой, параллельной оси x и имеющей ординату y, через $x_0(y)$, а абсциссу точки ее выхода через $x_1(y)$, причем $x_1 \geqslant x_0$; тогда внутренний интеграл будет

$$\int_{x_0(y)}^{x_1(y)} f_x(x, y)\,dx = f(x_1(y), y) - f(x_0(y), y).$$

В качестве пределов внешнего интеграла надо теперь поставить наименьшее и наибольшее значения ординаты y в области G — обозначим их через η_0 и η_1. Следовательно,

$$\iint_G f_x(x, y)\,dx\,dy = \int_{\eta_0}^{\eta_1} f(x_1(y), y)\,dy + \int_{\eta_1}^{\eta_0} f(x_0(y), y)\,dy = \oint_C f(x, y)\,dy.$$

Ясно, что для частного случая $g(x, y) \equiv 0$ это равенство равносильно формулированной выше теореме Гаусса. Очевидно также, что последняя формула охватывает и тот случай, когда контур C содержит прямолинейные отрезки, параллельные оси x, ибо на каждом таком отрезке $y = \text{const}$, а стало быть вклад этого отрезка в криволинейный интеграл в правой части равен нулю.

Пусть теперь $f(x, y) \equiv 0$ в G. Используя сделанное выше допущение, что любая прямая, параллельная оси y, пересекает контур C не более чем в двух точках, путем совершенно аналогичных рассуждений придем к формуле

$$\iint_G g_y(x, y)\,dx\,dy = \int_{\xi_0}^{\xi_1} g(x, y_1(x))\,dx - \int_{\xi_0}^{\xi_1} g(x, y_0(x))\,dx =$$
$$= -\oint_C g(x, y)\,dx.$$

Наличие в правой части знака минус не должно вызвать удивления, так как оси x и y на плоскости не вполне эквивалентны: ведь ось x переходит в ось y при *положительном* повороте на угол $\frac{\pi}{2}$, а ось y переходит в ось x при *отрицательном* повороте на $\frac{\pi}{2}$.

Сложение обеих полученных формул приводит к общей записи теоремы Гаусса:

$$\iint\limits_{G} \{f_x(x, y)+g_y(x, y)\}\, dx\, dy=\oint\limits_{C} \{f(x, y)\, dy-g(x, y)\, dx\},$$

которая таким образом доказана для областей описанного выше вида.

Полученную формулу нетрудно распространить на все области, которые возможно разрезать на конечное число составляющих областей такого вида, что граница каждой из них пересекается прямыми, параллельными осям координат, не более чем в двух точках. На рис. 87 показан пример такой невыпуклой области, составленной из выпуклых областей. Для каждой из этих составляющих областей теорема Гаусса доказана. В результате сложения формул Гаусса, выписанных для всех этих частичных областей, получается интегральная формула (1) для суммарной области, так как криволинейные интегралы вдоль внутренних границ (разрезов) взаимно уничтожаются: ведь каждая из них описывается два раза в двух противоположных направлениях!

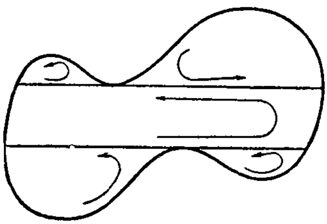


Рис. 87.

Добавим без доказательства, что теорема Гаусса фактически справедлива для *всякой* области, имеющей кусочно гладкую границу [1]. Доказательство можно получить путем предельного перехода.

Еще одно замечание. Допущение, что область можно разбить на конечное число составляющих областей такого вида, что граница каждой из них пересекается прямыми, параллельными осям координат не более чем в двух точках, можно заменить следующим условием: границу области можно разбить на конечное число частей, каждая из которых имеет однозначные проекции на обе координатные оси; при этом допускается, однако, чтобы проекция такой части на одну из осей координат состояла из одной точки, т. е. граница может содержать отрезки, параллельные осям.

---

[1]) Такая область может и не принадлежать рассмотренному только что классу областей: граница области может, например, содержать такой кусок кривой $y=x^2\sin\frac{1}{x}$, который пересекает ось x в бесконечном множестве точек.

В качестве примера применения теоремы Гаусса выведем с ее помощью знакомые нам формулы для площади S области G (т. I, стр. 317). Полагая $f(x, y) = x$ и $g(x, y) = 0$, сразу получим

$$S = \iint_G dx\,dy = \oint_C x\,dy.$$

Полагая $f(x, y) = 0$ и $g(x, y) = y$, имеем

$$S = -\oint_C y\,dx.$$

[Если же положить $f(x, y) = x$, $g(x, y) = y$, то получится третья формула:

$$2S = \oint_C (x\,dy - y\,dx).$$

Читателю, конечно, ясно, что интегралы, выведенные для площади (т. I, стр. 317), фактически являются криволинейными интегралами.] По поводу знака см. ниже § 4, n° 1.

2. Векторная запись теоремы Гаусса. Теорема Гаусса принимает особенно простой вид, если воспользоваться обозначениями и понятиями векторного анализа. Скалярные функции $f(x, y)$ и $g(x, y)$ мы будем рассматривать как координаты вектор-функции точки

$$\boldsymbol{F}(P) = \{F_1(x, y),\ F_2(x, y)\} = \{f(x, y),\ g(x, y)\}$$

в плоском векторном поле. Тогда подынтегральная функция двойного интеграла в формуле Гаусса

$$f_x(x, y) + g_y(x, y) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} = \operatorname{div}\boldsymbol{F},$$

т. е. является дивергенцией поля $\boldsymbol{F}$ (ср. гл. II, § 7, n° 7).

Для того чтобы получить векторную запись криволинейного интеграла в правой части формулы Гаусса, введем в качестве параметра на граничной кривой C ее длину дуги s, и притом так, чтобы s возрастала при положительном направлении обхода. Тогда вектор $\mathbf{t} = \{\dot{x}, \dot{y}\}$ есть *единичный касательный вектор* кривой C, направленный в сторону возрастания длины дуги s (это видно из сопоставления в т. I стр. 307, 308 и 324). Вектор $\boldsymbol{n} = \{\dot{y}(s), -\dot{x}(s)\}$ тоже имеет длину 1 и направлен по нормали к граничной кривой.

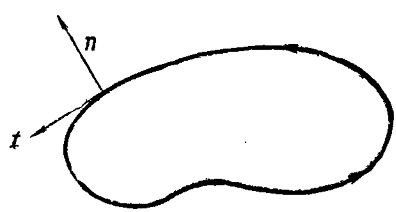


Рис. 88.

[Введем наряду с ортами $\boldsymbol{e}_1$ и $\boldsymbol{e}_2$ осей x и y третий орт $\boldsymbol{e}_3 = [\boldsymbol{e}_1, \boldsymbol{e}_2]$. Тогда векторное произведение

$$[\mathbf{t}\boldsymbol{e}_3] = [(\dot{x}\boldsymbol{e}_1 + \dot{y}\boldsymbol{e}_2)\,\boldsymbol{e}_3] = \dot{y}\boldsymbol{e}_1 - \dot{x}\boldsymbol{e}_2 = \{\dot{y}, -\dot{x}\} = \boldsymbol{n}.$$

Так как касательный вектор $\mathbf{t}$ направлен в сторону положительного обхода граничной кривой, то $\boldsymbol{n}$ есть единичный вектор *внешней* нормали по отношению к области G (рис. 88). Нетрудно сделать соответствующий рисунок для многосвязной области и убедиться, что для внутренних ее границ вектор $\boldsymbol{n}$ тоже будет направлен по *внешней* (к области) нормали.]

Теперь подынтегральное выражение криволинейного интеграла запишется так:

$$f(x, y)\,dy - g(x, y)\,dx = (F_1\dot{y} - F_2\dot{x})\,ds = \boldsymbol{Fn}\,ds = F_n\,ds,$$

где $\boldsymbol{Fn}$ есть скалярное произведение векторов $\boldsymbol{F}$ и $\boldsymbol{n}$, а F_n обозначает проекцию вектора F на внешнюю нормаль $\boldsymbol{n}$.

Отсюда получаем векторную запись теоремы Гаусса:

$$\iint\limits_{G} \operatorname{div} \boldsymbol{F}\,dS = \int\limits_{+C} \boldsymbol{Fn}\,ds = \int\limits_{+C} F_n\,ds.$$

Словесно она выражается так: *двойной интеграл от дивергенции плоского векторного поля по замкнутой области G равен криволинейному интегралу вдоль границы области (в положительном направлении обхода) от проекции вектора поля на внешнюю нормаль.*

3. Теорема Стокса для плоскости. Теореме Гаусса [теореме Остроградского для плоскости] можно дать еще одно, совершенно иное векторное толкование, носящее название *теоремы Стокса.* Снова рассматриваем плоское векторное поле $\boldsymbol{F}(P) = \{F_1(x, y), F_2(x, y)\}$, но на этот раз полагаем $f(x, y) = F_2(x, y)$ и $g(x, y) = -F_1(x, y)$. Тогда в формуле (1) Гаусса (стр. 384) подынтегральная функция двойного интеграла примет следующий вид:

$$f_x(x, y) + g_y(x, y) = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}.$$

С выражением такого типа мы уже встречались при построении понятия ротора векторного поля (гл. II, § 7, n° 7). Для того чтобы применить сюда понятие ротора, продолжим поле вектора $\boldsymbol{F}$ на трехмерное пространство, полагая вектор поля не зависящим от z, а его третью координату (его проекцию на ось z) постоянной:

$$\boldsymbol{F}(P) = \boldsymbol{F}(x, y) = \{F_1(x, y),\quad F_2(x, y),\ \text{const}\}.$$

Получим тогда

$$\operatorname{rot} \boldsymbol{F} = \left\{0,\ 0,\ \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right\},$$

и подынтегральная функция двойного интеграла

$$\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = (\operatorname{rot} \boldsymbol{F})_z$$

будет проекцией вектора $\operatorname{rot} \boldsymbol{F}$ на ось z. Это выражение удобно назвать *скалярной ротацией* плоского векторного поля $\boldsymbol{F}$ и обозначать через $(\mathrm{rot})F$, т. е. заключить символ ротора в круглые скобки:

$$(\mathrm{rot})F = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}.$$

Подынтегральное выражение криволинейного интеграла в формуле (1) выразится теперь так:

$$f(x,\ y)\,dy - g(x,\ y)\,dx = F_1\,dx + F_2\,dy = (F_1\dot{x} + F_2\dot{y})\,ds = \boldsymbol{F}\mathbf{t}\,ds = \\ = F_{\mathbf{t}}\,ds,$$

где на ds умножается скалярное произведение вектора поля на единичный касательный вектор граничной кривой или, что то же самое, проекция вектора поля на положительное направление касательной.

Теперь формула (1) приобретет следующий вид:

$$\iint\limits_G \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dx\,dy = \int\limits_{+C} (F_1\,dx + F_2\,dy) \tag{2}$$

или

$$\iint\limits_G (\mathrm{rot})F\,dS = \int\limits_{+C} F_{\mathbf{t}}\,ds = \int\limits_{+C} \boldsymbol{F}\,d\boldsymbol{r}. \tag{2a}$$

Эта формула и выражает теорему Стокса для плоскости:

Двойной интеграл скалярной ротации плоского векторного поля по замкнутой области равен криволинейному интегралу касательной составляющей вектора поля вдоль граничной кривой, пробегаемой в положительном направлении.

Однако можно использовать векторный характер ротора пространственного векторного поля и, учитывая тот факт, что в формулу (2) входят только составляющие векторного поля, лежащие в плоскости xy, освободить теорему Стокса для плоской области от предположения, что область G лежит именно в плоскости xy. Таким образом приходим к следующей более общей трактовке теоремы Стокса для плоской области:

$$\iint\limits_G (\mathrm{rot}\,F)_n\,dS = \int\limits_{+C} F_{\mathbf{t}}\,ds = \int\limits_{+C} F\,dr, \tag{3}$$

где G есть плоская область, лежащая как угодно в пространстве, ограниченная кривой C, а $(\mathrm{rot}\,\boldsymbol{F})_n$ есть проекция вектора $\mathrm{rot}\,\boldsymbol{F}$ на нормаль к плоскости области G [причем направление нормали должно быть согласовано с направлением обхода границы C по правилу правого винта].

Заметим кстати, что из теоремы Стокса для плоскости вытекает новое простое доказательство того факта, что условие $\frac{\partial F_2}{\partial x} = \frac{\partial F_1}{\partial y}$ обеспечивает независимость криволинейного интеграла $\int (F_1\,dx + F_2\,dy)$ от пути интегрирования (т. е. доказательство достаточности этого условия). Мы уже знаем, что независимость от пути равносильна обращению в нуль интеграла вдоль любой замкнутой кривой. Если такая замкнутая кривая C является границей области G рассмотренного выше типа, то теорема Стокса преобразовывает криволинейный интеграл $\oint (F_1\,dx + F_2\,dy)$ в двойной интеграл выражения $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$ по области G. Ясно, что из обращения в нуль этого выражения вытекает обращение в нуль криволинейного интеграла вдоль контура C области G.

4. Формулы Грина. С теоремой Гаусса тесно связаны некоторые другие преобразования интегралов, известные под названием теорем или формул Грина и часто применяемые в теории дифференциальных уравнений. Для вывода этих формул возьмем две функции $u(x, y)$ и $v(x, y)$, имеющие в области G непрерывные производные первого и второго порядка. Применяя теорему Гаусса к функциям $F_1 = uv_x$ и $F_2 = uv_y$ и учитывая, что

$$\frac{\partial}{\partial x}(uv_x) + \frac{\partial}{\partial y}(uv_y) = (u_xv_x + uv_{xx}) + (u_yv_y + uv_{yy}) = \\ = u_xv_x + u_yv_y + u\,\nabla^2 v,$$

где ∇^2 есть оператор Лапласа (гл. II, § 7, n° 7), получим

$$\iint\limits_G (u_xv_x + u_yv_y)\,dx\,dy + \iint\limits_G u\nabla^2 v\,dx\,dy = \int\limits_{+C} (-uv_y\,dx + uv_x\,dy).$$

Эта интегральная формула называется *первой формулой Грина.* Из условий теоремы Гаусса видно, что эта формула доказана, если функции u_x, v_x, u_y, v_y, v_{xx}, v_{yy} непрерывны в замкнутой области G. Если функции u_{xx} и u_{yy} тоже непрерывны, то справедлива также аналогичная формула

$$\iint\limits_G (u_xv_x + u_yv_y)\,dx\,dy + \iint\limits_G v\nabla^2 u\,dx\,dy = \int\limits_{+C} (-vu_y\,dx + vu_x\,dy).$$

Вычитая эту формулу из предыдущей, получим следующее соотношение:

$$\iint\limits_G (u\nabla^2 v - v\nabla^2 u)\,dx\,dy = \int\limits_{+C} \{(vu_y - uv_y)\,dx - (vu_x - uv_x)\,dy\},$$

которое носит название *второй формулы Грина.*

Криволинейные интегралы в формулах Грина можно записать в несколько ином виде, воспользовавшись тем, что производная от функции точки $f(x, y)$ по направлению внешней нормали $\boldsymbol{n} = \{\dot{y}, -\dot{x}\}$ выражается так:

$$D_n f(x, y) = \boldsymbol{n}\,\mathrm{grad}\,f = f_x\dot{y} - f_y\dot{x},$$

если длина дуги s возрастает в направлении положительного обхода границы. Отсюда видно, что в первой формуле Грина

$$-uv_y\,dx + uv_x\,dy = u(-v_y\dot{x} + v_x\dot{y})\,ds = uD_nv\,ds,$$

а во второй формуле Грина

$$(vu_y - uv_y)\,dx - (vu_x - uv_x)\,dy = \\ = u(v_x\dot{y} - v_y\dot{x})\,ds - v(u_x\dot{y} - u_y\dot{x})\,ds = (uD_nv - vD_nu)\,ds.$$

Стало быть, формулы Грина принимают следующий вид:

$$\iint_G (u_x v_x + u_y v_y)\,dx\,dy + \iint_G u\nabla^2 v\,dx\,dy = \int_{+C} u D_n v\,ds$$

и

$$\iint_G (u\nabla^2 v - v\nabla^2 u)\,dx\,dy = \int_{+C} (uD_n v - vD_n u)\,ds.$$

Заметив, что $u_x v_x + u_y v_y = \operatorname{grad} u \operatorname{grad} v$ (скалярному произведению двух градиентов), а $\nabla^2 v = \operatorname{grad} \operatorname{div} v$, можно первую формулу Грина представить и в векторной записи:

$$\iint_G \operatorname{grad} u \operatorname{grad} v\,dS + \iint_G u \operatorname{div} \operatorname{grad} v\,dS = \int_{+C} uD_n v\,ds.$$

5. Двойной интеграл от якобиана. Можно вывести еще одну замечательную формулу преобразования двойного интеграла следующим образом. Двойные интегралы от функций

$$\frac{\partial}{\partial x}(uv_y) = u_x v_y + uv_{yx} \quad \text{и} \quad \frac{\partial}{\partial y}(uv_x) = u_y v_x + uv_{xy}$$

преобразуем каждый в отдельности по теореме Гаусса в криволинейные интегралы, а затем вычтем из первого соотношения второе. Получится следующая формула:

$$\iint_G (u_x v_y - u_y v_x)\,dx\,dy = \int_{+C} (uv_x\,dx + uv_y\,dy). \qquad (*)$$

Эта формула освещает по-новому смысл якобиана. Подынтегральной функцией двойного интеграла является якобиан $\dfrac{\partial(u, v)}{\partial(x, y)}$. Предположим, что якобиан положителен в односвязной области G и что область G плоскости xy отображается на область G' плоскости uv с помощью преобразования

$$u = u(x, y), \qquad v = v(x, y)$$

с сохранением направления обхода $\left(\text{ибо } \dfrac{\partial(u, v)}{\partial(x, y)} > 0\right)$. Как известно, криволинейным интегралом

$$\oint_C (uv_x\,dx + uv_y\,dy) = \oint_{C'} u\,dv,$$

вычисленным вдоль граничной кривой C' области G' в положительном направлении, выражается площадь области G'. Стало быть, и двойной интеграл от якобиана равен площади отображающей области G'. Но, с другой стороны, площадь G' равна $\iint_{G'} du\,dv$. Поэтому

$$\iint_{G'} du\,dv = \iint_G \frac{\partial(u, v)}{\partial(x, y)}\,dx\,dy.$$

Таким образом, получен новый вывод формулы преобразования двойного интеграла (гл. IV, § 4, n° 1, стр. 275) для того частного случая, когда подынтегральная функция в левой части равна единице. Разделим теперь обе части последнего равенства на площадь области G, а затем заставим область G стягиваться в точку. Тогда в правой части в пределе получится производная двойного интеграла $\iint\limits_G \frac{\partial(u, v)}{\partial(x, y)}\, dx\, dy$ по области, равная подынтегральной функции, т. е. якобиану $\frac{\partial(u, v)}{\partial(x, y)}$. Стало быть, якобиан равен пределу отношения площади изображения G' к площади изображаемой области, когда область G стягивается в точку, или, как принято говорить, коэффициенту искажения площади в этой точке.

Так как, по теореме интегрального исчисления о среднем значении, отношение площади области G' к площади области G равно промежуточному значению якобиана, то уже из самого определения двойного интеграла почти сразу вытекает общая формула преобразования

$$\iint\limits_{G'} f(u, v)\, du\, dv = \iint\limits_{G} f \cdot \frac{\partial(u, v)}{\partial(x, y)}\, dx\, dy.$$

Детальное проведение этого рассуждения предоставляем читателю. Еще одно полное доказательство этой формулы будет дано в § 3, n° 3.

6. Преобразование плоского лапласиана к новым (в частности, полярным) координатам. Метод, примененный в конце n° 5, дает возможность преобразовать лапласиан $\nabla^2 v = v_{xx} + v_{yy}$ к новым, криволинейным, координатам. Для этой цели исходим из формулы

$$\iint\limits_G \nabla^2 v\, dx\, dy = \oint\limits_C D_n v\, ds,$$

получающейся из первой или из второй формулы Грина (в конце n° 4) при $u = 1$. Возьмем производные по области от обеих сторон этого равенства, т. е. разделим каждую из них на площадь области G и перейдем к пределу, когда область G стягивается в точку. Тогда в левой части получится $\nabla^2 v$. Для того чтобы преобразовать $\nabla^2 v$ к новым координатам, надо в правой части выполнить соответствующее преобразование криволинейного интеграла к новым координатам, разделить его на площадь области G и найти предел полученной дроби. Преимущество этого способа по сравнению с прямым вычислением лапласиана заключается в том, что мы избавляемся от трудоемкого преобразования вторых производных от v, так как криволинейный интеграл содержит лишь первые производные.

В качестве важного примера выполним преобразование $\nabla^2 v$ к полярным координатам (r, θ). За область G примем малую ячейку сети полярных координат, ограниченную дугами окружностей (с центрами в начале координат) радиусов r и $r+h$ и полупрямыми θ и $\theta+k$; площадь этой ячейки равна, как известно, $kh\rho$, где $\rho=r+\frac{1}{2}h$.

Согласно общему выводу, сделанному выше,

$$\nabla^2 v=\lim_{\substack{h\to 0\\ k\to 0}}\frac{1}{\rho kh}\oint\limits_C D_n v\,ds.$$

Вычислим теперь криволинейный интеграл вдоль границы нашей ячейки, учитывая, что производная по внешней нормали на окружности радиуса $r+h$ равна $v_r(r+h,\theta)$, на окружности радиуса r она равна $-v_r(r,\theta)$, на полупрямой θ имеем $D_n v=-\frac{1}{r}v_\theta(r,\theta)$, а на полупрямой $\theta+k$ она равна $\frac{1}{r}v_\theta(r,\theta+k)$. Тогда

$$\nabla^2 v=\lim_{\substack{h\to 0\\ k\to 0}}\frac{1}{\rho}\left(\frac{1}{k}\int\limits_\theta^{\theta+k}\frac{(r+h)\,v_r(r+h,\theta)-rv_r(r,\theta)}{h}\,d\theta+\right.$$
$$\left.+\frac{1}{h}\int\limits_r^{r+h}\frac{v_\theta(r,\theta+k)-v_\theta(r,\theta)}{kr}\,dr\right).$$

Применив последовательно теоремы о среднем значении дифференциального, а затем интегрального исчисления, приведем это выражение к следующему виду:

$$\nabla^2 v=\lim_{\substack{h\to 0\\ k\to 0}}\frac{1}{\rho}\left\{r_1 v_{rr}(r_1,\theta_1)+v_r(r_1,\theta_1)+\frac{1}{r}v_{\theta\theta}(r_2,\theta_2)\right\},$$

где r_1, r_2 и θ_1, θ_2 — промежуточные значения переменных r и θ, удовлетворяющие условиям $(r<r_1,\ r_2<r+h)$ и $(\theta<\theta_1,\ \theta_2<\theta+k)$. В пределе при $h\to 0$, $k\to 0$ сразу получим

$$\nabla^2 v=\frac{1}{r}(rv_r)_r+\frac{1}{r^2}v_{\theta\theta}.$$

Это и есть искомая формула преобразования.

§ 3. Наглядное истолкование интегральных теорем для плоскости и их приложения

1. Гидромеханическое истолкование теоремы Гаусса. Дивергенция и производительность источников. Изложенные выше интегральные теоремы для плоскости допускают наглядное истолкование с помощью представления о стационарном плоском потоке несжимаемой жидкости. Стационарным плоским потоком называется такое движение жидкости, распределенной на плоскости с постоянной поверхностной плотностью, в котором состояние движения, т. е. вектор-скорость, в каждой точке не зависит от времени. Стало быть, такой поток вполне определяется плоским полем вектора-скорости $v=\{v_1,\ v_2\}$. Постоянную поверхностную плотность жидкости примем для простоты равной единице.

Рассмотрим какую-либо дугу C и припишем ей в каждой ее точке единичный нормальный вектор $\boldsymbol{n}$, направление которого непрерывно изменяется. Тем самым в каждой точке дуги выбрано положительное направление нормали; какую выбрать из двух возможностей — безразлично. Тогда общее количество жидкости, протекающей в единицу времени через дугу C в направлении положительной нормали, выражается интегралом

$$\int\limits_C \boldsymbol{v}\boldsymbol{n}\,ds = \int\limits_C v_n ds,$$

где s — переменная длина дуги кривой C.

К этому интегралу приходят следующим рассуждением. Сначала заменяют кривую ломаной линией, звенья которой имеют длины $\Delta s_1, \Delta s_2, \ldots, \Delta s_n$, и предполагают, что вдоль каждого звена ломаной вектор $\boldsymbol{v}$ постоянен. Затем совершают обычный предельный переход от ломаной к дуге кривой C.

Если кривая C замкнута и является границей области G, а вектор $\boldsymbol{n}$ направлен в сторону внешней нормали, то теорема Гаусса

$$\oint\limits_C \boldsymbol{v}\boldsymbol{n}\,ds = \iint\limits_G \operatorname{div}\boldsymbol{v}\,dS$$

утверждает, что общее количество жидкости, *вытекающей* в единицу времени из области G, равно двойному интегралу от дивергенции поля скоростей по области G.

Из этого утверждения сразу получается наглядная интерпретация понятия дивергенции. Если криволинейный интеграл на левой стороне равенства обращается в нуль, то это значит, что количество жидкости, вытекающей через часть границы C, компенсируется притоком жидкости через остальную часть границы и в области G не происходит ни накопления, ни убыли жидкости, что естественно, так как процесс не зависит от времени. Если значение криволинейного интеграла положительно, то в общей сложности жидкость вытекает из области G, если же оно отрицательно, то имеет место приток жидкости в область G. Так как процесс — стационарный, т. е. не зависит от времени, то общее количество жидкости в области G не может ни увеличиваться, ни уменьшаться; стало быть, в самой области G жидкость должна либо создаваться, либо уничтожаться. В таком случае говорят, что в области G существуют *источники* или *стоки* (отрицательные источники), и стационарность потока выражается в том, что наличные источники и стоки регулируют вытекание и приток жидкости так, что плотность ее в каждой точке области G остается неизменной. Общее количество жидкости, вытекающее из области G в единицу времени, характеризует результирующую обильность или «дебит» всех источников и стоков области G; это количество мы будем называть *суммарной производительностью источников* области G. Она положительна, если преобладают источники, и отрицательна, если превалируют стоки. Отношение суммарной производительности источников к площади области можно назвать средней производительностью источников области G, а предел этого отношения, когда область G стягивается к ее точке P, есть *удельная производительность* в точке P. Стало быть, удельная производительность источников есть производная по области от суммарной производительности. Из теоремы Гаусса вытекает, что *дивергенция поля скоростей равна удельной производительности источников.* Таким образом, теорема Гаусса дает возможность приписать наглядный смысл понятию дивергенции, которое было первоначально введено чисто формальным определением.

Этот же гидромеханический смысл дивергенции можно получить с помощью следующего грубого рассуждения.

Данный поток разложим на два потока, один из которых течет параллельно оси x со скоростью $v_1 = v_1(x, y)$, а другой — параллельно оси y со скоростью $v_2 = v_2(x, y)$, и рассмотрим прямоугольник с вершинами $P_1(\xi, \eta)$, $P_2(\xi + h, \eta)$, $P_3(\xi, \eta + k)$, $P_4(\xi + h, \eta + k)$. Если считать, что скорость v_1 имеет на стороне P_1P_3 постоянное значение $v_1(\xi, \eta)$, а на стороне P_2P_4 постоянное значение $v_1(\xi + h, \eta)$, то результирующее количество жидкости, вытекающее из прямоугольника в положительном направлении оси x в единицу времени, выразится разностью

$$kv_1(\xi + h, \eta) - kv_1(\xi, \eta).$$

Отношение этого выражения к площади hk прямоугольника

$$\frac{v_1(\xi + h, \eta) - v_1(\xi, \eta)}{h}$$

дает приближенно для потока, параллельного оси x, среднее превышение оттока жидкости над ее притоком в расчете на единицу площади. Аналогично получим для второго составляющего потока, параллельного оси y, среднее превышение оттока жидкости над ее притоком

$$\frac{v_2(\xi, \eta + k) - v_2(\xi, \eta)}{k}.$$

Таким образом, для результирующего потока средний избыток оттока жидкости над ее притоком (на единицу площади) приближенно равен

$$\frac{v_1(\xi + h, \eta) - v_1(\xi, \eta)}{h} + \frac{v_2(\xi, \eta + k) - v_2(\xi, \eta)}{k},$$

а предел этой суммы при $h \to 0$ и $k \to 0$ равен как раз известному нам выражению $\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} = \operatorname{div} \boldsymbol{v}$. Вместе с тем это рассуждение придает дивергенции тот же смысл, который выяснился ранее.

Особый интерес представляет *поток, лишенный источников* в рассматриваемой области G, т. е. в области G жидкость не создается и не уничтожается. Такой поток характеризуется условием

$$\operatorname{div} v \equiv 0 \text{ в области } G.$$

Согласно теореме Гаусса, это условие эквивалентно утверждению что

$$\oint_C v_n\,ds = 0$$

вдоль любой замкнутой кривой, лежащей целиком в области G.

2. Интерпретация теоремы Стокса в поле скоростей и в силовом поле. Теорему Стокса тоже можно пояснить с помощью плоского потока несжимаемой жидкости, характеризуемого полем скоростей $\boldsymbol{v}(P) = \{v_1(x, y), v_2(x, y)\}$. Интеграл $\oint_C \boldsymbol{v}\,d\boldsymbol{r} = \oint_C v_t\,ds$ вдоль замкнутой кривой C мы будем называть *циркуляцией* вектора $\boldsymbol{v}$ вдоль этой кривой. По теореме Стокса

$$\oint_C \boldsymbol{v}\,d\boldsymbol{r} = \iint_G (\mathrm{rot})\,\boldsymbol{v}\,dS,$$

а из этого равенства вытекает, что скалярная ротация, $(\mathrm{rot})\,\boldsymbol{v}$, есть *удельная циркуляция* или *плотность* циркуляции в поле $\boldsymbol{v}(P)$; ее называют также *плотностью вихрей.* В этих терминах теорема Стокса утверждает, что циркуляция вектора-скорости вдоль замкнутой кривой C равна двойному интегралу плотности вихрей по области G, ограниченной кривой C.

И здесь особый интерес представляют такие векторные поля, в которых циркуляция вдоль *любой* замкнутой кривой равна нулю, так что по теореме Стокса плотность циркуляции равна нулю всюду в области G. Такие векторные поля или потоки называются *безвихревыми* или *лишенными вихрей*; они характеризуются соотношением

$$(\mathrm{rot})\,\boldsymbol{v} = 0.$$

Если стационарное плоское векторное поле лишено как источников, так и вихрей, то оно удовлетворяет системе двух уравнений

$$\operatorname{div}\boldsymbol{v} \equiv \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} = 0, \qquad (\mathrm{rot})\,\boldsymbol{v} \equiv \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} = 0.$$

Заметим, кстати, что эта система уравнений особенно интересна, потому что она играет роль и в других математических науках, главным образом в теории функций комплексной переменной (ср. гл. VIII, стр. 554). Тем самым эта система уравнений устанавливает связь этих математических дисциплин с гидромеханикой.

Теорема Стокса допускает еще и другую механическую интерпретацию. Рассмотрим вместо поля скоростей силовое поле $\boldsymbol{F}(P) = \{F_1(x, y), F_2(x, y)\}$. Тогда криволинейный интеграл

$$\int_C \boldsymbol{F}\,d\boldsymbol{r} = \int_C (F_1\,dx + F_2\,dy),$$

распространенный на любую линию (замкнутую или незамкнутую), дает работу, производимую силовым полем, когда материальная точка описывает линию C. Если кривая C — замкнутая и представляет собой границу области G, то теорема Стокса

$$\oint_C \boldsymbol{F}\,d\boldsymbol{r} = \iint_G (\mathrm{rot})\,\boldsymbol{F}\,dS$$

утверждает, что работа, совершаемая полем при однократном обходе материальной точкой кривой C, равна двойному интегралу скалярной ротации силы $\boldsymbol{F}$ по области G. Если работа силового поля всегда равна нулю, когда материальная точка описывает любую замкнутую кривую, то во всем поле

$$(\mathrm{rot})\,\boldsymbol{F} = 0 \quad \text{или} \quad \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = 0.$$

Обратно, если это условие выполнено во всем поле, то из теоремы Стокса вытекает, что интеграл

$$\oint_C \boldsymbol{F}\,d\boldsymbol{r} = \oint_C (F_1\,dx + F_2\,dy) = 0$$

вдоль любой лежащей в поле замкнутой кривой (ср. § 2, петит в конце n° 3).

Отсюда видно, в согласии с § 1, что *совершенная полем работа не зависит от пути в том и только в том случае, если* $(\mathrm{rot})\,\boldsymbol{F} = 0$ *во всей области* G.

3. Преобразование двойного интеграла. В качестве приложения теоремы Гаусса дадим здесь еще один вывод формулы преобразования двойного интеграла (гл. IV, § 4, n° 1 и петит на стр. 392). Пусть замкнутая область G плоскости xy с граничной кривой C отображена взаимно однозначно на область G' плоскости uv, ограниченную контуром C', с помощью преобразования $x = x(u, v)$, $y = y(u, v)$, — притом с сохранением направления обхода. Предположим еще, что обе области удовлетворяют условиям приложимости теоремы Гаусса. Для того чтобы преобразовать двойной интеграл

$$I = \iint_G f(x, y)\,dx\,dy$$

в интеграл по области G', преобразуем его сначала в криволинейный интеграл по граничной кривой C. Этот интеграл, приводящийся к обыкновенному определенному интегралу, преобразуется сразу в криволинейный интеграл по контуру C' области G', а этот последний интеграл преобразуется по теореме Гаусса опять в двойной интеграл по области G'. В итоге должна получиться формула преобразования двойного интеграла I по области G в интеграл по области G'.

Для того чтобы осуществить этот план, введем в качестве вспомогательной функции какую-нибудь первообразную $A(x, y)$ для функции $f(x, y)$ по переменной x, так что

$$A_x = f(x, y).$$

Тогда, по теореме Гаусса,

$$I=\iint_G f(x, y)\,dx\,dy=\iint_G A_x\,dx\,dy=\oint_C A(x, y)\,dy.$$

Подставим теперь в интеграл, стоящий справа, вместо x и y их выражения $x(u, v)$ и $y(u, v)$ через u и v; тогда интеграл вдоль границы C области G перейдет в интеграл вдоль границы C' области G' и, стало быть,

$$I=\oint_{C'} A(x, y)(y_u\,du+y_v\,dv).$$

Полученный интеграл по контуру C' мы вновь преобразуем по теореме Гаусса (на этот раз в обратном направлении) в двойной интеграл по области G':

$$\oint_{C'}(Ay_u\,du+Ay_v\,dv)=\iint_{G'}\{(Ay_v)_u-(Ay_u)_v\}\,du\,dv.$$

В силу равенств

$$(Ay_v)_u=A_u y_v+Ay_{vu}, \qquad (Ay_u)_v=A_v y_u+Ay_{uv},$$

а также

$$A_u=A_x x_u+A_y y_u, \qquad A_v=A_x x_v+A_y y_v, \qquad A_x=f(x, y),$$

подынтегральное выражение приводится к следующему виду:

$$(Ay_v)_u-(Ay_u)_v=(x_u y_v-x_v y_u)f(x(u, v), y(u, v)).$$

Итак, получаем окончательно искомую формулу:

$$I=\iint_G f(x, y)\,dx\,dy=\iint_{G'} f(x(u, v), y(u, v))\cdot(x_u y_v-x_v y_u)\,du\,dv.$$

§ 4. Интеграл по поверхности

Теория интегрирования функций от трех аргументов включает в себя, кроме тройных интегралов (интегралов по трехмерной области) и криволинейных интегралов (интегралов по одномерной области), еще один, третий, тип интегралов — интегралы по поверхности. Прежде чем определить это понятие, мы для большей ясности предпошлем некоторые общие соображения, которые вместе с тем помогут уточнить изложенные ранее идеи, в особенности те, что относятся к двойным интегралам.

1. Интегрирование по ориентированной области. Начнем с обыкновенного определенного интеграла $\int_a^b f(x)\,dx$ от функции одного аргумента. Областью интегрирования является промежуток от $x=a$

до $x=b$. На основании определения интеграла было установлено соотношение

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

которое можно выразить словесно так: области интегрирования C, т. е. рассматриваемому промежутку приписывается определенное направление или, как принято говорить, определенная ориентация. Если изменить ориентацию промежутка интегрирования на обратную, то значение интеграла умножается на -1. Это свойство интеграла можно записать и в следующем виде:

$$\int_{-C} f(x)\,dx = -\int_{+C} f(x)\,dx,$$

где область интегрирования обозначена символом $+C$, если она пробегается по направлению от a к b, и символом $-C$, если она пробегается по направлению от b к a.

В теории криволинейных интегралов (на плоскости и в пространстве) мы тоже сочли необходимым приписать дуге кривой C, вдоль которой производится интегрирование, определенное направление обхода или ориентацию таким образом, что при изменении ориентации на противоположную значение интеграла умножается на -1. Ясно, что для интегралов по многомерным областям целесообразно ввести аналогичные соглашения об ориентации и знаке и дополнить ими данные выше определения кратных интегралов.

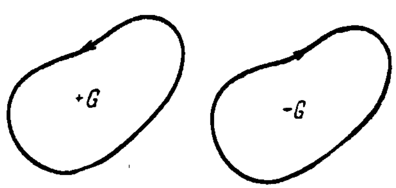


Рис. 89.

а) Двойной интеграл по ориентированной области на плоскости xy. В т. I, гл. V, § 2, n°1 мы ввели понятие ориентации плоской области G, связав эту ориентацию с направлением обхода границы C области, и условились приписывать площади области определенный знак — положительный или отрицательный, смотря по тому, положительно или отрицательно направление обхода ее границы C. Плоская область, границе которой присвоено определенное направление обхода, называется *ориентированной областью* (рис. 89); мы будем ее называть *положительно ориентированной*, если граничной кривой присвоено *положительное направление обхода*, и *отрицательно ориентированной* в противном случае. В свое время мы выразили площадь области G двойным интегралом $\iint\limits_G dx\,dy$.

В соответствии со сказанным выше, мы теперь припишем области определенную ориентацию, и если ее контуру C присвоено положительное направление обхода, то мы область будем обозначать через

$+G$, а ее граничную кривую — через $+C$; если же контуру присвоено отрицательное направление обхода, то область будем обозначать через $-G$, а ее граничную кривую через $-C$. Обозначая абсолютную величину площади области через $|F|$, мы теперь введем дополнительное определение:

$$\iint\limits_{+G} dx\,dy = |F|, \qquad \iint\limits_{-G} dx\,dy = -|F|.$$

С другой стороны, площадь области G выражается также криволинейным интегралом:

$$|F| = -\oint\limits_C y\,dx = \oint\limits_C x\,dy = \frac{1}{2}\oint\limits_C (x\,dy - y\,dx).$$

Полезно убедиться на конкретных примерах, что эти три выражения дают всегда положительный результат. Проверим, например, что $-\oint\limits_C y\,dx$ по контуру $+C$ квадрата $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ есть положительное число. Действительно, на обеих вертикальных сторонах $dx = 0$; интеграл вдоль стороны $y = 0$ тоже обращается в нуль; остается интеграл вдоль четвертой стороны, но на ней $y = 1$, а $dx < 0$.

В дальнейшем, если не сделано специальной оговорки, мы под областью G будем всегда подразумевать *положительно ориентированную область*.

В соответствии с этим мы вводим для всякого двойного интеграла следующее определение:

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_{+G} f(x, y)\,dx\,dy = -\iint\limits_{-G} f(x, y)\,dx\,dy.$$

Это определение в точности соответствует тем положениям, которые были ранее установлены для обыкновенных интегралов и для криволинейных интегралов. Оно представляет собой не доказанный факт, а только соглашение, оправдываемое соображениями целесообразности.

б) Преобразование двойного интеграла по ориентированной области. Поясним геометрически пользу этого соглашения. При взаимно однозначном отображении области G плоскости xy на область G' плоскости uv мы видели, что площадь области G выражается в новых координатах интегралом

$$\iint\limits_G dx\,dy = \iint\limits_{G'} \frac{\partial(x, y)}{\partial(u, v)}\,du\,dv \qquad (*)$$

при условии, что якобиан $\frac{\partial(x, y)}{\partial(u, v)}$ положителен во всех точках области G'. А как же обстоит дело, если якобиан отрицателен во всей области G'? Мы знаем, что при положительном якобиане ориентация (т. е. направление обхода границы у областей G и G') одинакова, между тем как при отрицательном якобиане обе эти области имеют противоположные ориентации. Поэтому в случае отрицательного якобиана формула (*) неверна, если рассматривать двойной интеграл без учета ориентации области интегрирования. Однако *эта формула сохраняет силу и в случае отрицательного якобиана*, если под G разуметь область *ориентированную* (положительно или отрицательно), а под G' ту *ориентированную* область, в которую переходит G в результате данного преобразования. Дело в том, что при отрицательном якобиане ориентация области изменяется и, в силу нашего соглашения о знаке двойного интеграла по ориентированной области, это обстоятельство погашает знак якобиана.

Ясно, что и общая формула преобразования

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_{G'} f(x, y)\frac{\partial(x, y)}{\partial(u, v)}\,du\,dv$$

верна независимо от того, является ли якобиан $\frac{\partial(x, y)}{\partial(u, v)}$ всюду положительным или всюду отрицательным в области интегрирования, если интегралы берутся по *ориентированным* областям и если наше преобразование переводит ориентированную область G в ориентированную область G'.

Стало быть, справедливость формулы преобразования двойного интеграла во всех без исключения случаях достигнута только благодаря введению понятий ориентации области и интегрирования по ориентированной области с соответствующим правилом знаков.

Однако, формула преобразования может оказаться неверной, если якобиан не сохраняет постоянного знака в области интегрирования; в этом случае не обеспечена однозначная обратимость отображения.

Геометрическое определение ориентации можно дать еще и другим путем, без использования границы области. Рассмотрим сначала какую-либо точку P области и припишем этой точке некоторое направление вращения, которое можно, например, представить как направление обхода малой окружности с центром в P. Область G называется ориентированной, если каждой ее точке отнесено в этом смысле направление вращения и если при непрерывном переходе от точки к точке направление вращения сохраняется.

в) Ориентация поверхности в пространстве. С помощью этой идеи можно также приписать ориентацию всякой связной поверхности или связному куску поверхности в пространстве xyz. Сначала приписывают направление вращения какой-либо точке P этой поверхности, окружая эту точку малым контуром, лежащим на поверхности и выбрав определенное направление обхода этого контура. Затем перемещают точку P вместе с присвоенным ей направлением вращения непрерывно по поверхности; таким путем можно отнести определенное направление вращения (или обхода) всякой точке поверхности (возможные исключительные случаи мы обсудим ниже). Поверхность, снабженная таким путем определенным направлением вращения (обхода), называется *ориентированной поверхностью* (рис. 90).

Процесс присвоения ориентации поверхности в пространстве можно себе лучше уяснить с помощью следующих соображений. Кусок поверхности имеет, как правило, две разные стороны, которые удобнее всего различать, называя одну из них положительной, а другую — отрицательной. Какую из двух сторон назвать положительной, а какую отрицательной, само по себе безразлично. Например, в качестве положительной стороны плоскости xy можно выбрать ту сторону, на которую указывает положительная ось z.

Положительную сторону поверхности S можно отметить, строя от каждой точки поверхности некоторый вектор, направленный в положительную сторону, например, нормальный вектор, если в этой точке существует единственная нормаль. *Если этот нормальный вектор* $\mathfrak{n}$

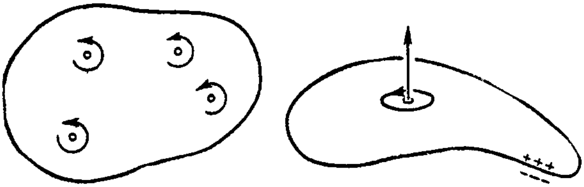


Рис. 90.

и ориентация поверхности, т. е. направление обхода, присвоенное начальной точке вектора $\mathfrak{n}$ *на поверхности, образуют правый винт* (ср. гл. I, § 1, n° 1), *то поверхность называется положительно ориентированной, если же они образуют левый винт, то поверхность называется отрицательно ориентированной.* Другими словами, поверхность называется положительно (отрицательно) ориентированной, если ее вместе с присвоенным ей направлением вращения можно перевести с помощью непрерывной деформации в положительно ориентированную плоскость xy и притом так, что направление ее положительной нормали перейдет в направление положительной (отрицательной) оси z. Отсюда видно, что, приписав заранее определенный знак каждой стороне поверхности, мы подготовили естественный путь для присвоения знака ориентации этой поверхности. Затруднения, которые начинающий, возможно, будет испытывать в этих рассуждениях, коренятся исключительно в том, что здесь речь идет не о доказательствах, а об *определениях* или *соглашениях*, которые найдут свое оправдание впоследствии — в успешном упрощении дальнейших рассуждений.

Присвоить куску поверхности S определенную ориентацию можно также и следующим путем, исходя из его параметрического представления $x=x(u, v)$, $y=y(u, v)$, $z=z(u, v)$. Эти уравнения отображают некоторую определенную область B плоскости uv на наш кусок поверхности S. Если мы установим на плоской области B какую-нибудь ориентацию, то в результате отображения эта ориентация переносится на поверхность S. Тем самым процесс ориентировки поверхности S осуществлен.

г) Поверхности, которым невозможно приписать ориентацию. Необходимо отметить, что в пространстве существуют такие поверхности, которым *принципиально невозможно приписать никакой ориентации*, так как на них нельзя отличить двух раздельных сторон. Простейшая поверхность этого типа была открыта Мёбиусом и называется *листом Мёбиуса* (рис. 91). Эту поверхность легко изготовить из продолговатой прямоугольной полоски бумаги. Для этого надо склеить оба узких конца полоски, повернув предварительно один из концов на 180° вокруг длинной средней линии прямоугольника. Этот лист Мёбиуса обладает следующим свойством: если выйти из какой-либо его точки, лежащей, например, на средней линии, и двигаться вдоль этой средней линии, то, совершив полный оборот, вернемся в исходную точку, но уже на *противоположной* стороне поверхности. Если движущаяся точка, совершая это перемещение, непрерывно переносит вместе с собой некоторую окружающую ее ориентированную замкнутую кривую, сохраняя ее ориентацию, то ясно, что возвращение в исходную точку сопровождается изменением ориентации кривой на противоположную. Стало быть, на такой поверхности можно перейти с одной ее стороны на другую, не переходя края, так что невозможно присвоить этой поверхности какую-либо ориентацию в описанном выше смысле. Такие поверхности, не поддающиеся ориентировке, исключаются из рассмотрения в последующем изложении.

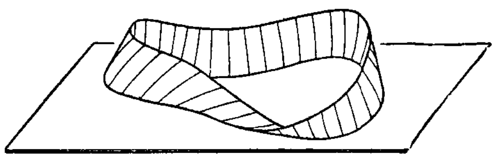


Рис. 91.

Нетрудно составить параметрические уравнения поверхности типа Мёбиуса. Для этой цели возьмем в плоскости xy окружность $x = r\cos u$, $y = r\sin u$. Через точку P окружности, соответствующую значению u параметра, проведем в плоскости, содержащей ось z и радиус-вектор точки P, прямую, образующую с осью z угол $\frac{1}{2}u$ и с радиус-вектором точки P угол $\frac{1}{2}\pi - \frac{1}{2}u$. Направляющий вектор этой прямой

$$l = \left\{\sin\frac{u}{2}\cos u,\ \sin\frac{u}{2}\sin u,\ \cos\frac{u}{2}\right\},$$

а ее параметрические уравнения с параметром v таковы:

$$x = r\cos u + v\sin\frac{u}{2}\cos u,\quad y = r\sin u + v\sin\frac{u}{2}\sin u,\quad z = v\cos\frac{u}{2}. \qquad (*)$$

Здесь u рассматривается как постоянная, а параметр v на прямой равен алгебраической величине отрезка от точки P окружности до переменной точки прямой.

Будем теперь изменять также и параметр u в промежутке $0 \leqslant u \leqslant 2\pi$. Тогда отрезок прямой $-1 \leqslant v \leqslant 1$ длиною 2 единицы опишет поверхность типа листа Мёбиуса и уравнения (*) дают ее параметрическое представление с параметрами u и v.

д) Ориентированная область пространства и тройной интеграл по такой области. Трехмерные области в пространстве тоже можно снабдить определенной ориентацией. Это можно сделать следующим образом. Рассмотрим пространственную область G, ограниченную замкнутой поверхностью S. Назовем положительной стороной поверхности ту из ее сторон, которая обращена *внутрь* области, а стало быть, нормальный вектор, направленный внутрь, определяет положительное направление нормали. Если мы присвоим поверхности такую ориентацию, т. е. такое направление вращения, которое образует со внутренней нормалью *правый винт*, то будем называть *пространственную область G положительно ориентированной* (на рис. 92 изображен положительно ориентированный шар); если же присвоить поверхности такую ориентацию, которая образует с внутренней нормалью левый винт, то область G будем считать отрицательно ориентированной. Например, куб $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$, $0 \leqslant z \leqslant 1$ будем иметь положительную ориентацию, если присвоить положительную ориентацию его основанию, лежащему в плоскости xy.

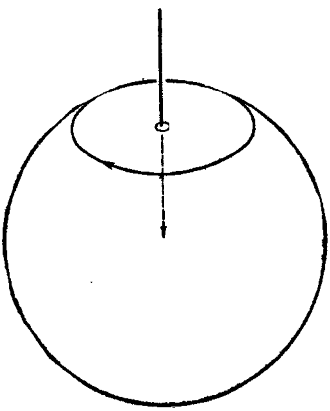


Рис. 92.

В пространстве целесообразно также приписывать объему области G положительный или отрицательный знак, смотря по тому, присвоена ли области положительная или отрицательная ориентация. Соответственно этому условимся, что тройной интеграл по ориентированной области изменяет свой знак, если ориентация области изменяется на противоположную:

$$\iiint\limits_{-G} f(x, y, z)\, dx\, dy\, dz = -\iiint\limits_{+G} f(x, y, z)\, dx\, dy\, dz.$$

Только при наличии такого соглашения формула преобразования

$$\iiint\limits_{G} f(x, y, z)\, dx\, dy\, dz = \iiint\limits_{G'} f(x, y, z) \frac{\partial (x, y, z)}{\partial (u, v, w)}\, du\, dv\, dw$$

становится справедливой и в том случае, если якобиан $\dfrac{\partial (x, y, z)}{\partial (u, v, w)}$ отрицателен по всей области G. Действительно, таким же методом каким это сделано для преобразования функции двух переменных (гл. III, § 3, конец n° 6), можно доказать, что в случае всюду отрицательного якобиана отображение области G на область G' приводит к изменению знака ориентации пространственной области. В силу соглашения о знаке интеграла по ориентированной области, тройной интеграл должен при этом изменить знак. Таким образом, отрицательный знак якобиана погашается изменением знака интеграла.

2. Определение интеграла по поверхности. После этих подготовительных соображений возможно уже дать общее определение интеграла по поверхности (поверхностного интеграла). Пусть в некоторой области G пространства xyz определены три функции $F_1(x, y, z)$, $F_2(x, y, z)$, $F_3(x, y, z)$ в качестве координат вектор-функции точки $\boldsymbol{F}=\boldsymbol{F}(P)$. Другими словами, пусть задано векторное поле

$$\boldsymbol{F}(P)=\{F_1(x, y, z),\ F_2(x, y, z),\ F_3(x, y, z)\}.$$

Рассмотрим кусок поверхности S, имеющий своей проекцией на плоскость xy замкнутую область B этой плоскости и заданный уравнением $z=z(x, y)$, где $z(x, y)$ — однозначная функция, и предположим, что на S установлена определенная ориентация, которая в результате проектирования на плоскость xy переносится на проекцию B куска поверхности S. Обозначим через $\boldsymbol{n}$ единичный нормальный вектор к поверхности S, направленный так, что в сочетании с ориентацией поверхности он образует правый винт. Разобьем кусок поверхности S на m частей (ячеек) $S_1, S_2, \ldots, S_m$ и обозначим площади этих ячеек через $\Delta S_1, \Delta S_2, \ldots, \Delta S_m$. Проектируя эти ячейки на плоскость xy, мы разобьем и область B на m ячеек $B_1, B_2, \ldots, B_m$, площади которых обозначим через $\Delta B_1, \Delta B_2, \ldots, \Delta B_m$, причем эти m ячеек покрывают всю область B однократно и без пробелов. Площади ΔS_k будем считать положительными; поэтому мы должны приписать площадям ΔB_k положительный или отрицательный знак, смотря по тому, какую ориентацию (положительную или отрицательную) получили при проектировании ячейки B_k и область B в плоскости xy. Площади $\Delta S_1, \Delta S_2, \ldots, \Delta S_m$ связаны с площадями $\Delta B_1, \Delta B_2, \ldots, \Delta B_m$ соотношениями вида

$$\Delta B_k = q_k \Delta S_k,$$

причем, когда ячейка S_k стягивается в точку (x, y, z) поверхности, то величина q_k имеет своим пределом косинус угла $\gamma=\gamma(x, y, z)$, образуемого нормальным вектором $\boldsymbol{n}$ с положительной осью z. Выберем на ячейке S_k поверхности какую-либо ее точку (x_k, y_k, z_k), так что $z_k=z(x_k, y_k)$, и построим сумму

$$\sum_{k=1}^{m} F_3(x_k, y_k, z_k)\Delta B_k = \sum_{k=1}^{m} F_3(x_k, y_k, z_k) q_k \Delta S_k.$$

Заставим теперь число m ячеек разбиения неограниченно возрастать, и притом так, чтобы наибольший поперечник ячеек S_k (а вместе с тем и ячеек B_k) стремился к нулю. Тогда написанная выше сумма будет стремиться к пределу, который обозначают символом

$$\iint_S F_3(x,\ y,\ z)\,dx\,dy \quad \text{или} \quad \iint_S F_3(x,\ y,\ z)\cos\gamma\,dS.$$

Этот предел называют *поверхностным интегралом, взятым по куску поверхности* S или, короче, *интегралом по поверхности* S. Сразу ясно, что этот предел действительно существует, так как получающийся интеграл можно рассматривать как обычный двойной интеграл по ориентированной плоской области B, а именно как двойной интеграл

$$\iint_B F_3(x,\ y,\ z(x,\ y))\,dx\,dy.$$

Переходя к обобщению введенного сейчас понятия интеграла по поверхности, подчеркнем, что для этого обобщения, а также и для приложений весьма существенно, чтобы область B была ориентированной областью.

Если кусок поверхности S может быть также задан с помощью однозначной функции $x=x(y,\ z)$ или однозначной функции $y=$ $=y(z,\ x)$, то таким же точно путем можно определить интегралы

$$\iint_S F_1(x,\ y,\ z)\,dy\,dz=\iint_{B'} F_1(x(y,\ z),\ y,\ z)\,dy\,dz=\iint_S F_1(x, y, z)\cos\alpha\,dS$$

и

$$\iint_S F_2(x,\ y,\ z)\,dz\,dx=\iint_{B''} F_2(x,\ y(z,\ x),\ z)\,dz\,dx=\iint_S F_2(x, y, z)\cos\beta\,dS,$$

где B' и B'' — ориентированные области координатных плоскостей yz и zx, полученные проектированием на них ориентированного куска поверхности S, а α и β — углы, образуемые нормальным вектором $\boldsymbol{n}$ поверхности с положительными осями x и y.

Теперь мы дадим определение интеграла по поверхности S от вектор-функции точки

$$\boldsymbol{F}(P)=\{F_1(x, y, z),\ F_2(x, y,\ z),\ F_3(x,\ y,\ z)\}$$

как суммы трех определенных выше поверхностных интегралов:

$$\iint_S \{F_1(x,\ y, z)\,dy\,dz+F_2(x,\ y,\ z)\,dz\,dx+F_3(x,\ y,\ z)\,dx\,dy\}=$$

$$=\iint_S \{F_1(x,\ y,\ z)\cos\alpha+F_2(x,\ y,\ z)\cos\beta+F_3(x,\ y,\ z)\cos\gamma\}\,dS=$$

$$=\iint_S \boldsymbol{Fn}\,dS=\iint_S F_n\,dS.$$

Здесь подынтегральная функция есть скалярное произведение $\boldsymbol{Fn}$ и равна проекции F_n вектора поля на направление положительной нормали к поверхности.

Если поверхность задана в параметрическом виде уравнениями $x=x(u, v)$, $y=y(u, v)$, $z=z(u, v)$, причем ориентированному куску поверхности S соответствует ориентированная область B плоскости uv, то наш поверхностный интеграл можно записать в следующей форме:

$$\iint\limits_{B}\left\{F_1(x, y, z)\frac{\partial(y, z)}{\partial(u, v)}+F_2(x, y, z)\frac{\partial(z, x)}{\partial(u, v)}+F_3(x, y, z)\frac{\partial(x, y)}{\partial(u, v)}\right\}du\,dv.$$

Стало быть, мы его привели к виду обыкновенного двойного интеграла по области B плоскости uv.

Теперь легко освободиться от сделанного выше специального допущения по поводу расположения поверхности относительно плоскостей координат. Это — допущение, что поверхность выражается уравнением $z=z(x, y)$, где $z(x, y)$ есть однозначная функция. Теперь мы только предположим, что ориентированную поверхность S можно разбить с помощью конечного числа гладких дуг на конечное число частей $S_1, S_2, \ldots, S_m$ таким образом, что каждая из этих частей может быть задана однозначными функциями $x=x(y, z)$ или $y=y(z, x)$, или $z=z(x, y)$. Тогда можно взять поверхностный интеграл, согласно данному выше определению, по каждой из частей S_k, и интеграл по ориентированной поверхности S определяем как сумму этих m поверхностных интегралов. В том исключительном случае, когда какая-либо часть поверхности S или вся она является куском цилиндрической поверхности с образующими, перпендикулярными к одной из плоскостей координат, так что ее проекция на эту плоскость будет уже не двумерной областью, а линией, то соответствующий член подынтегральной функции можно опустить, ибо, когда область интегрирования вырождается в линию, двойной интеграл обращается в нуль.

Если, например, S является *замкнутой* поверхностью, то проекции ее различных частей S_k накладываются частично друг на друга и имеют различные ориентации. Проще всего обстоит дело, если поверхность или заданный кусок поверхности отображается как целое взаимно однозначно на некоторую ориентированную область B плоскости uv с помощью параметрического задания $x=x(u, v)$, $y=y(u, v)$, $z=z(u, v)$. Тогда нет надобности разбивать поверхность S на части, и для определения интеграла по поверхности можно воспользоваться данным выше его параметрическим представлением, которое остается в силе.

3. Физическое истолкование интеграла по поверхности. Понятие интеграла по поверхности тоже можно истолковать физически, применяя его к изучению стационарного движения несжимаемой жидкости, на сей раз в пространстве, причем объемную плотность жидкости примем равной единице. Рассмотрим векторное поле скорости течения жидкости и в соответствии с этим будем писать $\boldsymbol{v}$ вместо $\boldsymbol{F}$:

$$\boldsymbol{v}(P)=\{v_1(x, y, z),\ v_2(x, y, z),\ v_3(x, y, z)\},$$

Тогда скалярное произведение $\boldsymbol{vn}$ дает в каждой точке поверхности S проекцию v_n скорости жидкости на направление положительной нормали к поверхности. Выражение

$$\boldsymbol{vn}\Delta S_k = \{v_1(x_k, y_k, z_k)\cos\alpha_k + v_2(x_k, y_k, z_k)\cos\beta_k + \\ + v_3(x_k, y_k, z_k)\cos\gamma_k\}\,\Delta S_k$$

приближенно равно количеству жидкости, протекающей через элемент поверхности ΔS_k с отрицательной стороны на положительную, причем это количество может, конечно, оказаться и отрицательным. Поэтому интеграл по поверхности S

$$\iint_S \boldsymbol{vn}\,dS = \iint_S (v_1 dy\,dz + v_2 dz\,dx + v_3 dx\,dy)$$

дает общее количество жидкости, протекающее в единицу времени через поверхность S с отрицательной ее стороны на положительную. Отсюда видно, какую важную роль играет при математическом описании движения жидкости различение между положительной и отрицательной сторонами поверхности. В гидромеханике интеграл $\iint_S \boldsymbol{vn}\,dS$ принято называть *потоком вектора* $\boldsymbol{v}$ *через поверхность* S; этот термин подсказан объясненным выше физическим смыслом этого интеграла.

По аналогии с этим в любом векторном поле $\boldsymbol{F}(P)$ поверхностный интеграл $\iint_S \boldsymbol{Fn}\,dS$ называют *потоком вектора поля через поверхность* S с отрицательной на положительную ее сторону. В частном случае *силового поля* вектор $\boldsymbol{F}$ указывает своим направлением направление так называемой силовой линии, а его модуль дает величину силы в каждой точке поля. Поверхностный интеграл $\iint_S \boldsymbol{Fn}\,dS$ называют тогда *силовым потоком.*

§ 5. Интегральные теоремы Гаусса и Грина в пространстве

1. Теорема Гаусса[1] **в пространстве.** С помощью понятия интеграла по поверхности можно теперь распространить на трехмерное пространство теорему Гаусса, доказанную для плоскости в § 2, п° 1. Суть теоремы Гаусса для плоскости заключается в том, что двойной интеграл по плоской области преобразуется в криволинейный интеграл вдоль граничной кривой этой области. Рассмотрим теперь замкнутую трехмерную область G в пространстве xyz и предположим, как

---

[1]) Эта теорема, как и теорема Гаусса на плоскости, есть частный случай общей теоремы Остроградского, дающей преобразование n-кратного интеграла по области G пространства n измерений в интеграл по $(n-1)$-мерной гиперповерхности, ограничивающей область G. (*Прим. перев.*)

всегда, что ограничивающая ее поверхность S может быть разбита на конечное число частей, из которых каждая имеет непрерывно изменяющуюся касательную плоскость. Кроме того, мы предположим сначала, что всякая прямая, параллельная какой-либо координатной оси и имеющая с областью G общие внутренние точки, пересекает ее граничную поверхность точно в двух точках; впоследствии мы освободимся от этого последнего допущения.

Пусть дана вектор-функция точки (векторное поле)

$$F(P)=\{F_1(x, y, z),\ F_2(x, y, z),\ F_3(x, y, z)\},$$

где функции $F_1(x, y, z)$, $F_2(x, y, z)$, $F_3(x, y, z)$ непрерывны вместе со своими частными производными первого порядка в области G и на ее границе. Рассмотрим сначала тройной интеграл по области G:

$$\iiint_G \frac{\partial F_3(x, y, z)}{\partial z}\,dx\,dy\,dz.$$

Спроектируем область G на плоскость xy; ее проекцией будет некоторая область B этой плоскости. В каждой точке (x, y) области B восставим перпендикуляр к плоскости xy; согласно сделанному выше допущению, он пересечет граничную поверхность в двух точках. Обозначим аппликату точки входа в область G через $z=z_0(x, y)$ и аппликату точки выхода из нее — через $z=z_1(x, y)$. Тогда любой объемный интеграл по области G можно представить в следующем виде:

$$\iiint_G f(x, y, z)\,dx\,dy\,dz=\iint_B dx\,dy\int_{z_0}^{z_1} f(x, y, z)\,dz.$$

Так как в нашем интеграле $f=\frac{\partial F_3}{\partial z}$, то интегрирование по z можно выполнить:

$$\int_{z_0}^{z_1}\frac{\partial F_3}{\partial z}\,dz=F_3(x, y, z_1)-F_3(x, y, z_0),$$

так что

$$\iiint_G \frac{\partial F_3(x, y, z)}{\partial z}\,dx\,dy\,dz=\iint_B F_3(x, y, z_1)\,dx\,dy-$$

$$-\iint_B F_3(x, y, z_0)\,dx\,dy.$$

Если поверхности S присвоена положительная ориентация относительно области G, то область B плоскости xy, рассматриваемая как проекция той части $z=z_0(x, y)$ поверхности S, которая состоит из точек входа, имеет положительную ориентацию, а та ее часть $z=z_1(x, y)$, которая состоит из точек выхода, имеет отрицательную ориентацию. Поэтому два интеграла по области B можно объединить в виде одного интеграла по всей замкнутой поверхности S, и в итоге получится следующая формула:

$$\iiint\limits_G \frac{\partial F_3(x, y, z)}{\partial z}\,dx\,dy\,dz = -\oiint\limits_S F_3(x, y, z)\,dx\,dy.$$

[Колечком, надетым на символ двойного интеграла, отмечено, что он распространяется на *замкнутую* поверхность, а стрелкой показано, что поверхность S ориентирована по своей внутренней нормали.]

Эта формула остается, очевидно, справедливой, если поверхность S содержит цилиндрические части с образующими, параллельными оси z; действительно, эти части вносят в поверхностный интеграл вклад, равный нулю, так как их проекции на плоскость xy состоят из линий.

Аналогичные две формулы получаются для тройных интегралов по области G от функций $\frac{\partial F_1}{\partial x}$ и $\frac{\partial F_2}{\partial y}$. Сложив все три формулы, придем к общей интегральной формуле

$$\iiint\limits_G \left\{\frac{\partial F_1(x, y, z)}{\partial x} + \frac{\partial F_2(x, y, z)}{\partial y} + \frac{\partial F_3(x, y, z)}{\partial z}\right\} dx\,dy\,dz =$$
$$= -\oiint\limits_S \{F_1(x, y, z)\,dy\,dz + F_2(x, y, z)\,dz\,dx + F_3(x, y, z)\,dx\,dy\},$$

которая и выражает *интегральную теорему Гаусса*. Пользуясь обозначениями стр. 406, можно формулу Гаусса записать и в следующем виде:

$$\iiint\limits_G \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\right) dx\,dy\,dz =$$
$$= -\oiint\limits_S (F_1 \cos\alpha' + F_2 \cos\beta' + F_3 \cos\gamma')\,dS.$$

Напомним, что мы поверхность S ориентировали положительно относительно объемной области G, так что α', β', γ' — углы *внутренней* нормали n' с положительными направлениями осей координат.

Формулу Гаусса нетрудно распространить на более общие области. Достаточно потребовать, чтобы область G можно было разбить с помощью конечного числа поверхностей с непрерывно изменяющейся касательной плоскостью на конечное число частичных областей (ячеек) G_k, обладающих нужным нам свойством, что всякая прямая, параллельная какой-либо из осей координат и проходящая через внутреннюю точку ячейки G_k, пересекает границу последней точно в двух точках. Для всякой ячейки G_k теорема Гаусса верна. Выпишем формулы Гаусса для всех G_k и сложим их. Тогда мы слева получим тройной интеграл по всей области G, в правой же части некоторые из слагаемых поверхностных интегралов соединятся в интеграл по всей поверхности S, другие же (взятые по тем поверхностям, с помощью которых выполнено разбиение области G) взаимно уничтожатся. Наконец, как и в плоском случае (стр. 386), достаточно потребовать, чтобы граничная поверхность области G состояла из конечного числа частей, имеющих однозначные проекции на все три координатные плоскости, причем, однако, допускаются такие цилиндрические части, проекциями которых являются линии.

В теореме Гаусса для пространства, как и для плоскости, удобно пользоваться единичным вектором *внешней* нормали $\boldsymbol{n}$ вместо нормального вектора $\boldsymbol{n}'$, направленного внутрь области. При внесении в интегральную формулу $\boldsymbol{n}' = -\boldsymbol{n}$ надо будет изменить знак в правой части, и формула Гаусса примет следующий вид:

$$\iiint\limits_G \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\right) dx\,dy\,dz =$$
$$= \oiint\limits_S (F_1 \cos\alpha + F_2 \cos\beta + F_3 \cos\gamma)\,dS,$$

где теперь α, β, γ — углы, образуемые *внешней* нормалью $\boldsymbol{n}$ с осями координат.

Нетрудно получить векторную запись формулы Гаусса, подобно тому, как это было сделано для плоскости. Вспомним, что функции $F_1(x, y, z)$, $F_2(x, y, z)$, $F_3(x, y, z)$ были заданы как проекции на оси координат вектора поля

$$\boldsymbol{F} = \{F_1, F_2, F_3\}.$$

На основании гл. II, § 7, n° 7 имеем

$$\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = \operatorname{div} \boldsymbol{F}.$$

С другой стороны,

$$F_1 \cos\alpha + F_2 \cos\beta + F_3 \cos\gamma = \boldsymbol{F}\boldsymbol{n} = F_n,$$

т. е. подынтегральная функция поверхностного интеграла равна проекции вектора поля на направление внешней нормали. Таким образом, получаем векторную запись теоремы Гаусса:

$$\iiint\limits_G \operatorname{div} \boldsymbol{F}\,dx\,dy\,dz = \oiint\limits_S \boldsymbol{F}\boldsymbol{n}\,dS = \oiint\limits_S F_n\,dS.$$

В качестве частного случая теоремы Гаусса можно получить формулу для объема области G, ограниченной замкнутой поверхностью, ориентированной по *внешней* нормали. Положим $F_1 = 0$, $F_2 = 0$, $F_3 = z$; тогда $\operatorname{div} \boldsymbol{F} = 1$ и $F_n\,dS = z\cos\gamma\,dS = z\,dx\,dy$, откуда объем

$$V = \iiint\limits_G dx\,dy\,dz = \oiint\limits_S z\,dx\,dy.$$

Аналогично получаем еще два выражения для этого объема:

$$U = \oint\!\!\oint_S x\,dy\,dz = \oint\!\!\oint_S y\,dz\,dx.$$

Замечательно, что круговая замена букв x, y, z не приводит в этих формулах к изменению знака, между тем как в соответствующих формулах для площади плоской фигуры B

$$S = \oint_B x\,dy = -\oint_B y\,dx$$

взаимная замена x и y влечет за собой перемену знака интеграла. Это связано с тем, что на плоскости xy взаимная замена осей x и y изменяет направление вращения на плоскости, между тем как в пространстве круговое замещение положительных направлений осей координат, т. е. замена x на y, y на z и z на x, не превращает правой системы координат в левую.

2. Физический смысл теоремы Гаусса в пространстве. Наглядное толкование пространственной теоремы Гаусса получается, как и на плоскости, при рассмотрении поля скоростей $\boldsymbol{v}(P)$ стационарного течения несжимаемой жидкости плотности 1. Масса жидкости, протекающая через малую площадку границы G изнутри наружу, приближенно равна $v_n\,dS$, где v_n есть проекция вектора-скорости $\boldsymbol{v}$ на направление внешней нормали $\boldsymbol{n}$ в какой-либо точке этой площадки. Следовательно, общая масса жидкости, протекающая в единицу времени через замкнутую поверхность S изнутри наружу, выражается интегралом $\oint\!\!\oint_S v_n\,dS$, взятым по поверхности S. Стало быть, в формуле Гаусса

$$\iiint_G \operatorname{div} \boldsymbol{v}\,dx\,dy\,dz = \oint\!\!\oint_S v_n\,dS,$$

написанной для этого поля, стоящий справа интеграл дает количество жидкости, вытекающей из области G в единицу времени. Этот интеграл преобразовывается в тройной интеграл от дивергенции скорости по объему области G. Отсюда нетрудно выяснить физический смысл $\operatorname{div} \boldsymbol{v}$. Так как, согласно условию, жидкость несжимаема, а ее движение стационарно, то теряемая областью G жидкость должна непрерывно возмещаться, а это значит, что внутри области должны существовать «источники», производящие положительное или отрицательное количество жидкости. Поверхностный интеграл в правой части дает поэтому общую производительность (мощность, дебит) источников области G. Отношение этого поверхностного интеграла к объему области выражает среднюю производительность области G, а предел этого отношения при условии, что область G стягивается к своей точке P, т. е. производная поверхностного интеграла по области, дает удельную производительность источника в точке P. Но эта производная равна производной по области от объемного интеграла $\iiint\limits_{G} \operatorname{div} \boldsymbol{v}\, dx\, dy\, dz$, а эта последняя равна подынтегральной функции, т. е. $\operatorname{div} \boldsymbol{v}$. Отсюда видно, что в поле стационарного потока несжимаемой жидкости $\operatorname{div} \boldsymbol{v}$ выражает удельную производительность источника (или стока) в рассматриваемой точке.

[Быть может, этот результат будет еще нагляднее, если рассматривать вектор поля $\boldsymbol{v}(P)$ как вектор скорости *теплового* потока, т. е. направление этого вектора указывает направление движения (распространения) тепла, а его модуль выражает количество тепла, проходящего в этом направлении в единицу времени через малую площадку, перпендикулярную к $\boldsymbol{v}$. Тогда выполненное выше рассуждение показывает, что в поле теплового потока $\operatorname{div} \boldsymbol{v}$ означает удельную мощность источника (или поглощателя) тепла в рассматриваемой точке.]

Особый интерес вызывает поле $\boldsymbol{F}(P)$, *лишенное источников.* Такое поле характеризуется тем, что во всех его точках

$$\operatorname{div} \boldsymbol{F} = 0.$$

Из теоремы Гаусса вытекает, что в таком поле $\oiint\limits_{S} F_n\, dS = 0$, т. е. в *векторном поле, лишенном источников, поток вектора поля через любую замкнутую поверхность равен нулю.*

Этому факту можно придать несколько иное выражение. Проведем через замкнутую ориентированную кривую C в поле, лишенном источников, две поверхности S_1 и S_2, для которых эта кривая служит общей границей, причем эти две поверхности совместно отграничивают односвязную область G пространства. Запишем интегральную формулу Гаусса для этой области G, но нормальные векторы мы направим, как показано на рис. 93: у поверхности S_1 внутрь области G, а у поверхности S_2 — от этой области наружу так, чтобы направление обхода кривой C составляло правый винт с нормальными векторами обеих поверхностей. Вследствие этого интеграл по S_2 надо взять со знаком плюс, а по S_1 — со знаком минус, и равенство $\oiint\limits_{S} F_n\, dS = 0$ принимает следующий вид:

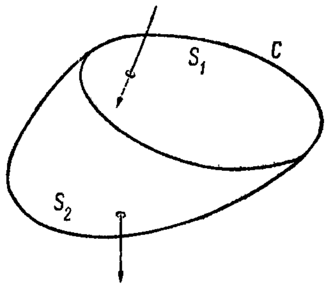


Рис. 93.

$$\iint\limits_{S_2} F_n\, dS - \iint\limits_{S_1} F_n\, dS = 0$$

или

$$\iint\limits_{S_1} F_n\, dS = \iint\limits_{S_2} F_n\, dS.$$

Это значит, что в векторном поле, лишенном источников, поток вектора через две поверхности, имеющие общую граничную кривую, имеет одинаковое значение или (в гидравлической картине) через две поверхности с общей граничной кривой протекает в единицу времени одинаковое количество жидкости. Стало быть, количество протекающей жидкости не зависит от выбора поверхности, «натянутой» на граничную замкнутую кривую C; оно может поэтому зависеть только от выбора контура C. Естественно возникает задача: как выразить этот поток вектора поля, или количество протекающей жидкости, через данные, определяющие контур C? Ответ на этот вопрос будет дан в следующем параграфе теоремой Стокса.

3. Теоремы Грина. Как и в случае двух независимых переменных (§ 2, n° 4), из теоремы Гаусса в пространстве можно вывести две формулы, которые принято называть формулами или теоремами Грина. Для их вывода применим теорему Гаусса (в ее векторной записи) к векторному полю

$$\boldsymbol{F} = u \operatorname{grad} v = \{uv_x,\ uv_y,\ uv_z\}.$$

Тогда в области G

$$\operatorname{div} \boldsymbol{F} = \frac{\partial}{\partial x}(uv_x) + \frac{\partial}{\partial y}(uv_y) + \frac{\partial}{\partial z}(uv_z),$$

а на ее граничной поверхности S

$$F_n = \boldsymbol{F}\boldsymbol{n} = u\,\boldsymbol{n} \operatorname{grad} v = uD_n v.$$

Пользуясь знакомым символом лапласиана $\nabla^2 v = v_{xx} + v_{yy} + v_{zz}$ сразу получаем из теоремы Гаусса первую формулу Грина

$$\iiint\limits_G (u_x v_x + u_y v_y + u_z v_z)\,dx\,dy\,dz + \iiint\limits_G u\nabla^2 v\,dx\,dy\,dz = \underset{S}{\oiint} u\,D_n v\,dS.$$

Проделав те же вычисления для векторного поля $\boldsymbol{F} = v \operatorname{grad} u$, получим параллельную формулу

$$\iiint\limits_G (u_x v_x + u_y v_y + u_z v_z)\,dx\,dy\,dz + \iiint\limits_G v\nabla^2 u\,dx\,dy\,dz = \underset{S}{\oiint} vD_n u\,dS.$$

Вычитая из первой формулы вторую, приходим ко второй формуле Грина:

$$\iiint\limits_G (u\nabla^2 v - v\nabla^2 u)\,dx\,dy\,dz = \underset{S}{\oiint} (uD_n v - vD_n u)\,dS.$$

4. Приложения теорем Гаусса и Грина в пространстве.

а) Преобразование лапласиана к сферическим координатам. Если подставить во вторую формулу Грина $u = 1$, то получится

$$\iiint\limits_G \nabla^2 v\,dx\,dy\,dz = \underset{S}{\oiint} D_n v\,dS.$$

С помощью этой формулы можно выполнить преобразование лапласиана $\nabla^2 v$ к сферическим координатам (r, φ, θ) тем же способом, которым в § 2, n° 6 двумерный лапласиан был преобразован к полярным координатам. Для этого надо выбрать в качестве области G ячейку сетки сферических координат, ограниченную координатными поверхностями r и $r+h$, φ и $\varphi+k$, θ и $\theta+l$. В результате получится

$$\nabla^2 u = \frac{1}{r^2 \sin\theta}\left\{\frac{\partial}{\partial r}(r^2 u_r \sin\theta) + \frac{\partial}{\partial\varphi}\left(\frac{u_\varphi}{\sin\theta}\right) + \frac{\partial}{\partial\theta}(u_\theta \sin\theta)\right\}.$$

Выполнение вычислений (аналогичное § 2, n° 6) предоставляем читателю.

б) Силы объемные и силы поверхностные. Силы, действующие в сплошной среде, можно трактовать двояко: либо как силы, распределенные по объему, либо как силы, действующие на поверхность. Связь между этими двумя точками зрения устанавливается интегральной теоремой Гаусса.

Мы рассмотрим только частный случай сил, действующих в жидкости постоянной плотности $\rho = 1$, если в ней существует скалярное поле давления $p(x, y, z)$. Это значит, что на всякую малую площадку, проходящую через точку (x, y, z), жидкость действует с силой, перпендикулярной к этой площадке и равной $p(x, y, z)$ в расчете на единицу площади. Рассмотрим некоторую область G, лежащую в жидкости и ограниченную поверхностью S. На элемент dS этой поверхности S, для которого $\boldsymbol{n}$ есть единичный вектор внешней нормали, жидкость давит с силой $d\boldsymbol{F} = -p\boldsymbol{n}\,dS$. Проекции этой силы на оси координат получим, умножая ее скалярно на орты осей $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$:

$$dF_1 = -p\boldsymbol{e}_1\boldsymbol{n}\,dS, \qquad dF_2 = -p\boldsymbol{e}_2\boldsymbol{n}\,dS, \qquad dF_3 = -p\boldsymbol{e}_3\boldsymbol{n}\,dS,$$

а проекции результирующей силы $\boldsymbol{F}$ равны соответственно

$$F_1 = -\iint\limits_S p\boldsymbol{e}_1\boldsymbol{n}\,dS, \qquad F_2 = -\iint\limits_S p\boldsymbol{e}_2\boldsymbol{n}\,dS, \qquad F_3 = -\iint\limits_S p\boldsymbol{e}_3\boldsymbol{n}\,dS.$$

Применим теорему Гаусса к каждому из этих поверхностных интегралов в отдельности. Так как $\operatorname{div}(p\boldsymbol{e}_1) = p_x$, $\operatorname{div}(p\boldsymbol{e}_2) = p_y$, $\operatorname{div}(p\boldsymbol{e}_3) = p_z$, то в итоге получим

$$F_1 = -\iiint\limits_G p_x\,dx\,dy\,dz, \quad F_2 = -\iiint\limits_G p_y\,dx\,dy\,dz, \quad F_3 = -\iiint\limits_G p_z\,dx\,dy\,dz,$$

а следовательно результирующая сила, действующая на область G, как целое, есть

$$\boldsymbol{F} = -\iiint\limits_G \operatorname{grad} p\,dx\,dy\,dz.$$

Полученный результат можно выразить следующим образом. Силы, действующие в жидкости из-за наличия скалярного поля давления $p(x, y, z)$, можно рассматривать двояко: 1) как поверхностные силы, которые производят давление на всякую площадку, проходящую через точку (x, y, z) перпендикулярно к ней с поверхностной плотностью $p(x, y, z)$, и 2) как объемные силы, действующие на каждый элемент объема с объемной плотностью $-\operatorname{grad} p$.

Упражнение

1*. Преобразование

$$x_i = x_i(p_1, p_2, p_3) \qquad (i = 1, 2, 3)$$

определяет криволинейную систему координат p_1, p_2, p_3. Известно, что эта система «ортогональна», т. е. если положить $a_{ik} = \dfrac{\partial x_i}{\partial p_k}$, то выполняются следующие соотношения:

$$a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} = 0,$$
$$a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33} = 0,$$
$$a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} = 0.$$

а) Доказать, что

$$\frac{\partial(x_1, x_2, x_3)}{\partial(p_1, p_2, p_3)} = \sqrt{e_1e_2e_3},$$

где

$$e_i = a_{1i}^2 + a_{2i}^2 + a_{3i}^2 \qquad (i = 1, 2, 3).$$

б) Доказать, что

$$\frac{\partial p_i}{\partial x_k} = \frac{1}{e_i}\frac{\partial x_k}{\partial p_i} = \frac{1}{e_i}a_{ki}.$$

в) Выразить лапласиан $\nabla^2 u = u_{x_1x_1} + u_{x_2x_2} + u_{x_3x_3}$ через p_1, p_2, p_3 (с помощью теоремы Гаусса).

г) Выразить $\nabla^2 u$ через фокальные координаты (t_1, t_2, t_3), определенные в упр. 6, гл. III, конец § 3, стр. 176.

§ 6. Теорема Стокса в пространстве

1. Формулировка и доказательство теоремы. Мы уже познакомились в § 2, n° 3 с теоремой Стокса для плоскости xy. В этом параграфе мы рассмотрим теорему Стокса для любой кривой поверхности.

Пусть C — замкнутая кусочно гладкая ориентированная кривая в пространстве, а S — какой-либо кусок поверхности, ограниченный кривой C. Мы будем предполагать, что единичный нормальный вектор $\boldsymbol{n}$ изменяется на S непрерывно или кусочно непрерывно и в сочетании с ориентацией граничной кривой C образует правый винт. Далее, пусть в окрестности поверхности S задано векторное поле

$$\boldsymbol{B}(P) = \{B_1(x, y, z), \quad B_2(x, y, z), \quad B_3(x, y, z)\}.$$

Теорема Стокса утверждает, что

$$\iint_S \boldsymbol{n} \operatorname{rot} \boldsymbol{B}\, dS = \oint_C \boldsymbol{t}\boldsymbol{B}\, dS \qquad \text{или} \qquad \iint_S (\operatorname{rot} \boldsymbol{B})_n\, dS = \oint_C B_{\mathbf{t}}\, ds,$$

где длина дуги s кривой C возрастает в направлении ее ориентации, $\mathbf{t}$ есть единичный касательный вектор этой кривой, а $B_{\mathbf{t}}$ есть проекция вектора $\boldsymbol{B}$ на положительное направление касательной. В координатной записи формула Стокса имеет следующий вид:

$$\iint_S \left\{\left(\frac{\partial B_3}{\partial y} - \frac{\partial B_2}{\partial z}\right) dy\, dz + \left(\frac{\partial B_1}{\partial z} - \frac{\partial B_3}{\partial x}\right) dz\, dx + \left(\frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y}\right) dx\, dy\right\} =$$
$$= \oint_C (B_1\, dx + B_2\, dy + B_3\, dz). \qquad (1)$$

Эта формула дает преобразование стоящего слева интеграла по ориентированной поверхности S в криволинейный интеграл по граничной кривой C этой поверхности, причем ориентация границы C согласована с ориентацией S по правилу правого винта.

Содержание теоремы Стокса можно сделать сразу понятным с помощью следующего рассуждения. Начнем с того, что для любой плоскости теорема уже доказана. Предположим теперь, что S есть многогранная поверхность, составленная из плоских граней-многоугольников, так что и граница C есть многоугольник. Теорему Стокса можно применить к каждой из граней и все полученные формулы сложить. Тогда все криволинейные интегралы, взятые вдоль внутренних ребер многогранной поверхности S, взаимно уничтожатся и мы получим формулу Стокса для многогранной поверхности. Для того чтобы доказать теорему для общего случая, остается только совершить предельный переход от многогранной к любой поверхности S и от многоугольной границы к произвольному кусочно гладкому контуру C. Однако строгое выполнение этого предельного перехода сложно и трудоемко. Поэтому мы ограничимся этим эвристическим замечанием и проведем доказательство с помощью прямого вычисления.

Для краткости введем обозначение

$$\operatorname{rot} \boldsymbol{B} = \boldsymbol{A} = \{A_1,\ A_2,\ A_3\},$$

так что

$$A_1(x,\ y,\ z) = \frac{\partial B_3}{\partial y} - \frac{\partial B_2}{\partial z}, \qquad A_2(x,\ y,\ z) = \frac{\partial B_1}{\partial z} - \frac{\partial B_3}{\partial x},$$
$$A_3(x,\ y,\ z) = \frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y}.$$

Согласно гл. II, § 7, n° 7,

$$\operatorname{div} \boldsymbol{A} = \operatorname{div} \operatorname{rot} \boldsymbol{B} = 0,$$

т. е. поле $A(P)$ лишено источников, а в § 5, n° 2 мы видели, что в таком поле $\iint\limits_S A_n\,dS$ вполне определяется граничной кривой C куска поверхности S и не зависит от выбора этой поверхности. В левой части формулы Стокса стоит именно этот интеграл, и задача состоит в том, чтобы его преобразовать в такое выражение, которое вполне определяется граничной кривой C.

Для этой цели будем исходить из обычного параметрического представления поверхности S с помощью двух параметров u и v, так что она отображается взаимно однозначно на область D плоскости uv. По общему правилу преобразования двойного интеграла имеем

$$\iint\limits_S A_n\,dS = \iint\limits_S (A_1\,dy\,dz + A_2\,dz\,dx + A_3\,dx\,dy) =$$
$$= \iint\limits_D \left\{\left(\frac{\partial B_3}{\partial y} - \frac{\partial B_2}{\partial z}\right)\frac{\partial(y, z)}{\partial(u, v)} + \left(\frac{\partial B_1}{\partial z} - \frac{\partial B_3}{\partial x}\right)\frac{\partial(z, x)}{\partial(u, v)} + \right.$$
$$\left. + \left(\frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y}\right)\frac{\partial(x, y)}{\partial(u, v)}\right\} du\,dv.$$

Подынтегральную функцию в правой части мы преобразуем, собирая вместе члены с B_1, потом с B_2 и, наконец, с B_3. Так, собирая члены, содержащие производные от B_1, получим

$$-\frac{\partial B_1}{\partial y}\left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right) - \frac{\partial B_1}{\partial z}\left(\frac{\partial x}{\partial u}\frac{\partial z}{\partial v} - \frac{\partial z}{\partial u}\frac{\partial x}{\partial v}\right).$$

К этому выражению мы прибавим еще член

$$-\frac{\partial B_1}{\partial x}\left(\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial x}{\partial u}\frac{\partial x}{\partial v}\right),$$

тождественно равный нулю, и тогда совокупность членов подынтегрального выражения, содержащих производные от B_1, запишется так:

$$\frac{\partial x}{\partial v}\left(\frac{\partial B_1}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial B_1}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial B_1}{\partial z}\frac{\partial z}{\partial u}\right) - \frac{\partial x}{\partial u}\left(\frac{\partial B_1}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial B_1}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial B_1}{\partial z}\frac{\partial z}{\partial v}\right) =$$
$$= \frac{\partial B_1}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial B_1}{\partial v}\frac{\partial x}{\partial u}.$$

Таким же путем получаем остальные два члена подынтегральной функции:

$$\frac{\partial B_2}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial B_2}{\partial v}\frac{\partial y}{\partial u} \quad \text{и} \quad \frac{\partial B_3}{\partial u}\frac{\partial z}{\partial v} - \frac{\partial B_3}{\partial v}\frac{\partial z}{\partial u}.$$

Итак,

$$\iint\limits_S A_n\,dS = \iint\limits_D \left(\frac{\partial(B_1, x)}{\partial(u, v)} + \frac{\partial(B_2, y)}{\partial(u, v)} + \frac{\partial(B_3, z)}{\partial(u, v)}\right) du\,dv$$

по ориентированной области D, контур K которой имеет ориентацию, соответствующую ориентации кривой C. Теперь мы воспользуемся формулой (*), выведенной в § 2, n° 5.

Заменив в этой формуле x на u, y на v, u на B_1 и v на x, получим

$$\iint\limits_D \frac{\partial(B_1, x)}{\partial(u, v)}\,du\,dv = \iint\limits_D \left(\frac{\partial B_1}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial B_1}{\partial v}\frac{\partial x}{\partial u}\right) du\,dv =$$
$$= \int\limits_K \left(B_1 \frac{\partial x}{\partial u}\,du + B_1 \frac{\partial x}{\partial v}\,dv\right) = \int\limits_C B_1\,dx = \int\limits_C B_1 \frac{dx}{ds}\,ds,$$

причем ориентации области D и граничных кривых K и C должны быть согласованы. С помощью той же формулы (*) (§ 2, n° 5) преобразуются остальные два двойных интеграла в криволинейные интегралы по контуру C, и сложив все три формулы, приходим к окончательному результату:

$$\iint\limits_S A_n\,dS = \int\limits_C \left(B_1 \frac{dx}{ds} + B_2 \frac{dy}{ds} + B_3 \frac{dz}{ds}\right) ds = \oint\limits_C \boldsymbol{B}\boldsymbol{t}\,ds = \oint\limits_C B_{\boldsymbol{t}}\,ds,$$

а так как $\boldsymbol{A} = \operatorname{rot} \boldsymbol{B}$, то формула Стокса

$$\iint\limits_S (\operatorname{rot} \boldsymbol{B})_n\,dS = \oint\limits_C B_{\boldsymbol{t}}\,ds$$

доказана. Эта формула верна, если вектор-функция точки $\boldsymbol{A} = \operatorname{rot} \boldsymbol{B}$ непрерывна в рассматриваемой области пространства, а поверхность S состоит из одного или нескольких кусков, которые могут быть заданы параметрическими уравнениями вида $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, причем эти функции имеют непрерывные производные первого порядка.

В § 5 в конце n° 1 было показано, что в векторном поле, дивергенция которого тождественно равна нулю, поток вектора поля через кусок поверхности зависит только от выбора контура (граничной кривой) этого куска и не зависит от выбора поверхности, натянутой на данный контур. Возник вопрос: как выразить зависимость поверхностного интеграла (потока) от граничной кривой куска поверхности? Теорема Стокса дает возможность ответить на этот вопрос. В дополнении к этой главе, в § 2 будет доказано, что всякое векторное поле $\boldsymbol{A}(P)$, дивергенция которого тождественно равна нулю, может быть представлено в виде

$$\boldsymbol{A} = \operatorname{rot} \boldsymbol{B}.$$

Стало быть, теорема Стокса и выражает эту зависимость интеграла по куску поверхности S от его граничной кривой.

2. Физический смысл теоремы Стокса. Физический смысл теоремы Стокса в пространстве аналогичен ее физическому истолкованию на плоскости[1]). Векторное поле $\boldsymbol{B}(P)$ и здесь истолковывают как поле скоростей стационарного течения несжимаемой жидкости и интеграл $\oint\limits_C B_t\,ds = \oint\limits_C \boldsymbol{B}\,d\boldsymbol{r}$ называют *циркуляцией* скорости потока вдоль замкнутой кривой C. Теорема Стокса утверждает, что циркуляция скорости вдоль ориентированной замкнутой кривой равна интегралу от ротора скорости по любой поверхности, ограниченной этой кривой, причем ориентация поверхности и ориентация граничной кривой должны быть согласованы по правилу правого винта.

Применим теорему Стокса к элементу поверхности S с непрерывно изменяющейся касательной плоскостью. Разделим обе части формулы Стокса для этого элемента поверхности на его площадь и перейдем к пределу, заставляя этот элемент и его граничную кривую стягиваться к точке P, оставаясь неизменно на взятой поверхности S. Тогда мы в левой части формулы получим производную поверхностного интеграла от $\operatorname{rot}\boldsymbol{B}$ по области интегрирования, равную проекции вектора $\operatorname{rot}\boldsymbol{B}$ на положительную нормаль в точке P поверхности S. В правой части получится величина, которую естественно назвать *удельной циркуляцией* или *плотностью циркуляции* скорости $\boldsymbol{B}$ на поверхности S в ее точке P. Следовательно, проекция вектора $\operatorname{rot}\boldsymbol{B}$ на положительную нормаль поверхности в ее точке P равна плотности циркуляции вектора $\boldsymbol{B}$ на этой поверхности в той же точке P. При этом направление циркуляции и положительная нормаль должны совместно образовать правый винт.

Вместе с тем это рассуждение показывает, что ротор векторного поля имеет смысл, не зависящий от системы координат и, стало быть, действительно является вектором.

Векторное поле $\boldsymbol{B}(P)$ можно также интерпретировать как силовое поле механической или электрической природы. Тогда криволинейный интеграл в правой части формулы Стокса выражает работу, совершаемую полем, когда частица, испытывая его воздействие, описывает кривую C. Теорема Стокса дает преобразование выражения для этой работы в интеграл по поверхности S, ограниченной кривой C, причем подынтегральная функция этого интеграла равна проекции ротора силы поля $\boldsymbol{B}$ на положительное направление нормали.

Из теоремы Стокса можно извлечь новый вывод условия независимости криволинейного интеграла в пространстве от пути интегрирования (ср. также замечание, напечатанное петитом, в конце § 2, n° 3). Основной вопрос заключался в следующем: каким свойством должно обладать векторное поле $\boldsymbol{B}(P)$, чтобы криволинейный $\oint \boldsymbol{B}\,d\boldsymbol{r}$ вдоль любой замкнутой кривой обращался в нуль? Так вот из теоремы

[1]) Заслуживает внимания тот факт, что на плоскости теоремы Гаусса и Стокса отличаются между собой формально лишь знаком, между тем как в пространстве существенно различно не только физическое содержание этих двух теорем, но и их формальное выражение.

Стокса сразу видно, что условие $\operatorname{rot} \boldsymbol{B}=0$ обеспечивает обращение в нуль этого интеграла[1]). Поэтому обращение в нуль ротора или, как говорят, *безвихревой характер* векторного поля является достаточным условием обращения в нуль интеграла $\oint \boldsymbol{B}\, d\boldsymbol{r}$ вдоль любой замкнутой кривой, а следовательно, и достаточным условием независимости криволинейного интеграла от пути. (В § 1 мы видели, что оно является и необходимым условием.) Из § 1 мы знаем, что при этом условии, т. е. в поле, лишенном вихрей, вектор поля $\boldsymbol{B}(P)$ может быть представлен в виде градиента некоторой скалярной функции $U(x, y, z)$, называемой *потенциалом*

$$\boldsymbol{B}=\operatorname{grad} U.$$

Если наше безвихревое поле не имеет к тому же и источников, т. е. и $\operatorname{div} \boldsymbol{B}=0$, то функция $U(x, y, z)$ удовлетворяет дифференциальному уравнению

$$\operatorname{div} \operatorname{grad} U=0$$

или, в координатной записи,

$$\nabla^2 U \equiv \frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}=0 .$$

Стало быть, векторное поле $\boldsymbol{B}(P)$, лишенное вихрей и источников, имеет потенциал, удовлетворяющий уравнению Лапласа $\nabla^2 U=0$, с которым мы уже встречались в т. I, гл. X и которое принадлежит к числу важнейших дифференциальных уравнений анализа.

§ 7. Принципиальные соображения о связи между дифференцированием и интегрированием в пространстве многих переменных

Полезно резюмировать факты, установленные в этой главе, с единой, общей точки зрения.

В случае одной независимой переменной мы в гл. II первого тома сочли основной теоремой дифференциального и интегрального исчисления взаимную связь между дифференцированием и интегрированием. Эта основная теорема (формула Ньютона — Лейбница) выражается для функции одного аргумента следующим образом: если функция $f(x)$ непрерывна в замкнутой области $a \leqslant x \leqslant b$, а $F(x)$ — ее первообразная, то

$$\int_a^b f(x)\, dx=F(b)-F(a). \tag{1}$$

---

[1]) Впрочем, здесь содержится молчаливое предположение, что на эту кривую можно действительно «натянуть» поверхность описанного выше типа. Так как это может приводить к затруднениям (например, для кривых, имеющих узловые точки), то доказательство этой теоремы, данное в § 1, следует предпочесть.

Обратно, для всякой функции $F(x)$, имеющей непрерывную производную, возможно построить соответствующую ей функцию $f(x)=F'(x)$, удовлетворяющую равенству (1). Для той цели, которую мы себе теперь ставим, существенной является первая часть этой основной теоремы, а именно возможность приведения интеграла по одномерной области к выражению $F(b)-F(a)$, зависящему только от граничных точек области интегрирования, которые, можно сказать, образуют область нулевого измерения. Другими словами, если подынтегральная функция является производной от некоторой функции $F(x)$, то интеграл по одномерной области приводится к выражению, вычисляемому с помощью функции $F(x)$ и зависящему только от нульмерной границы области интегрирования.

Нечто вполне аналогичное формуле Ньютона — Лейбница для одномерного интеграла дают различные интегральные теоремы, доказанные выше для многомерных областей. В каждой такой теореме интеграл, взятый по какой-либо области, лежащей в заданном пространстве, будь то по дуге кривой, куску поверхности или по объему, приводится к выражению, зависящему только от границы области интегрирования. Возьмем, например, теорему Гаусса для двух измерений:

$$\iint\limits_G \left(\frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}\right) dx\,dy=\oint\limits_C (F_1\,dy-F_2\,dx).$$

Эта теорема устанавливает следующее: если подынтегральная функция интеграла $\iint\limits_G f(x, y)\,dx\,dy$ по замкнутой области G имеет вид

$$f(x, y)=\frac{\partial}{\partial x}F_1(x, y)+\frac{\partial}{\partial y}F_2(x, y),$$

то двойной интеграл можно преобразовать в выражение, зависящее только лишь от одномерной границы двумерной области G, а именно в криволинейный интеграл вдоль граничной кривой C. Таким образом, теорема Гаусса дает возможность уменьшить число измерений области интегрирования на единицу. Вместо полученного выше выражения $F(b)-F(a)$, определяемого границей промежутка интегрирования, здесь возникает криволинейный интеграл вдоль границы плоской области G. Конечно, здесь не может быть и речи о первообразной функции F. Этой единой первообразной функции здесь в некотором смысле соответствует вектор-функция точки $F(P)=\{F_1(x, y), F_2(x, y)\}$. С другой стороны, применение теоремы Гаусса требует, чтобы подынтегральная функция $f(x, y)$ двойного интеграла была представлена в виде дивергенции $\frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}$. Это требование оставляет еще большой простор для выбора первообразного векторного поля $F=\{F_1, F_2\}$, тогда как для обычного одномерного интеграла первообразная функция допускает только произвол в выборе аддитивной постоянной.

Нетрудно показать, что для заданной подынтегральной функции $f(x, y)$ двойного интеграла действительно существует большой произвол в выборе функций $F_1(x, y)$ и $F_2(x, y)$, удовлетворяющих уравнению $\frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}=f(x, y)$. В самом деле, можно, например, положить $F_2(x, y)$ равной тождественно нулю или равной любой известной функции f, а затем определить соответствующую функцию $F_1(x, y)$ из уравнения $\frac{\partial F_1}{\partial x}=f-\frac{\partial F_2}{\partial y}$. Тогда $F_1=\int\left(f-\frac{\partial F_2}{\partial y}\right)dx$, где y при интегрировании по x рассматривается как постоянный параметр. Полученное таким путем векторное поле можно еще сложить с произвольным векторным полем, дивергенция которого равна нулю, и мы опять получим первообразное векторное поле.

В двумерном пространстве, помимо теоремы Гаусса и теоремы Стокса, которые в этом случае по существу равносильны, существует еще одно обобщение формулы Ньютона — Лейбница, а именно теорема, выражающая условие независимости криволинейного интеграла от пути интегрирования. Дело в том, что в двумерном пространстве существуют замкнутые одномерные многообразия, имеющие нульмерные границы; это дуги кривых с двумя граничными точками, и задача состоит в том, чтобы криволинейный интеграл $\int \boldsymbol{F}\,d\boldsymbol{r}$ по такой дуге привести к выражению, зависящему только от граничных точек. Мы уже знаем из § 1, что такое приведение возможно в том и только в том случае, если существует такая скалярная «первообразная» функция (потенциал) $U(x, y)$, что вектор $\boldsymbol{F}$ поля является ее градиентом

$$\boldsymbol{F}=\operatorname{grad} U.$$

Тогда

$$\int\limits_{(P_0)}^{(P)} \boldsymbol{F}\,d\boldsymbol{r}=U(P)-U(P_0)=U(\xi, \eta)-U(\xi_0, \eta_0).$$

Сравнивая этот результат с формулой Ньютона — Лейбница для обычного интеграла, убеждаемся, что выражение подынтегральной скалярной функции в виде производной заменяется здесь представлением подынтегральной векторной функции в виде градиента, и роль первообразной функции играет потенциал этого градиента. Однако, есть и существенное различие в этом вопросе между обычным интегралом и криволинейным: не у всякого криволинейного интеграла подынтегральная вектор-функция точки $\boldsymbol{F}(P)$ может быть представлена как градиент; такое представление возможно в том и только в том случае, если выполняется условие интегрируемости $\frac{\partial F_2}{\partial x}=\frac{\partial F_1}{\partial y}$.

В пространстве трех измерений дело обстоит вполне аналогично. Теорема Гаусса преобразует тройной интеграл по трехмерной области, ограниченной замкнутой поверхностью, в интеграл по этой граничной поверхности, которая представляет собой замкнутую двумерную область, вложенную в трехмерное пространство и не имеющую граничной кривой. Это преобразование связано с тем, что подынтегральную функцию тройного интеграла представляют в виде дивергенции векторного поля

$$\boldsymbol{F}(P)=\{F_1(x,\ y,\ z),\quad F_2(x,\ y,\ z),\quad F_3(x,\ y,\ z)\}$$

и эта вектор-функция точки играет здесь в известном смысле роль первообразной функции.

Как и в теореме Гаусса для двумерной области, и здесь для заданной подынтегральной функции $f(x,\ y,\ z)$ всегда можно построить соответствующее ей первообразное векторное поле $\boldsymbol{F}(P)$, так что $f=\operatorname{div}\boldsymbol{F}$, и притом бесконечным числом способов.

С криволинейными интегралами в трехмерном пространстве дело обстоит совершенно так же, как на плоскости; поэтому нет надобности говорить о них особо. Но в трехмерном пространстве надо еще рассмотреть поверхностный интеграл по двумерной области, т. е. по куску поверхности, ограниченному пространственной кривой; поверхностный интеграл занимает там промежуточное место между криволинейным и тройным интегралом. Приведение интеграла по куску поверхности к выражению, зависящему только от границы этого куска (в данном случае этим выражением является криволинейный интеграл вдоль граничной кривой), дается теоремой Стокса (§ 6).

Такое положение дел у интегралов по поверхности аналогично ситуации у криволинейных интегралов. Для того чтобы подынтегральная функция поверхностного интеграла

$$\iint_S f(x,\ y,\ z)\,dS=\iint_S F_n\,dS=\iint_S (F_1\,dy\,dz+F_2\,dz\,dx+F_3\,dx\,dy)$$

могла быть представлена в виде $f(x,\ y,\ z)=(\operatorname{rot}\boldsymbol{B})_n=\boldsymbol{n}\operatorname{rot}\boldsymbol{B}$, необходимо выполнение условия

$$\operatorname{div}\boldsymbol{F}=0,\quad \text{т. е.}\quad \frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}+\frac{\partial F_3}{\partial z}=0.$$

Стало быть, преобразование интеграла по куску поверхности S в интеграл вдоль граничной кривой не всегда возможно. На деле это условие оказывается не только необходимым, но и достаточным (доказательство будет дано в Дополнениях к этой главе, § 2).

Упражнения

1. Вычислить поверхностный интеграл $\oiint \frac{z}{p}\,dS$ по той половине поверхности эллипсоида $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$, для которой $z\geqslant 0$, если

$$\frac{1}{p}=\frac{lx}{a^2}+\frac{my}{b^2}+\frac{nz}{c^2},$$

где $l,\ m,\ n$ — направляющие косинусы внешней нормали.

2. Вычислить поверхностный интеграл $\oiint H\,dS$ по шаровой поверхности $x^2+y^2+z^2=1$, если

$$H=a_1x^4+a_2y^4+a_3z^4+3a_4x^2y^2+3a_5y^2z^2+3a_6x^2z^2.$$

3.* Доказать обобщение теоремы Гаусса на n-мерное пространство [т. е. теорему Остроградского]. Пусть G есть некоторая область в n-мерном пространстве $(x_1, x_2, \ldots, x_n)$, и пусть ее граничная поверхность S задана уравнением

$$\varphi(x_1, x_2, \ldots, x_n)=0,$$

причем в области G вообще $\varphi\leqslant 0$. Даны n функций $a_i(x_1, x_2, \ldots, x_n)$, где $i=1, 2, \ldots, n$, и все они непрерывно дифференцируемы в G. Тогда

$$\overbrace{\int\int_G\ldots\int}^{n\text{ раз}}\left(\frac{\partial a_1}{\partial x_1}+\frac{\partial a_2}{\partial x_2}+\ldots+\frac{\partial a_n}{\partial x_n}\right)dx_1dx_2\ldots dx_n=$$

$$=\overbrace{\int_S\ldots\int}^{n-1\text{ раз}}(a_1\nu_1+a_2\nu_2+\ldots+a_n\nu_n)\,dS,$$

где dS есть элемент площади поверхности S, определенный в гл. IV, стр. 323, а $\nu_1, \nu_2, \ldots, \nu_n$ — координаты единичного вектора внешней нормали, т. е.

$$\nu_i=\frac{\varphi_{x_i}}{\sqrt{\varphi_{x_1}^2+\varphi_{x_2}^2+\ldots+\varphi_{x_n}^2}}.$$

ДОПОЛНЕНИЯ К ГЛАВЕ V

§ 1. Замечания к теоремам Гаусса и Стокса

В основном материале этой главы, при доказательстве теорем Гаусса и Стокса, мы исходили каждый раз из кратного интеграла по исходной многомерной области и по выполнении простого интегрирования приводили его к интегралу по граничному многообразию этой области. Однако, формулы, выражающие эти теоремы, можно также получить, двигаясь обратным путем. Соответствующие преобразования, поучительные и сами по себе, мы здесь изложим вкратце для вывода теоремы Стокса на плоскости.

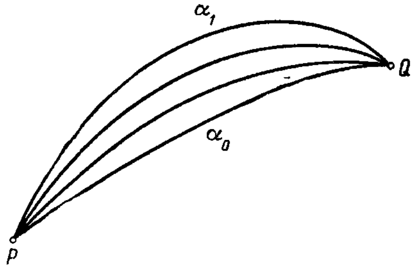


Рис. 94.

Возьмем на плоскости xy две постоянные точки P и Q и соединим их кривой C, представленной своими параметрическими уравнениями с параметром t. Будем непрерывно деформировать эту кривую таким образом, чтобы в процессе деформации, переходя из своего начального положения в конечное, кривая C покрывала однократно область G (рис. 94). Эту мысль мы уточним аналитически следующим образом:

различные положения кривой C пусть определяются различными значениями некоторого параметра α, так что семейство кривых C задается параметрическими уравнениями

$$x = x(t, \alpha), \quad y = y(t, \alpha), \quad t_0 \leqslant t \leqslant t_1,$$

где $x(t_0, \alpha)$, $y(t_0, \alpha)$ — координаты точки P, а $x(t_1, \alpha)$, $y(t_1, \alpha)$ — координаты точки Q. Обе эти точки неподвижны, и их координаты не должны зависеть от α. Итак, α есть параметр, каждое значение которого определяет одну из кривых семейства, а t есть параметр, изменение которого определяет движение точки вдоль выбранной кривой. Когда α пробегает интервал $\alpha_0 \leqslant \alpha \leqslant \alpha_1$, C описывает область G. Мы предполагаем, что функции $x(t, \alpha)$ и $y(t, \alpha)$ имеют непрерывные производные первого порядка по t и по α и непрерывные смешанные производные второго порядка $\frac{\partial^2 x}{\partial t\, \partial \alpha} = x_{t\alpha}$ и $\frac{\partial^2 y}{\partial t\, \partial \alpha} = y_{t\alpha}$, а также что всюду в области G, за исключением точек P и Q, якобиан $\frac{\partial(x, y)}{\partial(t, \alpha)}$ отличен от нуля, например положителен. Тогда область G, если не считать точек P и Q, отображена взаимно однозначно на прямоугольник $\alpha_0 \leqslant \alpha \leqslant \alpha_1$, $t_0 \leqslant t \leqslant t_1$ плоскости α, t.

Допустим теперь, что в замкнутой области G заданы две функции $a(x, y)$ и $b(x, y)$, имеющие непрерывные производные, и рассмотрим криволинейный интеграл

$$I(\alpha) = \int\limits_{C_\alpha} \{a(x, y)\, dx + b(x, y)\, dy\} = \int\limits_{t_0}^{t_1} (a x_t + b y_t)\, dt$$

вдоль кривой C_α семейства, соответствующей значению α параметра. Исследуем зависимость этого интеграла от α. Для этой цели составим производную

$$\frac{dI(\alpha)}{d\alpha} = \int\limits_{t_0}^{t_1} \{(a_x x_\alpha + a_y y_\alpha) x_t + (b_x x_\alpha + b_y y_\alpha) y_t + a x_{t\alpha} + b y_{t\alpha}\}\, dt$$

по правилу дифференцирования определенного интеграла по параметру. Интеграл от суммы последних двух членов преобразуем по правилу интегрирования произведения:

$$\int\limits_{t_0}^{t_1} (a x_{\alpha t} + b y_{\alpha t})\, dt = [a x_\alpha + b y_\alpha]_{t_0}^{t_1} - \int\limits_{t_0}^{t_1} (a_t x_\alpha + b_t y_\alpha)\, dt =$$

$$= -\int\limits_{t_0}^{t_1} [(a_x x_t + a_y y_t) x_\alpha + (b_x x_t + b_y y_t) y_\alpha]\, dt;$$

проинтегрированный член обратился в нуль, ибо $x_\alpha = y_\alpha = 0$ при

$t=t_0$ и при $t=t_1$. В результате имеем

$$\frac{dI(\alpha)}{d\alpha}=\int_{t_0}^{t_1}[a_y(y_\alpha x_t-y_t x_\alpha)+b_x(x_\alpha y_t-x_t y_\alpha)]\,dt,$$

или

$$\frac{dI(\alpha)}{d\alpha}=\int_{t_0}^{t_1}(a_y-b_x)\frac{\partial(x,\,y)}{\partial(t,\,\alpha)}\,dt.$$

Это равенство мы проинтегрируем по α по промежутку от α_0 до α_1:

$$I(\alpha_1)-I(\alpha_0)=\int_{\alpha_0}^{\alpha_1}\int_{t_0}^{t_1}(a_y-b_x)\frac{\partial(x,\,y)}{\partial(t,\,\alpha)}\,dt\,d\alpha.$$

Здесь справа стоит двойной интеграл по прямоугольнику $t_0\leqslant t\leqslant t_1$, $\alpha_0\leqslant\alpha\leqslant\alpha_1$ плоскости αt, равный двойному интегралу по соответствующей области G плоскости xy. Следовательно,

$$I(\alpha_0)-I(\alpha_1)=\iint_G(b_x-a_y)\,dx\,dy,$$

но теперь мы в левой части имеем криволинейный интеграл $\int(a\,dx+b\,dy)$ вдоль контура области G, состоящего из кривой C_α, пробегаемой от P до Q, и из кривой C_{α_1}, пробегаемой от Q до P. Таким образом, коль скоро наши условия выполнены, теорема Стокса на плоскости доказана.

Предоставляем читателю вывести тем же методом теорему Стокса в пространстве. Теорему Гаусса в пространстве тоже можно доказать, исходя из интеграла по куску поверхности, ограниченному замкнутой кривой, и деформируя эту поверхность с сохранением граничной кривой так, чтобы она описала пространственную область G.

Следует, однако, заметить, что этот способ вывода интегральных теорем сам по себе не дает полностью того, что было получено с помощью прежних доказательств. Для того чтобы достигнуть той же степени общности в отношении, например, теоремы Стокса на плоскости, надо еще доказать, что всякая область G плоскости того типа, который рассматривался в прежнем доказательстве, может быть покрыта семейством кривых G_α, обладающих перечисленными свойствами непрерывности и дифференцируемости. Такое доказательство возможно, но оно столь сложно, что прежний метод доказательства заслуживает предпочтения.

§ 2. Представление векторного поля, лишенного источников, в виде ротора

В связи с замечанием, сделанным на стр. 419 в конце n° 1, исследуем теперь вопрос, можно ли всякое векторное поле, лишенное источников, представить в виде ротора. Другими словами: для всякого ли векторного поля $A(P)$, в котором $\operatorname{div}A\equiv 0$ в некоторой замкнутой области G пространства xyz, существует такое другое векторное поле $\boldsymbol{B}(P)$, что

$$\boldsymbol{A} = \operatorname{rot} \boldsymbol{B}$$

всюду в G? Мы сейчас докажем, что это действительно так.

Наша задача состоит в том, чтобы по данному векторному полю $\boldsymbol{A}(P) = \{A_1(x, y, z), A_2(x, y, z), A_3(x, y, z)\}$ найти векторное поле $\boldsymbol{B}(P) = \{B_1(x, y, z), B_2(x, y, z), B_3(x, y, z)\}$, удовлетворяющее в области G трем дифференциальным уравнениям с частными производными:

$$A_1 = \frac{\partial B_3}{\partial y} - \frac{\partial B_2}{\partial z}, \qquad A_2 = \frac{\partial B_1}{\partial z} - \frac{\partial B_3}{\partial x}, \qquad A_3 = \frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y}.$$

Для упрощения мы предположим, что область G, в которой определено векторное поле $\boldsymbol{A}(P)$ и выполняется условие $\operatorname{div} \boldsymbol{A} = \frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z} = 0$, является прямоугольным параллелепипедом. Наша система уравнений определяет решение неоднозначно, и при нахождении B_1, B_2, B_3 остается еще большой произвол. Поэтому мы наложим дополнительное условие $B_3 \equiv 0$ в G. Тогда система уравнений упрощается:

$$\frac{\partial B_2}{\partial z} = -A_1, \qquad \frac{\partial B_1}{\partial z} = A_2, \qquad \frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y} = A_3.$$

Из первого уравнения находим интегрированием по z одно из решений

$$B_2 = -\int_{z_0}^{z} A_1(x, y, \zeta)\, d\zeta,$$

причем x и y играют роль постоянных величин, а z_0 и встречающееся ниже y_0, означают аппликату и ординату произвольной постоянной точки P_0 области G. Из второго уравнения определяем

$$B_1 = \int_{z_0}^{z} A_2(x, y, \zeta)\, d\zeta + \alpha(x, y),$$

где $\alpha(x, y)$ есть произвольная функция от x и y, пока еще совершенно неопределенная.

Подставляя найденные выражения для B_2 и B_1 в третье уравнение, получаем уравнение для определения функции $\alpha(x, y)$:

$$-\int_{z_0}^{z} \left(\frac{\partial A_1(x, y, \zeta)}{\partial x} + \frac{\partial A_2(x, y, \zeta)}{\partial y}\right) d\zeta - \alpha_y(x, y) = A_3(x, y, z).$$

В силу условия $\operatorname{div} \boldsymbol{A} = 0$, имеем $\frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} = -\frac{\partial A_3}{\partial z}$, и последнее уравнение приводится к следующему виду:

$$\alpha_y(x, y) = \int_{z_0}^{z} \frac{\partial A_3(x, y, \zeta)}{\partial \zeta} d\zeta - A_3(x, y, z) = -A_3(x, y, z_0).$$

Из этого уравнения получаем интегрированием по y функцию $\alpha(x, y)$, которая до сих пор оставалась неопределенной:

$$\alpha(x, y) = -\int_{y_0}^{y} A_3(x, \eta, z_0)\, d\eta$$

(берем лишь одну из первообразованных). Теперь искомая вектор-функция $\boldsymbol{B} = \{B_1, B_2, B_3\}$ определилась так:

$$B_1 = \int_{z_0}^{z} A_2(x, y, \zeta)\, d\zeta - \int_{y_0}^{y} A_3(x, \eta, z_0)\, d\eta,$$

$$B_2 = -\int_{z_0}^{z} A_1(x, y, \zeta)\, d\zeta,$$

$$B_3 = 0,$$

и мы нашли такое векторное поле $\boldsymbol{B}(P)$, что $\boldsymbol{A} = \operatorname{rot} \boldsymbol{B}$. Но найденное поле $\boldsymbol{B}$ является не единственным ответом задачи, что видно уже из хода решения. И действительно, если составить векторную сумму $\boldsymbol{B} + \operatorname{grad} U$, где $U(x, y, z)$ — произвольная дважды непрерывно дифференцируемая скалярная функция, то также будем иметь $\operatorname{rot}(\boldsymbol{B} + \operatorname{grad} U) = \boldsymbol{A}$, ибо всегда $\operatorname{rot} \operatorname{grad} U = 0$ (гл. II, § 7, n° 7). С другой стороны, обратно, если $\boldsymbol{B}^*$ есть любая вектор-функция точки, для которой $\operatorname{rot} \boldsymbol{B}^* = \boldsymbol{A}$, то $\operatorname{rot}(\boldsymbol{B}^* - \boldsymbol{B}) = 0$, в силу чего (гл. V, § 1, n° 7, стр. 378) вектор $\boldsymbol{B}^* - \boldsymbol{B}$ может быть представлен в виде градиента некоторой скалярной функции $U(x, y, z)$. Таким образом, векторное поле $\boldsymbol{B}^* = \boldsymbol{B} + \operatorname{grad} U$ является самым общим решением нашей задачи.

Упражнения

1. Пусть $f(x, y)$ — непрерывная функция, имеющая непрерывные первые и вторые производные. Доказать, что если

$$f_{xx} f_{yy} - f_{xy}^2 \neq 0,$$

то преобразование

$$u = f_x(x, y), \qquad v = f_y(x, y), \qquad w = -z + x f_x(x, y) + y f_y(x, y)$$

однозначно обратимо, и обратное преобразование имеет следующий вид:

$$x = g_u(u, v), \qquad y = g_v(u, v), \qquad z = -w + u g_u(u, v) + v g_v(u, v).$$

2. Представить гравитационное векторное поле

$$A = \left\{ \frac{x}{\sqrt{(x^2+y^2+z^2)^3}}, \quad \frac{y}{\sqrt{(x^2+y^2+z^2)^3}}, \quad \frac{z}{\sqrt{(x^2+y^2+z^2)^3}} \right\}$$

в виде $A = \operatorname{rot} B$.

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ V

1. Пусть $\varphi(t)$, $a(t)$ и $b(t)$ — непрерывно дифференцируемые функции параметра t в интервале $0 \leqslant t \leqslant 2\pi$, причем $a(2\pi) = a(0)$, $b(2\pi) = b(0)$, $\varphi(2\pi) = \varphi(0) + 2n\pi$ (где n — целое число), и пусть x и y — постоянные. Рассматривая уравнения

$$\xi = x\cos\varphi - y\sin\varphi + a, \qquad \eta = x\sin\varphi + y\cos\varphi + b$$

как параметрические уравнения замкнутой плоской кривой Γ (с параметром t), доказать, что

$$\frac{1}{2}\oint_\Gamma (\xi\, d\eta - \eta\, d\xi) = A(x^2+y^2) + Bx + Cy + D,$$

где

$$A = \frac{1}{2}\oint_\Gamma d\varphi, \quad B = \oint_\Gamma (a\cos\varphi + b\sin\varphi)\, d\varphi,$$

$$C = \oint_\Gamma (-a\sin\varphi + b\cos\varphi)\, d\varphi, \qquad D = \frac{1}{2}\oint_\Gamma (a\, db - b\, da).$$

2.* Пусть плоскость Q как целое движется относительно неподвижной плоскости Π, с которой она совпадает, таким образом, что каждая ее точка P описывает замкнутую кривую плоскости Π, ограничивающую площадь алгебраической величины $S(P)$. Обозначим через $2n\pi$ (n — целое число) величину полного вращения плоскости Q относительно неподвижной плоскости Π. Доказать следующие результаты:

а) Если $n \neq 0$, то на плоскости Q существует такая точка M, что для всякой другой точки P этой плоскости

$$S(P) = \pi n MP^2 + S(M).$$

б) Если $n = 0$, то возможны два случая: б$_1$) на плоскости Q существует такая ориентированная прямая Δ, что для всякой точки P этой плоскости

$$S(P) = \lambda d(P),$$

где $d(P)$ есть расстояние точки P от прямой Δ, а λ — постоянный положительный коэффициент, либо б$_2$) $S(P)$ имеет одинаковое значение для всех точек P плоскости Q (*теорема Штейнера*).

3*. Прямолинейный отрезок AB совершает в плоскости Π периодическое движение наподобие шатуна: точка B движется против часовой стрелки по окружности с центром C, когда точка A движется периодически вдоль прямой, проходящей через C. Применить результаты упр. 2 для определения площади, ограниченной замкнутой кривой, которую описывает в плоскости Π точка K, жестко скрепленная с отрезком AB.

4. Конечные точки A и B прямолинейного стержня AB описывают полный оборот вдоль замкнутой выпуклой кривой Γ. Точка K стержня AB, для которой $AK = a$, $KB = b$, описывает в результате этого движения замкнутую кривую Γ'. Доказать, что площадь, заключенная между кривыми Γ и Γ', равна πab (*теорема Хольдича* — Holditch).

5*. Доказать, что если к каждому элементу ds жесткой гладкой замкнутой пространственной кривой Г приложить силу величины $\frac{ds}{\rho}$ по направлению главного нормального вектора (стр. 106), то кривая Г останется в равновесии. Кривизну $1/\rho$ кривой Г в точке элемента ds будем предполагать конечной и непрерывной на всей этой кривой. (На основании принципов статики твердого тела задача сводится к доказательству того, что

$$\oint_{\Gamma} \frac{\boldsymbol{n}}{\rho}\, ds = 0 \quad \text{и} \quad \oint_{\Gamma} \frac{[\boldsymbol{rn}]}{\rho}\, ds = 0,$$

где $\boldsymbol{n}$ — единичный главный нормальный вектор кривой Г для элемента ds, а $\boldsymbol{r}$ — его радиус-вектор.)

6. Доказать, что замкнутая жесткая поверхность Σ остается в равновесии под действием равномерно по ней распределенного давления, направленного внутрь. (Если обозначить через $\boldsymbol{n}_i$ единичный нормальный вектор, направленный внутрь, и через $\boldsymbol{r}$ радиус-вектор элемента поверхности dS, то задача сводится к доказательству векторных равенств

$$\oint\!\!\oint_{\Sigma} \boldsymbol{n}_i\, dS = 0 \quad \text{и} \quad \oint\!\!\oint_{\Sigma} [\boldsymbol{rn}_i]\, dS = 0.)$$

7*. Твердое тело объема V, ограниченное поверхностью Σ, полностью погружено в жидкость удельного веса 1. Доказать, что статический эффект давления жидкости на тело эквивалентен силе $\boldsymbol{F}$ величины V, направленной вертикально вверх и приложенной к центру массы C объемной области, занимаемой телом.

8*. Дан эллипсоид Σ: $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$. Обозначим через p расстояние от центра эллипсоида до касательной плоскости в его точке $P(x, y, z)$ и через dS — элемент площади в этой точке. Доказать следующие соотношения:

а) $\oint\!\!\oint_{\Sigma} p\, dS = 4\pi abc$, б) $\oint\!\!\oint_{\Sigma} \frac{1}{p}\, dS = \frac{4\pi}{3abc}(b^2c^2 + c^2a^2 + a^2b^2)$.

9. Обыкновенный плоский угол измеряется длиной дуги, вырезаемой его сторонами из единичной окружности с центром в вершине угла. Эту идею можно использовать и для измерения *телесного угла*, ограниченного конической поверхностью с вершиной A, следующим образом. В порядке определения величина телесного угла принимается равной площади, вырезаемой им из единичной сферы с центром A. Например, мера телесного угла октанта $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$ равна $4\pi/8 = \pi/2$.

Обозначим теперь через Г замкнутую пространственную кривую, через Σ — поверхность, «натянутую» на эту кривую и ограниченную ею, и через A — фиксированную точку, не лежащую ни на Г, ни на Σ. Элемент площади dS в точке M поверхности Σ определяет элементарный конус с вершиной в A. Простое рассуждение показывает, что телесный угол этого конуса

$$d\Omega = \frac{\cos\theta}{r^2}\, dS,$$

где $r = AM$, а θ есть угол между вектором $\overline{MA}$ и нормальным вектором поверхности Σ в точке M. Этот элементарный телесный угол будет положительным или отрицательным, смотря по тому, каков угол θ — острый или тупой. Истолковать поверхностный интеграл

$$\Omega = \iint_{\Sigma} \frac{\cos\theta}{r^2}\, dS$$

геометрически как телесный угол и показать, что

$$\Omega=\iint_{\Sigma}\frac{(a-x)\,dy\,dz+(b-y)\,dz\,dx+(c-z)\,dx\,dy}{[(a-x)^2+(b-y)^2+(c-z)^2]^{3/2}},$$

где (a, b, c) и (x, y, z) — декартовы координаты точек A и M.

10. Доказать, сначала прямым вычислением, а затем с помощью интерпретации интеграла как телесного угла, что

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{dx\,dy}{(x^2+y^2+1)^{3/2}}=2\pi.$$

11*. Доказать, что телесный угол с вершиной $(0, 0, 0)$, стягиваемый всей поверхностью однополостного гиперболоида

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1,$$

равен

$$8c\int_0^{\frac{\pi}{2}}\sqrt{\frac{b^2\cos^2\varphi+a^2\sin^2\varphi}{a^2b^2+b^2c^2\cos^2\varphi+a^2c^2\sin^2\varphi}}\,d\varphi.$$

12. Показать, что значение поверхностного интеграла

$$\Omega=\iint_{\Sigma}\frac{(a-x)\,dy\,dz+(b-y)\,dz\,dx+(c-z)\,dx\,dy}{[(a-x)^2+(b-y)^2+(c-z)^2]^{3/2}}$$

не зависит от выбора поверхности Σ при условии, что граничная кривая Γ остается неизменной. Пользуясь этим результатом, доказать, что если Σ есть замкнутая поверхность, то $\Omega=4\pi$ или 0, смотря по тому, где находится точка $A\,(a, b, c)$: внутри объема, ограниченного поверхностью Σ, или вне его. (Для этой цели произвести интегрирование по всей наружной стороне замкнутой поверхности Σ.)

13*. На поверхности взят кусок ее Σ, ограниченный кривой Γ. Рассмотрим поверхностный интеграл

$$\Omega\,(a, b, c)=\iint_{\Sigma}\frac{(a-x)\,dy\,dz+(b-y)\,dz\,dx+(c-z)\,dx\,dy}{R^3},$$

где

$$R^2=(a-x)^2+(b-y)^2+(c-z)^2,$$

как функцию точки $A\,(a, b, c)$. Доказать, что координаты градиента этой функции $\Omega\,(A)$ выражаются следующими криволинейными интегралами:

$$\frac{\partial\Omega}{\partial a}=\oint_{\Gamma}\frac{(z-c)\,dy-(y-b)\,dz}{R^3},\qquad \frac{\partial\Omega}{\partial b}=\oint_{\Gamma}\frac{(x-a)\,dz-(z-c)\,dx}{R^3},$$

$$\frac{\partial\Omega}{\partial c}=\oint_{\Gamma}\frac{(y-b)\,dx-(x-a)\,dy}{R^3}.$$

Эти формулы, играющие важную роль в электромагнетизме, можно объединить в одну векторную формулу

$$\operatorname{grad}\Omega = -\oint_{\Gamma} \frac{[\boldsymbol{R}\,d\boldsymbol{R}]}{R^3},$$

где

$$\boldsymbol{R} = \overline{AM} = \{x-a,\ y-b,\ z-c\} \quad \text{и} \quad R = |\boldsymbol{R}|.$$

14*. Проверить, что выражение

$$\frac{-4xy\,dx + 2(x^2-y^2-1)\,dy}{(x^2+y^2-1)^2+4y^2}$$

является полным дифференциалом угла, под которым из точки (x, y) виден отрезок $-1 \leqslant x \leqslant 1$, $y=0$. Пользуясь этим фактом, доказать с помощью геометрического рассуждения следующий результат:

В плоскости xy дана ориентированная замкнутая кривая Γ, не проходящая через точки $(-1;\ 0)$, $(1;\ 0)$. Пусть кривая Γ пересекает отрезок $-1 < x < 1$, $y=0$ p раз, переходя из верхней полуплоскости $y>0$ в нижнюю, где $y<0$, и n раз, переходя из нижней полуплоскости в верхнюю. Тогда

$$\Theta = \oint_{\Gamma} \frac{-4xy\,dx + 2(x^2-y^2-1)\,dy}{(x^2+y^2-1)^2+4y^2} = 2\pi(p-n).$$

Таким образом, если Γ есть кривая

$$r = 2\cos 2\theta, \qquad 0 \leqslant \theta \leqslant 2$$

(в полярных координатах), то $\Theta = 0$.

15.** Рассмотрим в плоскости xy единичную окружность C

$$x' = \cos\varphi,$$
$$y' = \sin\varphi,$$
$$z' = 0$$
$$(0 \leqslant \varphi \leqslant 2\pi).$$

Обозначим через Ω телесный угол с вершиной в точке $P(x, y, z)$, стягиваемый круговым диском $x^2+y^2 \leqslant 1$, $z=0$. Пусть теперь точка P описывает в пространстве ориентированную замкнутую кривую Γ, не встречающую окружности C, причем кривая Γ пересекает круговой диск $x^2+y^2<1$, $z=0$ p раз, переходя из верхнего полупространства $z>0$ в нижнее, где $z<0$, и n раз, переходя из нижнего полупространства в верхнее. Если переменная точка P начинает свое движение с точки P_0 кривой Γ со значением $\Omega=\Omega_0$, то, обойдя Γ один раз (при этом движении телесный угол Ω изменяется непрерывно), точка P вернется в P_0 со значением $\Omega=\Omega_1$. Доказать с помощью геометрического рассуждения, что

$$\Omega_1 - \Omega_0 = \oint_{\Gamma} d\Omega = 4\pi(p-n).$$

Перепишем найденное выше (упр. 13) векторное равенство так:

$$\operatorname{grad}\Omega = -\oint_{C} \frac{[\overline{PP'}\,dP']}{|\overline{PP'}|^3}.$$

[При этом положено $\boldsymbol{R} = \overline{PP'}$, $d\boldsymbol{R} = dP'$, где dP' есть своеобразное обозначение: это «дифференциал точки», т. е. вектор-смещение точки P'.]

Пользуясь последней формулировкой, доказать, что

$$\oint_C \oint_\Gamma \frac{1}{|\overline{PP'}|^3} \begin{vmatrix} x'-x & dx & dx' \\ y'-y & dy & dy' \\ z'-z & dz & dz' \end{vmatrix} =$$

$$= \oint_\Gamma \oint_C \frac{(x'-x)(dy\,dz'-dz\,dy') + (y'-y)(dz\,dx'-dx\,dz') + (z'-z)(dx\,dy'-dy\,dx')}{\left\{(x'-x)^2+(y'-y)^2+(z'-z)^2\right\}^{3/2}} =$$

$$= 4\pi\,(p-n).$$

(Этот повторный криволинейный интеграл, принадлежащий Гауссу, дает число зацеплений, которые кривая Г образует с окружностью C. Следует заметить, что его обращение в нуль является необходимым условием отделимости кривых Г и C (реализуемых в виде двух нитей). Однако, пример, изображенный на рис. 95, показывает, что это условие является недостаточным; на этом рисунке $p=n=1$, однако кривую Г и окружность C невозможно отделить друг от друга.)

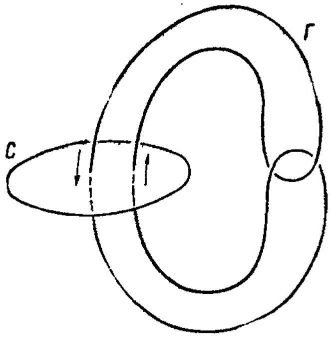


Рис. 95.

16*. В пространстве дана замкнутая кривая Г, на которой выбрано определенное направление обхода. Доказать, что существует вектор a, обладающий следующим характеристическим свойством: каков бы ни был единичный вектор n°, скалярное произведение an° равно алгебраическому значению площади, ограниченной ортогональной проекцией кривой Г на плоскость П, перпендикулярную к вектору n°. (Заметьте: вектор n° определяет положительное направление вращения на плоскости П, а заданная ориентация кривой Г определяет направление обхода ее проекции на плоскость П.) В частности, проекция кривой Г на любую плоскость, параллельную вектору a, имеет алгебраическую площадь нуль. (Вектор a можно назвать *вектором площади* кривой Г.)

ГЛАВА VI

ДОПОЛНИТЕЛЬНЫЕ СВЕДЕНИЯ О ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЯХ

Некоторые частные типы обыкновенных дифференциальных уравнений мы уже изучали в т. I, гл. XI. В рамках этой книги мы не имеем возможности развить общую теорию таких уравнений. Мы дадим, однако, в этой главе хотя бы очерк основных сведений из этой дисциплины, отправляясь от некоторых примеров из механики.

§ 1. Дифференциальные уравнения движения точки в пространстве

1. Уравнения движения. В т. I, гл. V, §§ 4 и 5 и в гл. XI мы уже занимались изучением движения материальной точки, но только движения специального типа: предполагалось, что оно происходит по заранее заданной кривой. Теперь мы откажемся от этого ограничения и рассмотрим массу m, которую представляем себе сосредоточенной в точке $P(x, y, z)$, движущейся в пространстве. Радиус-вектор движущейся точки мы обозначим через $\boldsymbol{r}$; координаты x, y, z точки P являются вместе с тем координатами ее радиус-вектора. Движение материальной точки будет представлено математически, если удастся найти выражение радиус-вектора $\boldsymbol{r}$ или координат x, y, z в виде функций от времени t. Производную по времени t мы будем обозначать, как и раньше, точкой, поставленной над символом функции. Тогда вектор $\dot{\boldsymbol{r}} = \{\dot{x}, \dot{y}, \dot{z}\}$ с модулем $v = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$ представляет скорость, а вектор $\ddot{\boldsymbol{r}} = \{\ddot{x}, \ddot{y}, \ddot{z}\}$ — ускорение точки P.

Мы не будем заниматься рассмотрением *принципов механики* по существу, и только примем за исходный пункт следующие определения и факты. Произведение вектора-ускорения на массу m мы будем называть *вектором-силой* или просто силой $\boldsymbol{F} = \{X, Y, Z\}$. Следовательно,

$$m\ddot{\boldsymbol{r}} = \boldsymbol{F}. \tag{1}$$

Это уравнение называется *ньютоновым основным уравнением механики*. Оно равносильно трем координатным уравнениям:

$$m\ddot{x} = X, \qquad m\ddot{y} = Y, \qquad m\ddot{z} = Z. \tag{1a}$$

Уравнение (1) или система (1a) представляют собой покамест лишь определение термина *сила*. Оказывается, однако, что во многих случаях этот вектор-силу возможно определить без ссылки на подлежащее изучению конкретное движение, так как в пространстве, на основании физических соображений, заранее задано некоторое силовое поле. Тогда основные уравнения механики можно уже рассматривать не как определения, а как дифференциальные уравнения, которым должны удовлетворять радиус-вектор и координаты материальной точки при всяком ее движении под воздействием заданного силового поля.

Примером такого силового поля является поле силы тяжести; оно заранее известно до постановки любой конкретной задачи на движение точки в этом поле. Если направить ось z вертикально вверх, то сила тяжести

$$\boldsymbol{F}=\{0,\ 0,\ -mg\}=-mg\,\mathrm{grad}\,z,$$

где g есть постоянное ускорение силы тяжести.

Другой пример дает поле тяготения, порождаемое массой μ, сосредоточенной в начале координат и притягивающей по закону Ньютона. Если $r=|\boldsymbol{r}|=\sqrt{x^2+y^2+z^2}$ есть расстояние притягиваемой точки (x, y, z) массы m от начала, то в этом случае силовое поле задается выражением

$$\boldsymbol{F}=\gamma\mu m\,\mathrm{grad}\,\frac{1}{r}=-\gamma\mu m\,\frac{\boldsymbol{r}^{\circ}}{r^2}=-\gamma\mu m\,\frac{\boldsymbol{r}}{r^3}$$

(ср. стр. 112), и уравнение движения в этом поле будет

$$\ddot{\boldsymbol{r}}=-\gamma\mu\,\frac{\boldsymbol{r}}{r^3}$$

или, в координатах,

$$\ddot{x}=-\gamma\mu\,\frac{x}{r^3},\qquad \ddot{y}=-\gamma\mu\,\frac{y}{r^3},\qquad \ddot{z}=-\gamma\mu\,\frac{z}{r^3}.$$

Вообще, если $\boldsymbol{F}=\boldsymbol{F}(P)=\{X(x, y, z),\ Y(x, y, z),\ Z(x, y, z)\}$ есть заданное силовое поле, то уравнение движения в этом поле будет

$$m\ddot{\boldsymbol{r}}=\boldsymbol{F}. \tag{1}$$

Координатная запись этого уравнения

$$m\ddot{x}=X,\qquad m\ddot{y}=Y,\qquad m\ddot{z}=Z \tag{1a}$$

образует систему трех линейных дифференциальных уравнений второго порядка для трех неизвестных функций $x(t)$, $y(t)$, $z(t)$. Основная задача механики точки состоит в том, чтобы определить из этих дифференциальных уравнений действительное движение материальной точки, если в «начальный момент», скажем при $t=0$, заданы положение движущейся точки, т. е. ее координаты $x(0)=x_0$, $y(0)=y_0$, $z(0)=z_0$, и ее скорость, т. е. величины $\dot{x}(0)=\dot{x}_0$, $\dot{y}(0)=\dot{y}_0$, $\dot{z}(0)=\dot{z}_0$, или, как говорят, если заданы *начальное положение* и *начальная скорость*. Эта задача называется задачей *решения* или *интегрирования* системы (1a) при заданных начальных условиях.

Применение слова *интегрирование* в смысле процесса решения дифференциальных уравнений объясняется тем, что этот процесс является в известном смысле обобщением обычного интегрирования функций (ср. т. I, стр. 597).

2. Закон сохранения энергии. Собираясь в дальнейшем выполнить интегрирование этой системы дифференциальных уравнений для конкретных задач, мы предварительно выведем из уравнений движения несколько общих теорем. Из гл. V, § 1, n° 4 мы уже знакомы с понятием работы, совершаемой силовым полем при движении в нем материальной точки, и знаем, что эта работа выражается криволинейным интегралом

$$\int \boldsymbol{F}\,d\boldsymbol{r} = \int \boldsymbol{F}\dot{\boldsymbol{r}}\,dt = \int (X\,dx + Y\,dy + Z\,dz)$$

вдоль пути, описанного движущейся точкой.

Если силовое поле $\boldsymbol{F}(P)$ может быть представлено в виде градиента силовой функции Φ:

$$\boldsymbol{F} = \operatorname{grad}\Phi,$$

то работа, совершаемая полем, не зависит от пути и вполне определяется его начальной и конечной точками (гл. V, § 1, n° 5). Силовое поле, которое может быть представлено как градиент, называют, следуя Гельмгольцу, *консервативным*[1]) *силовым полем.* В таком поле векторное уравнение движения принимает следующий вид:

$$m\ddot{\boldsymbol{r}} = -\operatorname{grad} U, \tag{A}$$

если вместо силовой функции Φ ввести *потенциал* или *потенциальную энергию* $U = -\Phi$ (ср. гл. V, § 1, n° 5). Необходимо заметить, что силовая функция и потенциал определяются по данному полю неоднозначно, так как они содержат произвольное постоянное слагаемое.

В координатной записи мы получаем для консервативного силового поля систему дифференциальных уравнений

$$m\ddot{x} = -U_x, \quad m\ddot{y} = -U_y, \quad m\ddot{z} = -U_z. \tag{A_1}$$

Эту систему дифференциальных уравнений невозможно решить в общем виде, но из нее выводится новое уравнение, которое уже не содержит вторых производных, но в котором зато появляются первые производные от искомых функций $x(t)$, $y(t)$, $z(t)$. Для этой цели помножим векторное уравнение (A) скалярно на $\dot{\boldsymbol{r}}$. Тогда в левой части получится производная по t от выражения $\frac{1}{2}m\dot{\boldsymbol{r}}^2 = \frac{1}{2}mv^2$, а в правой части — производная по t от скалярной функции $-U$:

$$\frac{d}{dt}\left(\frac{m}{2}\dot{\boldsymbol{r}}^2\right) = \frac{d}{dt}(-U).$$

---

[1]) От латинского conservare — сохранять, в связи с теоремой о сохранении энергии, которую мы сейчас докажем.

Интегрируя это уравнение, получаем

$$\frac{m}{2}\dot{r}^2 = -U + C,$$

где C — произвольная постоянная, т. е. величина, не зависящая от времени t. Тот же результат можно вывести из эквивалентной координатной системы дифференциальных уравнений (A_1). Три уравнения системы надо умножить последовательно на $\dot{x}$, $\dot{y}$, $\dot{z}$ и результаты сложить. Получится уравнение

$$\frac{m}{2}\frac{d}{dt}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = -\frac{dU}{dt},$$

интегрирование которого приводит к тому же окончательному уравнению, но в координатной записи:

$$\frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + U = C. \qquad (B)$$

Выражение

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{m}{2}v^2$$

называется *кинетической энергией* или *энергией движения* материальной точки, а величина U — ее *потенциальной энергией* или *энергией положения*. Не имея возможности подробно заняться физическим содержанием этих понятий, мы лишь устанавливаем, что полученное уравнение (B) выражает так называемый *закон сохранения энергии: при движении в консервативном силовом поле полная энергия, т. е. сумма кинетической и потенциальной энергии, остается постоянной.*

В следующем параграфе (§ 2) мы увидим на конкретных задачах, как использовать закон сохранения энергии для интегрирования уравнений движения.

3. Равновесие. Устойчивость. Уравнения движения в сочетании с допущением, что $F = -\operatorname{grad} U$, дают возможность исследовать проблему равновесия в консервативном поле. Говорят, что материальная точка находится в равновесии в данном силовом поле, если она остается в покое. Это возможно лишь в том случае, если скорость и ускорение материальной точки равны нулю в течение всего рассматриваемого промежутка времени. Поэтому из уравнений движения вытекают следующие необходимые условия равновесия:

$$\operatorname{grad} U = 0 \quad \text{или} \quad U_x = 0, \quad U_y = 0, \quad U_z = 0.$$

Но этими же уравнениями определяются точки, в которых потенциальная энергия U имеет стационарное значение.

Особый интерес представляет тот факт, что *точка, в которой потенциальная энергия имеет строгий минимум, является точкой* *устойчивого равновесия.* Равновесие называется устойчивым при следующем условии: если вызвать достаточно малое нарушение равновесия, то все последующее движение будет сколь угодно мало отличаться от состояния покоя. Точнее, по двум наперед заданным произвольно малым положительным числам ρ и ε можно найти два положительных числа δ и η, столь малых, что если отклонить материальную точку от ее положения покоя меньше чем на δ, сообщив ей при этом еще скорость, меньшую по модулю чем η, то во всем дальнейшем движении отклонение точки от положения равновесия не превысит расстояния ρ, а модуль ее скорости не превзойдет начального значения ε.

Замечательно, что эту теорему об устойчивости равновесия можно доказать, не доведя до конца интегрирование уравнений движения. Доказательство опирается только на допущение, что в рассматриваемом положении равновесия потенциальная энергия имеет собственный (строгий) минимум. Для упрощения мы предположим, что положение равновесия, в котором потенциальная энергия U имеет минимум, находится в начале координат; в противном случае эту точку можно привести в начало с помощью параллельного переноса осей координат. Мы знаем, что потенциальная энергия U, согласно своему определению, содержит произвольную аддитивную постоянную, ибо функция U и функция $U+\text{const}$ дают одно и то же силовое поле, так как при дифференцировании постоянная пропадает. Поэтому можно принять без уменьшения общности, что минимальное значение $U(0, 0, 0)=0$.

Опишем вокруг начала координат сферу S_r радиуса r; в силу нашего предположения, что функция U имеет в начале координат минимум нуль, значение $r<\rho$ можно выбрать столь малым, что на сфере S_r и внутри ее всюду будет $U>0$. Обозначим наименьшее значение функции U на сфере S_r через a; по условию, $a>0$. Поэтому ясно, что материальная точка никогда не достигнет поверхности сферы S_r, пока ее потенциальная энергия остается меньше чем a. Так как функция U непрерывна, то можно найти число δ, зависящее от a и столь малое, что в сфере S_δ радиуса δ с центром в начале значение U не превышает $\frac{a}{2}$. Выберем теперь начальное положение материальной точки внутри сферы S_δ и сообщим ей начальную скорость со столь малым модулем v_0, что начальное значение кинетической энергии

$$T_0=\frac{1}{2}mv_0^2<\frac{a}{2},$$

другими словами, $v_0<\sqrt{\frac{a}{m}}$; тогда, в силу закона сохранения энергии, всегда будет

$$T+U=T_0+U_0<a,$$

Так как всегда $T \geqslant 0$, то всегда будет $U < a$, и поэтому материальная точка не может никогда удалиться от положения равновесия на расстояние, большее числа r. Далее, так как $U \geqslant 0$, то $T < a$ в продолжение всего движения, и для модуля v скорости всегда будет $v < \sqrt{\frac{2a}{m}}$.

В силу непрерывности функции U, число a стремится к нулю вместе с r. Поэтому r можно выбрать столь малым, что $\sqrt{\frac{2a}{m}} < \varepsilon$, т. е. $a < \frac{1}{2} m\varepsilon^2$, так что модуль скорости всегда меньше чем ε. Таким образом, если начальное положение материальной точки находится внутри S_δ, т. е. ее начальное отклонение меньше чем δ, а начальная ее скорость имеет модуль $v_0 < \sqrt{\frac{a}{m}} = \eta$, то она всегда останется внутри сферы S_r радиуса $r < \rho$ и всегда модуль ее скорости будет $v < \varepsilon$.

§ 2. Примеры из механики точки

1. Движение материальной точки, брошенной под углом к горизонту. В качестве простейшего примера рассмотрим движение материальной точки под действием силы тяжести. Положительную ось z направим вертикально вверх. Тогда уравнения движения (после сокращения на m) примут следующий вид:

$$\frac{d^2x}{dt^2} = 0, \qquad \frac{d^2y}{dt^2} = 0, \qquad \frac{d^2z}{dt^2} = -g.$$

Интегрируя эти уравнения, получим координаты вектора-скорости:

$$\frac{dx}{dt} = a_1, \qquad \frac{dy}{dt} = b_1, \qquad \frac{dz}{dt} = -gt + c_1,$$

где a_1, b_1 и c_1 — постоянные; интегрируя еще раз, получим координаты движущейся точки как функции времени:

$$x = a_1t + a_2, \qquad y = b_1t + b_2, \qquad z = -\frac{g}{2}t^2 + c_1t + c_2,$$

где a_2, b_2 и c_2 — тоже постоянные. Значения шести постоянных интегрирования получаются из начальных условий. Не ограничивая общности механической задачи, можно выбрать систему координат таким образом, чтобы в момент времени $t = 0$ материальная точка находилась в ее начале. Полагая в последних трех уравнениях $t = 0$ и $x = y = z = 0$, сразу получаем $a_2 = b_2 = c_2 = 0$. Далее можно принять, без потери общности, что начальная скорость лежит в плоскости zx, так что проекция b_1 начальной скорости на ось y равна нулю. В силу этих допущений, $y(t) = 0$ при всех значениях t, и траектория движущейся точки является плоской кривой, — она лежит в плоскости zx. Остальные два уравнения

$$x = a_1t, \qquad z = -\frac{g}{2}t^2 + c_1t$$

дают параметрическое представление траектории. Исключив из них время t, получим явное уравнение траектории

$$z = -\frac{g}{2a_1^2}x^2 + \frac{c_1}{a_1}x.$$

Это — парабола, обращенная выпуклостью вверх, с осью, параллельной оси z. Вершина параболы соответствует максимуму функции $z(x)$, и ее координаты (x_1, z_1) находим, решая уравнение $z'(x) = 0$:

$$x_1 = \frac{a_1c_1}{g}, \qquad z_1 = -\frac{g}{2a_1^2}\frac{a_1^2c_1^2}{g^2} + \frac{c_1}{a_1}\frac{a_1c_1}{g} = \frac{c_1^2}{2g}.$$

[Начальная скорость $v_0 = \{a_1, 0, c_1\}$. Всегда можно принять $a_1 > 0$; для этого надо направить положительную ось x в ту сторону, куда указывает горизонтальная компонента начальной скорости. Если при этом $c_1 > 0$, то начальная скорость направлена (вообще говоря, не вертикально) вверх.] Для достижения вершины параболы движущейся точке потребуется тогда время

$$T = \frac{x_1}{a_1} = \frac{c_1}{g}.$$

Через двойной такой промежуток времени (от начала движения) $2T = \frac{2c_1}{g}$ падающая масса достигнет точки с координатами $x_2 = \frac{2a_1c_1}{g} = 2x_1$, $z_2 = 0$, т. е. вернется на горизонтальную прямую $y = z = 0$, проходящую через начальное положение.

Теперь нетрудно решить следующую задачу: под каким углом к горизонту надо бросить материальную точку с заданной начальной скоростью $v_0 = \sqrt{a_1^2 + c_1^2} > 0$, чтобы получить возможно большую дальность полета $\frac{2a_1c_1}{g}$? При этом предполагается, что $a_1 \geqslant 0$ и $c_1 \geqslant 0$. Существование такого угла очевидно, см. примеры гл. III, § 6.

Поставленная задача равносильна задаче нахождения максимума функции $f(a_1, c_1) = a_1c_1$ при дополнительном условии $a_1^2 + c_1^2 = v_0^2 > 0$. Правило множителей Лагранжа дает для определения неизвестных a_1 и c_1 систему уравнений

$$a_1 + 2\lambda c_1 = 0, \qquad c_1 + 2\lambda a_1 = 0, \qquad a_1^2 + c_1^2 = v_0^2,$$

из которых получаем $a_1 = c_1$ [1]. Таким образом, углом метания, обеспечивающим наибольшую дальность полета, является угол в 45°.

2. Малые колебания около положения равновесия. В § 1, n° 3 мы исследовали вопрос об устойчивости равновесия. Теперь мы поставим задачу о движении материальной точки вокруг положения устойчивого равновесия, соответствующего минимуму потенциальной энергии. Эту задачу можно решить приближенно следующим путем. Ради краткости мы ограничимся изучением движения в плоскости xy при допущении, что параллельно оси z не действуют никакие силы.

---

[1]) Кстати, этот результат непосредственно вытекает из очевидного соотношения $|a_1c_1| \leqslant \frac{a_1^2 + c_1^2}{2}$, в котором знак равенства имеет место в том и только в том случае, если $|a_1| = |c_1|$.

Будем считать, что положением устойчивого равновесия является начало координат, и представим потенциальную энергию U в окрестности начала по формуле Тэйлора в следующем виде:

$$U = U_0 + px + qy + \frac{1}{2}(ax^2 + 2bxy + cy^2) + \ldots, \tag{1}$$

где $p = U_x(0, 0)$, $q = U_y(0, 0)$, $a = U_{xx}(0, 0)$, $b = U_{xy}(0, 0)$, $c = U_{yy}(0, 0)$, а U_0 есть произвольное постоянное, которое можно положить равным нулю. Так как начало координат соответствует минимуму потенциальной энергии, то $p = U_x(0, 0) = 0$ и $q = U_y(0, 0) = 0$. Стало быть, в выражении (1) отсутствуют члены нулевой и первой степени. В соответствии с тем, что начало координат дает минимум функции $U(x, y)$, группа членов второй степени

$$Q(x, y) = \frac{1}{2}(ax^2 + 2bxy + cy^2)$$

образует *положительно определенную* квадратичную форму (ср. стр. 222); к тому же мы предположим, что в достаточно малой окрестности положения равновесия потенциальную энергию U можно с достаточной точностью заменить этой квадратичной формой $Q(x, y)$. При сделанных предположениях уравнения движения принимают следующий вид:

$$m\ddot{\mathbf{r}} = -\operatorname{grad} Q$$

или

$$m\ddot{x} = -ax - by, \quad m\ddot{y} = -bx - cy.$$

Эти уравнения легко будет решить, если предварительно повернуть систему координат на подходящий угол φ. Из аналитической геометрии известно, что, при надлежащем выборе угла поворота φ осей координат, преобразование

$$x = \xi\cos\varphi - \eta\sin\varphi, \quad y = \xi\sin\varphi + \eta\cos\varphi$$

переводит положительно определенную квадратичную форму $2Q(x, y) = ax^2 + 2bxy + cy^2$ в выражение вида

$$2Q = \alpha\xi^2 + \beta\eta^2,$$

где ξ и η — новые прямоугольные координаты, а α и β — положительные постоянные. (Дело в том, что уравнение $Q(x, y) = 1$ представляет эллипс, а при подходящем выборе угла φ член с произведением xy можно уничтожить.) В этих новых координатах уравнение движения $m\ddot{\mathbf{r}} = -\operatorname{grad} Q$ эквивалентно следующей системе дифференциальных уравнений:

$$m\ddot{\xi} = -\alpha\xi, \quad m\ddot{\eta} = -\beta\eta,$$

где ξ, η — новые координаты радиус-вектора $\mathbf{r}$. В этой системе переменные отделены: первое уравнение содержит только ξ, второе —

только η, причем оба уравнения однотипны. В т. I, гл. V, § 4, n° 3 мы уже нашли общее решение уравнения такого типа. Имеем

$$\xi = A_1 \sin \sqrt{\frac{\alpha}{m}}\,(t - c_1), \qquad \eta = A_2 \sin \sqrt{\frac{\beta}{m}}\,(t - c_2),$$

где c_1, c_2, A_1, A_2 являются постоянными интегрирования, которые дают возможность конкретизировать процесс движения при произвольно заданных начальных условиях.

Форма решений показывает, что движение вокруг устойчивого положения равновесия представляет собой наложение (суперпозицию) гармонических колебаний по двум «главным направлениям» — по оси ξ и по оси η с круговыми частотами $\sqrt{\frac{\alpha}{m}}$ и $\sqrt{\frac{\beta}{m}}$. Общее исследование этих колебаний, которого мы здесь проводить не будем, показывает, что результирующее движение может иметь весьма разнообразные формы.

Приведем лишь несколько частных случаев таких составных колебаний. Сначала рассмотрим движение, определяемое уравнениями

$$\xi = \sin (t + c), \qquad \eta = \sin (t - c).$$

Исключив время t, получим уравнение траектории

$$(\xi + \eta)^2 \sin^2 c + (\xi - \eta)^2 \cos^2 c = 4 \sin^2 c \cos^2 c;$$

траектория оказалась эллипсом. Оба составляющих колебания имеют одинаковую круговую частоту 1 и одинаковую амплитуду 1, но разность фаз $2c$. Если придавать этой разности фаз все значения от $2c = 0$ до $2c = \frac{\pi}{2}$,

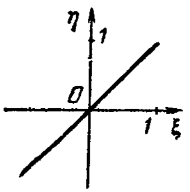


Рис. 96.

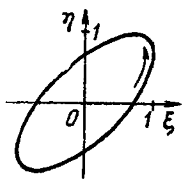


Рис. 97.

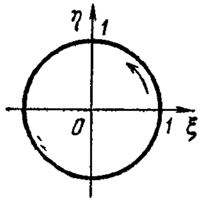


Рис. 98.

то получаются эллипсы всех возможных эксцентриситетов ε от $\varepsilon = 1$ при $c = 0$, когда эллипс вырождается в отрезок прямой $\xi - \eta = 0$ между точками $(-1; -1)$ и $(1; 1)$, до $\varepsilon = 0$ при $c = \frac{\pi}{4}$, когда эллипс переходит в окружность $\xi^2 + \eta^2 = 1$. Таким образом результирующее движение может иметь разнообразные формы от колебания вдоль прямой до вращения по окружности (рис. 96—98).

Рассмотрим теперь движение, определяемое уравнениями

$$\xi = \sin t, \qquad \eta = \sin 2 (t - c).$$

Теперь частоты составляющих колебаний уже не равны и фигуры колебания получаются значительно более сложные. На рис. 99—101 изображены фигуры колебания для $c=0$, $c=\frac{\pi}{8}$ и $c=\frac{\pi}{4}$. В первых двух случаях материальная точка пробегает раз за разом замкнутую кривую в одном и том же направлении, а в третьем случае она колеблется туда и обратно вдоль дуги параболы $\eta=2\xi^2-1$.

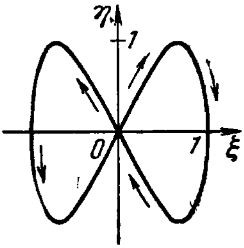


Рис. 99.

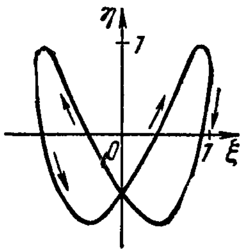


Рис. 100.

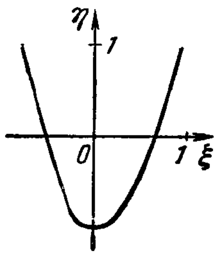


Рис. 101.

Все кривые, служащие траекториями движения при сложении всевозможных гармонических колебаний по двум взаимно перпендикулярным направлениям, носят общее название *фигур Лиссажу*.

3. Движение планет. В рассмотренных выше задачах дифференциальные уравнения движения получались (либо сразу, либо после некоторого преобразования) в таком виде, что каждая координата удовлетворяла своему отдельному дифференциальному уравнению, решая которое, мы определяли эту координату как функцию времени. Теперь же мы поставим и решим важнейшую задачу такого типа, когда систему уравнений движения уже не удается разбить простым путем на отдельные уравнения для каждой из искомых функций, и интегрирование такой системы требует более сложных вычислений. Мы выведем *законы Кеплера о движении планет из закона всемирного тяготения Ньютона.* Пусть в начале координат покоится масса μ (например, Солнце), порождающая гравитационное поле, в котором сила, действующая на единицу массы, дана как вектор-функция точки

$$f(P)=\gamma\mu\,\mathrm{grad}\,\frac{1}{r}=-\gamma\mu\frac{r}{r^3}.$$

Каково движение материальной точки (планеты) массы m под воздействием этого силового поля? Уравнение движения

$$m\ddot{r}=mf \quad \text{или} \quad \ddot{r}=-\gamma\mu\frac{r}{r^3} \tag{1}$$

эквивалентно в координатной записи системе

$$\ddot{x}=-\gamma\mu\frac{x}{r^3},\qquad \ddot{y}=-\gamma\mu\frac{y}{r^3},\qquad \ddot{z}=-\gamma\mu\frac{z}{r^3}. \tag{2}$$

Для того чтобы решить эту систему дифференциальных уравнений, запишем сначала для искомого движения уравнение, выражающее закон сохранения энергии:

$$\frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)-\frac{\gamma\mu m}{r}=C, \tag{3}$$

где C есть число, остающееся постоянным в продолжение всего движения и вполне определяемое начальными условиями.

Из уравнений движения (2) можно вывести следующую систему уравнений, которые, как и уравнение энергии, уже не содержат вторых производных, а содержат лишь первые производные:

$$y\dot{z}-z\dot{y}=c_1, \qquad z\dot{x}-x\dot{z}=c_2, \qquad x\dot{y}-y\dot{x}=c_3. \tag{4}$$

Для вывода, например, первого из этих уравнений умножим третье уравнение системы (2) на y, второе уравнение на z и из первого результата вычтем второй; получится

$$y\ddot{z}-z\ddot{y}=0 \quad \text{или} \quad \frac{d}{dt}(y\dot{z}-z\dot{y})=0,$$

откуда интегрированием получим $y\dot{z}-z\dot{y}=c_1$. Аналогично выводятся остальные два уравнения.

Уравнения (4) можно также вывести векторным путем. Для этого помножим векторное уравнение движения $\ddot{r}=f$ слева векторно на r; получим $[r\ddot{r}]=0$, так как вектор f коллинеарен r. Легко убедиться, что $[r\ddot{r}]$ равно производной по времени от вектор-функции $[r\dot{r}]$; действительно, $\frac{d}{dt}[r\dot{r}]=[\dot{r}\dot{r}]+[r\ddot{r}]=[r\ddot{r}]$. Стало быть, $\frac{d}{dt}[r\dot{r}]=0$, откуда

$$[r\dot{r}]=c, \tag{5}$$

где $c=\{c_1, c_2, c_3\}$ есть произвольный вектор, не зависящий от времени. Ясно, что система уравнений (4) есть лишь координатная запись уравнения (5). Из этого вывода ясно, что уравнение (5) справедливо не только для ньютонова, но и для всякого центрального силового поля, т. е. такого поля, в котором сила коллинеарна радиус-вектору точки.

Вектор $[r\dot{r}]$ называется *моментом скорости*, а вектор $m[r\dot{r}]$ — *моментом количества движения* относительно начала координат.

Соотношения (4) дают возможность значительно упростить нашу задачу. Не ограничивая этим общности, выберем систему координат таким образом, чтобы в начальный момент (скажем, при $t=0$) материальная точка находилась в плоскости xy и чтобы ее скорость в этот же момент тоже лежала в этой плоскости, так что $z(0)=0$ и $\dot{z}(0)=0$. Поэтому при $t=0$ выражения $y\dot{z}-z\dot{y}$ и $z\dot{x}-x\dot{z}$ обращаются в нуль, а так как эти выражения, согласно уравнениям (4), не зависят от времени, то при сделанном выборе системы координат уравнения (4) принимают следующий вид:

$$y\dot{z}-z\dot{y}=0, \qquad z\dot{x}-x\dot{z}=0, \qquad x\dot{y}-y\dot{x}=c_3=k.$$

Из этих уравнений прежде всего вытекает, что все движение происходит в плоскости $z=0$. В самом деле, так как мы естественно исключаем из рассмотрения возможность столкновения планеты с Солнцем, то можно предположить, что три координаты x, y и z никогда не обращаются одновременно все в нуль. Поэтому при $t=0$, когда $z(0)=0$, по крайней мере одна из двух других координат не обращается в нуль; пусть, например, $x(0)\neq 0$. Из равенства $z\dot{x}-x\dot{z}=0$ вытекает, что

$$\frac{d}{dt}\left(\frac{z}{x}\right)=\frac{x\dot{z}-z\dot{x}}{x^2}=0,$$

откуда при всех значениях t имеем $z=ax$, где a — постоянная. Полагая $t=0$, получим $z(0)=ax(0)$, так как $z(0)=0$ и $x(0)\neq 0$, заключаем, что $a=0$. Следовательно, всегда $z=0$.

[Что движение происходит в одной плоскости, проходящей через притягивающий центр, вытекает сразу из векторного уравнения (5), показывающего, что если постоянный вектор $c\neq 0$, то радиус-вектор r всегда перпендикулярен вектору c. Умножив уравнение (5) скалярно на r, получим уравнение плоскости орбиты $cr=0$. Эта плоскость определяется начальным положением r_0 и начальной скоростью $\dot{r}_0$: ее нормальный вектор $c=[r_0\dot{r}_0]$. Если же $c=0$, то из уравнения (5) следует, что скорость $\dot{r}$ всегда коллинеарна r, и движение происходит по радиус-вектору, так что плоскость траектории становится неопределенной.]

Теперь наша задача интегрирования уравнений движения привелась к решению системы двух дифференциальных уравнений первого порядка:

$$\frac{1}{2}m(\dot{x}^2+\dot{y}^2)-\frac{\gamma\mu m}{r}=C, \qquad x\dot{y}-y\dot{x}=k.$$

Введем теперь в эти уравнения вместо x и y полярные координаты r и θ с помощью преобразования

$$x=r\cos\theta, \qquad y=r\sin\theta.$$

Так как

$$\dot{x}^2+\dot{y}^2=\dot{r}^2+r^2\dot{\theta}^2 \quad \text{и} \quad x\dot{y}-y\dot{x}=r^2\dot{\theta},$$

то для полярных координат r и θ, как функций времени t, получается следующая система дифференциальных уравнений:

$$\left.\begin{aligned} \frac{m}{2}(\dot{r}^2+r^2\dot{\theta}^2)-\frac{\gamma\mu m}{r}=C,\\ r^2\dot{\theta}=k. \end{aligned}\right\} \tag{6}$$

Первое из этих уравнений выражает закон сохранения энергии, второе — кеплеровский закон площадей. Из т. I, гл. V, § 2, n° 5 известно, что выражение $\frac{1}{2}r^2\dot{\theta}$ равно производной по времени от площади сектора, описанного радиус-вектором движущейся точки, т. е. равно скорости изменения площади этого сектора. Второе из наших уравнений утверждает, что эта скорость постоянна. Этот результат и называется *первым законом Кеплера* или *законом площадей: радиус-вектор планеты описывает в равные промежутки времени равные площади.*

[Закон площадей сразу вытекает из векторного уравнения (5); в силу геометрического смысла векторного произведения, из (5) получается, что скорость изменения площади сектора равна $\frac{1}{2}|[r\dot{r}]| = \frac{1}{2}|c|$, т. е. она постоянна.]

Если «постоянная площадей» k равна нулю, то $\dot{\theta} = 0$ и $\theta = \text{const}$, т. е. материальная точка должна двигаться по прямой, проходящей через начало координат. Этот частный случай мы не станем рассматривать и будем определенно предполагать, что $k \neq 0$.

Займемся теперь определением геометрической формы орбиты, отложив покамест вопрос о зависимости движения от времени. Для этой цели рассматриваем угол θ как функцию от r или r как функцию от θ и из уравнений (6) определим $dr/d\theta$ как функцию от r. Выполняем это следующим путем. Выражение $\dot{\theta} = \frac{k}{r^2}$ из закона площадей и выражение $\dot{r} = \frac{dr}{dt} = \frac{dr}{d\theta}\dot{\theta} = \frac{k}{r^2}\frac{dr}{d\theta}$ подставляем в первое из уравнений (6) и сразу получаем дифференциальное уравнение орбиты

$$\frac{m}{2}\left\{\frac{k^2}{r^4}\left(\frac{dr}{d\theta}\right)^2 + \frac{k^2}{r^2}\right\} - \frac{\gamma\mu m}{r} = C$$

или

$$\left(\frac{dr}{d\theta}\right)^2 = r^4\left(\frac{2C}{mk^2} + \frac{2\gamma\mu}{k^2}\frac{1}{r} - \frac{1}{r^2}\right).$$

Выражение в скобках в правой части можно преобразовать так:

$$\frac{2C}{mk^2} + \frac{2\gamma\mu}{k^2}\frac{1}{r} - \frac{1}{r^2} = -\left(\frac{1}{r} - \frac{\gamma\mu}{k^2}\right)^2 + \frac{\gamma^2\mu^2}{k^4} + \frac{2C}{mk^2} =$$
$$= -\left(\frac{1}{r} - \frac{\gamma\mu}{k^2}\right)^2 + \frac{\gamma^2\mu^2}{k^4}\left(1 + \frac{2Ck^2}{m\gamma^2\mu^2}\right).$$

Для сокращения введем следующие обозначения:

$$\frac{\gamma\mu}{k^2} = \frac{1}{p}, \qquad 1 + \frac{2Ck^2}{m\gamma^2\mu^2} = \varepsilon^2.$$

Тогда дифференциальное уравнение орбиты примет следующий вид:

$$\left(\frac{dr}{d\theta}\right)^2 = r^4\left\{\frac{\varepsilon^2}{p^2} - \left(\frac{1}{r} - \frac{1}{p}\right)^2\right\}.$$

Оно упростится, если ввести вместо r новую неизвестную функцию

$$u = \frac{1}{r} - \frac{1}{p}, \qquad \text{так что} \qquad \frac{du}{d\theta} = -\frac{1}{r^2}\frac{dr}{d\theta}.$$

Новая искомая функция u удовлетворяет дифференциальному уравнению

$$\left(\frac{du}{d\theta}\right)^2=\frac{\varepsilon^2}{p^2}-u^2 \quad \text{или} \quad \frac{du}{d\theta}=\sqrt{\frac{\varepsilon^2}{p^2}-u^2}.$$

Это дифференциальное уравнение первого порядка с отделяющимися переменными. Оно интегрируется квадратурой

$$\theta-\theta_0=\int\frac{du}{\sqrt{\frac{\varepsilon^2}{p^2}-u^2}}=\arcsin\frac{pu}{\varepsilon}$$

(см. т. I, гл. IV, стр. 242). Таким образом, уравнение орбиты получается в следующем виде:

$$u=\frac{1}{r}-\frac{1}{p}=\frac{\varepsilon}{p}\sin(\theta-\theta_0).$$

Угол θ_0 можно выбрать произвольно, так как безразлично, от какого неподвижного радиус-вектора отсчитывать угол θ. Примем $\theta_0=\frac{\pi}{2}$, так что значению $u=0$ соответствует значение $\theta=\frac{\pi}{2}$. Тогда $\frac{1}{r}-\frac{1}{p}=-\frac{\varepsilon}{p}\cos\theta$ и окончательно уравнение орбиты будет

$$r=\frac{p}{1-\varepsilon\cos\theta}.$$

Из аналитической геометрии известно, что это полярное уравнение кривой второго порядка, один из фокусов которой находится в начале координат. Полученный результат дает *второй закон Кеплера: планеты* [и кометы] *движутся по коническим сечениям, в одном из фокусов которых находится Солнце.*

Интересна связь между постоянными интегрирования

$$p=\frac{k^2}{\gamma\mu},\qquad \varepsilon^2=1+\frac{2Ck^2}{m\gamma^2\mu^2}$$

и константами, характеризующими вид орбиты. Величина p называется *параметром* конического сечения; у эллипса и гиперболы параметр p связан с полуосями a и b простой формулой $p=\frac{b^2}{a}$.

Эксцентриситет ε (>0) определяет вид и форму конического сечения: оно является эллипсом при $\varepsilon<1$, параболой при $\varepsilon=1$ и гиперболой, если $\varepsilon>1$.

Из равенства $\varepsilon^2=1+\frac{2Ck^2}{m\gamma^2\mu^2}$ сразу видно, что эти три различные возможности можно поставить в связь с постоянной C, характеризующей энергию: орбита будет эллипсом, если $C<0$, параболой, если $C=0$, и гиперболой, если $C>0$.

Пусть материальная точка в момент $t=0$ помещена в силовом поле в положении r_0 с начальной скоростью v_0. Тогда из соотношения

$$C=\frac{m}{2}v_0^2-\frac{\gamma\mu m}{r_0}$$

вытекает поразительный вывод, что [при $k\neq 0$] вид орбиты (эллипс, парабола или гипербола) нисколько не зависит от направления начальной скорости и зависит только лишь от ее модуля v_0.

Теперь, когда найдена функция $r(\theta)=\frac{p}{1-\varepsilon\cos\theta}$, можно уже определить протекание движения во времени. Из закона площадей $r^2\dot\theta=k$ имеем

$$\int_{\theta_0}^{\theta} r^2 d\theta=\int_{\theta_0}^{\theta}\frac{p^2\,d\theta}{(1-\varepsilon\cos\theta)^2}=k(t-t_0).$$

Таким образом определена зависимость от времени полярного угла θ, а следовательно, и полярного радиуса r.

Третий закон Кеплера является простым следствием первых двух. Как известно, этот закон утверждает, что *у эллиптических орбит отношение квадрата периода обращения к кубу большой полуоси есть величина постоянная, одинаковая для всех планет, и вполне определяется заданным гравитационным полем.*

Обозначим через T период обращения и через a — большую полуось; тогда

$$\frac{T^2}{a^3}=\text{const}$$

и эта постоянная одинакова для всех планет данной планетной системы и вполне определяется массой μ, порождающей гравитационное поле, и постоянной тяготения

Для доказательства воспользуемся законом площадей в проинтегрированном виде (см. выше):

$$\int_{\theta_0}^{\theta} r^2\,d\theta=k(t-t_0).$$

Возьмем интеграл слева по промежутку от $\theta_0=0$ до $\theta=2\pi$; тогда в левой части получится удвоенная площадь орбитального эллипса, равная $2\pi ab$, а справа $t-t_0=T$ (период обращения). Следовательно,

$$2\pi ab=kT.$$

С другой стороны, постоянная k связана с полуосями a и b соотношением $\frac{k^2}{\gamma\mu}=\frac{b^2}{a}(=p)$. Исключив из этих двух равенств букву k, получим соотношение

$$\frac{T^2}{a^3}=\frac{4\pi^2}{\gamma\mu},$$

выражающее третий закон Кеплера.

Упражнения

1. Доказать, что при $t \to \infty$ скорость небесного тела стремится к нулю, если оно движется по параболической орбите, и к положительному пределу, если его орбитой является гипербола.

2*. Планета движется по эллипсу. Построим на большой оси эллипса, как на диаметре, вспомогательную окружность. Пусть P_0 — перигелий, т. е. вершина эллипса, ближайшая к Солнцу, P — положение планеты в момент времени t, P' — точка вспомогательной окружности, лежащая на одном с точкой P перпендикуляре к большой оси и с той же стороны от нее, что и P. Обозначим через $\omega = \omega(t)$ эксцентрическую аномалию, т. е. угол P_0MP, где M — центр эллипса. Доказать, что ω и t связаны уравнением Кеплера

$$k(t - t_0) = ab(\omega - \varepsilon \sin \omega),$$

где t_0 — момент прохождения планеты через перигелий.

3. Доказать, что материальная точка, притягиваемая к центру O силой, величина которой пропорциональна расстоянию от O, движется по эллипсу с центром O.

4. Доказать, что орбита материальной точки, отталкиваемой от центра O силой, величина которой есть $f(r)$, где f — заданная функция, имеет в полярных координатах (r, θ) уравнение

$$\theta - \theta_0 = \int \frac{dr}{r^2 \sqrt{\frac{2c}{k^2} + \frac{2}{k^2} \int f(r)\, dr - \frac{1}{r^2}}}.$$

5. Доказать, что уравнение орбиты материальной точки, отталкиваемой от центра O силой, величина которой равна $\frac{\mu}{r^3}$, есть

$$\frac{1}{r} = \begin{cases} \frac{2c}{k^2\gamma} \cos(\gamma\theta + \varepsilon) & \text{при } \mu < k^2, \\ \frac{2c}{k^2\gamma} \operatorname{ch}(\gamma\theta + \varepsilon) & \text{при } \mu > k^2, \end{cases}$$

где $\gamma = \sqrt{\left|1 - \frac{\mu}{k^2}\right|}$, а ε — постоянная интегрирования.

6. Доказать, что в орбите, описываемой в центральном силовом поле вида $\mathbf{f} = -f(r)\, \mathbf{r}^\circ$, модуль силы притяжения (на единицу массы) есть

$$f = |\mathbf{f}| = \frac{k^2}{q^3} \frac{dq}{dr},$$

где q есть расстояние от центра притяжения до касательной к орбите, а k — постоянная площадей.

Пользуясь этим, доказать, что орбита может быть кардиоидой $r = a(1 + \cos\theta)$, если сила притяжения на единицу массы пропорциональна r^{-4}.

§ 3. Некоторые сведения из общей теории дифференциальных уравнений первого порядка

Рассмотрение общей теории обыкновенных дифференциальных уравнений вышло бы далеко за рамки этой книги. Однако мы здесь наметим хотя бы очерк элементов этой теории, и начнем с дифференциального уравнения первого порядка.

1. Геометрический смысл дифференциального уравнения первого порядка. Общий вид дифференциального уравнения первого порядка есть

$$F(x, y, y') = 0, \qquad (A)$$

причем мы будем предполагать, что функция F непрерывно дифференцируема по всем своим аргументам x, y, y'. Важно уяснить себе наглядно геометрический смысл этого уравнения. В некоторой плоской области это уравнение налагает условие на направление касательной к любой кривой $y = y(x)$, проходящей через точку (x, y) этой области. Предположим, что в некоторой области G плоскоси xy уравнение (A) можно однозначно разрешить относительно y', т. е привести его к виду

$$y' = f(x, y), \qquad (B)$$

где функция $f(x, y)$ непрерывно дифференцируема по x и y. Тогда дифференциальное уравнение (B), а вместе с ним и (A) приводит в соответствие каждой точке (x, y) области G определенное *направление* [или *линейный элемент*; не говорят «элементарный вектор» только потому, что, в отличие от вектора, на этом линейном элементе не выделено какое-либо одностороннее направление.] Таким образом, дифференциальное уравнение изображается геометрически в виде *поля направлений*, и следовательно, задача решения дифференциального уравнения истолковывается геометрически как задача о нахождении таких кривых, которые, так сказать, вписываются в это поле направлений, т. е. в каждой своей точке имеют касательную именно того направления, которое предписано в этой точке уравнением $y' = f(x, y)$. Эти кривые называются *интегральными кривыми дифференциального уравнения*.

[С геометрической точки зрения естественно и целесообразно не исключать из рассмотрения те точки (x, y) области G, в которых $f(x, y)$ обращается в бесконечность; можно считать, что таким точкам дифференциальное уравнение (B) относит направление, параллельное оси y.]

Наглядному представлению кажется сразу ясным, что через каждую точку (x, y) области G проходит одна и только одна интегральная кривая дифференциального уравнения $y' = f(x, y)$. Точное положение дела устанавливается следующей основной *теоремой существования*:

Если в дифференциальном уравнении $y' = f(x, y)$ функция f непрерывна в области G и имеет в ней непрерывную производную по y, то через каждую точку (x_0, y_0) области G проходит одна и только одна интегральная кривая, т. е. существует одно и только одно решение $y(x)$ дифференциального уравнения, для которого $y(x_0) = y_0$.

Доказательство этой теоремы будет дано в n° 4, здесь же мы ограничимся рассмотрением нескольких примеров.

а) Дано дифференциальное уравнение

$$y' = -\frac{x}{y}.$$

За область G примем, например, полуплоскость $y<0$. В каждой точке (x, y) поля направлений элемент касательной перпендикулярен радиус-вектору этой точки. Отсюда вытекает геометрически, что полуокружности с центрами в начале координат должны быть интегральными кривыми нашего дифференциального уравнения. Это нетрудно проверить аналитически: из уравнений этих полуокружностей $y=-\sqrt{c^2-x^2}$ получаем $y'=\frac{x}{\sqrt{c^2-x^2}}$, так что для них действительно $y'=-\frac{x}{y}$.

б) Поле направлений дифференциального уравнения

$$y' = \frac{y}{x}$$

имеет, очевидно, в каждой своей точке направление прямой, соединяющей эту точку с началом координат. Отсюда ясно, что полупрямые, выходящие из начала координат, «вписываются» в это поле и являются поэтому интегральными кривыми. И действительно, сразу видно, что функции $y=cx$ удовлетворяют дифференциальному уравнению при любом значении постоянной c.

В начале координат поле направлений этого дифференциального уравнения уже не определено однозначно [вернее сказать, оно вовсе не определено], и с этим связан тот факт, что начало координат является для дифференциального уравнения «особой точкой», через которую проходит бесконечное множество интегральных кривых.

в) Легко проверить аналитически, что дифференциальное уравнение

$$y' = \frac{x}{y} \qquad (y \neq 0)$$

имеет своими интегральными кривыми семейство гипербол

$$y=\sqrt{c^2+x^2},$$

а интегральные кривые дифференциального уравнения

$$y' = -\frac{y}{x} \qquad (x \neq 0)$$

составляют семейство гипербол

$$y=\frac{c}{x};$$

в обоих случаях c есть постоянная, различные значения которой определяют отдельные кривые семейства.

Сформулированная выше теорема существования показывает, что совокупность решений дифференциального уравнения первого порядка $y'=f(x, y)$ представляет собой однопараметрическое семейство функций вида $y=\varphi(x, c)$, содержащих, помимо x, еще и параметр c — *произвольную постоянную интегрирования* (ср. т. I, стр. 597). Таким параметром может, например, служить $c=y_0=y(0)$. Обычное интегрирование функции $f(x)$ является только частным случаем интегрирования нашего дифференциального уравнения, именно тем частным случаем, когда $f(x, y)$ фактически не содержит y. В этом случае все линейные элементы поля направлений полностью определяются одной лишь абсциссой x, и сразу ясно, что интегральные кривые получаются одна из другой параллельным перенесением по направлению оси y. Аналитически это выражается тем хорошо известным фактом, что при неопределенном интегрировании, т. е. при решении дифференциального уравнения $y' = f(x)$, результат содержит произвольное постоянное слагаемое.

Геометрическое истолкование дифференциального уравнения первого порядка подсказывает метод его *приближенного графического решения*, т. е. графического построения интегральных кривых. Этот метод совершенно подобен частному случаю графического интегрирования функции $f(x)$ (т. I, гл. II, § 5) и заключается в том, что интегральную кривую заменяют ломаной линией, каждое звено которой имеет направление, приписанное дифференциальным уравнением его начальной точке (или какой-нибудь другой его точке). Такую ломаную можно построить, исходя из любой точки области G. Чем меньше выбранная длина каждого звена ломаной, тем точнее звенья ломаной «вписываются» или «укладываются» в поле направлений дифференциального уравнения, не только в своих начальных точках, но и на всем своем протяжении. Отметим здесь без доказательства важный факт, что построенная таким способом ломаная, при неограниченном уменьшении длин ее сторон, имеет своим пределом интегральную кривую, проходящую через исходную точку ломаной. Для применения графического метода необязательно, чтобы функция $f(x, y)$ была задана аналитически; достаточно, чтобы она была задана графически или таблицей.

Практическое выполнение графического интегрирования делается систематичнее и легче, если для построения поля направлений пользоваться методом *изоклин*. Для этого предварительно строят геометрические места точек, которым дифференциальное уравнение приписывает одинаковые направления, т. е. строят семейство кривых $f(x, y) = c = \text{const}$. Эти геометрические места и называются *изоклинами* (кривыми равного наклона). Каждому значению постоянной c соответствует определенное направление, которое можно, например, нанести на вспомогательный чертеж рядом с картиной изоклин. Проводимая интегральная кривая должна пересекать каждую изоклину, имея в точке пересечения направление, которое берут из вспомогательного чертежа; поэтому построение ломаной, аппроксимирующей интегральную кривую, легко выполняется путем проведения параллельных прямых.

На рис. 103 показано графическое интегрирование уравнения $y' = \frac{\sqrt{x^2 + y^2}}{x}$. Изоклинами являются здесь полупрямые, выходящие из начала координат. Направления $y' = c$, соответствующие перенумерованным изоклинам, даны на вспомогательном рис. 102 под теми же номерами.

Это дифференциальное уравнение можно решить в квадратурах (как однородное уравнение (см. т. I, стр. 599) или путем перехода

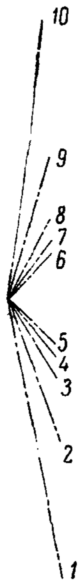


Рис. 102.

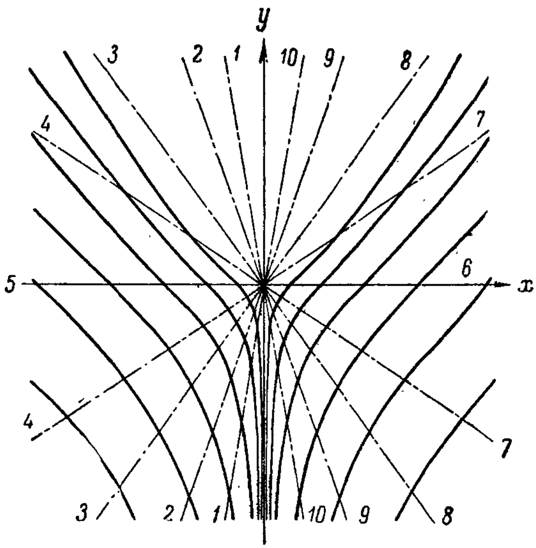


Рис. 103.

к полярным координатам), но полученное аналитически общее решение отнюдь не столь ясно и не так легко поддается исследованию.

2. Дифференциальное уравнение семейства кривых. Особые решения. Ортогональные траектории. Согласно теореме существования, каждому дифференциальному уравнению первого порядка соответствует однопараметрическое семейство кривых. Естественно возникает вопрос: обратимо ли это соответствие? Другими словами: для всякого ли однопараметрического семейства кривых $\varphi(x, y; c)=0$ или $y=g(x, c)$ существует свое дифференциальное уравнение

$$F(x, y, y')=0,$$

которому удовлетворяют все кривые семейства? И далее, как найти это дифференциальное уравнение? При этом существенно, что дифференциальное уравнение семейства кривых уже не содержит параметра c этого семейства. Таким образом, это дифференциальное уравнение осуществляет задание семейства кривых без использования параметра.

Такое дифференциальное уравнение легко получить фактически. Дифференцируя по x уравнение семейства

$$\varphi(x, y; c) = 0, \tag{1}$$

имеем

$$\varphi_x + \varphi_y y' = 0. \tag{2}$$

Если φ_y не равно тождественно нулю, то из уравнений (1) и (2) исключим параметр c, и в результате получится искомое дифференциальное уравнение. Это исключение параметра c выполнимо во всякой области плоскости xy, в которой из уравнения (1) можно выразить параметр c через x и y. Действительно, тогда достаточно заменить в уравнении (2) параметр c, входящий в φ_x и φ_y, полученным выражением $c = c(x, y)$, чтобы прийти к искомому дифференциальному уравнению семейства кривых. Рассмотрим несколько примеров.

а) Возьмем семейство концентрических окружностей $x^2 + y^2 - c^2 = 0$. При дифференцировании этого уравнения по x параметр c исключается и сразу получается дифференциальное уравнение семейства

$$x + yy' = 0$$

в согласии с примером а) из n° 1.

б) Рассмотрим семейство окружностей радиуса 1 с центрами на оси абсцисс; уравнение этого семейства

$$(x - c)^2 + y^2 = 1.$$

Дифференцирование по x дает

$$(x - c) + yy' = 0.$$

Исключив c из обоих уравнений, приходим к дифференциальному уравнению семейства:

$$y^2(y'^2 + 1) = 1.$$

в) Уравнение $y = (x - c)^2$ семейства парабол, касающихся оси x, дает при дифференцировании уравнение $y' = 2(x - c)$. Исключение параметра c приводит к дифференциальному уравнению

$$(y')^2 = 4y.$$

Внимательное рассмотрение последних двух примеров показывает, что соответствующему дифференциальному уравнению удовлетворяют не только кривые своего семейства, но и другие линии; в примере б) это прямые $y = -1$ и $y = 1$, в примере в) это ось абсцисс $y = 0$, что легко проверить подстановкой в дифференциальное уравнение. Впрочем, эти факты непосредственно вытекают без вычислений из геометрического смысла дифференциального уравнения. В самом деле, прямые $y = 1$ и $y = -1$ являются огибающими семейства окружностей примера б), а прямая $y = 0$ является огибающей семейства парабол примера в), а так как огибающая в каждой своей точке касается одной из кривых семейства, то она в этой точке должна иметь именно то направление, которое предписано полем направлений. Поэтому всякая огибающая семейства интегральных кривых тоже должна удовлетворять дифференциальному уравнению. Такие решения дифференциального уравнения, которые получаются путем построения огибающей однопараметрического семейства интегральных кривых, называются *особыми решениями.*

Замечательно, что особые решения дифференциального уравнения $F(x, y, y')=0$ можно получить и не зная еще его общего решения, т. е. не находя предварительно однопараметрического семейства частных решений. Вспомним, что по теореме существования решение дифференциального уравнения определено *однозначно* в окрестности точки (x, y), если в этой окрестности возможно привести дифференциальное уравнение к виду $y'=f(x, y)$, где $f(x, y)$ есть непрерывно дифференцируемая функция. Отсюда вытекает, что в точке, через которую помимо частного решения (т. е. кривой семейства) проходит еще и особое решение, приведение дифференциального уравнения к такому виду невозможно. Поэтому в окрестности такой точки (x, y) уравнение $F(x, y, y')=0$ не может быть разрешено однозначно относительно y'. Однако теорема существования неявной функции (гл. III, § 1) гарантирует существование такого явного выражения вида $y'=f(x, y)$, если в рассматриваемой точке $F_{y'}\neq 0$. Стало быть, необходимое (хотя отнюдь не достаточное) условие существования в точке (x, y) *особого решения* выражается уравнением

$$\frac{\partial}{\partial y'}F(x, y, y')=0.$$

Исключив y' из этого уравнения и заданного дифференциального уравнения, получим уравнение, связывающее только x и y, которому должно удовлетворять особое решение, если оно существует.

Данные выше примеры иллюстрируют это правило. Так, из дифференциального уравнения $y^2(y'^2+1)=1$ примера б) дифференцированием по y' получаем второе уравнение $y^2y'=0$. Исключение из этих двух уравнений величины y' дает уравнение $y^2=1$ или $y=\pm 1$, т. е. обнаруженные выше особые решения.

Пусть область G однозначно покрывается однопараметрическим семейством кривых $\Phi(x, y)=c$. Отнесем каждой точке P области G направление касательной той кривой семейства, которая проходит через эту точку. Тогда получится поле направлений, описываемое дифференциальным уравнением $y'=-\frac{\Phi_x}{\Phi_y}$. Если же каждой точке P отнесем направление нормали кривой семейства, проходящей через эту точку, то получится другое поле направлений, описываемое дифференциальным уравнением

$$y'=\frac{\Phi_y}{\Phi_x}.$$

Решения этого дифференциального уравнения называются *ортогональными траекториями* исходного семейства кривых $\Phi(x, y)=c$. Кривые данного семейства и его ортогональные траектории пересекаются всюду под прямыми углами.

Отсюда ясно, что если семейство кривых задано своим дифференциальным уравнением $y'=f(x, y)$, то дифференциальное уравнение семейства ортогональных траекторий можно написать сразу, не нуждаясь в предварительном интегрировании дифференциального уравнения, определяющего заданное семейство кривых; искомое дифференциальное уравнение ортогональных траекторий есть, очевидно,

$$y' = -\frac{1}{f(x, y)}.$$

В данном выше примере а) семейства окружностей $x^2 + y^2 - c^2 = 0$ находим из дифференциального уравнения $x + yy' = 0$ этого семейства дифференциальное уравнение его ортогональных траекторий: $y' = \frac{y}{x}$. Следовательно, ортогональные траектории составляют пучок прямых, проходящих через начало координат [см. n° 1, пример б)].

Рассмотрим теперь семейство софокусных парабол (гл. III, § 2, в конце n° 1, стр. 148, ср. рис. 36 на стр. 157)

$$y^2 - 2p\left(x + \frac{p}{2}\right) = 0 \quad \text{при} \quad p > 0.$$

Дифференциальное уравнение этого семейства таково:

$$y' = \frac{-x + \sqrt{x^2 + y^2}}{y}.$$

Следовательно, дифференциальное уравнение ортогональных траекторий будет

$$y' = \frac{-1}{\frac{1}{y}\left(-x + \sqrt{x^2 + y^2}\right)} = \frac{1}{y}\left(-x - \sqrt{x^2 + y^2}\right),$$

а общим решением этого уравнения является семейство парабол $y^2 - 2p\left(x + \frac{p}{2}\right) = 0$ со значениями $p < 0$; это тоже семейство софокусных парабол, имеющее к тому же общие фокусы с исходным семейством.

3. Интегрирующий множитель. Запишем дифференциальное уравнение $y' = f(x, y)$ в следующем виде:

$$dy - f(x, y)\,dx = 0,$$

где dx — дифференциал независимой переменной, а dy — дифференциал искомой функции. Если умножить это уравнение на любой отличный от нуля множитель $b(x, y)$, то получим дифференциальное уравнение вида

$$a(x, y)\,dx + b(x, y)\,dy = 0,$$

равносильное исходному, и задача заключается в том, чтобы найти все те функции $y(x)$, для которых дифференциалы dx и dy удовлетворяют этому дифференциальному уравнению тождественно относительно x.

Существует случай, когда общее решение можно сразу написать. Это тот случай, когда выражение $a\,dx+b\,dy$ является полным дифференциалом некоторой функции $F(x, y)$, т. е. когда существует такая функция $F(x, y)$, что $a=\frac{\partial F}{\partial x}$ и $b=\frac{\partial F}{\partial y}$. При этом условии дифференциальное уравнение $a\,dx+b\,dy=0$ можно записать в виде

$$dF=0,$$

где dF обозначает полный дифференциал функции $F(x, y)$. Ясно, что общее решение этого уравнения есть

$$F(x, y)=c;$$

из этого неявного уравнения можно выразить y как функцию от x, содержащую еще и постоянную интегрирования c.

Согласно гл. V, § 1, стр. 378, необходимым и достаточным условием того, чтобы выражение $a\,dx+b\,dy$ было полным дифференциалом некоторой функции $F(x, y)$, является выполнение критерия интегрируемости $\frac{\partial a}{\partial y}=\frac{\partial b}{\partial x}$ (если притом эти частные производные непрерывны). Если этот критерий выполняется, то криволинейный интеграл выражения $a\,dx+b\,dy$ не зависит от пути интегрирования и, при фиксированной начальной точке P_0 этого пути, представляет собой функцию $F(x, y)$ конечной точки $P(x, y)$, и при помощи этой функции F и получается данное выше общее решение $F(x, y)=c$.

Однако, как правило, коэффициенты $a(x, y)$ и $b(x, y)$ дифференциального уравнения $a\,dx+b\,dy=0$ не удовлетворяют критерию интегрируемости. Не выполняется, например, этот критерий для дифференциального уравнения $dx+\frac{y}{x}dy=0$. В таких случаях можно попытаться помножить дифференциальное уравнение на множитель $\mu(x, y)$, выбранный с таким расчетом, чтобы после умножения новые коэффициенты μa и μb удовлетворяли условию интегрируемости. Тогда общее решение дифференциального уравнения будет иметь вид

$$\int\limits_L (\mu a\,dx+\mu b\,dy)=c,$$

где путь интегрирования L криволинейного интеграла можно выбрать таким образом, чтобы возможно более упростить вычисления; в итоге решение задачи приведется к обыкновенному интегрированию. Выражение $\mu(x, y)$, обладающее описанным выше свойством, называется *интегрирующим множителем* дифференциального уравнения. Для уравнения $dx+\frac{y}{x}dy=0$ таким множителем является $\mu(x, y)=x$. Действительно, после умножения на x уравнение примет вид $x\,dx+$ $+y\,dy=0$, в котором левая часть есть полный дифференциал функции $\frac{x^2+y^2}{2}$. Итак, $d\frac{x^2+y^2}{2}=0$, откуда $x^2+y^2=2C$, т. е. общее решение дифференциального уравнения представляет семейство окружностей с общим центром в начале координат, в согласии с результатом примера а) в n° 1.

Как видно из его определения, интегрирующий множитель удовлетворяет дифференциальному уравнению

$$\frac{\partial}{\partial y}(\mu a)=\frac{\partial}{\partial x}(\mu b) \qquad \text{или} \qquad a\mu_y-b\mu_x+(a_y-b_x)\mu=0,$$

а это уже дифференциальное уравнение *с частными производными* (см. т. I, стр. 596). Стало быть, нахождение интегрирующего множителя есть задача отнюдь не более легкая, чем задача решения исходного дифференциального уравнения; однако во многих случаях удается найти интегрирующий множитель по соображению, как, например, в приведенном выше примере. Впрочем, интегрирующий множитель сохраняет преимущественно теоретический интерес, и мы не имеем возможности входить здесь в дальнейшие подробности.

4. Теорема существования и единственности решения. Займемся теперь доказательством формулированной в n° 1 теоремы существования и единственности решения дифференциального уравнения $y'=f(x, y)$. При этом можно допустить без потери общности, что речь идет о том решении $y(x)$, для которого $y(0)=0$. В самом деле, если поставить вопрос о решении $y(x)$, удовлетворяющем более общему начальному условию $y(x_0)=y_0$, то можно ввести вместо x и y новые переменные $\xi=x-x_0$ и $\eta=y-y_0$, и тогда получится для ξ и η дифференциальное уравнение

$$\frac{d\eta}{d\xi}=f(\xi+x_0,\ \eta+y_0)=f_1(\xi,\ \eta)$$

того же типа, что и первоначальное, но уже с начальным условием $\eta(0)=0$.

Для доказательства можно ограничиться достаточно малой окрестностью точки $x=0$, ибо, когда будут доказаны существование и единственность решения для такого интервала, окружающего точку $x=0$, то можно будет затем применить это доказательство к достаточно малой окрестности граничной точки первого интервала и таким путем продолжить решение на примыкающий интервал и т. д.

Сперва убедимся в том, что двух различных решений дифференциального уравнения, удовлетворяющих заданному начальному условию, быть не может. Действительно, если бы существовали два решения $y_1(x)$ и $y_2(x)$, то их разность $\varphi(x)=y_1-y_2$ удовлетворяла бы уравнению

$$\varphi'(x)=f\{x,\ y_1(x)\}-f\{x,\ y_2(x)\}.$$

Применив к правой части теорему о среднем значении, получим

$$\varphi'(x)=(y_1-y_2)f_y(x,\ \bar{y})=\varphi(x)f_y(x,\ \bar{y}),$$

где $\bar{y}$ есть некоторое значение, промежуточное между y_1 и y_2. Функции $y_1(x)$ и $y_2(x)$ непрерывны в окрестности $|x| \leqslant a$ начала координат, и при $x=0$ обе они обращаются в нуль. Поэтому обе функции ограничены в этой окрестности, и существует такое число b, что $|y_1(x)| \leqslant b$ и $|y_2(x)| \leqslant b$, а стало быть, и $|\bar{y}| \leqslant b$ при $|x| \leqslant a$. Обозначим через M какую-нибудь верхнюю границу величины $|f_y|$ в области $|x| \leqslant a$, $|y| \leqslant b$ и через D наибольшее значение величины $|\varphi(x)|$ в интервале $|x| \leqslant a$; предположим, что это значение D принимается при $x=\xi$. Тогда

$$|\varphi'(x)| = |\varphi(x) f_y(x, \bar{y})| \leqslant DM.$$

Следовательно,

$$D = |\varphi(\xi)| = \left|\int_0^{\xi} \varphi'(x)\,dx\right| \leqslant |\xi|\,DM \leqslant aDM.$$

Число a можно выбрать настолько малым, что $aM < 1$, потому что, если $|f_y(x, y)| < M$ в некоторой области R: $|x| \leqslant a$, $|y| \leqslant b$, то это неравенство остается в силе в любой области, получающейся из R в результате уменьшения числа a. Но если $aM < 1$, то из неравенства $D \leqslant aMD$ вытекает, что $D=0$. Это значит, что в таком интервале $|x| \leqslant a$ оба решения равны: $y_1(x) = y_2(x)$.

Руководящей идеей этого доказательства является известный факт, что если подынтегральная функция ограничена, то значение интеграла имеет тот же порядок малости, что и промежуток интегрирования, когда этот промежуток стремится к нулю.

Перейдем теперь к доказательству существования решения. Мы *построим* это решение с помощью *метода итерации* или *последовательного приближения*; кстати, этот метод очень важен и для приложений, в частности для численного решения дифференциальных уравнений. Решение получается при этом в виде предельной функции последовательности приближенных решений $y_0(x)$, $y_1(x)$, $y_2(x)$, ... В качестве исходного или нулевого приближения берем $y_0(x)=0$. Первое приближение мы строим в соответствии с дифференциальным уравнением так:

$$y_1(x) = \int_0^x f(\xi, 0)\,d\xi;$$

с его помощью получаем второе приближение

$$y_2(x) = \int_0^x f(\xi, y_1(\xi))\,d\xi$$

и, продолжая так шаг за шагом, строим вообще по $(n-1)$-му приближению n-е приближение

$$y_n(x) = \int_0^x f(\xi, y_{n-1}(\xi))\,d\xi.$$

Если в интервале $|x| \leqslant a$ эти приближающие функции сходятся равномерно к предельной функции $y(x)$, то можно выполнить предельный переход под знаком интеграла, и для предельной функции получится уравнение

$$y(x) = \int_0^x f(\xi, y(\xi))\, d\xi;$$

дифференцирование этого уравнения по x даст $y' = f(x, y)$, так что предельная функция $y(x)$ действительно является искомым решением.

Итак, надо только доказать равномерную сходимость последовательности функций $y_0(x)$, $y_1(x)$, $y_2(x)$, ... Это доказательство мы проведем для достаточно малого интервала $|x| \leqslant a$. Введем вспомогательную функцию $\varphi_n(x) = y_{n+1}(x) - y_n(x)$ и обозначим через D_n наибольшее значение абсолютной величины $|\varphi_n(x)|$ этой функции в интервале $|x| \leqslant a$.

Имеем

$$\varphi_n'(x) = y_{n+1}'(x) - y_n'(x) = f(x, y_n) - f(x, y_{n-1}).$$

Применив сюда теорему о среднем значении, получим

$$\varphi_n'(x) = (y_n - y_{n-1}) f_y(x, \bar{y}_{n-1}) = \varphi_{n-1}(x) f_y(x, \bar{y}_{n-1}(x)),$$

где $\bar{y}_{n-1}$ есть некоторое промежуточное значение между y_{n-1} и y_n. Пусть в прямоугольнике $|x| \leqslant a$, $|y| \leqslant b$ выполняются неравенства $|f_y(x, y)| \leqslant M$, $|f(x, y)| \leqslant M_1$. Предположим еще, что функции $y_n(x)$ удовлетворяют в интервале $|x| \leqslant a$ неравенству $|y_n| \leqslant b$. Тогда, в силу определения функции $y_{n+1}(x)$, в этом интервале

$$|y_{n+1}(x)| = \left| \int_0^x f(\xi, y_n(\xi))\, d\xi \right| \leqslant |x| M_1 \leqslant a M_1.$$

Поэтому мы выберем верхнюю границу a для $|x|$ настолько малой, чтобы было $aM_1 \leqslant b$. Тогда в интервале $|x| \leqslant a$ непременно будет $|y_{n+1}(x)| \leqslant b$. Так как для $y_0(x) = 0$ очевидно, что $|y_0(x)| \leqslant b$, то во всем интервале $|x| \leqslant a$ при всех значениях n будет $|y_n(x)| \leqslant b$. Следовательно, можно в равенстве

$$\varphi_{n+1}(x) = \int_0^x f_y(\xi, \bar{y}_n(\xi))\, \varphi_n(\xi)\, d\xi$$

оценить интеграл справа, пользуясь неравенством $|f_y| \leqslant M$, и тогда для D_{n+1} (наибольшего значения величины $|\varphi_{n+1}(x)|$ в интервале $|x| \leqslant a$) получится соотношение

$$D_{n+1} \leqslant a M D_n.$$

Теперь, если нужно, уменьшим еще a настолько, чтобы было

$aM \leqslant q < 1$, где q — фиксированная положительная правильная дробь, например $q = 3/4$. Тогда $D_{n+1} \leqslant qD_n \leqslant q^n D_0$.

Рассмотрим теперь ряд

$$\varphi_0(x) + \varphi_1(x) + \varphi_2(x) + \ldots + \varphi_n(x) + \ldots,$$

n-я частичная сумма которого есть $\sum\limits_{k=0}^{n-1} \varphi_k(x) = y_n(x)$. Абсолютная величина общего члена этого ряда, $|\varphi_n(x)| \leqslant D_0\, q^{n-1}$. Стало быть, абсолютные величины членов нашего ряда не превышают при $|x| \leqslant a$ соответствующих членов сходящегося геометрического ряда с положительными постоянными членами. На основании критерия, данного в т. I, гл. VIII, § 4, n° 2, стр. 453, частичные суммы нашего ряда, т. е. функции $y_n(x)$, сходятся в интервале $|x| \leqslant a$ равномерно к предельной функции $y(x)$. Отсюда видно, что существует такой интервал $|x| \leqslant a$, в котором дифференциальное уравнение $y' = f(x, y)$ имеет одно и только одно решение, удовлетворяющее заданному выше начальному условию.

Остается показать, что это решение можно шаг за шагом продолжить до границы (замкнутой ограниченной) области G, в которой определена функция $f(x, y)$. Проведенное выше доказательство показывает, что если решение $y(x)$ уже продолжено до некоторой точки $x = \xi$, то его можно продолжить вдоль оси x за эту точку еще на некоторый отрезок длины a, но a зависит от координат ξ, $y(\xi)$ той точки интегральной кривой, до которой было доведено ее построение. Поэтому можно опасаться, как бы эта величина a не убывала с каждым шагом столь быстро, что продолжение решения окажется возможным лишь в очень небольшой области, сколько бы шагов ни предпринимать. Покажем, что такая ситуация невозможна.

Допустим, что G' есть замкнутая ограниченная область, лежащая целиком внутри G. Тогда можно найти столь малое $b > 0$, что для всякой точки (x_0, y_0) в G' весь квадрат $x_0 - b \leqslant x \leqslant x_0 + b$, $y_0 - b \leqslant y \leqslant y_0 + b$ лежит в G. Пусть в области G выполняются неравенства $|f_y(x, y)| \leqslant M$ и $|f(x, y)| \leqslant M_1$; тогда нетрудно убедиться, что все условия, наложенные на a в данном выше доказательстве, непременно будут выполнены, если принять за a наименьшее из чисел b, $1/(2M)$ и b/M_1. Это значение a уже не зависит от точки (x_0, y_0); стало быть, мы можем с каждым шагом продвигаться на постоянный отрезок a, и поэтому последовательное продолжение решения может быть доведено до границы области G'. Но ведь G' есть *любая* замкнутая область, содержащаяся в G; следовательно, решение может быть продолжено до границы области G.

5. Системы дифференциальных уравнений первого порядка и дифференциальные уравнения высшего порядка. Значительная часть проведенных выше рассуждений распространяется на систему дифференциальных уравнений первого порядка, в которой число уравнений равно числу неизвестных функций одного аргумента. Мы рассмотрим здесь достаточно общий пример системы двух дифференциальных уравнений с двумя искомыми функциями $y(x)$ и $z(x)$:

$$y' = f(x, y, z), \qquad z' = g(x, y, z). \tag{1}$$

И здесь мы предполагаем, что функции f и g непрерывно дифференцируемы. Такую систему дифференциальных уравнений тоже можно изобразить наглядно в виде *поля направлений*, но в пространстве xyz: каждой точке (x, y, z) той области пространства, в которой заданы функции f и g, наша система предписывает определенное направление, направляющие косинусы которого относятся друг к другу как $dx : dy : dz = 1 : f : g$. С геометрической точки зрения задача интегрирования нашей системы состоит в том, чтобы найти такие пространственные кривые, которые, так сказать, «вписываются» в это поле направлений. Аналогично теории одного дифференциального уравнения, и здесь имеет место следующая *теорема существования и единственности: через каждую точку области G, в которой функции f и g непрерывно дифференцируемы, проходит одна и только одна интегральная кривая системы* (1). Таким образом, область G покрывается двупараметрическим семейством пространственных кривых; это семейство кривых дает общее решение системы (1) в виде двух функций $y = y(x;\ c_1,\ c_2)$ и $z = z(x;\ c_1,\ c_2)$, содержащих, помимо независимой переменной x, еще два произвольных параметра c_1 и c_2, которые появляются в качестве постоянных интегрирования. Доказательство теоремы существования можно провести тем же методом итерации (последовательного приближения), что и в n° 4.

Особая важность систем дифференциальных уравнений первого порядка проистекает от того, что всякое дифференциальное уравнение высшего порядка, скажем, порядка n, можно заменить системой n дифференциальных уравнений первого порядка с n неизвестными функциями.

Например, дифференциальное уравнение второго порядка

$$y'' = \varphi(x, y, y') \tag{2}$$

можно заменить системой двух дифференциальных уравнений первого порядка. Для этого надо производную y' принять за новую искомую функцию z, и тогда наше дифференциальное уравнение запишется в виде системы двух уравнений

$$\left.\begin{aligned} y' &= z, \\ z' &= \varphi(x, y, z). \end{aligned}\right\} \tag{3}$$

Эта система равносильна данному дифференциальному уравнению второго порядка в том смысле, что всякое решение уравнения (2) является также решением системы (3), и обратно.

6. Интегрирование с помощью степенного ряда (метод неопределенных коэффициентов). В заключение этого параграфа сообщим один общий метод, который часто применяется с успехом для решения дифференциальных уравнений. Это метод интегрирования с помощью степенного ряда.

Предположим, что в дифференциальном уравнении

$$y' = f(x, y)$$

функция $f(x, y)$ разлагается в степенной ряд по обоим аргументам x и y и, стало быть, имеет производные любого порядка по x и по y. Тогда естественно искать решения дифференциального уравнения тоже в виде степенного ряда

$$y = c_0 + c_1 x + c_2 x^2 + \dots,$$

определяя коэффициенты этого ряда с помощью заданного дифференциального уравнения.

Первые n членов этого ряда представляют собой аппроксимирующий многочлен степени n искомого решения. Таким образом, этот метод является аналитическим аналогом приближенного графического метода интегрирования, намеченного в n° 1.

Эту программу можно выполнить двумя путями.

а) Дифференцируем наше гипотетическое решение по x:

$$y' = c_1 + 2c_2 x + 3c_3 x^2 + \dots,$$

и этот производный ряд пишем вместо левой части данного дифференциального уравнения. В правой части пишем разложение функции $f(x, y)$ в степенной ряд, причем вместо y подставляем ряд $c_0 + c_1 x + c_2 x^2 + \dots$ Затем приравниваем друг другу коэффициенты при одинаковых степенях переменной x слева и справа. Этот путь нахождения решения называется *методом неопределенных коэффициентов*; мы им уже не раз пользовались для решения дифференциальных уравнений в первом томе, гл. VIII, § 6 и Доп. к гл. VIII, § 5. Член c_0 останется произвольным в качестве постоянной интегрирования, а коэффициенты $c_1, c_2, c_3, \dots$ надо определять последовательно из полученных уравнений.

б) Однако, часто оказывается более простым и наглядным следующий путь. Допустим, например, что требуется найти решение дифференциального уравнения, удовлетворяющее начальному условию $y(0) = 0$ или, на геометрическом языке, найти интегральную кривую, проходящую через начало координат. Очевидно, $c_0 = 0$. Остальные коэффициенты легко найти по известным формулам для коэффициентов степенного ряда (ряда Тэйлора):

$$c_k = \frac{1}{k!} y^{(k)}(0).$$

Сначала находим $c_1 = y'(0) = f(0, 0)$. Для определения c_2 нужно иметь выражение для второй производной $y''(x)$. Это выражение мы найдем, дифференцируя наше дифференциальное уравнение по x:

$$y''(x)=f_x+f_y y'.$$

Подставив сюда $x=0$ и уже известные значения $y(0)=0$ и $y'(0)=$ $=f(0, 0)$, получим $y''(0)$ и $c_2=\frac{1}{2}y''(0)$. Этот процесс можно продолжать неограниченно и определить один за другим дальнейшие коэффициенты $c_3, c_4, \ldots$

Можно показать, что этот процесс всегда дает решение, т. е. *сходящийся* степенной ряд для $y(x)$, если степенной ряд для $f(x, y)$ сходится абсолютно внутри некоторого круга с центром в начале координат. Доказательство мы здесь приводить не будем.

Упражнения

1. У нижеследующих дифференциальных уравнений проверить, что левая часть есть полный дифференциал, и проинтегрировать эти уравнения:

а) $(3x^2+6xy^2)\,dx+(6x^2y+4y^3)\,dy=0$;

б) $\dfrac{x\,dx+y\,dy}{\sqrt{1+x^2+y^2}}+\dfrac{y\,dx-x\,dy}{x^2+y^2}=0.$

2. Решить дифференциальное уравнение

$$(xy^2-y^3)\,dx+(1-xy^2)\,dy=0,$$

имеющее интегрирующий множитель, не зависящий от x.

[3. Решить дифференциальное уравнение

$$(x^2y+y+1)\,dx+(x+x^3)\,dy=0,$$

имеющее интегрирующий множитель, не зависящий от y.]

4. Найти общее решение дифференциального уравнения

$$2y^3\,dx+(3xy^2-1)\,dy=0$$

и с его помощью определить интегрирующий множитель этого уравнения.

5. Дано семейство плоских кривых

$$f(x, y, c)=0.$$

Исключив параметр c из этого уравнения и из уравнения

$$f_x+f_y y'=0,$$

получим дифференциальное уравнение этого семейства

$$F(x, y, y')=0$$

(ср. n° 2). Пусть $\varphi(p)$ есть заданная функция переменной p; всякая кривая K, удовлетворяющая дифференциальному уравнению

$$F(x, y, \varphi(y'))=0,$$

называется *траекторией* семейства кривых $f(x, y, c)=0$. Сопоставление этого дифференциального уравнения с дифференциальным уравнением нашего семейства кривых показывает, что

$$y'=\varphi(Y')$$

дает соотношение между наклоном Y' кривой K в любой ее точке и наклоном y' кривой семейства $f(x, y, c)=0$, проходящей через эту точку. Самым важным является тот случай, когда $\varphi(p) = -\frac{1}{p}$, приводящий к дифференциальному уравнению

$$F\left(x, y, -\frac{1}{y'}\right) = 0$$

ортогональных траекторий нашего семейства кривых (ср. конец n° 2).

С помощью этого метода найти ортогональные траектории следующих семейств кривых:

$$\text{а) } x^2 + y^2 + cy - 1 = 0; \qquad \text{б) } y = cx^2;$$

$$\text{в) } \frac{x^2}{a^2 + c} + \frac{y^2}{b^2 + c} = 1 \qquad (a > b > 0, \quad -b^2 < c < \infty);$$

$$\text{г) } y = \cos x + c; \qquad \text{д) } (x - c)^2 + y^2 = a^2.$$

Для каждого случая построить графики обоих ортогональных между собой семейств кривых.

6. Для пучка прямых $y = cx$ найти два семейства траекторий, у которых: а) угловой коэффициент касательной к траектории в два раза больше углового коэффициента прямой; б) угловой коэффициент касательной к траектории равен по абсолютной величине и противоположен по знаку угловому коэффициенту прямой.

7. *Дифференциальное уравнение* вида

$$y = xp + \psi(p), \qquad p = y'$$

было впервые исследовано *Клеро*. Дифференцируя по x, имеем

$$[x + \psi'(p)] \cdot \frac{dp}{dx} = 0.$$

Отсюда: либо

а) $p = c = \text{const}$, так что

$$y = cx + \psi(c)$$

есть общее решение дифференциального уравнения Клеро, представляющее семейство прямых;

либо б)

$$x = -\psi'(p),$$

что вместе с уравнением

$$y = -p\psi'(p) + \psi(p)$$

дает параметрическое представление так называемого *особого решения*. Заметим, что кривая, определяемая особым решением, является огибающей семейства прямых, дающего общее решение.

Найти этим методом особые решения следующих уравнений Клеро:

$$\text{а) } y = xy' - \frac{(y')^2}{4}; \qquad \text{б) } y = xy' + e^{y'}.$$

8. Найти дифференциальное уравнение семейства касательных к цепной линии $y = a \operatorname{ch} \frac{x}{a}$.

9. Лагранж исследовал более общее дифференциальное уравнение, линейное относительно пары переменных x и y, а именно

$$y = x\varphi(p) + \psi(p), \qquad p = y'. \tag{1}$$

Если $\varphi(p) \equiv p$, то получается как частный случай уравнение Клеро, которое уже было рассмотрено в упр. 7. Поэтому мы будем предполагать, что $\varphi(p) - p$ не равно тождественно нулю. Дифференцируя уравнение (1) по x, получим

$$p = \varphi(p) + [x\varphi'(p) + \psi'(p)]\frac{dp}{dx}. \tag{2}$$

Это дифференциальное уравнение, связывающее переменные x и p, можно привести к виду

$$\frac{dx}{dp} + \frac{\varphi'(p)}{\varphi(p) - p}x = -\frac{\psi'(p)}{\varphi(p) - p}, \tag{3}$$

в котором p принято за независимую переменную, а x за искомую функцию. Это линейное дифференциальное уравнение первого порядка; его общее решение (см. т. I, гл. XI, § 1, n° 3) находится с помощью квадратур и имеет следующий вид:

$$x = cv(p) + w(p).$$

Это уравнение дает совместно с исходным уравнением

$$y = x\varphi(p) + \psi(p)$$

параметрическое представление общего решения уравнения Лагранжа (1).

Мы предположили, что $\varphi(p) - p$ не равно тождественно нулю; но возможен случай, что уравнение $\varphi(p) - p = 0$ имеет некоторое число корней. Если $p = \alpha$ есть такой корень, то $\varphi(\alpha) - \alpha = 0$; сразу видно, что в таком случае $p = \alpha$ удовлетворяет уравнению (2). Подставив $p = \alpha$ в уравнение (1), получим

$$y = x\varphi(\alpha) + \psi(\alpha),$$

и нетрудно убедиться, что эта функция является решением уравнения (1) и что она не входит в состав общего решения. Итак, каждому корню уравнения $\varphi(p) - p = 0$ соответствует *особое решение* уравнения Лагранжа, изображаемое прямой линией.

Эти решения можно истолковать геометрически следующим образом. Рассмотрим в качестве вспомогательного уравнение Клеро

$$y = xp + \psi[\varphi^{-1}(p)],$$

где $\varphi^{-1}(p)$ есть обратная функция для $\varphi(p)$, т. е. $\varphi^{-1}[\varphi(p)] \equiv p$. Отсюда видно, что общее решение дифференциального уравнения Лагранжа представляет собой семейство траекторий семейства прямых

$$y = xc + \psi[\varphi^{-1}(c)]$$

или

$$y = x\varphi(C) + \psi(C).$$

Так, например,

$$y = \frac{x}{p} + \psi(p), \qquad p = y'$$

есть дифференциальное уравнение эвольвент, т. е. ортогональных траекторий семейства касательных кривой, представляющей особое решение следующего уравнения Клеро:

$$y = xp + \psi\left(-\frac{1}{p}\right), \qquad p = y'.$$

Пользуясь этим методом, проинтегрировать уравнение

$$y = x(p+a) - \frac{1}{4}(p+a)^2, \qquad p = y'.$$

10. Выразить, где это возможно, решения следующих дифференциальных уравнений с помощью элементарных функций:

$$\text{а)}\ \left(\frac{dy}{dx}\right)^2 = 1 - y^2; \qquad \text{б)}\ \left(\frac{dy}{dx}\right)^2 = \frac{1}{1-y^2};$$

$$\text{в)}\ \left(\frac{dy}{dx}\right)^2 = \frac{2a-y}{y}; \qquad \text{г)}\ \left(\frac{dy}{dx}\right)^2 = \frac{1-y^2}{1+y^2}.$$

В каждом из этих примеров построить график семейства интегральных кривых и из полученного чертежа обнаружить особые решения, если они существуют.

11. В переменной точке M плоской кривой Γ построить нормаль к этой кривой; отметить на нормали точку N ее пересечения с осью абсцисс и центр кривизны C, соответствующий точке M. Найти все кривые, обладающие следующим свойством:

$$MN \cdot MC = k = \text{const}.$$

Рассмотреть различные возможные случаи при $k > 0$ и $k < 0$ и построить графики.

12*. Найти дифференциальное уравнение трехпараметрического семейства всех окружностей $x^2 + y^2 + 2ax + 2by + c = 0$ на плоскости.

13. Решить однородное дифференциальное уравнение

$$(xy' - y)^2 = (x^2 - y^2)\left(\arcsin\frac{y}{x}\right)^2$$

и найти также особые решения. (См. т. I, гл. XI, § 1, n° 2.)

14*. Найти решение дифференциального уравнения

$$y'' + \frac{1}{x}y' + y = 0,$$

удовлетворяющее начальным условиям $y(0) = 1$, $y'(0) = 0$ в виде степенного ряда. Доказать, что эта функция тождественна функции Бесселя $J_0(x)$, определение которой дано на стр. 245, упр. 4.

15. Касательная в точке P плоской кривой пересекает отрицательную полуось y в точке T; эта кривая обладает тем свойством, что $|OP| = n|OT|$, где O — начало координат. Доказать, что уравнение кривой в полярных координатах имеет следующий вид:

$$r = a\frac{(1+\sin\theta)^n}{\cos^{n+1}\theta}.$$

§ 4. Линейные дифференциальные уравнения любого порядка

1. Определение. Теорема существования и единственности решения. Принцип суперпозиции. Дифференциальное уравнение колебания, изученное в т. I, гл. XI, §§ 3—5, и некоторые другие рассмотренные в предыдущем примеры принадлежат к общему типу линейных дифференциальных уравнений. *Линейным дифференциаль-*

ным уравнением (сокращенно: л. д. у.) для неизвестной функции $u(x)$ называется уравнение вида

$$a_0(x)u^{(n)}(x)+a_1(x)u^{(n-1)}(x)+\ldots+a_{n-1}(x)u'(x)+ \\ +a_n(x)u(x)=\varphi(x), \qquad (A)$$

где коэффициенты $a_0, a_1, a_2, \ldots, a_n$ и свободный член (правая часть) $\varphi(x)$ — заданные функции одной лишь независимой переменной x. Выражение, стоящее в левой части, называется линейным (однородным) дифференциальным выражением и обозначается символом $L[u]$.

[В уравнении (A) предполагается, что $a_0(x)$ не обращается тождественно в нуль. Тогда (A) будет л. д. у. порядка n. Если ограничить независимое переменное x интервалом $\alpha<x<\beta$, в котором коэффициент $a_0(x)$ нигде не обращается в нуль и в котором все функции $a_k(x)$ ($k=0, 1, 2, \ldots, n$) и правая часть $\varphi(x)$ непрерывны, то уравнение (A) можно привести к виду

$$u^{(n)}+p_1(x)u^{(n-1)}+\ldots+p_n(x)u=f(x), \qquad (B)$$

где $p_k(x)=\frac{a_k(x)}{a_0(x)}$ и $f(x)=\frac{\varphi(x)}{a_0(x)}$ являются в (α, β) известными непрерывными функциями от x.

Введем в дополнение к $u(x)$ еще $n-1$ неизвестных функций

$$u'=u_1(x), \quad u''=u_1'=u_2(x), \quad \ldots, \quad u^{(n-1)}=u_{n-2}'=u_{n-1}(x).$$

Тогда л. д. у. (B) будет равносильно системе n дифференциальных уравнений

$$\left.\begin{aligned} &\frac{du}{dx}=u_1, \quad \frac{du_1}{dx}=u_2, \quad \ldots, \quad \frac{du_{n-2}}{dx}=u_{n-1}, \\ &\frac{du_{n-1}}{dx}=-p_1(x)u_{n-1}-p_2(x)u_{n-2}-\ldots-p_{n-1}(x)u_1- \\ &\qquad\qquad -p_n(x)u+f(x) \end{aligned}\right\} \qquad (C)$$

с n искомыми функциями $u, u_1, \ldots, u_{n-1}$. (Ср. § 3, n° 5, стр. 463.)

В курсах дифференциальных уравнений для этой системы доказывается следующая *теорема существования и единственности:*

При сделанных выше предположениях относительно функций $p_k(x)$ и $f(x)$ существует одна и только одна система непрерывных в (α, β) функций $u(x), u_1(x), \ldots, u_{n-1}(x)$, удовлетворяющих системе л. д. у. (C) и принимающих в произвольно заданной точке $x=x_0$ промежутка (α, β) любые наперед заданные значения $u=b$, $u_1=b_1, u_2=b_2, \ldots, u_{n-1}=b_{n-1}$.

Отсюда вытекает, что *л. д. у. n-го порядка (B), а следовательно, и исходное уравнение (A) имеет в (α, β) одно и только одно решение $u(x)$, обладающее в этом интервале непрерывными производными до n-го порядка включительно и удовлетворяющее начальным условиям*

$$u(x_0)=b, \quad u'(x_0)=b_1, \quad u''(x_0)=b_2, \quad \ldots, \quad u^{(n-1)}(x_0)=b_{n-1},$$

В дальнейшем изложении при рассмотрении системы функций $\varphi_i(x)$ всегда предполагается, что среди этих функций нет тождественного нуля.

2) Если $n-1$ функций $\varphi_1(x), \varphi_2(x), \ldots, \varphi_{n-1}(x)$ линейно зависимы в заданном интервале, то присоединив еще одну функцию $\varphi_n(x)$, получим вновь линейно зависимую систему функций. Действительно, из существования тождества (относительно x)

$$c_1\varphi_1(x)+c_2\varphi_2(x)+\ldots+c_{n-1}\varphi_{n-1}(x)\equiv 0,$$

где не все c_i равны нулю, вытекает тождество

$$c_1\varphi_1(x)+c_2\varphi_2(x)+\ldots+c_{n-1}\varphi_{n-1}(x)+c_n\varphi_n(x)\equiv 0$$

с коэффициентом $c_n=0$, причем не все c_i равны нулю.]

Пример 1. Функции $1, x, x^2, \ldots, x^{n-1}$ линейно независимы в любом интервале. Действительно, равенство

$$c_0+c_1x+\ldots+c_{n-1}x^{n-1}=0$$

может выполняться при всех значениях x из рассматриваемого интервала лишь в том случае, если все коэффициенты c_i многочлена равны нулю.

Пример 2. Функции $e^{\alpha_1 x}, e^{\alpha_2 x}, \ldots, e^{\alpha_n x}$ линейно независимы, если все числа α_i различны. Доказательство мы проведем методом полной индукции. Сначала докажем, что две функции $e^{\alpha_1 x}$ и $e^{\alpha_2 x}$ при $\alpha_1\neq\alpha_2$ линейно независимы. В самом деле, из предположения, что

$$c_1e^{\alpha_1 x}+c_2e^{\alpha_2 x}=0$$

и, скажем, $c_2\neq 0$, вытекало бы, что $\dfrac{e^{\alpha_2 x}}{e^{\alpha_1 x}}=e^{(\alpha_2-\alpha_1)x}=-\dfrac{c_1}{c_2}$, т. е. функция, очевидным образом зависящая от x, равна постоянной, что невозможно.

Докажем теперь, что если утверждение, подлежащее доказательству, справедливо для $n-1$ функций вида $e^{\alpha_i x}$, то оно справедливо и для n таких функций. Предположим противное, что

$$c_1e^{\alpha_1 x}+c_2e^{\alpha_2 x}+\ldots+c_ne^{\alpha_n x}=0$$

тождественно относительно x. Не нарушая общности, можно считать, что неравный нулю коэффициент c_k содержится среди коэффициентов $c_1, c_2, \ldots, c_{n-1}$. Разделив это тождество на $e^{\alpha_n x}$, получим

$$c_1e^{(\alpha_1-\alpha_n)x}+c_2e^{(\alpha_2-\alpha_n)x}+\ldots+c_{n-1}e^{(\alpha_{n-1}-\alpha_n)x}+c_n=0.$$

Продифференцируем это равенство по x и тем самым исключим c_n, а затем помножим полученное равенство на $e^{\alpha_n x}$. В результате получится новое тождество:

$$(\alpha_1-\alpha_n)c_1e^{\alpha_1 x}+(\alpha_2-\alpha_n)c_2e^{\alpha_2 x}+\ldots+(\alpha_{n-1}-\alpha_n)c_{n-1}e^{\alpha_{n-1}x}=0$$

из которого вытекает, что $n-1$ функций $e^{\alpha_1 x}, e^{\alpha_2 x}, \ldots, e^{\alpha_{n-1}x}$ тоже линейно зависимы, что противоречит предположению. Так как линейная независимость *двух* функций вида $e^{\alpha_i x}$ была ранее доказана, то отсюда вытекает линейная независимость любого числа n таких функций.

Пример 3. Функции $\sin x$, $\sin 2x$, ..., $\sin nx$ линейно независимы в интервале $0 \leqslant x \leqslant \pi$. Доказательство с помощью интеграла

$$\int_{-\pi}^{\pi} \sin mx \sin nx\, dx = \begin{cases} 0 \text{ при } m \neq n, \\ \pi \text{ при } m = n, \end{cases}$$

мы предоставляем читателю (ср. т. I, стр. 247).

[Замечание. Если система n функций линейно зависима в некотором интервале (a, b) или $[a, b]$, то она, очевидно, линейно зависима и в любом частичном интервале, содержащемся в исходном. Однако обратная теорема неверна: система функций, линейно зависимая, скажем на $[a, b]$, может оказаться линейно независимой в более широком промежутке, объемлющем $[a, b]$, как это показывают следующие примеры.

Пример А. Система функций $\varphi_1(x) = x^m$, $\varphi_2(x) = x^m \operatorname{sgn} x$, где m — любое целое число, не меньшее единицы, линейно зависима при $x \leqslant 0$, где $\varphi_1 + \varphi_2 = 0$, она также линейно зависима при $x \geqslant 0$, где $\varphi_1 - \varphi_2 = 0$; между тем, в любом интервале, содержащем внутри себя начало $x = 0$, эти функции линейно независимы. Действительно, из гипотетического тождества $\alpha\varphi_1(x) + \beta\varphi_2(x) \equiv 0$ вытекает $\alpha - \beta = 0$ (при $x < 0$) и $\alpha + \beta = 0$ (при $x > 0$), так что это тождество возможно только при $\alpha = \beta = 0$. (Определение символа $\operatorname{sgn} x$ см. т. I, стр. 69.)

Пример Б. Система функций

$$\psi_1(x) = \begin{cases} x^m & \text{при} \quad x \leqslant 0, \\ 0 & \text{при} \quad x \geqslant 0, \end{cases} \qquad \psi_2(x) = \begin{cases} 0 & \text{при} \quad x \leqslant 0, \\ x^m & \text{при} \quad x \geqslant 0, \end{cases}$$

где $m \geqslant 1$, целое, линейно зависима при $x \leqslant 0$, ибо там $0 \cdot \psi_1 + 1 \cdot \psi_2 = 0$; при $x \geqslant 0$ она тоже линейно зависима, так как там $1 \cdot \psi_1 + 0 \cdot \psi_2 = 0$. Однако в любом промежутке, окружающем начало $x = 0$, функции $\psi_1(x)$ и $\psi_2(x)$ линейно независимы.]

3. Необходимое условие линейной зависимости n функций. Если функции $\varphi_i(x)$ имеют непрерывные производные до $(n-1)$-го порядка включительно, то справедлива следующая

Теорема 1. *Необходимым условием линейной зависимости n функций $\varphi_i(x)$ является выполнение тождества*

$$W(x) = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \dots & \varphi_n'(x) \\ \dots & \dots & \dots & \dots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \dots & \varphi_n^{(n-1)}(x) \end{vmatrix} \equiv 0.$$

Функция $W(x)$ называется *определителем Вронского* системы n функций $\varphi_i(x)$. (Мы предполагаем, что читатель знаком с элементами теории определителей любого порядка.)

Эта теорема легко доказывается следующим образом. Если наши функции линейно зависимы, то существует тождество (относительно x)

$$\sum c_i \varphi_i(x) = 0,$$

в котором не все коэффициенты c_i равны нулю. Дифференцируя это равенство последовательно $n-1$ раз, получим еще $n-1$ равенств:

$$\sum c_i\varphi_i'(x)=0,$$
$$\cdots\cdots\cdots$$
$$\sum c_i\varphi_i^{(n-1)}(x)=0.$$

Все эти n равенств можно рассматривать как линейную однородную систему n уравнений, которой заведомо удовлетворяют значения неизвестных $c_1, c_2, \ldots, c_n$, не все равные нулю. Следовательно, определитель системы $W(x)$ должен обратиться в нуль.

Итак, *если n функций линейно зависимы в некотором интервале, то их определитель Вронского тождественно равен нулю в этом интервале.*

[Ясно, что невыполнение необходимого условия линейной зависимости n функций является достаточным условием их линейной независимости, т. е. справедлива следующая

Теорема 2. *Если определитель Вронского $W(x)$ системы n функций не равен тождественно нулю в некотором интервале, то эти функции линейно независимы в этом интервале.*

Каждая из теорем 1 и 2 выводится из другой рассуждением от противного.

Теоремы, обратные теоремам 1 и 2, не верны: линейная независимость системы функций в некотором интервале может сопровождаться тождественным обращением в нуль ее определителя Вронского в этом же интервале. Для того чтобы убедиться в этом, достаточно привести один противоречащий пример, а мы уже фактически встретились с двумя такими примерами (см. замечание в конце n° 2):

1) В n° 2, пример А. Функции $\varphi_1(x)=x^m$ и $\varphi_2(x)=x^m\operatorname{sgn}x$ при целом $m\geqslant 1$, как мы видели, линейно независимы в любом промежутке, содержащем внутри себя начало $x=0$, а между тем их определитель Вронского тождественно равен нулю.

2) Там же в примере Б мы видели, что функции

$$\psi_1(x)=\begin{cases}x^m & \text{при } x\leqslant 0,\\ 0 & \text{при } x\geqslant 0,\end{cases}\qquad \psi_2(x)=\begin{cases}0 & \text{при } x\leqslant 0,\\ x^m & \text{при } x\geqslant 0\end{cases}$$

(целое $m\geqslant 1$) линейно независимы в любом промежутке, содержащем начало $x=0$ внутри себя, а вместе с тем легко проверить, что их определитель Вронского тождественно равен нулю.

Можно также построить опровергающий пример системы любого числа n функций. Положим

$$f_1(x)=1,\quad f_2(x)=x,\ \ldots,\ f_{n-2}(x)=x^{n-3},\quad f_{n-1}(x)=x^m,\quad f_n(x)=x^m\operatorname{sgn}x$$

с целым показателем $m>n-3$. Хотя эта система линейно зависима при $x\leqslant 0$, где

$$0\cdot f_1+0\cdot f_2+\ldots+0\cdot f_{n-2}+1\cdot f_{n-1}+1\cdot f_n\equiv 0,$$

и линейно зависимы при $x\geqslant 0$, где

$$0\cdot f_1+0\cdot f_2+\ldots+0\cdot f_{n-2}+1\cdot f_{n-1}-1\cdot f_n\equiv 0,$$

но в любом промежутке, окружающем точку $x = 0$, она линейно независима. Однако определитель Вронского этой системы функций тождественно равен нулю, так как элементы последних двух его столбцов всегда пропорциональны.

Заметим, что у функций $\varphi_2(x)$, $\psi_2(x)$ и $f_n(x)$ все производные до $(m-1)$-го порядка включительно всюду непрерывны и лишь m-е производные имеют при $x=0$ конечный разрыв. Выбирая показатель m достаточно большим, мы во всех трех контрпримерах будем иметь системы функций, обладающих непрерывными производными до любого желательного порядка. Впрочем, возможно построить и такой контрпример, в котором *n линейно независимых функций* имеют в данном интервале производные *любого* порядка и определитель Вронского, тождественно равный нулю.]

[4. Необходимое и достаточное условие линейной независимости n решений л. д. у. n-го порядка без правой части. Итак, мы не имеем решающего критерия линейной зависимости (или, напротив, независимости) n функций. Однако если n функций $u_1(x)$, $u_2(x), \ldots, u_n(x)$ являются решениями л. д. у. n-го порядка $L[u]=0$, то такое необходимое и достаточное условие можно установить с помощью следующей теоремы.

Теорема. *Если решения $u_1(x)$, $u_2(x), \ldots, u_n(x)$ л. д. у. n-го порядка*

$$L[u] \equiv a_0(x)u^{(n)} + a_1(x)u^{(n-1)} + \ldots + a_n(x)u = 0 \tag{1}$$

линейно независимы в интервале (α, β), *то их определитель Вронского не обращается в нуль ни в одной точке этого интервала.*

Напомним, что (α, β) есть интервал, в котором все $a_k(x)$ непрерывны ($k=0, 1, 2, \ldots, n$) и $a_0(x)$ нигде не обращается в нуль.

Доказательство. На основании принципа суперпозиции (n° 1) функция

$$u(x) = c_1u_1(x) + c_2u_2(x) + \ldots + c_nu_n(x) \tag{2}$$

является решением нашего л. д. у. (1) при любых значениях постоянных $c_1, c_2, \ldots, c_n$. Возьмем теперь любое значение ξ из интервала (α, β) и найдем частное решение $\bar{u}(x)$, определяемое начальными условиями

$$u(\xi)=0, \quad u'(\xi)=0, \quad u''(\xi)=0, \ldots, u^{(n-1)}(\xi)=0. \tag{3}$$

Согласно теореме существования (n° 1), л. д. у. (1) имеет одно и только одно решение, удовлетворяющее этим условиям, причем ясно, что это — тривиальное решение $u(x)\equiv 0$. Из линейной комбинации (2) оно получается при значениях $c_1=c_2=\ldots=c_n=0$ и только при этих значениях, в силу линейной независимости функций $u_1(x), u_2(x), \ldots$ $\ldots, u_n(x)$.

С другой стороны, попытаемся выделить это же решение из формулы (2) шаблонным приемом нахождения неизвестных коэффициентов $c_1, c_2, \ldots, c_n$ по начальным условиям (3). Для этого продифференцируем (2) по x последовательно $n-1$ раз, а затем подставим во все полученные вместе с (2) n равенств значение $x=\xi$. Тогда начальные данные (3) дадут для искомых n коэффициентов c_i следующую систему n линейных однородных уравнений:

$$\left.\begin{aligned} u(\xi) &= c_1u_1(\xi) + c_2u_2(\xi) + \ldots + c_nu_n(\xi) = 0,\\ u'(\xi) &= c_1u_1'(\xi) + c_2u_2'(\xi) + \ldots + c_nu_n'(\xi) = 0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ u^{(n-1)}(\xi) &= c_1u_1^{(n-1)}(\xi) + c_2u_2^{(n-1)}(\xi) + \ldots + c_nu_n^{(n-1)}(\xi) = 0. \end{aligned}\right\} \quad (4)$$

Определитель этой системы уравнений есть $W(\xi)$, т. е. значение определителя Вронского $W(x)$ системы функций $u_1, u_2, \ldots, u_n$ в точке ξ. Нам уже известно, что система уравнений (4) должна иметь единственное решение $c_1=c_2=\ldots=c_n=0$. Следовательно, $W(\xi)\neq 0$. Так как ξ есть произвольная точка интервала (α, β), то определитель Вронского $W(x)\neq 0$ нигде в этом интервале, и теорема доказана.

Из этой теоремы вытекает, что для системы *любых* n частных решений л. д. у. (1) существует только две взаимно исключающие друг друга возможности: либо ее определитель Вронского $W(x)$ тождественно равен нулю в (α, β), и тогда наши n решений линейно зависимы, либо $W(x)$ не обращается в нуль ни в одной точке этого интервала, и тогда наши n решений линейно независимы.

Другими словами, если n функций $u_1(x), u_2(x), \ldots, u_n(x)$ удовлетворяют какому-либо л. д. у. n-го порядка без правой части, то

1) *тождественное обращение в нуль определителя Вронского этих n функций является не только необходимым, но и достаточным условием их линейной зависимости и*

2) *отличие определителя Вронского от нуля хотя бы в одной точке интервала (α, β) является не только достаточным, но и необходимым условием линейной независимости системы функций $u_1(x), u_2(x), \ldots, u_n(x)$.*

Напомним, что, согласно доказанной выше теореме, в этом случае определитель Вронского не обращается в нуль ни в одной точке интервала (α, β).]

[**5. Фундаментальные системы решений л. д. у. без правой части. Структура его общего решения.** Всякая система n линейно независимых частных решений л. д. у. n-го порядка $L[u]=0$ называется *фундаментальной системой решений.* Докажем, что любое л. д. у. n-го порядка без правой части — (1) n° 4 — имеет фундаментальные системы решений. Для этого выберем какое-либо число ξ из интервала $\alpha<x<\beta$, в котором все $a_k(x)$ непрерывны ($k=0, 1, 2, \ldots, n$) и $a_0(x)$ ни разу не обращается в нуль, и возьмем систему решений $u_1(x), u_2(x), \ldots, u_n(x)$, определяемых начальными данными

$$u_k(\xi)=b_k^{(0)},\quad u_k'(\xi)=b_k^{(1)},\quad u_k''(\xi)=b_k^{(2)},\ \ldots,\ u_k^{(n-1)}(\xi)=b_k^{(n-1)}$$
$$(k=1, 2, \ldots, n),$$

где $b_k^{(0)}, b_k^{(1)}, \ldots, b_k^{(n-1)}$ — произвольные числа, подчиненные единственному условию, чтобы определитель

$$D = \begin{vmatrix} b_1^{(0)} & b_1^{(1)} & \ldots & b_1^{(n-1)} \\ b_2^{(0)} & b_2^{(1)} & \ldots & b_2^{(n-1)} \\ \cdot & \cdot & \cdot & \cdot \\ b_n^{(0)} & b_n^{(1)} & \ldots & b_n^{(n-1)} \end{vmatrix}$$

матрицы этих чисел был отличен от нуля. Этот определитель D равен, очевидно, значению $W(\xi)$ определителя Вронского $W(x)$ построенной нами системы частных решений в точке $x=\xi$. Согласно теореме существования (n° 1), эта система решений существует и однозначно определена, а так как $W(\xi) \neq 0$, наша система решений $u_1(x), u_2(x), \ldots, u_n(x)$ линейно независима. Тем самым существование фундаментальных систем решений доказано.

Особенно просто строится фундаментальная система, определяемая следующими начальными условиями. В предписанной точке, скажем $x=\xi$, должно быть:

$$\begin{array}{lllll} u_1 = 1, & u_1' = 0, & u_1'' = 0, & \ldots, & u_1^{(n-1)} = 0, \\ u_2 = 0, & u_2' = 1, & u_2'' = 0, & \ldots, & u_2^{(n-1)} = 0, \\ u_3 = 0, & u_3' = 0, & u_3'' = 1, & \ldots, & u_3^{(n-1)} = 0, \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n = 0, & u_n' = 0, & u_n'' = 0, & \ldots, & u_n^{(n-1)} = 1. \end{array}$$

Такая фундаментальная система решений называется *нормальной*. Для нее, очевидно, $W(\xi) = 1$.

Знание какой-либо фундаментальной системы данного л. д. у. без правой части дает возможность построить множество всех его решений на основании следующей теоремы:

Всякое решение $u(x)$ л. д. у. n-го порядка $L[u]=0$ можно представить в виде линейной комбинации

$$u(x) = c_1u_1(x) + c_2u_2(x) + \ldots + c_nu_n(x) \qquad (*)$$

всех n частных решений какой-либо фундаментальной системы данного л. д. у., с постоянными коэффициентами $c_1, c_2, \ldots, c_n$. Выражение (*) принято называть общим решением уравнения $L[u]=0$.

Доказательство. Тривиальное решение $u(x) \equiv 0$, очевидно, содержится в формуле (*); содержится в ней и любое из решений $u_l(x)$ выбранной фундаментальной системы: при $c_l = 1$ и $c_k = 0$, когда $k \neq l$. Пусть теперь $u(x)$ есть любое нетривиальное решение, отличное от $u_1, u_2, \ldots, u_n$. Так как $n+1$ функций $u(x), u_1(x), u_2(x), \ldots, u_n(x)$

удовлетворяют данному л. д. у. $L[u]=0$, то имеем $n+1$ тождеств относительно x:

$$\begin{aligned}
&a_0u^{(n)}+a_1u^{(n-1)}+\ldots+a_nu\ =0,\\
&a_0u_1^{(n)}+a_1u_1^{(n-1)}+\ldots+a_nu_1=0,\\
&a_0u_2^{(n)}+a_1u_2^{(n-1)}+\ldots+a_nu_2=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_0u_n^{(n)}+a_1u_n^{(n-1)}+\ldots+a_nu_n=0,
\end{aligned}$$

которые можно рассматривать как систему $n+1$ линейных однородных алгебраических уравнений для $n+1$ неизвестных коэффициентов a_i, причем заведомо известно, что эта система уравнений имеет нетривиальное решение $a_0, a_1, \ldots, a_n$. Поэтому определитель системы, равный, очевидно, определителю Вронского $n+1$ функций $u(x)$, $u_1(x)$, $u_2(x), \ldots, u_n(x)$, должен быть тождественно равен нулю в интервале (α, β).

С другой стороны, дифференцируя по x наше л. д. у.

$$a_0u^{(n)}+a_1u^{(n-1)}+\ldots+a_{n-1}u'+a_nu=0,$$

найдем, что каждая из $n+1$ функций $u, u_1, u_2, \ldots, u_n$ является решением л. д. у. $n+1$-го порядка без правой части

$$a_0u^{(n+1)}+(a_0'+a_1)u^{(n)}+(a_1'+a_2)u^{(n-1)}+\ldots+(a_{n-1}'+a_n)u'+a_n'u=0,$$

имеющего тот же основной интервал (α, β), в котором все коэффициенты непрерывны, и тот же коэффициент при высшей производной $a_0(x)$, не обращающийся в нуль нигде в (α, β). Поэтому, на основании n° 4 (конец), тождественное обращение в нуль определителя Вронского системы $n+1$ функций $u(x), u_1(x), \ldots, u_n(x)$ является достаточным критерием их линейной зависимости и, следовательно, существует тождество

$$\lambda_0u(x)+\lambda_1u_1(x)+\lambda_2u_2(x)+\ldots+\lambda_nu_n(x)=0,$$

в котором не все коэффициенты $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_n$ равны нулю. Ясно, что $\lambda_0\neq 0$, ибо в противном случае, функции $u_1, u_2, \ldots, u_n$ оказались бы линейно зависимыми, вопреки предположению, что они составляют фундаментальную систему. Стало быть,

$$\begin{aligned}
u(x)=-\frac{\lambda_1}{\lambda_0}u_1(x)-\frac{\lambda_2}{\lambda_0}u_2(x)-\ldots-\frac{\lambda_n}{\lambda_0}u_n(x)&=\\
=c_1u_1(x)+c_2u_2(x)+\ldots+c_nu_n(x)&,
\end{aligned}$$

что и требовалось доказать.]

В заключение заметим, что две различные фундаментальные системы решений $u_1, u_2, \ldots, u_n$ и $U_1, U_2, \ldots, U_n$ выражаются одна через другую с помощью линейного преобразования

$$U_i=\sum_{k=1}^{n}\gamma_{ik}u_k$$

с постоянными коэффициентами γ_{ik} и с определителем преобразования, не равным нулю.

6. Частный случай л. д. у. второго порядка. Такое дифференциальное уравнение имеет следующий вид:

$$L[u] \equiv a(x)u'' + b(x)u' + c(x)u = 0. \qquad (A)$$

Мы рассмотрим его более подробно, так как оно имеет очень важные приложения. Пусть $u_1(x)$ и $u_2(x)$ — фундаментальная система решений этого уравнения. Ее определитель Вронского $W = u_1u_2' - u_2u_1'$, а его производная $W' = u_1u_2'' - u_2u_1''$. Так как

$$L[u_1] \equiv 0 \quad \text{и} \quad L[u_2] \equiv 0,$$

то

$$u_1L[u_2] - u_2L[u_1] = aW' + bW = 0.$$

Интегрируя, получим

$$\ln|W| = -\int \frac{b(x)}{a(x)}\,dx + \ln|C|,$$

откуда

$$W = Ce^{-\int \frac{b}{a}dx},$$

где C — постоянная интегрирования. Этой формулой часто пользуются в более полной теории дифференциальных уравнений второго порядка.

Достойно упоминания другое свойство уравнения (A): л. д. у. второго порядка без правой части всегда можно преобразовать в (нелинейное) дифференциальное уравнение первого порядка типа Риккати. *Уравнением Риккати* называется уравнение вида

$$v' + p(x)v^2 + q(x)v + r(x) = 0.$$

Наше уравнение (A) преобразуется в уравнение Риккати введением новой искомой функции z с помощью подстановки $u' = uz$, так что $u'' = u'z + uz' = uz^2 + uz'$. В результате уравнение (A) приводится к следующему виду:

$$az' + az^2 + bz + c = 0.$$

Третье свойство уравнения (A). Если известно *одно* частное решение $v(x)$ л. д. у. второго порядка без правой части, то задачу его интегрирования можно привести к решению л. д. у. первого порядка и, стало быть, выполнить в квадратурах. Действительно, полагая $u = zv$, где $z(x)$ есть новая искомая функция, получим новое уравнение

$$az''v + 2az'v' + bz'v + zL[v] = 0;$$

так как v является решением уравнения (A), то $L[v] \equiv 0$ и уравнение (A) приводится к виду

$$avz'' + (2av' + bv)z' = 0,$$

Это л. д. у. первого порядка относительно неизвестной функции $w = z' = z'(x)$. Вместе с тем это уравнение с отделяющимися переменными. Из него можно найти w простым интегрированием, а затем из $z' = w$ получим z вторичным интегрированием и, наконец, найдем общее решение $u = zv$; оно содержит полагающиеся две произвольные постоянные, появившиеся в итоге двух квадратур.

Пример. Л. д. у. второго порядка

$$y'' - \frac{2}{x}y' + \frac{2}{x^2}y = 0$$

равносильно уравнению Риккати

$$z' + z^2 - \frac{2}{x}z + \frac{2}{x^2} = 0,$$

где $z = \frac{y'}{y}$. Исходное уравнение имеет частное решение $y = x$ или, в прежнем обозначении, $v = x$. Следовательно, оно приводится преобразованием $y = zv = zx$ к виду

$$xz'' = 0.$$

Отсюда $z = Cx + C_1$ и общее решение первоначального уравнения будет $y = Cx^2 + C_1x$.

Подчеркнем, что если известно одно частное решение $v(x)$ линейного дифференциального уравнения n-го порядка $L[u] = 0$, то преобразованием $u = vz$ уравнение приведется к виду такого же л. д. у. $(n-1)$-го порядка для новой искомой функции z.

Упражнения

1. Доказать, что если числа $a_1, a_2, \ldots, a_k$ все различны, а $P_1(x)$, $P_2(x), \ldots, P_k(x)$ — произвольные многочлены, не равные тождественно нулю, то функции $\varphi_1(x) = P_1(x)e^{a_1x}$, $\varphi_2(x) = P_2(x)e^{a_2x}, \ldots, \varphi_k(x) = P_k(x)e^{a_kx}$ линейно независимы.

2. Показать, что уравнение Бернулли (т. I, гл. XI, § 1, n° 4)

$$y' + a(x)y = b(x)y^n \qquad (n \neq 1)$$

можно преобразовать в линейное дифференциальное уравнение для новой искомой функции $z = y^{1-n}$. Воспользоваться этим свойством для решения следующих уравнений:

а) $xy' + y = y^2 \ln x$; б) $xy^2(xy' + y) = a^2$;
в) $(1 - x^2)y' - xy = axy^2$.

3. Показать, что дифференциальное уравнение Риккати

$$y' + P(x)y^2 + Q(x)y + R(x) = 0$$

можно преобразовать в дифференциальное уравнение Бернулли, если известно его частное решение $y_1 = y_1(x)$, а следовательно (см. упр. 2), его тогда можно преобразовать и в линейное дифференциальное уравнение.

Пользуясь этим, найти общее решение уравнения

$$y' - x^2y^2 + x^4 - 1 = 0,$$

зная его частное решение $y_1 = x$.

4. Найти совпадающие решения двух дифференциальных уравнений

а) $y' = y^2 + 2x - x^4$ и б) $y' = -y^2 - y + 2x + x^2 + x^4$.

5*. Выразить общее решение дифференциального уравнения

$$y' = y^2 + 2x - x^4$$

с помощью определенных интегралов, используя частное решение, найденное в упр. 4. Построить на плоскости xy эскиз интегральных кривых этого дифференциального уравнения.

6*. Пусть y_1, y_2, y_3, y_4 — четыре частных решения уравнения Риккати (ср. упр. 3). Доказать, что выражение

$$\frac{y_1 - y_3}{y_1 - y_4} : \frac{y_2 - y_3}{y_2 - y_4}$$

имеет постоянное значение.

7. Показать, что если известны два решения, $y_1(x)$ и $y_2(x)$, уравнения Риккати, то его общее решение будет

$$y - y_1 = c\,(y - y_2)\,e^{\int (y_2 - y_1)\,p\,dx}$$

где c — произвольная постоянная.

Пользуясь этим, найти общее решение дифференциального уравнения

$$y' - y \operatorname{tg} x = y^2 \cos x - \frac{1}{\cos x},$$

зная, что оно имеет частные решения вида $a \cos^n x$.

8. Доказать, что уравнения

$$\text{а) } (1 - x)\,y'' + xy' - y = 0$$

и

$$\text{б) } 2x\,(2x - 1)\,y'' - (4x^2 + 1)\,y' + y\,(2x + 1) = 0$$

имеют совпадающее частное решение; найти его, а затем определить об. решение каждого из этих уравнений.

7. Л. д. у. n-го порядка без правой части с постоянными коэффициентами. Мы видели в т. I, гл. XI, § 3, что простейшие колебательные процессы, протекающие без участия внешней силы, описываются л. д. у. второго порядка вида $y'' + a_1 y' + a_2 y = 0$ с постоянными коэффициентами. Более сложные колебательные проблемы приводят к л. д. у. n-го порядка

$$L[y] \equiv y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + \ldots + a_{n-1} y' + a_n y = 0$$

с постоянными действительными коэффициентами. Это уравнение можно решить тем же методом, что и в частном случае $n = 2$ (т. I, гл. XI, § 4, n°2). Положим $y = e^{\lambda x}$. Подставив эту функцию и ее производные в дифференциальное уравнение, получим

$$L[e^{\lambda x}] \equiv e^{\lambda x} (\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \ldots + a_{n-1} \lambda + a_n) = 0.$$

Так как $e^{\lambda x}$ нигде не обращается в нуль, то число λ должно удовлетворять следующему алгебраическому уравнению степени n:

$$f(\lambda) \equiv \lambda^n + a_1 \lambda^{n-1} + \ldots + a_{n-1} \lambda + a_n = 0.$$

Оно называется *характеристическим уравнением.* Если λ есть корень этого уравнения, то $e^{\lambda x}$ удовлетворяет дифференциальному уравнению. Рассмотрим различные возможности.

а) *Пусть все корни характеристического уравнения* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *различны.* Тогда в нашем распоряжении n линейно независимых частных решений $e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_n x}$ (см. n° 2, пример 2), и *общее решение* будет линейной комбинацией этих частных решений

$$y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} + \ldots + c_n e^{\lambda_n x}$$

с произвольными постоянными коэффициентами $c_1, c_2, \ldots, c_n$ [Как поступить в случае наличия среди λ_i комплексных корней, см. т. I, стр. 615, пункт в).]

Константы c_i можно так подобрать, чтобы функция $y(x)$ удовлетворяла начальным условиям, скажем, при $x=0$:

$$y(0)=y_0, \quad y'(0)=y_0', \ldots, y^{(n-1)}(0)=y_0^{(n-1)},$$

где $y_0, y_0', y_0'', \ldots, y_0^{(n-1)}$ — произвольно заданные действительные числа. В самом деле, для нахождения коэффициентов c_i надо решить следующую систему n линейных алгебраических уравнений:

$$\begin{aligned} c_1 + c_2 + \ldots + c_n &= y_0, \\ \lambda_1 c_1 + \lambda_2 c_2 + \ldots + \lambda_n c_n &= y_0', \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1^{n-1} c_1 + \lambda_2^{n-1} c_2 + \ldots + \lambda_n^{n-1} c_n &= y_0^{(n-1)}. \end{aligned}$$

Так как определитель этой системы не равен нулю, то она имеет одно и только одно решение $c_1, c_2, \ldots, c_n$.

б) *Пусть среди корней* λ_i *характеристического уравнения имеются кратные.* Тогда число частных решений вида $e^{\lambda_i x}$ меньше, чем необходимое число n для получения фундаментальной системы решений. Оказывается, однако, что если характеристическое уравнение $f(\lambda)=0$ имеет корень λ кратности k, то л. д. у. $L[y]=0$ имеет k частных решений $e^{\lambda x}, xe^{\lambda x}, x^2 e^{\lambda x}, \ldots, x^{k-1} e^{\lambda x}$. Для проверки этого утверждения заметим, что $x^m e^{\lambda x} = \frac{\partial^m}{\partial \lambda^m} e^{\lambda x}$ и что если $y = y(x, \lambda)$ зависит кроме x еще от параметра λ, как это и есть в данном случае, то операции L и $\frac{\partial^m}{\partial \lambda^m}$ коммутативны, т. е.

$$L\left[\frac{\partial^m y}{\partial \lambda^m}\right] = \frac{\partial^m}{\partial \lambda^m} L[y].$$

Подставим теперь в наше л. д. у. $L[y]=0$ последовательно $e^{\lambda x}$, $xe^{\lambda x}, \ldots, x^{k-1}e^{\lambda x}$. Пользуясь правилом Лейбница для производной

m-го порядка от произведения двух функций (т. I, стр. 228), имеем:

$$L[e^{\lambda x}] = e^{\lambda x} f(\lambda),$$

$$L[xe^{\lambda x}] = \frac{\partial}{\partial\lambda} L[e^{\lambda x}] = e^{\lambda x}[f'(\lambda) + xf(\lambda)],$$

$$L[x^2 e^{\lambda x}] = \frac{\partial^2}{\partial\lambda^2} L[e^{\lambda x}] = e^{\lambda x}[f''(\lambda) + 2xf'(\lambda) + x^2 f(\lambda)],$$

. . . .

$$L[x^{k-1}e^{\lambda x}] = \frac{\partial^{k-1}}{\partial\lambda^{k-1}} L[e^{\lambda x}] = \\ = e^{\lambda x}[f^{(k-1)}(\lambda) + (k-1)xf^{(k-2)}(\lambda) + \ldots + x^{k-1}f(\lambda)].$$

Так как, согласно предположению, λ есть корень характеристического уравнения $f(\lambda)=0$ кратности k, то

$$f(\lambda)=0,\ f'(\lambda)=0,\ f''(\lambda)=0,\ \ldots,\ f^{(k-1)}(\lambda)=0$$

(ср. т. I, стр. 173, упр. 14); поэтому

$$L[e^{\lambda x}]=0,\ L[xe^{\lambda x}]=0,\ \ldots,\ L[x^{k-1}e^{\lambda x}]=0.$$

Таким образом наше утверждение доказано.

Предположим теперь, что характеристическое уравнение имеет корни $\lambda_1, \lambda_2, \ldots, \lambda_\nu$, кратности которых соответственно $k_1, k_2, \ldots, k_\nu$, причем $k_1 + k_2 + \ldots + k_\nu = n$ (степени уравнения). Тогда наше л. д. у. $L[y]=0$ имеет следующие n частных решений:

$$\begin{array}{cccc} e^{\lambda_1 x}, & xe^{\lambda_1 x}, & \ldots, & x^{k_1-1}e^{\lambda_1 x}, \\ e^{\lambda_2 x}, & xe^{\lambda_2 x}, & \ldots, & x^{k_2-1}e^{\lambda_2 x}, \\ \ldots & \ldots & \ldots & \ldots \\ e^{\lambda_\nu x}, & xe^{\lambda_\nu x}, & \ldots, & x^{k_\nu-1}e^{\lambda_\nu x}. \end{array}$$

Все эти функции, согласно упр. 1 на стр. 479, линейно независимы и, следовательно, образуют фундаментальную систему решений.

Пример 1. $y''' + 2y'' - y' - 2y = 0$. Характеристическое уравнение

$$f(\lambda) \equiv \lambda^3 + 2\lambda^2 - \lambda - 2 = 0.$$

Общее решение

$$y = c_1 e^{-x} + c_2 e^{x} + c_3 e^{-2x}.$$

Частное решение, удовлетворяющее условию: $y(0)=2$, $y'(0)=0$, $y''(0)=2$, есть $y = e^x + e^{-x}$.

Пример 2. $y''' - y'' - y' + y = 0$. Общее решение

$$y = c_1 e^{x} + c_2 x e^{x} + c_3 e^{-x}.$$

Пример 3. $y''' - 2y' + 4y = 0$. Характеристическое уравнение

$$f(\lambda) \equiv \lambda^3 - 2\lambda + 4 \equiv (\lambda+2)(\lambda - 1 + i)(\lambda - 1 - i) = 0.$$

Общее решение

$$y = c_1 e^{-2x} + c_2 e^{x}\cos x + c_3 e^{x}\sin x.$$

Упражнения

1. Найти общие решения следующих уравнений:

$$\text{а) } y'''-y=0; \qquad \text{б) } y'''-4y''+5y'-2y=0;$$
$$\text{в) } y'''-3y''+3y'-y=0; \qquad \text{г) } y^{\mathrm{IV}}-3y''+2y=0.$$

[2. Нижеследующие л. д. у. с переменными коэффициентами, называемые *уравнениями Эйлера*, можно преобразовать в такие же уравнения с постоянными коэффициентами с помощью подходящей замены независимой переменной. Найти эту замену переменной, а затем и общие решения этих уравнений:

$$\text{а) } x^2y''+xy'-n^2y=0; \qquad \text{б) } x^3y'''+2x^2y''-4xy'+4y=0.]$$

3*. Материальная точка P массы 1 движется под действием двух сил

$$\boldsymbol{F}=-\lambda^2\boldsymbol{r} \quad \text{и} \quad \boldsymbol{f}=2\mu\,[\boldsymbol{kv}],$$

где $\boldsymbol{k}$ есть орт оси z, $\boldsymbol{r}$ — радиус-вектор точки P, а вектор $\boldsymbol{v}$ — ее скорость. Доказать, что если в момент $t=0$ точке P, находившейся в этот момент в начале координат, сообщена начальная скорость величины v_0 по направлению положительной оси x, то последующее движение описывается следующими уравнениями:

$$x=\frac{v_0}{\sqrt{\lambda^2+\mu^2}}\sin\sqrt{\lambda^2+\mu^2}\,t\cos\mu t,$$
$$y=\frac{v_0}{\sqrt{\lambda^2+\mu^2}}\sin\sqrt{\lambda^2+\mu^2}\,t\sin\mu t.$$

4. а) Пусть u, v — два линейно независимых решения л. д. у. без правой части

$$f(x)y'''-f'(x)y''+\varphi(x)y'+\lambda(x)y=0.$$

Доказать, что общее решение есть $y=Au+Bv+Cw$, где

$$w=u\int\frac{vf(x)\,dx}{(uv'-u'v)^2}-v\int\frac{uf(x)\,dx}{(uv'-u'v)^2}$$

и A, B, C — произвольные постоянные.

б) Найти общее решение уравнения

$$x^2(x^2+5)y'''-x(7x^2+25)y''+(22x^2+40)y'-30xy=0,$$

имеющего частные решения вида x^n.

8. Л. д. у. с правой частью и с переменными коэффициентами. Метод вариации произвольных постоянных. Прежде чем приступить к решению л. д. у. с правой частью

$$L[u]\equiv a_0(x)u^{(n)}+a_1(x)u^{(n-1)}+\ldots+a_{n-1}(x)u'+a_n(x)u=\varphi(x), \quad (1)$$

установим следующие важные факты (ср. т. I, гл. XI, § 5, n° 1). Во-первых, сумма какого-либо частного решения $v(x)$ уравнения (1) и общего решения $w(x, c_1, c_2, \ldots, c_n)$, соответствующего л. д. у. без правой части $L[w]=0$, является решением уравнения (1), содержащим n линейно независимых произвольных постоянных. Во-вторых, если ввести в уравнение (1), имеющее частное решение $v(x)$, новую неизвестную функцию $w(x)$ с помощью равенства

$$u(x)=v(x)+w(x),$$

то $w(x)$ удовлетворяет соответствующему уравнению $L[w]=0$. Оба эти положения легко проверить подстановкой в уравнение (1). Отсюда вытекает, что для получения общего решения л. д. у. с правой частью достаточно найти одно его частное решение. Этого можно добиться следующим путем.

Сначала определяем подходящим выбором постоянных $c_1, c_2, \ldots, c_n$ такое решение соответствующего однородного уравнения $L[u]=0$, которое удовлетворяет условиям

$$u(\xi)=0, \quad u'(\xi)=0, \ \ldots, \ u^{(n-2)}(\xi)=0, \quad u^{(n-1)}(\xi)=1.$$

Это решение, содержащее кроме x еще и параметр ξ, мы обозначим через $u(x, \xi)$. Функция $u(x, \xi)$ и ее первые n производных по x являются при фиксированном x непрерывными функциями от ξ. Например, для уравнения $u''+\omega^2 u=0$ такое решение $u(x, \xi)=\frac{1}{\omega}\sin\omega(x-\xi)$.

Так вот, мы утверждаем, что формула

$$v(x)=\int_0^x \varphi(\xi)\, u(x, \xi)\, d\xi \qquad (A)$$

дает частное решение $v(x)$ уравнения (1), которое в точке $x=0$ обращается в нуль вместе со своими $n-1$ первыми производными.

Для проверки этого утверждения продифференцируем функцию $v(x)$ n раз по x по правилу дифференцирования интеграла по параметру (гл. IV, § 1, п° 2). Принимая во внимание соотношения

$$u(x, x)=0, \ u'(x, x)=0, \ \ldots, \ u^{(n-2)}(x, x)=0, \ u^{(n-1)}(x, x)=1$$

$\left[\text{смысл участвующих здесь символов таков: } u^{(k)}(x, x)=\frac{\partial^k u(x, \xi)}{\partial x^k} \text{ при } \xi=x\right]$, получим:

$$v'(x)=\varphi(\xi)\, u(x, \xi)\Big|_{\xi=x}+\int_0^x \varphi(\xi)\, u'(x, \xi)\, d\xi=\int_0^x \varphi(\xi)\, u'(x, \xi)\, d\xi,$$

$$v''(x)=\varphi(\xi)\, u'(x, \xi)\Big|_{\xi=x}+\int_0^x \varphi(\xi)\, u''(x, \xi)\, d\xi=\int_0^x \varphi(\xi)\, u''(x, \xi)\, d\xi,$$

. . . .

$$v^{(n-1)}(x)=\varphi(\xi)\, u^{(n-2)}(x, \xi)\Big|_{\xi=x}+\int_0^x \varphi(\xi)\, u^{(n-1)}(x, \xi)\, d\xi=$$

$$=\int_0^x \varphi(\xi)\, u^{(n-1)}(x, \xi)\, d\xi,$$

$$v^{(n)}(x)=\varphi(\xi)\, u^{(n-1)}(x, \xi)\Big|_{\xi=x}+\int_0^x \varphi(\xi)\, u^{(n)}(x, \xi)\, d\xi=$$

$$=\varphi(x)+\int_0^x \varphi(\xi)\, u^{(n)}(x, \xi)\, d\xi.$$

Так как $L[u(x,\xi)]=0$, то нетрудно убедиться, что $L[v]=\varphi(x)$. Кроме того, ясно, что $v(0)=v'(0)=\ldots=v^{(n-1)}(0)=0$. Сложив частное решение $v(x)$ уравнения (1) с общим решением соответствующего л. д. у. без правой части, получим общее решение уравнения (1) с правой частью.

Решение $v(x)$ вида (A) имеет следующий физический смысл. Пусть независимая переменная $x=t$ означает время, а функция u есть переменная координата точки, движущейся вдоль прямой под влиянием силы $\varphi(x)$. Тогда действие этой силы можно рассматривать как суперпозицию малых действий элементарных импульсов. Построенная выше функция $u(x,\xi)$ соответствует импульсу величины 1, приложенному в момент времени $x=\xi$, а решение $v(x)$ суммирует действия импульсов $\varphi(\xi)$ за промежуток времени от 0 до x. Мы не имеем здесь возможности входить в дальнейшие подробности.

Уравнение $L[u]=\varphi(x)$ можно также решить следующим методом, который кажется на первый взгляд совершенно отличным от только что изложенного. Мы ищем решение $u(x)$ уравнения (1) в виде линейной комбинации n решений фундаментальной системы соответствующего л. д. у. без правой части:

$$u=\gamma_1(x)u_1(x)+\gamma_2(x)u_2(x)+\ldots+\gamma_n(x)u_n(x),$$

но допускаем при этом, что коэффициенты γ_i могут быть функциями от x. Так как вместо одной искомой функции введено n неизвестных функций $\gamma_i(x)$, то мы вправе наложить на них $n-1$ условие. Мы подчиним эти функции следующим условиям:

$$\left.\begin{array}{l} \gamma_1'u_1+\gamma_2'u_2+\ldots+\gamma_n'u_n=0, \\ \gamma_1'u_1'+\gamma_2'u_2'+\ldots+\gamma_n'u_n'=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \gamma_1'u_1^{(n-2)}+\gamma_2'u_2^{(n-2)}+\ldots+\gamma_n'u_n^{(n-2)}=0. \end{array}\right\} \qquad (B)$$

В силу этих условий выражения для производных $u', u'', \ldots, u^{(n)}$ искомого решения $u(x)$ упрощаются и принимают следующий вид:

$$\begin{aligned} u'&=\sum\gamma_i u_i', \\ u''&=\sum\gamma_i u_i'', \\ &\ldots\ldots\ldots\ldots \\ u^{(n-1)}&=\sum\gamma_i u_i^{(n-1)}, \\ u^{(n)}&=\sum\gamma_i u_i^{(n)}+\sum\gamma_i' u_i^{(n-1)}. \end{aligned}$$

Подставив выражения для $u, u', u'', \ldots, u^{(n)}$ в л. д. у. $L[u]=\varphi(x)$ и учтя, что $L[u_i]=0$ $(i=1, 2, \ldots, n)$, получим

$$\sum\gamma_i' u_i^{(n-1)}=\varphi(x).$$

Условия (B) совместно с этим последним уравнением составляют систему n линейных алгебраических уравнений для производных γ_i'.

Определитель этой системы есть детерминант Вронского фундаментальной системы решений u_i и в силу этого не обращается в нуль. Поэтому величины γ_i' определяются из этой системы уравнений однозначно, а затем находятся и коэффициенты γ_i при помощи квадратур, вносящих n произвольных постоянных интегрирования. Из хода рассуждений ясно, что мы получили самое общее решение нашего уравнения $L[u]=\varphi(x)$.

Этот способ решения называется *методом вариации* или *изменения произвольных постоянных*, так как общее решение получается, как и для л. д. у., без правой части, в виде линейной комбинации функций, но с заменой постоянных коэффициентов переменными.

Можно показать, что оба метода фактически совпадают. Для этого надо частное решение $u(x, \xi)$ уравнения $L[u]=0$, введенное в первом методе, представить в следующем виде:

$$u(x, \xi)=\sum \alpha_i(\xi)u_i(x).$$

Проведение доказательства предоставляем читателю.

Пример. Рассмотрим уравнение

$$u''-\frac{2}{x}u'+\frac{2}{x^2}u=xe^x.$$

В конце п°6 мы нашли для соответствующего л. д. у. без правой части $u''-\frac{2}{x}u'+\frac{2}{x^2}u=0$ фундаментальную систему решений $u_1=x$, $u_2=x^2$. Поэтому ищем решение предложенного уравнения в виде

$$u=\gamma_1(x)x+\gamma_2(x)x^2$$

и подчиняем коэффициенты γ_1, γ_2 условию

$$\gamma_1'x+\gamma_2'x^2=0.$$

К этому условию присоединяется уравнение

$$\gamma_1'+2\gamma_2'x=xe^x.$$

Из этих двух уравнений находим

$$\gamma_1'=-xe^x, \qquad \gamma_2'=e^x,$$

откуда

$$\gamma_1=-xe^x+e^x+c_1 \quad \text{и} \quad \gamma_2=e^x+c_2,$$

и общее решение предложенного л. д. у. с правой частью есть

$$u=xe^x+c_1x+c_2x^2.$$

9. Вынужденное движение простейшей колебательной системы. В первом томе, гл. XI, § 5 мы интегрировали л. д. у. простейших вынужденных колебаний

$$m\ddot{x}(t)+r\dot{x}(t)+kx(t)=f(t)$$

для частного случая, когда внешняя сила $f(t)$ есть периодическая функция времени. В качестве иллюстрации первого метода предыдущего номера мы решим теперь это уравнение при общем предположении, что $f(t)$ есть произвольная непрерывная функция времени. Простоты ради мы ограничимся случаем отсутствия трения ($r=0$) и разделим уравнение на m. Введя обозначения $\frac{k}{m}=\omega^2$ и $\frac{1}{m}f(t)=\varphi(t)$, представим л. д. у. вынужденного колебания в следующем виде:

$$\ddot{x}(t)+\omega^2 x(t)=\varphi(t).$$

Согласно первому методу из n°8, находим сперва зависящее от параметра λ частное решение $x(t, \lambda)$ соответствующего л. д. у. без правой части (вместо u, x, ξ пишем соответственно x, t, λ). В нашем случае, как мы показали в n°8, имеем

$$x(t, \lambda)=\frac{1}{\omega}\sin\omega(t-\lambda).$$

Тогда функция

$$v(t)=\frac{1}{\omega}\int_0^t \varphi(\lambda)\sin\omega(t-\lambda)\,d\lambda$$

является частным решением нашего л. д. у. $\ddot{x}+\omega^2 x=\varphi(t)$, удовлетворяющим начальным условиям $v(0)=0$, $\dot{v}(0)=0$. Согласно общему положению (§ 8, начало), общее решение будет

$$x(t)=\frac{1}{\omega}\int_0^t \varphi(\lambda)\sin\omega(t-\lambda)\,d\lambda+c_1\cos\omega t+c_2\sin\omega t,$$

где c_1 и c_2 — произвольные постоянные. [Тот же результат читатель может получить методом вариации постоянных.]

Если, в частности, правая часть $f(t)=c\cos\omega t$ или $c\sin\omega t$, то простое вычисление приведет к тем же результатам, которые получены в т. I, гл. XI, § 5.

Упражнения

1. В л. д. у. с постоянными коэффициентами

$$a_0y+a_1y'+\ldots+a_ny^{(n)}=P(x)$$

правая часть $P(x)$ есть многочлен.

а) Если $a_0\neq 0$, то рассмотрим формальное тождество

$$\frac{1}{a_0+a_1t+\ldots+a_nt^n}=b_0+b_1t+b_2t^2+\ldots$$

Доказать, что

$$y=b_0P(x)+b_1P'(x)+b_2P''(x)+\ldots$$

есть частное решение дифференциального уравнения.

б) Если $a_0=0$, но $a_1\neq 0$, то существует разложение в степенной ряд

$$\frac{1}{a_1t+a_2t^2+\ldots+a_nt^n}=bt^{-1}+b_0+b_1t+b_2t^2+\ldots$$

Доказать, что теперь

$$y = b \int P(x)\,dx + b_0 P(x) + b_1 P'(x) + b_2 P''(x) + \ldots$$

есть частное решение дифференциального уравнения.

2. С помощью метода упр. 1 найти частные решения дифференциальных уравнений

а) $y'' + y = 3x^2 - 5x$; б) $y'' + y' = (1 + x)^2$.

3. Дано л. д. у. с правой частью

$$a_0 y + a_1 y' + \ldots + a_n y^{(n)} = e^{\alpha x} P(x),$$

где α, a_0, a_1, ..., a_n — действительные постоянные, а $P(x)$ — многочлен. Для того чтобы найти частное решение этого уравнения, можно ввести новую искомую функцию $z = z(x)$ равенством $y = ze^{\alpha x}$, а затем найти частное решение полученного уравнения для z методом упр. 1.

Получить этим путем частные решения следующих уравнений:

а) $y'' + 4y' + 3y = 3e^x$; б) $y'' - 2y' + y = xe^x$.

4. Найти общее решение дифференциального уравнения

$$y'' - 5y' + 6y = e^x (x^2 - 3).$$

10. Определение частного решения по краевым условиям. Нагруженный канат и нагруженная балка. В задачах механики и в рассмотренных выше примерах мы всегда выделяли из многообразия решений дифференциального уравнения (из так называемого общего решения) таксе частное решение, которое удовлетворяет известным начальным условиям, а именно выбирали постоянные интегрирования с таким расчетом, чтобы искомая функция (для уравнения первого порядка) или искомая функция и ее $n - 1$ первых производных (для уравнения n-го порядка) принимали в заданной точке предписанные в условии значения. Однако во многих приложениях требуется найти не общее решение и не частное решение, удовлетворяющее определенным начальным условиям, а решение, определяемое так называемыми *краевыми* условиями; в таких случаях говорят о *краевой задаче*. Краевой задачей обыкновенного дифференциального уравнения называется задача нахождения такого решения этого уравнения, которое удовлетворяет в *нескольких* точках заданным условиям и которое надлежит рассматривать в интервалах между этими точками. [Эти точки являются краевыми (граничными) точками указанных интервалов — отсюда и название задачи.] Не вдаваясь в общую теорию краевых задач, мы рассмотрим здесь некоторые типичные примеры.

Пример 1. *Нагруженный канат.* В вертикальной плоскости xy (ось y направлена по вертикали) висит канат, закрепленный в двух точках: в начале координат и в точке $M(a, b)$ (рис. 104). Горизонтальная проекция натяжения каната имеет постоянную величину T. На канат действует нагрузка, погонная плотность которой на единицу длины горизонтальной проекции задана как кусочно непрерывная функция $p(x)$. Тогда ордината (прогиб) $y(x)$ переменной точки каната удовлетворяет дифференциальному уравнению $y''(x) = \varphi(x)$, где $\varphi(x) = \frac{p(x)}{T}$. Форма каната определится тем решением дифференциального уравнения, которое удовлетворяет краевым условиям $y(0)=0$, $y(a)=b$.

Решим эту краевую задачу. Общим решением соответствующего л. д. у. без правой части $y''=0$ является линейная функция c_0+c_1x. По первому методу n°8 ищем сначала частное решение $y(x, \xi)$ этого уравнения, удовлетворяющее условиям: $y(\xi)=0$, $y'(\xi)=1$; таким решением является функция $y(x, \xi)=x-\xi$. Затем составляем интеграл $\int_0^x \varphi(\xi)(x-\xi)\,d\xi$ — это будет частное решение данного л. д. у. $y''=\varphi(x)$, удовлетворяющее условиям: $y(0)=0$ и $y'(0)=0$. Общее решение этого уравнения есть

$$y(x)=c_0+c_1x+\int_0^x \varphi(\xi)(x-\xi)\,d\xi.$$

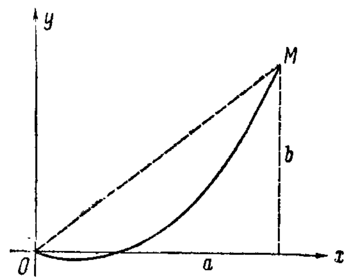


Рис. 104.

Теперь краевое условие $y(0)=0$ сразу дает $c_0=0$, а затем второе краевое условие $y(a)=b$ приводит к уравнению

$$b=c_1a+\int_0^a \varphi(\xi)(a-\xi)\,d\xi$$

для определения коэффициента c_1.

На практике, наряду с этой простейшей формой, встречается более сложная краевая задача, в которой на канат действуют, помимо непрерывно распределенной нагрузки, еще и сосредоточенные силы, например сила, приложенная в точке $x=x_0$. Такую сосредоточенную силу P рассматривают как идеальный предельный случай нагрузки $p(x)$, непрерывно распределенной только по интервалу от $x_0-\varepsilon$ до $x_0+\varepsilon$, когда $\varepsilon\to 0$, причем

$$\int_{x_0-\varepsilon}^{x_0+\varepsilon} p(x)\,dx=P,$$

т. е. полная нагрузка, действующая на участок каната $[x_0-\varepsilon, x_0+\varepsilon]$, остается постоянной в процессе предельного перехода $\varepsilon\to 0$. Число P считается тогда величиной сосредоточенной силы, приложенной в точке x_0.

Если мы предварительно, до перехода к пределу $\varepsilon\to 0$, проинтегрируем обе части нашего дифференциального уравнения $y''=p(x)/T$ от $x_0-\varepsilon$ до $x_0+\varepsilon$, то получим равенство

$$y'(x_0+\varepsilon)-y'(x_0-\varepsilon)=\frac{P}{T},$$

из которого в результате предельного перехода $\varepsilon\to 0$ вытекает, что

$$y'(x_0+0)-y'(x_0-0)=\frac{P}{T},$$

т. е. сосредоточенной силе P, приложенной в точке x_0, соответствует при переходе точки приложения x_0 скачок производной $y'(x)$ на величину $\frac{P}{T}$.

Приведем пример, из которого будет достаточно ясно, как наличие сосредоточенной силы видоизменяет нашу краевую задачу. Предположим, что канат натянут между точками $O(0; 0)$ и $M(1; 1)$ и что единственной нагрузкой является сосредоточенная сила величины P, приложенная в точке

$x=\frac{1}{2}$. [Другими словами, весом каната можно пренебречь по сравнению с силой P.] Эта физическая задача формулируется математически следующим образом. Требуется найти непрерывную функцию $y(x)$ по следующим данным: 1) она является решением дифференциального уравнения $y''=0$ всюду в интервале $0\leqslant x\leqslant 1$, за исключением точки $x_0=\frac{1}{2}$, 2) она удовлетворяет краевым условиям $y(0)=0$, $y(1)=1$ и 3) ее производная $y'(x)$ совершает в точке $x_0=\frac{1}{2}$ скачок величины P/T.

Для нахождения этого решения целесообразно представить его в следующем виде:

$$y(x)=ax+b \quad \text{при } 0\leqslant x\leqslant \frac{1}{2}$$

и

$$y(x)=c(1-x)+d \quad \text{при } \frac{1}{2}\leqslant x\leqslant 1.$$

Условия $y(0)=0$ и $y(1)=1$ сразу дают $b=0$ и $d=1$. Условие непрерывности функции $y(x)$ в точке $x=\frac{1}{2}$ дает

$$\frac{1}{2}a=\frac{1}{2}c+1.$$

Наконец, требование, чтобы производная $y'(x)$ при переходе точки $x_0=\frac{1}{2}$ слева направо совершила скачок вверх на величину P/T, приводит к уравнению

$$-c-a=\frac{P}{T}.$$

Из этих двух уравнений находим остальные коэффициенты $a=1-\frac{P}{2T}$, $c=-1-\frac{P}{2T}$, и следовательно, искомое решение определено.

Пример 2. *Нагруженная балка.* Совершенно аналогично обстоит дело у нагруженной балки (рис. 105). Пусть в случае отсутствия изгиба средняя линия балки совпадает с отрезком $[0, a]$ оси x. Тогда прогиб балки $y(x)$ под влиянием силы, действующей по вертикали параллельно оси y, удовлетворяет л. д. у. четвертого порядка

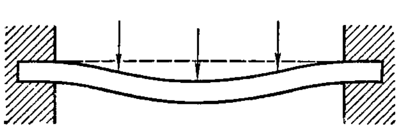


Рис. 105.

$$y^{IV}=\varphi(x),$$

причем $\varphi(x)=\frac{p(x)}{EI}$, где $p(x)$ — линейная плотность нагрузки, E — модуль упругости материала балки (равный напряжению, деленному на вызванное им удлинение), I — момент инерции поперечного сечения балки относительно горизонтальной прямой, проходящей через центр массы поперечного сечения. Теперь (см. n° 8) $y(x, \xi)=\frac{(x-\xi)^3}{3!}$ и общее решение нашего л. д. у.

с правой частью можно записать в таком виде:

$$y(x)=c_0+c_1x+c_2x^2+c_3x^3+\int_0^x \varphi(\xi)\frac{(x-\xi)^3}{3!}\,d\xi, \qquad (A)$$

где c_0, c_1, c_2, c_3 — произвольные постоянные интегрирования. Но реальная физическая задача состоит не в нахождении общего решения, а в отыскании частного решения, удовлетворяющего определенным краевым условиям. Если, например, концы балки заделаны, то краевые условия таковы:

$$y(0)=0, \qquad y(a)=0, \qquad y'(0)=0, \qquad y'(a)=0.$$

Тогда сразу видно, что $c_0=c_1=0$, а коэффициенты c_2 и c_3 определятся из условий

$$c_2a^2+c_3a^3+\int_0^a \varphi(\xi)\frac{(a-\xi)^3}{3!}\,d\xi=0,$$

$$2c_2a+3c_3a^2+\int_0^a \varphi(\xi)\frac{(a-\xi)^2}{2!}\,d\xi=0.$$

Для балки случай наличия сосредоточенных сил тоже вызывает особый интерес. И здесь сосредоточенную силу, приложенную в точке $x=x_0$, представляют себе возникшей из нагрузки $p(x)$, распределенной непрерывно по интервалу от $x_0-\varepsilon$ до $x_0+\varepsilon$; как и в задаче с канатом, ε устремляют к нулю, а $p(x)$ заставляют при этом возрастать таким образом, что суммарная нагрузка $\int\limits_{x_0-\varepsilon}^{x_0+\varepsilon} p(\xi)\,d\xi=P$ остается постоянной. Это число P называют тогда величиной сосредоточенной силы, приложенной в точке $x=x_0$. Как и в примере 1, интегрируем обе части нашего дифференциального уравнения от $x_0-\varepsilon$ до $x_0+\varepsilon$, а затем совершаем предельный переход $\varepsilon\to 0$. В итоге получаем, что наличие в точке x_0 сосредоточенной силы P эквивалентно математическому условию, что третья производная $y'''(x)$ искомого решения должна совершить в точке $x=x_0$ скачок, величина которого

$$y'''(x_0+0)-y'''(x_0-0)=\frac{P}{EI}.$$

Таким образом, возникает следующая математическая задача. Требуется найти решение $y(x)$ уравнения $y^{\mathrm{IV}}=0$, обладающее следующими свойствами: 1) $y(x)$ непрерывно на отрезке [0; 1] вместе со своими первой и второй производными, 2) оно удовлетворяет краевым условиям $y(0)=y(1)=y'(0)=y'(1)=0$ и 3) третья производная $y'''(x)$ совершает при переходе через точку $x=x_0$ скачок величины P/EI, а в остальных точках отрезка [0; 1] она непрерывна.

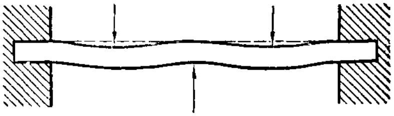


Рис. 106.

Если весомая балка оперта в точке $x=x_0$ (рис. 106), т. е. в этой точке прогибу предписано значение $y(x_0)=0$, то наличие опоры можно заменить действием сосредоточенной в x_0 силы (реакции опоры). Согласно закону механики: действие равно противодействию, реакция опоры равна и противоположна силе, с которой балка давит на опору. Величина этой силы сразу найдется из формулы

$$P=EI\,[y'''(x_0+0)-y'''(x_0-0)],$$

причем $y(x)$ удовлетворяет дифференциальному уравнению $y^{\mathrm{IV}}=\frac{p(x)}{EI}$ всюду на $[0;\ 1]$, за исключением точки $x=x_0$, и краевым условиям $y(0)=y(1)=y'(0)=y'(1)=0$ и $y(x_0)=0$. Кроме того, y, y' и y'' остаются непрерывными также и в точке $x=x_0$.

Для иллюстрации рассмотрим балку, простирающуюся от точки $x=0$ до точки $x=1$, причем эти ее концы заделаны; балка несет равномерную нагрузку линейной плотности $p(x)=1$ и оперта в точке $x=\frac{1}{2}$ (рис. 106). Для упрощения примем $EI=1$; тогда прогиб балки удовлетворяет дифференциальному уравнению

$$y^{\mathrm{IV}}=1.$$

Общее решение этого л. д. у., по формуле (A), есть многочлен четвертой степени, у которого коэффициент при x^4 равен $\frac{1}{4!}$. В виде такого многочлена его следует записать для каждой половины рассматриваемого интервала. В левой половине мы представим общее решение в виде

$$y=b_0+b_1x+b_2x^2+b_3x^3+\frac{1}{4!}x^4,$$

а в правой половине — в виде

$$y=c_0+c_1(x-1)+c_2(x-1)^2+c_3(x-1)^3+\frac{1}{4!}(x-1)^4.$$

Так как балка на концах $x=0$ и $x=1$ заделана, то

$$y(0)=y(1)=y'(0)=y'(1)=0,$$

откуда вытекает, что $b_0=b_1=c_0=c_1=0$. Далее, $y(x)$, $y'(x)$ и $y''(x)$ должны быть непрерывны в точке $x=\frac{1}{2}$, т. е. значения $y\left(\frac{1}{2}\right)$, $y'\left(\frac{1}{2}\right)$ и $y''\left(\frac{1}{2}\right)$, вычисленные из обоих многочленов, должны совпадать, а значение $y\left(\frac{1}{2}\right)$ должно равняться нулю. Это дает

$$\frac{1}{4}b_2+\frac{1}{8}b_3+\frac{1}{384}=\frac{1}{4}c_2-\frac{1}{8}c_3+\frac{1}{384}=0,$$

$$b_2+\frac{3}{4}b_3+\frac{1}{48}=-c_2+\frac{3}{4}c_3-\frac{1}{48},$$

$$2b_2+3b_3+\frac{1}{8}=2c_2-3c_3+\frac{1}{8}.$$

Из этих уравнений находим остальные коэффициенты:

$$b_2=c_2=\frac{1}{96},\qquad b_3=-c_3=-\frac{1}{24}.$$

Сила же реакции опоры в точке $x=\frac{1}{2}$ равна

$$y'''\left(\frac{1}{2}+0\right)-y'''\left(\frac{1}{2}-0\right)=\left(6c_3-\frac{1}{2}\right)-\left(6b_3+\frac{1}{2}\right)=-\frac{1}{2}.$$

§ 5. Потенциал гравитационного и электростатического поля. Уравнение Лапласа

До сих пор мы занимались только обыкновенными дифференциальными уравнениями. Теперь мы познакомимся с несколькими типичными примерами дифференциальных уравнений с частными производными (см. т. I, начало главы XI). Напомним, что дифференциальное уравнение с частными производными служит для нахождения функции от нескольких переменных; оно связывает независимые переменные, зависящую от них искомую функцию и частные производные от этой функции. Мы начнем с теории потенциала и дифференциального уравнения Лапласа, которому он удовлетворяет.

1. Потенциал непрерывного распределения массы или заряда. В гл. IV, § 7, n° 4 мы уже рассматривали силовые поля, порождаемые массами по закону тяготения Ньютона, и представили в них силу поля как градиент потенциала. В качестве обобщения рассмотренных ранее случаев мы теперь будем считать массу или заряд μ положительной или отрицательной величиной. Правда, отрицательные массы не предусматриваются в обычной теории тяготения Ньютона, но в теории электричества роль массы играет заряд или количество электричества, которое может быть как положительным, так и отрицательным, а закон взаимодействия зарядов Кулона имеет тот же вид, что и закон Ньютона. Если заряд μ сосредоточен в одной точке $M(\xi, \eta, \zeta)$, то выражение

$$\frac{\mu}{r}=\frac{\mu}{\sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}}$$

называется потенциалом этого точечного заряда в точке $P(x, y, z)$.

Можно было бы также сказать, что $\frac{\mu}{r}$ есть одно из выражений потенциала этого заряда. Дело в том, что функцию $\frac{\mu}{r}+C$, при любом значении постоянной C, можно с равным правом называть потенциалом заряда μ, так как она имеет тот же градиент и дает то же самое силовое поле.

Если имеется дискретное число n таких зарядов μ_k, сосредоточенных в точках $M_k(\xi_k, \eta_k, \zeta_k)$, которые часто называют *источниками* или *полюсами*, то потенциал этой системы зарядов равен (см. гл. IV, § 7, n° 4)

$$U=\sum_{k=1}^{n}\frac{\mu_k}{r_k}.$$

Напряженность соответствующего электростатического поля дается формулой $E=\operatorname{grad} U$. В поле тяготения Ньютона, когда μ_k означают не заряды, а массы, напряженность поля выражается формулой $f=$ $=\gamma \operatorname{grad} U$, где γ — гравитационная постоянная.

[В физике эту функцию $U(x, y, z)$ принято называть силовой функцией, а потенциалом V называют силовую функцию с обратным знаком, так что потенциал $V = -U$ и напряженность электростатического поля $E = -\operatorname{grad} V$.]

Если заряды или массы не сосредоточены в отдельных точках («источниках»), а распределены по некоторой области G пространства $\xi\eta\zeta$ с объемной плотностью $\mu(\xi, \eta, \zeta)$, то мы уже ввели в гл. IV, § 7, n° 4 в качестве потенциала такого пространственного распределения зарядов выражение

$$U = \iiint\limits_G \frac{\mu}{r}\, d\xi\, d\eta\, d\zeta.$$

Если заряды или массы распределены по куску Σ поверхности с поверхностной плотностью μ, то потенциал выражается так:

$$U = \iint\limits_\Sigma \frac{\mu(u, v)}{r}\, d\sigma,$$

где $d\sigma$ — элемент площади. Поверхность предполагается заданной в параметрическом виде, с параметрами u и v.

Аналогично потенциал заряженной линии (или материальной линии) выражается формулой

$$U = \int \frac{\mu(s)}{r}\, ds,$$

где s — переменная длина дуги, $\mu(s)$ — линейная плотность заряда или массы, а r — расстояние от точки s линии до точки наблюдения $P(x, y, z)$.

Во всех рассмотренных случаях поверхности

$$U(x, y, z) = C$$

являются геометрическими местами постоянного потенциала и называются *поверхностями уровня* или *эквипотенциальными поверхностями*.

Кривые, которые в каждой своей точке касаются соответствующего вектора поля, приложенного в этой точке, называются *векторными* или *силовыми линиями*. Стало быть, силовые линии это кривые, которые пересекают поверхности уровня под прямым углом. Отсюда видно, что в поле, порождаемом одним полюсом или конечным числом полюсов, семейство силовых линий исходит из этих полюсов, имея их, говоря образно, своими истоками. Например, при наличии одного лишь полюса семейство силовых линий состоит из всех полупрямых, выходящих из этого полюса.

Пример потенциала материальной (или заряженной) линии. Отрезок $-l \leqslant z \leqslant l$ оси z несет массу или заряд, распределенный с постоянной линейной плотностью μ. Рассмотрим на плоскости $z = 0$ точку $P(x, y)$. Тогда потенциал заряженного отрезка в точках плоскости xy будет

$$U(x, y) = \mu \int\limits_{-l}^{l} \frac{dz}{\sqrt{\rho^2 + z^2}} + C,$$

где $\rho=\sqrt{x^2+y^2}$ есть расстояние точки P от начала координат. Мы здесь добавили постоянное слагаемое, что, как мы уже знаем, не отразится на силовом поле, которое определяется как градиент потенциала. Первообразная для нашей подынтегральной функции $\frac{1}{\sqrt{\rho^2+z^2}}$ есть

$$\operatorname{arsh}\frac{z}{\rho}=\ln\frac{z+\sqrt{z^2+\rho^2}}{\rho},$$

так что в плоскости xy потенциал выражается так:

$$U(x, y)=2\mu\ln\frac{l+\sqrt{l^2+\rho^2}}{\rho}+C.$$

Для того чтобы вычислить отсюда потенциал заряженной линии, простирающейся вдоль оси z в обе стороны в бесконечность, припишем сначала произвольной постоянной значение $C=-2\mu\ln 2l$ (это делается с той целью, чтобы при предельном переходе $l\to\infty$ потенциал оставался конечным). Тогда потенциал отрезка $-l\leqslant z\leqslant l$ перепишется так:

$$U(x, y)=2\mu\ln\frac{l+\sqrt{l^2+\rho^2}}{\rho}-2\mu\ln 2l=2\mu\ln\frac{l+\sqrt{l^2+\rho^2}}{2l}-2\mu\ln\rho.$$

Теперь заставим l неограниченно возрастать; тогда $\lim\limits_{l\to\infty}\ln\frac{l+\sqrt{l^2+\rho^2}}{2l}=$ $=\ln 1=0$ и пределом потенциала $U(x, y)$ будет выражение

$$U_1(x, y)=-2\mu\ln\rho.$$

Таков, следовательно, *потенциал бесконечной прямой линии, по которой распределен заряд или масса с постоянной плотностью* μ. Это выражение пригодно для всех точек пространства. Здесь ρ есть расстояние точки наблюдения от заряженной линии.

2. Двойной слой и его потенциал. Наряду с поверхностным распределением заряда (простым слоем) в теории потенциала рассматривается также и так называемый *двойной слой*, который вводится следующим образом. Пусть в точке (ξ, η, ζ) сосредоточен заряд M, а в точке $(\xi+h, \eta, \zeta)$ равнопротивоположный заряд $-M$; тогда потенциал этой пары зарядов будет

$$\Phi=\frac{M}{\sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}}-\frac{M}{\sqrt{(x-\xi-h)^2+(y-\eta)^2+(z-\zeta)^2}}.$$

Заставим теперь расстояние h между обоими полюсами стремиться к нулю, а величину M безгранично возрастать, и притом так, что $M=-\frac{\mu}{h}$, где μ — постоянное число. Тогда в пределе получим для потенциала выражение

$$\mu\frac{\partial}{\partial\xi}\left(\frac{1}{r}\right).$$

Это выражение называется *потенциалом диполя* (*двойного полюса*) *с «моментом»* μ *и осью, направленной по оси* ξ. Физический смысл этого термина таков: это потенциал пары равнопротивоположных зарядов, расположенных весьма близко друг от друга. Ось диполя может иметь направление любого вектора $\mathbf{v}$, и тогда потенциал диполя будет равен выражению

$$\mu \frac{\partial}{\partial \mathbf{v}}\left(\frac{1}{r}\right),$$

где $\frac{\partial}{\partial \mathbf{v}} = D_{\mathbf{v}}$ есть символ дифференцирования по направлению вектора $\mathbf{v}$ оси диполя.

Теперь представим себе, что по поверхности $\sum$ распределены диполи с поверхностной плотностью момента μ, причем в каждой точке ось диполя направлена по нормали к поверхности; такое поверхностное распределение диполей и называется *двойным слоем*, и его потенциал равен интегралу по поверхности

$$\iint_{\Sigma} \mu(\xi, \eta, \zeta) \frac{\partial}{\partial \mathbf{v}}\left(\frac{1}{r}\right) d\sigma,$$

где $\frac{\partial}{\partial \mathbf{v}}$ обозначает дифференцирование по тому направлению нормали, которое совпадает с направлением оси элементарного диполя, r есть расстояние от точки поверхности (ξ, η, ζ) до точки поля (x, y, z), причем в процессе интегрирования точка (ξ, η, ζ) пробегает поверхность $\sum$. Можно себе представить, что такой *потенциал двойного слоя* возникает следующим образом. По обе стороны от поверхности $\sum$, на расстояниях h от нее, построим две параллельные ей поверхности; одной из этих поверхностей сообщим заряд с поверхностной плотностью $\mu(\xi, \eta, \zeta)/2h$, а на другой распределим заряд с поверхностной плотностью $-\mu(\xi, \eta, \zeta)/2h$. Эти два слоя совместно порождают во внешних точках потенциал, который при $h \to 0$ имеет своим пределом написанный выше поверхностный интеграл.

Для простоты мы предполагаем, что во всех написанных выше интегралах для потенциалов точка наблюдения (x, y, z) находится в такой части поля, где нет массы (заряда), так что подынтегральные функции и их производные непрерывны.

3. Дифференциальное уравнение потенциала. Теперь нетрудно установить, что в части поля, где нет масс (зарядов), все наши выражения для потенциала удовлетворяют дифференциальному уравнению Лапласа

$$\nabla^2 \Phi = 0 \quad \text{или} \quad \Phi_{xx} + \Phi_{yy} + \Phi_{zz} = 0.$$

Действительно, прямым вычислением можно установить, что функция $\frac{1}{r} = \frac{1}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}}$ удовлетворяет этому уравнению (в т. I, стр. 565, это проверено для частного случая $\xi = \eta = \zeta = 0$). Но отсюда вытекает, что и все остальные наши выражения, получающиеся путем суммирования или интегрирования функций вида $\frac{c}{r}$, тоже удовлетворяют уравнению Лапласа, так как дифференцирование по x, y, z можно производить под знаком интеграла. Удовлетворяет уравнению Лапласа и потенциал двойного слоя; это следует из того, что потенциал отдельного диполя удовлетворяет этому уравнению, в силу переместительности операций ∇^2 и $\frac{\partial}{\partial \nu}$:

$$\nabla^2 \frac{\partial}{\partial \nu}\left(\frac{1}{r}\right) = \frac{\partial}{\partial \nu} \nabla^2 \left(\frac{1}{r}\right) = 0.$$

Конечно, дифференцирование $\frac{\partial}{\partial \nu}$ относится к переменным (ξ, η, ζ), между тем как операция ∇^2 — к переменным (x, y, z). Но дело в том, что выражение $\frac{1}{r}$ как функция шести переменных $(x, y, z; \xi, \eta, \zeta)$ симметрично относительно обеих троек аргументов, а потому удовлетворяет также и дифференциальному уравнению

$$\Phi_{\xi\xi} + \Phi_{\eta\eta} + \Phi_{\zeta\zeta} = 0.$$

Нетрудно проверить, что и выражение $-2\mu \ln \sqrt{x^2+y^2}$, найденное выше для потенциала заряженной вертикальной прямой, тоже удовлетворяет уравнению Лапласа; так как этот потенциал не зависит от z, то он удовлетворяет более простому «двумерному» уравнению Лапласа:

$$\Phi_{xx} + \Phi_{yy} = 0.$$

Изучение этих и родственных им дифференциальных уравнений с частными производными составляет одну из самых важных глав анализа. Следует, однако, подчеркнуть, что главной задачей теории потенциала отнюдь не является отыскание возможно более общих решений таких уравнений. Главный интерес сосредоточен на вопросе о существовании решений, удовлетворяющих некоторым заранее заданным условиям. Так, центральной задачей теории потенциала является следующая *краевая задача*: требуется найти в области G такое решение Φ уравнения Лапласа, которое непрерывно вместе со своими частными производными первого и второго порядка и на границе области G принимает заданные непрерывные значения.

4. Однородный двойной слой. Мы не имеем возможности заняться здесь подробным изучением гармонических функций, т. е. функций, удовлетворяющих уравнению Лапласа. Достаточно будет показать на нескольких примерах, как производится такое исследование, главными орудиями которого являются теоремы Гаусса и Грина, доказанные в гл. V, § 5.

Сначала рассмотрим потенциал двойного слоя с постоянной плотностью момента $\mu = 1$:

$$V = \iint_{\Sigma} \frac{\partial}{\partial \nu}\left(\frac{1}{r}\right) d\sigma.$$

Этот интеграл имеет простой геометрический смысл. Допустим, что всякая точка куска поверхности Σ, несущего двойной слой, видна из точки $P(x, y, z)$, т. е. может быть соединена с P отрезком прямой, не встречающим Σ ни в какой другой точке. Тогда кусок поверхности Σ вместе со всеми полупрямыми, идущими от точки P к его границе, образует конусовидную область G пространства. Мы утверждаем, *что потенциал нашего однородного двойного слоя равен с точностью до знака телесному углу, под которым граница куска поверхности Σ видна из точки P.* Поясним, что под этим телесным углом мы разумеем площадь того куска шаровой поверхности радиуса 1 с центром в точке P, который вырезается образующими конуса G. Этому телесному углу приписывают положительный знак, если образующие конуса, идущие от точки P, пересекают поверхность Σ в том же направлении, что и положительная нормаль $\mathbf{v}$, и отрицательный знак в противном случае (ср. упр. 9, стр. 431).

Для доказательства вспомним, что выражение $u=\frac{1}{r}$ не только как функция точки (x, y, z), но и как функция точки (ξ, η, ζ) удовлетворяет уравнению Лапласа

$$\nabla^2 u = u_{\xi\xi} + u_{\eta\eta} + u_{\zeta\zeta} = 0.$$

Точку $P(x, y, z)$ будем считать неподвижной; прямоугольные координаты точек конусовидной области G обозначим через ξ, η, ζ. Вершину P конуса вырежем из него, описав вокруг P небольшую сферу радиуса ρ; остаток конусовидной области обозначим через G_ρ. Применим теперь к функции $u=\frac{1}{r}$, рассматриваемой как функция точки (ξ, η, ζ) в области G_ρ, теорему Грина (гл. V, § 5, п°4, a) в следующей форме:

$$\iiint\limits_{G_\rho} \nabla^2 u \, d\xi \, d\eta \, d\zeta = \oint\!\!\!\oint\limits_{S} D_n u \, d\sigma,$$

где S означает полную поверхность области G_ρ, а D_n означает дифференцирование по направлению внешней нормали. Так как $\nabla^2 u = 0$, то левая часть равна нулю [1]).

Стоящий в правой части интеграл по поверхности S распадается на три интеграла: 1) интеграл по поверхности Σ, несущей двойной слой, 2) интеграл по боковой поверхности, составленной образующими

---

[1]) Из этой формы теоремы Грина вытекает общий вывод, что если функция u удовлетворяет уравнению Лапласа всюду внутри некоторой замкнутой поверхности S, то $\oint\!\!\!\oint\limits_{S} D_n u \, d\sigma = 0$.

конуса, и 3) интеграл по куску Γ_ρ поверхности малой сферы радиуса ρ. Первый из этих интегралов есть

$$\iint_\Sigma D_n\left(\frac{1}{r}\right)d\sigma = \begin{cases} V, & \text{если} \quad \nu = n, \\ -V, & \text{если} \quad \nu = -n, \end{cases}$$

т. е. этот интеграл по поверхности Σ равен потенциалу V двойного слоя, если на Σ направление оси плотности момента совпадает с направлением внешней нормали поверхности S, и равен $-V$, если оно совпадает с направлением внутренней нормали. Второй интеграл равен нулю, так как на боковой поверхности нормальный вектор n перпендикулярен к образующим конуса, а следовательно, направлен по касательным к сферам $r = \text{const}$. В третьем интеграле, по куску Γ_ρ малой сферы радиуса ρ, имеем $r = \rho$, а $D_n\left(\frac{1}{r}\right) = \left[-\frac{\partial}{\partial r}\frac{1}{r}\right]_{r=\rho} = \frac{1}{\rho^2}$, так как внешняя нормаль направлена в сторону наибыстрейшего уменьшения r. Поэтому

$$\iint_{\Gamma_\rho} D_n\left(\frac{1}{r}\right)d\sigma = \frac{1}{\rho^2}\iint_{\Gamma_\rho} d\sigma = \Omega,$$

где Ω есть телесный угол, под которым из точки P виден кусок поверхности Σ. Следовательно, формула Грина, примененная к области G_ρ, дает

$$\pm V + \Omega = 0,$$

откуда

$$V = \mp \Omega.$$

Это и есть сформулированное выше утверждение. Знак потенциала окажется положительным или отрицательным, смотря по тому, какая сторона двойного слоя лежит ближе к точке наблюдения P, заряженная положительно или отрицательно.

Если двойной слой не имеет такого простого расположения относительно точки P, как описано выше, так что некоторая часть лучей, выходящих из P, пересекает поверхность Σ в нескольких точках, то достаточно разбить эту поверхность на подходящее число частей более простого вида и расположения относительно точки P, чтобы убедиться, что наше утверждение остается в силе. Потенциал однородного двойного слоя (с плотностью момента $\mu = 1$), расположенного на куске поверхности, имеющем границу, равен, с точностью до знака, телесному углу, под которым эта граница видна из точки наблюдения $P(x, y, z)$.

Если двойным слоем покрыта *замкнутая* поверхность, то, разбив ее на две части, имеющие (общую) границу, найдем, что потенциал равен нулю, если точка P лежит вне поверхности, и $\pm 4\pi$, если P лежит внутри ее.

В случае двух независимых переменных аналогичное рассуждение показывает, что интеграл

$$\int_C D_\nu(\ln r)\,ds,$$

взятый вдоль дуги C, равен, с точностью до знака, углу, под которым хорда, соединяющая конечные точки дуги, видна из точки $P(x, y)$.

Этот результат можно также вывести прямым геометрическим путем. Пусть точка $Q(\xi, \eta)$ лежит на дуге C. Тогда производная от $\ln r$ в точке Q по направлению нормали к кривой C выразится так:

$$D_\nu \ln r = \frac{d}{dr}\ln r \cdot \cos(\boldsymbol{\nu}, \boldsymbol{r}) = \frac{1}{r}\cos(\boldsymbol{\nu}, \boldsymbol{r}),$$

где символом $(\boldsymbol{\nu}, \boldsymbol{r})$ обозначен угол между нормальным вектором $\boldsymbol{\nu}$ и радиус-вектором $\boldsymbol{r}$. С другой стороны, элемент дуги ds выражается в полярных координатах (r, θ) формулой

$$ds = \frac{r\,d\theta}{\cos(\boldsymbol{\nu}, \boldsymbol{r})}$$

(получается сравнением формулы для $\operatorname{tg}\mu$, т. I, стр. 310, с формулами т. I, стр. 324). Таким образом,

$$\int D_\nu(\ln r)\,ds = \int \frac{1}{r}\cos(\boldsymbol{\nu}, \boldsymbol{r})\frac{r\,d\theta}{\cos(\boldsymbol{\nu}, \boldsymbol{r})} = \int d\theta.$$

Следовательно, данный интеграл равен углу, под которым дуга C видна из точки $P(x, y)$. Кстати, соответствующий пространственный результат, полученный выше, тоже можно вывести геометрическим путем.

5. Теорема о среднем значении. В качестве второго применения преобразования Грина докажем, что всякая гармоническая функция, т. е. всякая функция $u(x, y, z)$, удовлетворяющая в некоторой области G уравнению Лапласа, обладает свойством, выражаемым следующей теоремой о среднем значении:

Значение гармонической функции $u(x, y, z)$ в центре P произвольного шара, лежащего полностью в области G, равно среднему значению этой функции на поверхности упомянутого шара, т. е.

$$u(x, y, z) = \frac{1}{4\pi r^2}\iint_{S_r} \bar{u}\,d\sigma,$$

где $u(x, y, z)$ есть значение функции в центре P, а $\bar{u}$ — ее значение на поверхности S_r шара радиуса r.

Для доказательства построим шаровую поверхность S_ρ с тем же центром P и радиусом $\rho < r$. Так как внутри S_ρ всюду $\nabla^2 u = 0$, то на основании сноски на стр. 498 поверхностный интеграл

$$\iint_{S_\rho} D_n u\,d\sigma = 0.$$

Обозначим через ξ, η, ζ текущие прямоугольные координаты и введем систему сферических координат с началом в точке $P(x, y, z)$ с помощью формул

$$\xi - x = \rho\cos\varphi\sin\theta, \qquad \eta - y = \rho\sin\varphi\sin\theta, \qquad \zeta - z = \rho\cos\theta.$$

Тогда последнее равенство примет такой вид:

$$\oiint\limits_{S_\rho} \frac{\partial u(\rho, \theta, \varphi)}{\partial\rho}\, d\sigma = 0.$$

Так как элемент площади шаровой поверхности S_ρ есть $d\sigma = \rho^2 d\bar{\sigma}$, где $d\bar{\sigma}$ есть элемент площади *единичной* сферы S (ср. гл. IV, стр. 295, петит), то при $\rho > 0$

$$\oiint\limits_{S} \frac{\partial u}{\partial\rho}\, d\bar{\sigma} = 0,$$

причем $d\bar{\sigma}$ и область интегрирования уже не зависят от ρ. Следовательно, и

$$\int\limits_0^r d\rho \oiint\limits_{S} \frac{\partial u}{\partial\rho}\, d\bar{\sigma} = 0.$$

Переменив порядок интегрирования и выполнив интегрирование по ρ, получим

$$\oiint\limits_{S} \{u(r, \theta, \varphi) - u(0, \theta, \varphi)\}\, d\bar{\sigma} = 0.$$

Так как $u(0, \theta, \varphi) = u(x, y, z)$ не зависит от θ и φ, а

$$\oiint\limits_{S} u(r, \theta, \varphi)\, d\bar{\sigma} = \frac{1}{r^2}\oiint\limits_{S_r} u(r, \theta, \varphi)\, d\sigma = \frac{1}{r^2}\oiint\limits_{S_r} \bar{u}\, d\sigma,$$

то из предыдущего равенства вытекает, что

$$\frac{1}{r^2}\oiint\limits_{S_r} \bar{u}\, d\sigma - 4\pi u(x, y, z) = 0,$$

откуда

$$u(x, y, z) = \frac{1}{4\pi r^2}\oiint\limits_{S_r} \bar{u}\, d\sigma,$$

и теорема о среднем значении доказана. [Еще более простое доказательство теоремы о среднем значении для гармонических функций дал Б. П. Демидович (УМН, т. IX, вып. 3 (61), стр. 213, 1954 г.).]

Совершенно аналогичным путем можно доказать для всякой функции $u(x, y)$, зависящей от двух переменных и удовлетворяющей двумерному уравнению Лапласа $u_{xx}+u_{yy}=0$, *теорему о среднем значении на окружности*, выраженную формулой

$$u(x, y)=\frac{1}{2\pi r}\oint_{C_r} \bar{u}\, ds,$$

где $\bar{u}$ — значение гармонической функции u на окружности C_r радиуса r с центром в точке $P(x, y)$, а ds есть элемент дуги этой окружности.

6. Краевая задача для окружности. Интеграл Пуассона. В качестве примера краевой задачи рассмотрим уравнение Лапласа на плоскости для круговой области $x^2+y^2 \leqslant R^2$. Введем полярные координаты (r, θ). Требуется найти такую функцию $u(x, y)$, которая непрерывна в указанном круге и на его границе, имеет в этой области непрерывные производные первого и второго порядка, удовлетворяет уравнению Лапласа $u_{xx}+u_{yy}=0$ и на границе области, т. е. на окружности круга, принимает заданные краевые значения $u(R, \theta)=f(\theta)$. Мы предполагаем, что $f(\theta)$ есть заданная непрерывная периодическая функция от θ (с периодом 2π), имеющая кусочно непрерывную производную по θ.

Решение этой задачи дается в полярных координатах так называемым *интегралом Пуассона*

$$u=\frac{R^2-r^2}{2\pi}\int_0^{2\pi}\frac{f(\alpha)}{R^2-2Rr\cos(\theta-\alpha)+r^2}\, d\alpha.$$

Для доказательства преобразуем уравнение Лапласа к полярным координатам

$$\nabla^2 u \equiv \frac{1}{r}(ru_r)_r+\frac{1}{r^2}u_{\theta\theta}=0$$

и построим последовательность его частных решений следующим путем. Будем искать такие решения, которые можно выразить как произведение функции от r на функцию от θ, т. е. решения вида

$$u=\varphi(r)\psi(\theta).$$

Подставив это выражение в уравнение Лапласа, получим

$$r\cdot\frac{r\varphi''(r)+\varphi'(r)}{\varphi(r)}=-\frac{\psi''(\theta)}{\psi(\theta)}.$$

Так как левая часть не зависит от θ, а правая не зависит от r, то их равенство возможно лишь в том случае, если обе части не зависят ни от r, ни от θ, т. е. равны одной и той же постоянной k. Отсюда получается для $\psi(\theta)$ дифференциальное уравнение

$$\psi''+k\psi=0.$$

Так как искомое решение u должно быть периодической функцией от θ с периодом 2π, то и функция $\psi(\theta)$ должна быть периодической с периодом 2π. Отсюда вытекает, что постоянная k должна иметь вид $k=n^2$, где n — целое число, и тогда

$$\psi(\theta)=a\cos n\theta+b\sin n\theta,$$

где a и b — произвольные постоянные.

Для функции $\varphi(r)$ получается линейное дифференциальное уравнение

$$r^2\varphi''(r)+r\varphi'(r)-n^2\varphi(r)=0,$$

имеющее линейно независимые частные решения r^n и r^{-n} (см. упр. 2а на стр. 483), что, впрочем, нетрудно проверить подстановкой. Так как r^{-n} обращается в бесконечность в начале координат, тогда как функция u должна там сохранить непрерывность, то остается только первое решение $\varphi(r)=r^n$; умножая это выражение $\varphi(r)$ на найденное выше $\psi(\theta)$, получаем последовательность решений уравнения Лапласа:

$$r^n(a\cos n\theta+b\sin n\theta).$$

Согласно принципу суперпозиции (§ 4, n° 1), мы можем теперь построить новые решения уравнения Лапласа, представляющие собой линейные комбинации таких решений:

$$\frac{1}{2}a_0+\sum r^n(a_n\cos n\theta+b_n\sin n\theta).$$

Даже бесконечный ряд такого вида будет решением, если он только равномерно сходится внутри круга и допускает в нем двукратное почленное дифференцирование.

Представим себе теперь, что заданная в условии задачи краевая функция $f(\theta)$ разложена в ряд Фурье

$$f(\theta)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos n\theta+b_n\sin n\theta);$$

тогда этот ряд непременно сходится абсолютно и равномерно (т. I, гл. IX, стр. 529). Следовательно, ряд

$$u(r,\ \theta)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\frac{r^n}{R^n}(a_n\cos n\theta+b_n\sin n\theta) \tag{*}$$

подавно сходится равномерно и абсолютно внутри нашего круга. Но этот ряд можно почленно дифференцировать, если $r<R$, так как ряд, полученный дифференцированием, тоже сходится равномерно (по поводу дифференцирования степенного ряда ср. т. I, гл. VIII, § 5, n° 2). Стало быть, функция $u(r,\ \theta)$, представленная бесконечным рядом (*), является потенциальной функцией и принимает заданные значения на границе круга; следовательно, она дает решение нашей краевой задачи.

Остается привести полученное решение к тому интегральному виду, который мы назвали выше интегралом Пуассона. Для этого подставим вместо коэффициентов a_n и b_n их интеральные выражения

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(\alpha)\cos n\alpha\, d\alpha, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(\alpha)\sin n\alpha\, d\alpha.$$

Так как сходимость ряда равномерна, то можно перетавить порядок интегрирования и суммирования, и мы получим

$$u(r, \theta) = \frac{1}{\pi}\int_0^{2\pi} f(\alpha)\left\{\frac{1}{2} + \sum_{n=1}^{\infty} \frac{r^n}{R^n}\cos n(\theta - \alpha)\right\} d\alpha.$$

Предоставляем читателю доказать тождество

$$\frac{1}{2} + \sum_{n=1}^{\infty} \frac{r^n}{R^n}\cos n\tau = \frac{1}{2}\,\frac{R^2 - r^2}{R^2 - 2Rr\cos\tau + r^2},$$

пользуясь методом т. I, гл. IX, § 2, n° 4 (стр. 509), и интегральная формула Пуассона будет доказана.

Упражнения

1. Применив к формуле Пуассона преобразование инверсии, найти потенциальную функцию $u(x, y)$, ограниченную во *внешней* области единичного круга и принимающую заданные значения на его границе (это так называемая *внешняя* краевая задача для круга).

2*. Для потенциала отрезка $x = y = 0$, $-l \leqslant z \leqslant l$, несущего заряд постоянной линейной плотности μ, найти: а) эквипотенциальные поверхности и б) силовые линии.

3*. Доказать, что если заданы значения потенциальной функции $u(x, y, z)$ и ее производной по нормали $D_n u = \frac{\partial u}{\partial n}$ на замкнутой поверхности S, то значение функции u в любой внутренней точке дается следующим выражением:

$$u(x, y, z) = \frac{1}{4\pi}\iint_S \left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial(1/r)}{\partial n}\right) d\sigma,$$

где r есть расстояние от переменной точки интегрирования до точки (x, y, z). $\left(\text{Применить теорему Грина к функциям } u \text{ и } \frac{1}{r}\right)$

§ 6. Дальнейшие примеры дифференциальных уравнений с частными производными

Рассмотрим теперь вкратце некоторые часто встречающиеся дифференциальные уравнения с частными производными.

[Мы позволяем себе использовать здесь в nn° 1 и 2 небольшой отрывок из книги: Р. Курант и Д. Гильберт, Методы математической физики, т. II, гл. I, § 1, 1951, в нашем переводе.]

1. Некоторые сведения о многообразии решений. Совокупность всех решений *обыкновенного* дифференциального уравнения порядка n (за исключением, быть может, так называемых *особых* решений) выражается в виде функции u от независимой переменной x, содержащей кроме x еще n произвольных постоянных $c_1, c_2, \ldots, c_n$:

$$u = \varphi(x;\ c_1,\ c_2,\ \ldots,\ c_n).$$

У дифференциальных уравнений *с частными производными* дело обстоит сложнее. И здесь можно ставить вопрос о нахождении всей совокупности решений или *общего решения*, из которого, путем конкретизации некоторых содержащихся в нем произвольных элементов, можно получить все индивидуальные или частные решения (опять-таки за исключением возможных *особых* решений). Однако, у дифференциальных уравнений с частными производными такие произвольные элементы фигурируют уже не в виде произвольных постоянных, но в виде произвольных функций, и их количество равно, вообще говоря, порядку дифференциального уравнения. Число же независимых переменных, от которых эти произвольные функции зависят, на единицу меньше числа аргументов общего решения.

Пример 1. Найти общее решение $u(x, y)$ дифференциального уравнения $u_y = 0$.

Из этого уравнения вытекает, что искомая функция u не зависит от y; стало быть,

$$u = w(x),$$

где $w(x)$ есть произвольная функция от x.

В этом номере мы буквами w и v будем обозначать произвольные функции.

Пример 2. Аналогичным рассуждением нетрудно получить общее решение уравнения

$$u_{xy} = 0.$$

Действительно, прежде всего имеем

$$u_x = w_1(x),$$

откуда интегрированием получим

$$u = \int_{x_0}^{x} w_1(\xi)\, d\xi + v(y) = w(x) + v(y).$$

Пример 3. Общее решение дифференциального уравнения с правой частью

$$u_{xy} = a(x, y)$$

получается в виде суммы одного частного решения этого уравнения и общего решения соответствующего уравнения без правой части $u_{xy} = 0$:

$$u = \int_{x_0}^{x} d\xi \int_{y_0}^{y} a(\xi, \eta)\, d\eta + w(x) + v(y).$$

Пример 4. Дифференциальное уравнение с частными производными

$$u_x = u_y$$

приводится преобразованием независимых переменных

$$x + y = \xi, \qquad x - y = \eta$$

к виду

$$2u_\eta = 0.$$

На основании примера 1 общее решение этого уравнения есть $u = w(\xi)$, а стало быть, общее решение предложенного уравнения есть

$$u = w(x + y).$$

Пример 5. Дифференциальное уравнение

$$\alpha u_x + \beta u_y = 0$$

приводится аналогичным преобразованием

$$\xi = \alpha x + \beta y, \qquad \eta = \beta x - \alpha y$$

к виду

$$(\alpha^2 + \beta^2)\, u_\xi = 0.$$

Следовательно, общее решение данного уравнения есть

$$u = w(\eta) = w(\beta x - \alpha y).$$

Пример 6. Если искомая функция зависит от трех или от большего числа независимых переменных, то общее решение содержит произвольные функции от двух или большего числа переменных. Например, дифференциальное уравнение

$$u_z = 0$$

для функции $u(x, y, z)$ имеет общее решение

$$u = w(x, y).$$

2. Одномерное волновое уравнение. Явления распространения волн, например света или звука, описываются так называемым *волновым уравнением*. Мы начнем с рассмотрения простого идеализированного случая «одномерной волны». Такая волна состоит в изменении какого-либо свойства u, например давления, положения частицы или напряженности электрического поля, причем это свойство u зависит не только от координаты x (направление распространения мы принимаем за ось x), но и от времени t.

Такая функция $u(x, t)$ удовлетворяет дифференциальному уравнению вида

$$u_{xx} = \frac{1}{a^2} u_{tt}, \tag{1}$$

где a есть постоянная, определяемая физической природой среды, в которой происходит распространение волны.

Для нахождения общего решения вводим новые независимые переменные

$$\xi = x - at, \qquad \eta = x + at.$$

Тогда

$$u_x = u_\xi + u_\eta, \qquad u_t = -au_\xi + au_\eta,$$
$$u_{xx} = u_{\xi\xi} + 2u_{\xi\eta} + u_{\eta\eta}, \qquad u_{tt} = a^2u_{\xi\xi} - 2a^2u_{\xi\eta} + a^2u_{\eta\eta},$$

в силу чего наше волновое уравнение преобразовывается к виду

$$u_{\xi\eta} = 0,$$

общее решение которого (см. n° 1, пример 2) есть $u = f(\xi) + g(\eta)$, где $f(\xi)$ и $g(\eta)$ — произвольные дважды дифференцируемые функции. Так как уравнение $u_{\xi\eta} = 0$ есть следствие уравнения (1), то все решения последнего должны содержаться в формуле

$$u = f(x - at) + g(x + at). \tag{2}$$

С другой стороны, обратно, при любом выборе функций $f(\xi)$ и $g(\eta)$, имеющих непрерывные производные первого и второго порядка, выражение (2) удовлетворяет волновому уравнению. В самом деле, полагая в (2) $\xi = x - at$ и $\eta = x + at$, имеем

$$u_{xx} = f''(\xi) + g''(\eta) \quad \text{и} \quad u_{tt} = a^2f''(\xi) + a^2g''(\eta),$$

и сразу видно, что $u_{xx} = \frac{1}{a^2}u_{tt}$. Отсюда вытекает, что формула (2) дает общее решение уравнения (1) при любом выборе дважды непрерывно дифференцируемых функций $f(\xi)$ и $g(\eta)$.

Каждый член общего решения (2) в отдельности представляет волновой процесс, распространяющийся со скоростью a вдоль оси x: первый описывает волну, бегущую в положительном направлении оси x, второй — волну, движущуюся в отрицательном ее направлении. В самом деле, если в первой волне функция $f(x - at)$ имеет значение $f(x_1 - at_1)$ в какой-либо точке x_1 в момент t_1, то эта функция принимает такое же значение в более поздний момент t в точке $x = x_1 + a(t - t_1)$, так как $x - at = x_1 - at_1$, так что $f(x - at) = f(x_1 - at_1)$. Таким же рассуждением можно проверить, что функция $g(x + at)$ представляет волну, движущуюся со скоростью a в отрицательном направлении оси x.

Для волнового уравнения мы теперь поставим и решим следующую задачу с начальными условиями. Из всех возможных решений дифференциального уравнения (1) требуется выделить такое решение, для которого начальное состояние при $t = 0$

$$u(x, 0) = \varphi(x) \quad \text{и} \quad u_t(x, 0) = \psi(x)$$

определяется двумя заранее заданными функциями $\varphi(x)$ и $\psi(x)$.

Для решения этой задачи надо только написать

$$u = f(x - at) + g(x + at)$$

и определить функции f и g из системы двух уравнений:

$$f(x)+g(x)=\varphi(x),$$
$$-f'(x)+g'(x)=\frac{1}{a}\psi(x).$$

Второе из этих уравнений дает

$$-f(x)+g(x)=\frac{1}{a}\int_0^x \psi(\tau)\,d\tau+c,$$

где c — постоянная интегрирования. Комбинируя этот результат с первым уравнением системы, получим функции f и g, а затем найдем искомое решение

$$u(x, t)=\frac{\varphi(x+at)+\varphi(x-at)}{2}+\frac{1}{2a}\int_{x-at}^{x+at}\psi(\tau)\,d\tau.$$

Предоставляем читателю доказать самостоятельно (путем введения вместо x и t новых переменных $\xi=x-at$ и $\eta=x+at$), что других решений поставленной задачи быть не может.

3. Волновое уравнение в трехмерном пространстве. Функция u, удовлетворяющая волновому уравнению в трехмерном пространстве, зависит от четырех аргументов, а именно от трех пространственных координат x, y, z и от времени t. Само волновое уравнение в пространстве имеет следующий вид:

$$u_{xx}+u_{yy}+u_{zz}=\frac{1}{a^2}u_{tt} \qquad (1)$$

или, короче,

$$\nabla^2 u=\frac{1}{a^2}u_{tt}.$$

И здесь нетрудно найти такие решения, которые описывают распространение плоской волны в физическом смысле.

Действительно, всякая функция $f(\xi)$, имеющая непрерывные производные первого и второго порядка, дает решение дифференциального уравнения (1), если положить ξ равным линейному выражению вида

$$\xi=\alpha x+\beta y+\gamma z\pm at,$$

коэффициенты которого удовлетворяют соотношению

$$\alpha^2+\beta^2+\gamma^2=1.$$

В самом деле, при этих обстоятельствах

$$\nabla^2 u=(\alpha^2+\beta^2+\gamma^2)f''(\xi)=f''(\xi)$$

и

$$u_{tt}=a^2 f''(\xi),$$

откуда видно, что функция $u=f(\alpha x+\beta y+\gamma z\pm at)$ действительно является решением волнового уравнения (1).

Из этого вида функции u ясно, что во всех точках плоскости $\alpha x+\beta y+\gamma z-p=0$, имеющей единичный нормальный вектор $n^{\circ}=\{\alpha, \beta, \gamma\}$ и отстоящей на расстоянии p от начала координат, распространяющееся свойство или «возбуждение», представляемое функцией u, имеет в любой данный момент t одинаковое значение. Это возбуждение распространяется в пространстве таким образом, что всякая плоскость, перпендикулярная к вектору n°, является геометрическим местом одинакового состояния возбуждения. Скорость распространения по направлению общей нормали n° этих плоскостей равна a.

В теоретической физике процесс, распространяющийся описанным здесь способом, называется *плоской волной.*

Особенно важен тот случай, когда возбуждение изменяется со временем периодически. Если частота колебания есть ω, то такое явление может быть представлено функцией

$$u=Ae^{ik(\alpha x+\beta y+\gamma z)}e^{i\omega t},$$

где k обозначает, как обычно, число, обратное длине волны λ, $k=\frac{1}{\lambda}=\frac{\omega}{a}$.

Для волнового уравнения в трехмерном пространстве можно найти и другие решения, представляющие *сферическую волну*, распространяющуюся из заданной точки, например из начала координат. Сферическая волна определяется тем свойством, что в любой фиксированный момент времени геометрическим местом постоянного значения возбуждения является сфера с центром в источнике возбуждения, скажем, в начале координат, т. е. во всех точках каждой такой сферы возбуждение имеет одинаковое значение.

Для того чтобы найти решения, обладающие этим свойством, надо преобразовать лапласиан $\nabla^2 u$ к сферическим координатам (r, θ, φ) (см. гл. V, § 5, n° 3), а затем сделать допущение, что искомая функция u зависит только от r и t и уже не зависит от θ и φ. Полагая в волновом уравнении, преобразованном к сферическим координатам, производные от u по θ и по φ равными нулю, приведем его к следующему виду:

$$u_{rr}+\frac{2}{r}u_r=\frac{1}{a^2}u_{tt} \quad \text{или} \quad (ru)_{rr}=\frac{1}{a^2}(ru)_{tt}.$$

Отсюда ясно, что функция $w=ru$ является решением одномерного волнового уравнения $w_{rr}=\frac{1}{a^2}w_{tt}$, рассмотренного выше. Следовательно,

$$w=f(r-at)+g(r+at),$$

откуда

$$u=\frac{f(r-at)+g(r+at)}{r},$$

Читатель сам легко проверит, что функция такого типа действительно удовлетворяет уравнению $\nabla^2 u = \frac{1}{a^2} u_{tt}$.

Физически функция $u = \frac{1}{r} f(r - at)$ представляет волну, распространяющуюся из начала координат в окружающее пространство со скоростью a.

4. Уравнения Максвелла в вакууме. В качестве заключительного примера рассмотрим систему уравнений с частными производными, известных под названием *уравнений Максвелла* и лежащих в основе электродинамики. Мы не станем пытаться подойти к этим уравнениям с физической точки зрения и только воспользуемся ими для иллюстрации различных математических понятий, развитых выше.

Электромагнитное состояние в вакууме определяется двумя векторами: электрическим вектором $E = \{E_1, E_2, E_3\}$ и магнитным вектором $\boldsymbol{H} = \{H_1, H_2, H_3\}$. Эти векторы (точнее, вектор-функции точки) удовлетворяют уравнениям Максвелла:

$$\left.\begin{aligned} \operatorname{rot} \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{H}}{\partial t} &= 0, \\ \operatorname{rot} \boldsymbol{H} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} &= 0, \end{aligned}\right\} \qquad (M)$$

где c есть скорость света в вакууме. В координатной записи уравнения Максвелла имеют следующий вид:

$$\begin{aligned} \frac{\partial E_3}{\partial y} - \frac{\partial E_2}{\partial z} + \frac{1}{c}\frac{\partial H_1}{\partial t} &= 0, \\ \frac{\partial E_1}{\partial z} - \frac{\partial E_3}{\partial x} + \frac{1}{c}\frac{\partial H_2}{\partial t} &= 0, \\ \frac{\partial E_2}{\partial x} - \frac{\partial E_1}{\partial y} + \frac{1}{c}\frac{\partial H_3}{\partial t} &= 0 \end{aligned}$$

и

$$\begin{aligned} \frac{\partial H_3}{\partial y} - \frac{\partial H_2}{\partial z} - \frac{1}{c}\frac{\partial E_1}{\partial t} &= 0, \\ \frac{\partial H_1}{\partial z} - \frac{\partial H_3}{\partial x} - \frac{1}{c}\frac{\partial E_2}{\partial t} &= 0, \\ \frac{\partial H_2}{\partial x} - \frac{\partial H_1}{\partial y} - \frac{1}{c}\frac{\partial E_3}{\partial t} &= 0. \end{aligned}$$

Таким образом, для координат векторов $\boldsymbol{E}$ и $\boldsymbol{H}$, рассматриваемых как функции точки и времени, это — система шести дифференциальных уравнений с частными производными первого порядка.

Выведем теперь несколько важных следствий уравнений Максвелла.

1) Вычислим дивергенцию от обоих уравнений (M). Так как $\operatorname{div} \operatorname{rot} \boldsymbol{F} = 0$ (см. конец гл. II), а порядок дифференцирования по времени и вычисления дивергенции (которое состоит в дифферен-

цировании по координатам x, y, z) можно переставить, то получим

$$\operatorname{div} \boldsymbol{E} = \text{const}, \qquad \operatorname{div} \boldsymbol{H} = \text{const},$$

т. е. $\operatorname{div} \boldsymbol{E}$ и $\operatorname{div} \boldsymbol{H}$ не зависят от времени. Если допустить, что эти постоянные интегрирования равны нулю в начальный момент, то они всегда останутся равными нулю.

2) Рассмотрим какую-либо замкнутую поверхность S, лежащую в электромагнитном поле, и возьмем объемные интегралы по области G, ограниченной этой поверхностью:

$$\iiint \operatorname{div} \boldsymbol{E}\, dV \quad \text{и} \quad \iiint \operatorname{div} \boldsymbol{H}\, dV.$$

Применяя теорему Гаусса (стр. 411) к этим тройным интегралам, мы их преобразуем в интегралы по поверхности S от проекций E_n и H_n на нормаль к этой поверхности. Так как $\operatorname{div} \boldsymbol{E} = 0$ и $\operatorname{div} \boldsymbol{H} = 0$, то отсюда вытекает, что

$$\oint\!\!\oint_S \boldsymbol{E}_n\, d\sigma = 0 \quad \text{и} \quad \oint\!\!\oint_S \boldsymbol{H}_n\, d\sigma = 0.$$

Эти поверхностные интегралы называются потоком электрической и магнитной напряженности через поверхность S. Полученный результат можно поэтому формулировать так: при сделанных выше допущениях поток электрической напряженности и поток магнитной напряженности через любую замкнутую поверхность равны нулю.

3) Мы получим еще одно следствие из уравнений Максвелла, рассматривая кусок поверхности Σ, ограниченный кривой Γ, лежащей на этой поверхности.

Спроектируем обе части векторных уравнений (M) на нормаль $\boldsymbol{n}$ к поверхности Σ. Вспомнив, что проекция любого вектора $\boldsymbol{A}$ на направление $\boldsymbol{n}$ обозначается через A_n, получим

$$(\operatorname{rot} \boldsymbol{E})_n = -\frac{1}{c}\frac{\partial H_n}{\partial t}, \qquad (\operatorname{rot} \boldsymbol{H})_n = \frac{1}{c}\frac{\partial E_n}{\partial t}.$$

Эти уравнения мы проинтегрируем по куску поверхности Σ с элементом площади $d\sigma$ и левые части преобразуем по теореме Стокса (гл. V, § 6, n° 1) в криволинейный интеграл по граничной кривой Γ, а в правой части вынесем дифференцирование по t за знак поверхностного интеграла. В результате получим

$$\oint_\Gamma E_s\, ds = -\frac{1}{c}\frac{d}{dt}\iint_\Sigma H_n\, d\sigma,$$

$$\oint_\Gamma H_s\, ds = \frac{1}{c}\frac{d}{dt}\iint_\Sigma E_n\, d\sigma,$$

где в левой части символы E_s и H_s обозначают проекции векторов $\boldsymbol{E}$ и $\boldsymbol{H}$ на направление касательной к кривой Γ в сторону ее обхода, а направление нормального вектора $\boldsymbol{n}$ на Σ образует с направлением обхода границы Γ правый винт.

Содержание последних двух интегральных формул можно выразить словесно так: криволинейный интеграл напряженности электрического (магнитного) поля вдоль замкнутой кривой пропорционален скорости изменения магнитного (электрического) потока через поверхность, натянутую на эту кривую, причем коэффициент пропорциональности равен $-\frac{1}{c}\left(\frac{1}{c}\right)$.

4) В заключение установим связь между уравнениями Максвелла и волновым уравнением. Из уравнений (M) можно исключить вектор-функцию $\boldsymbol{H}$. Для этой цели дифференцируем второе уравнение по t, а затем подставляем $\frac{\partial \boldsymbol{H}}{\partial t}$ из первого уравнения. Получится

$$-\operatorname{rot}\operatorname{rot}\boldsymbol{E}=\frac{1}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2}.$$

Так как $\operatorname{rot}\operatorname{rot}\boldsymbol{u}=\operatorname{grad}\operatorname{div}\boldsymbol{u}-\nabla^2\boldsymbol{u}$ (см. конец гл. II), а в нашем случае $\operatorname{div}\boldsymbol{E}=0$, то окончательно имеем

$$\nabla^2\boldsymbol{E}=\frac{1}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2}.$$

Аналогично можно из уравнений (M) исключить $\boldsymbol{E}$, причем для $\boldsymbol{H}$ получится уравнение

$$\nabla^2\boldsymbol{H}=\frac{1}{c^2}\frac{\partial^2\boldsymbol{H}}{\partial t^2}.$$

Таким образом, обе вектор-функции $\boldsymbol{E}$ и $\boldsymbol{H}$ удовлетворяют волновому уравнению

$$\nabla^2\boldsymbol{u}=\frac{1}{c^2}\boldsymbol{u}_{tt};$$

это значит, что любая координата каждого из векторов $\boldsymbol{E}$ и $\boldsymbol{H}$ удовлетворяет волновому уравнению.

Упражнения

1. Проинтегрировать следующие дифференциальные уравнения с частными производными:

а) $u_{xy}=0$; б) $u_{xyz}=0$.

2*. Решить уравнение $u_{xx}+5u_{xy}+6u_{yy}=e^{x+y}$ путем приведения его к виду $u_{\xi\eta}=a(\xi,\eta)$.

3. Вывести дифференциальное уравнение с частными производными для двупараметрического семейства шаровых поверхностей

$$z^2=1-(x-a)^2-(y-b)^2.$$

4*. Обозначим через $u(x, t)$ дважды непрерывно дифференцируемое решение волнового уравнения

$$u_{xx}=\frac{1}{a^2}u_{tt} \qquad (a>0).$$

Пусть $\varphi(x)$ — заданная дважды непрерывно дифференцируемая функция, удовлетворяющая условиям

$$\varphi(0)=\varphi'(0)=\varphi''(0)=0.$$

Найти для $x \geqslant 0$ и $t \geqslant 0$ то решение $u(x, t)$, которое определяется краевым условием

$$u(0, t)=\varphi(t) \quad \text{при} \quad t \geqslant 0$$

и начальными условиями

$$u(x, 0)=u_t(x, 0)=0 \quad \text{при} \quad x \geqslant 0.$$

5. Найти решение уравнения

$$u_{xy}=u,$$

удовлетворяющее условиям $u(x, 0)=u(0, y)=1$, в виде степенного ряда.

6. а) Для дифференциального уравнения

$$u_x^2+u_y^2=1$$

найти частные решения вида $u=f(x)+g(y)$.

б) Найти частные решения уравнения

$$u_x u_y=1$$

вида $u=f(x)+g(y)$ и вида $u=f(x)\,g(y)$.

7*. Доказать, что если $z=u(x, y; a, b)$ является решением дифференциального уравнения с частными производными

$$F(x, y, z, z_x, z_y)=0,$$

зависящим от двух параметров a и b, то огибающая любого однопараметрического семейства решений, выделенного из $z=u(x, y; a, b)$, тоже является решением.

8. Воспользоваться результатом упр. 7 для нахождения новых решений уравнения из 6б, полагая $b=ka$ в его решении $u=ax+\frac{1}{a}y+b$ (где k — постоянная).

9. Найти решения уравнения

$$\frac{\partial^2 z}{\partial t^2}=a^2\frac{\partial^2 z}{\partial x^2},$$

удовлетворяющие также уравнению

$$\left(\frac{\partial z}{\partial t}\right)^2=a^2\left(\frac{\partial z}{\partial x}\right)^2.$$

10. Доказать, что если K есть однородная функция от x, y и z, то уравнение

$$\frac{\partial}{\partial x}\left(K\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(K\frac{\partial u}{\partial y}\right)+\frac{\partial}{\partial z}\left(K\frac{\partial u}{\partial z}\right)=0$$

имеет решение, являющееся степенью выражения $(x^2+y^2+z^2)$.

ГЛАВА VII
ЭЛЕМЕНТЫ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ

§ 1. Введение

1. Постановка задачи. Вспомним теорию обычных максимумов и минимумов дифференцируемой функции $f(x_1, x_2, \ldots, x_n)$ от n независимых переменных. Необходимое условие (гл. III, § 6, n° 2) для существования экстремума такой функции в некоторой области изменения независимых переменных записывается так:

$$df = 0 \quad \text{или} \quad \operatorname{grad} f = 0 \quad \text{или} \quad f_{x_i} = 0 \qquad (i = 1, 2, \ldots, n).$$

Эти уравнения выражают *стационарный характер* функции f в соответствующей точке. Вопрос о том, являются ли эти стационарные точки действительно точками максимума или минимума, может быть решен только после дальнейшего исследования. В противоположность необходимым условиям, которые являются *уравнениями*, *достаточные условия* экстремума выражаются *неравенствами*.

Вариационное исчисление тоже занимается нахождением экстремумов (и вообще стационарных значений). Однако здесь мы встречаемся с совершенно новой ситуацией. Дело в том, что функции, об экстремумах которых теперь пойдет речь, будут зависеть уже не от одной независимой переменной или от конечного числа независимых переменных; эти своеобразные функции, называемые *функционалами*, зависят от выбора одной или нескольких *функций*. Это значит, что для определения функционала требуется знание поведения одной или нескольких функций или кривых (или, порою, поверхностей); те функции, от выбора которых зависит функционал, можно назвать *аргументными функциями* или *функциональными аргументами*.

Всеобщее внимание к задачам этого типа привлек в 1696 г. Иоанн Бернулли постановкой *задачи о брахистохроне*.

В вертикальной плоскости xy (ось y направлена вертикально вниз) даны две точки $A(x_0, y_0)$ и $B(x_1, y_1)$, причем $x_1 > x_0$, $y_1 > y_0$. Требуется соединить эти точки такой гладкой кривой $y = \varphi(x)$, что материальной точке, скользящей без трения по этой кривой, под действием одной только силы тяжести понадобится на путь от A до B наименьшее возможное время.

Математическая формулировка задачи основывается на физическом положении, что при движении по кривой $y = \varphi(x)$ скорость $\frac{ds}{dt}$

(s — длина дуги кривой) равна $\sqrt{2g(y-y_0)}$, т. е. пропорциональна корню квадратному из высоты падения. Поэтому время, затраченное на путь от A до B, будет

$$T=\int_{x_0}^{x_1}\frac{dt}{ds}\,\frac{ds}{dx}\,dx=\frac{1}{\sqrt{2g}}\int_{x_0}^{x_1}\frac{\sqrt{1+y'^2}}{\sqrt{y-y_0}}\,dx$$

(ср. т. I, стр. 343). Если опустить несущественный множитель $\sqrt{2g}$ и положить $y_0=0$ (что не влечет за собой утраты общности), то получится следующая постановка задачи.

Среди всех непрерывно дифференцируемых функций $y=\varphi(x)$, $y\geqslant 0$, для которых $\varphi(x_0)=0$, $\varphi(x_1)=y_1$, найти такую, которая сообщает интегралу

$$I\{\varphi\}=\int_{x_0}^{x_1}\sqrt{\frac{1+y'^2}{y}}\,dx$$

наименьшее возможное значение.

Ниже (§ 2, в конце n° 3) мы решим эту задачу и получим ответ, что кривая $y=\varphi(x)$ должна быть *циклоидой*. Этот результат весьма поразил современников Бернулли. Здесь мы желаем только подчеркнуть, что задача Бернулли и элементарные задачи на максимумы и минимумы являются совершенно различными проблемами. Выражение $I\{\varphi\}$ зависит от всего поведения функции φ; его значений нельзя определить, устанавливая значения конечного числа независимых переменных, т. е. это выражение нельзя рассматривать как функцию в обычном смысле. Это функционал I, зависящий от функции $\varphi(x)$, и характер этого выражения как функционала мы и отмечаем фигурными скобками.

А вот еще одна задача того же типа.

Две точки $A(x_0,\ y_0)$ и $B(x_1,\ y_1)$ (где $x_1>x_0$, $y_0>0$, $y_1>0$) требуется соединить кривой $y=\varphi(x)$, лежащей выше оси x и обладающей тем свойством, что поверхность, образуемая вращением этой кривой вокруг оси x, имеет *наименьшую возможную площадь*. Впоследствии мы узнаем, что решением является *цепная линия*.

Пользуясь выражением, полученным в т. I, стр. 329, для площади поверхности вращения, и опуская несущественный множитель 2π, приходим к следующей математической постановке задачи.

Среди всех непрерывно дифференцируемых функций $y=\varphi(x)$, для которых $\varphi(x_0)=y_0$, $\varphi(x_1)=y_1$, $\varphi(x)>0$, найти такую, которая сообщает интегралу

$$I\{\varphi\}=\int_{x_0}^{x_1}y\,\sqrt{1+y'^2}\,dx$$

наименьшее возможное значение.

К тому же типу принадлежит и элементарная геометрическая задача о нахождении кратчайшей линии, соединяющей две точки A и B на плоскости. Действительно, аналитическое содержание задачи состоит в том, что требуется найти такие две функции $x(t)$ и $y(t)$ параметра t в интервале $t_0 \leqslant t \leqslant t_1$, для которых заданы значения $x(t_0) = x_0$, $x(t_1) = x_1$ и $y(t_0) = y_0$, $y(t_1) = y_1$ и которые обладают тем свойством, что интеграл

$$\int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2}\, dt \qquad \left(\dot{x} = \frac{dx}{dt},\ \dot{y} = \frac{dy}{dt}\right)$$

принимает для них наименьшее возможное значение. Решением является, конечно, прямая линия.

В отличие от этой задачи, о нахождении кратчайшей линии, соединяющей две точки на плоскости, соответствующая задача о нахождении *геодезических, т. е. кратчайших линий на заданной поверхности* $G(x, y, z) = 0$, не является тривиальной. Задача об определении геодезических линий состоит в следующем: требуется соединить две точки (x_0, y_0, z_0) и (x_1, y_1, z_1) заданной поверхности кратчайшей возможной кривой, лежащей на этой поверхности.

На аналитическом языке задача ставится так. Среди всех троек функций $x(t)$, $y(t)$, $z(t)$ параметра t, удовлетворяющих уравнению

$$G(x, y, z) = 0$$

тождественно относительно t, причем

$$x(t_0) = x_0, \quad y(t_0) = y_0, \quad z(t_0) = z_0; \qquad x(t_1) = x_1, \quad y(t_1) = y_1, \quad z(t_1) = z_1,$$

найти такую тройку функций, которая сообщает интегралу

$$\int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt$$

наименьшее возможное значение.

К тому же типу задач принадлежит также *изопериметрическая задача* о нахождении замкнутой кривой заданной длины, заключающей *наибольшую возможную площадь*; эта задача была уже рассмотрена в гл. III, Дополнения, § 5, и мы там доказали, что искомой кривой является окружность. [Буквальный смысл слов изопериметрическая задача — задача с заданным (постоянным) периметром.]

Доказательство, данное в гл. V, Дополнения, § 5, приложимо только к выпуклым кривым; однако следующее рассуждение позволяет распространить результат непосредственно на любую кривую. Рассмотрим выпуклую кривую K, «охватывающую» кривую C (ср. гл. II, Дополнения, стр. 121, упр. 2), т. е. выпуклую кривую наименьшей площади, заключающую в себе внутреннюю область кривой C. Эта кривая K состоит из выпуклых дуг кривой C и из отрезков касательных к C, которые касаются этой кривой C в двух точках и, так сказать, срезают вогнутые части кривой C. Очевидно, что площадь «охватывающей» кривой K превосходит площадь кривой C, если последняя не выпукла, а периметр кривой K, напротив, меньше периметра C. Заставим теперь K равномерно расширяться, так что она все время сохраняет свою форму, пока из нее не получится кривая K', имеющая заданный периметр. Тогда K' будет кривой, имеющей тот же периметр, что и C, но заключающей бо́льшую площадь. Отсюда вытекает, что в изопериметрической задаче о нахождении замкнутой кривой, содержащей наибольшую площадь, можно с самого начала ограничиться рассмотрением только выпуклых кривых.

Простейшая общая постановка того типа задач, о котором здесь шла речь, такова:

Дана функция $F(x, \varphi, \varphi')$ от трех аргументов x, φ и φ', которая непрерывна и имеет непрерывные производные первого и второго порядка в рассматриваемой области своих аргументов. Если в этой функции F заменить φ функцией $y=\varphi(x)$, а φ' — ее производной $y'=\varphi'(x)$, то F станет функцией от x, а интеграл

$$I\{\varphi\}=\int\limits_{x_0}^{x_1} F(x, y, y')\,dx$$

будет определенным числом, зависящим от всего поведения функции $y=\varphi(x)$, т. е. $I\{\varphi\}$ является «функцией от функции $\varphi(x)$» или функционалом. Так вот основная задача вариационного исчисления формулируется так:

Среди всех функций $\varphi(x)$, определенных и непрерывных в интервале $x_0 \leqslant x \leqslant x_1$, имеющих в этом интервале непрерывные первую и вторую производные и удовлетворяющих краевым условиям $\varphi(x_0)=y_0$ и $\varphi(x_1)=y_1$, найти такую, для которой интеграл $I\{\varphi\}$ имеет наименьшее возможное значение (или наибольшее возможное значение).

При рассмотрении этой задачи весьма существенным пунктом является *класс допустимых функций* $\varphi(x)$. Сама задача требует лишь, чтобы после подстановки $\varphi(x)$ подынтегральная функция F была кусочно непрерывной функцией от x, а это обеспечено, если производная $\varphi'(x)$ кусочно непрерывна. Мы же сделали «условия допустимости» более строгими, потребовав, чтобы первые и даже вторые производные были непрерывны. Тем самым поле, в котором надлежит искать максимум или минимум, конечно, суживается. Мы увидим, однако, что это ограничение в действительности не повлияет на искомое решение, т. е. функция, решающая задачу в более широком поле, всегда найдется в суженном, ограниченном поле функций, имеющих непрерывные первую и вторую производные.

Задачи описанного типа встречаются очень часто в геометрии и физике. Мы отметим здесь лишь один пример. Основной принцип геометрической оптики можно формулировать как вариационную задачу этого типа. Рассмотрим луч света в плоскости xy и предположим, что скорость света является заданной функцией $v=v(x, y, y')$, зависящей от точки (x, y) и от направления y', где $y=\varphi(x)$ есть уравнение светового пути, а $y'=\varphi'(x)$ — соответствующая производная.

[Это значит, что свет распространяется в неоднородной и не изотропной среде.] Тогда *принцип Ферма́ о наименьшем времени распространения света* формулируется так:

Действительный путь, по которому свет проходит от данной точки A до данной точки B, всегда таков, что время, затраченное на него, меньше того времени, которое потребовалось бы свету на преодоление любого другого пути от A до B.

Обозначим через t время, а через s переменную длину дуги любой кривой $y=\varphi(x)$, соединяющей точки A и B. Тогда время, которое требуется свету для прохождения дуги этой кривой от A до B, выражается интегралом

$$I\{\varphi\}=\int_{x_0}^{x_1}\frac{ds}{v}=\int_{x_0}^{x_1}\frac{\sqrt{1+y'^2}}{v(x, y, y')}\,dx.$$

Для того чтобы по принципу Ферма́ определить действительный путь света, надлежит решить задачу о нахождении функции $y=\varphi(x)$, для которой этот интеграл имеет наименьшее возможное значение.

Ясно, что в этом виде оптическая задача действительно равносильна той общей задаче, которая была поставлена выше, причем фигурирующая там функция F выражается через скорость света v так: $F=\dfrac{\sqrt{1+y'^2}}{v(x, y, y')}$. Впрочем, в большинстве задач оптики скорость света v не зависит от направления и является только функцией точки: $v=v(x, y)$ [Среда — изотропная, но *не* однородная.]

2. Необходимые условия экстремума. Поставим себе целью найти необходимые условия того, чтобы функция $u=\varphi(x)$ могла дать максимум или минимум указанного выше интеграла $I\{\varphi\}$ или, пользуясь общим термином, экстремум этого интеграла. Мы применим для этого метод, аналогичный тому, которым мы пользовались в элементарной задаче об экстремуме функции одной или нескольких переменных. Предположим, что $y=\varphi=u(x)$ есть уже найденная искомая функция. Тогда надо как-то выразить тот факт, что (в случае минимума) значение интеграла I должно *возрасти*, если заменить u другой функцией из класса допустимых функций φ. Кроме того, поскольку нас интересует получение только необходимых условий, мы можем здесь ограничиться рассмотрением функций φ, близких к u, т. е. таких функций φ, для которых $|\varphi-u|$ остается между заранее предписанными границами.

Представим себе функцию $u(x)$ членом однопараметрического семейства функций с параметром ε, которое мы построим следующим образом. Возьмем какую-либо функцию $\eta(x)$, обращающуюся в нуль на границах нашего интервала, так что $\eta(x_0)=\eta(x_1)=0$, и имеющую непрерывные производные первого и второго порядка в замкнутом интервале $[x_0, x_1]$. С помощью $\eta(x)$ мы и построим семейство функций

$$\varphi(x, \varepsilon) = u(x) + \varepsilon\eta(x).$$

Выражение $\varepsilon\eta(x) = \delta u$ называется *вариацией функции u*. $\Big($Так как $\eta(x) = \frac{\partial \varphi}{\partial \varepsilon}$, то символ δ обозначает дифференциал от $\varphi(x, \varepsilon)$ при $\varepsilon = 0$, когда ε рассматривается как независимое переменное, а x как параметр.$\Big)$ Тогда, если функции u и η выбраны, то интеграл

$$I(u + \varepsilon\eta) = \Phi(\varepsilon) = \int\limits_{x_0}^{x_1} F(x, u + \varepsilon\eta, u' + \varepsilon\eta')\,dx$$

будет функцией от ε, и предположение, что $u(x)$ дает минимум функционала $I\{\varphi\}$, означает, что $\Phi(\varepsilon)$ имеет минимум при $\varepsilon = 0$, так что в качестве необходимых условий имеем равенство

$$\Phi'(0) = 0$$

и затем неравенство

$$\Phi''(0) \geqslant 0.$$

Аналогично, если бы мы искали максимум, то получили бы в качестве необходимых условий, то же самое равенство $\Phi'(0) = 0$ и неравенство $\Phi''(0) \leqslant 0$. Условие $\Phi'(0) = 0$ должно выполняться при любом выборе функции $\eta(x)$, обладающей перечисленными выше свойствами, а в остальном произвольной.

Отбрасывая вопрос о различении между максимумом и минимумом, мы скажем, что если функция $u(x)$ удовлетворяет уравнению $\Phi'(0) = 0$ при любом выборе функции $\eta(x)$, то интеграл $I\{\varphi\}$ *стационарен* при $\varphi = u$. Пользуясь, как и выше, символом δ для обозначения дифференциала от I (как функции от ε) при $\varepsilon = 0$, можно также сказать, что равенство

$$\delta I = \varepsilon\Phi'(0) = 0,$$

если оно удовлетворяется функцией $\varphi = u$ при произвольной функции η, выражает стационарный характер интеграла I. Выражение

$$\varepsilon\Phi'(0) = \varepsilon \left[\frac{d}{d\varepsilon} \int\limits_{x_0}^{x_1} F(x, u + \varepsilon\eta, u' + \varepsilon\eta')\,dx \right]_{\varepsilon = 0}$$

называется *вариацией* или, точнее, *первой вариацией* интеграла I. Поэтому *стационарный характер интеграла и обращение в нуль первой вариации означает совершенно то же самое.* (От термина *вариация* произошло название дисциплины: *вариационное исчисление.*)

Стационарность функционала является *необходимым* условием наличия максимума или минимума, но (как и в теории обычных максимумов и минимумов) отнюдь не является *достаточным* для осуществления одной из этих возможностей. Мы не можем здесь остановиться подробнее на проблеме достаточных условий и в дальнейшем ограничимся проблемой стационарности.

Нашей главной целью будет преобразовать условие стационарности $\Phi'(0)=0$ таким образом, чтобы оно стало условием, налагаемым на одну лишь функцию u и уже не содержащим произвольной функции η.

Упражнения

1. В связи с задачей о брахистохроне (см. n° 1, начало) вычислить время падения, если точки A и B соединены прямой линией.

2. Пусть положение частицы, движущейся в трехмерном пространстве, задается сферическими координатами (r, θ, φ), и пусть ее скорость $v=\frac{1}{f(r)}$. Сколько времени потребуется частице для прохождения дуги кривой между ее точками A и B, если эта кривая задана в сферических координатах параметрическими уравнениями $r=r(\sigma)$, $\theta=\theta(\sigma)$, $\varphi=\varphi(\sigma)$, где σ — параметр.

§ 2. Дифференциальное уравнение Эйлера для простейшего случая

1. Вывод дифференциального уравнения Эйлера. Основной критерий вариационного исчисления формулируется так:

Для того чтобы функционал

$$I\{\varphi\}=\int_{x_0}^{x_1} F(x, \varphi, \varphi')\,dx$$

имел стационарное значение при $\varphi=u$, необходимо и достаточно, чтобы $\varphi=u$ была допустимой функцией, удовлетворяющей дифференциальному уравнению Эйлера

$$L[u]\equiv F_u-\frac{d}{dx}F_{u'}=0$$

или, в развернутом виде,

$$-L[u]\equiv F_{u'u'}u''+F_{uu'}u'+F_{xu'}-F_u=0.$$

Для доказательства заметим, что выражение

$$\Phi(\varepsilon)=\int_{x_0}^{x_1} F(x, u+\varepsilon\eta, u'+\varepsilon\eta')\,dx$$

можно дифференцировать по ε под знаком интеграла (ср. гл. IV, § 1, n° 2), если, после дифференцирования подынтегральной функции, под знаком интеграла получится непрерывная или по крайней мере кусочно непрерывная функция от x. В нашем случае, полагая $u+\varepsilon\eta=y$ и дифференцируя, получим под знаком интеграла выражение $\eta F_y + \eta' F_{y'}$, которое, в силу сделанных относительно F, u и η допущений, удовлетворяет требуемому условию. Поэтому мы сразу получаем

$$\Phi'(0) = \int_{x_0}^{x_1} (\eta F_u + \eta' F_{u'})\,dx, \qquad \text{где} \qquad F = F(x,\ u,\ u').$$

Для наших последующих целей (ср. следующую страницу), заметим, что при выводе этого уравнения мы воспользовались только лишь непрерывностью функций u и η и кусочной непрерывностью их первых производных. В полученном уравнении произвольная функция фигурирует в двояком виде, а именно как η и в виде η'. Но от η' легко освободиться, если применить ко второму члену правило интегрирования произведения:

$$\int_{x_0}^{x_1} \eta' F_{u'}\,dx = \eta F_{u'}\Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta\left(\frac{d}{dx}F_{u'}\right)dx = -\int_{x_0}^{x_1} \eta\left(\frac{d}{dx}F_{u'}\right)dx,$$

так как, по условию, $\eta(x_0) = \eta(x_1) = 0$. При этом интегрировании по частям предполагается, что выражение $\frac{d}{dx}F_{u'}$ может быть составлено, но это действительно так, ибо мы с самого начала предположили непрерывность вторых производных. Поэтому, введя под знаком интеграла сокращенное обозначение

$$L[u] = F_u - \frac{d}{dx}F_{u'},$$

получим уравнение

$$\Phi'(0) = \int_{x_0}^{x_1} \eta L[u]\,dx = 0.$$

Так вот функция u должна удовлетворять этому уравнению при любом выборе функции η, обладающей требуемыми свойствами, а в остальном совершенно произвольной. На основании формулируемой ниже леммы I отсюда будет вытекать, что

$$L[u] = 0.$$

Лемма I. *Если функция $C(x)$, непрерывная в рассматриваемом интервале, удовлетворяет соотношению*

$$\int_{x_0}^{x_1} \eta(x)\,C(x)\,dx = 0,$$

где $\eta(x)$ есть любая дважды непрерывно дифференцируемая функция, причем $\eta(x_0) = \eta(x_1) = 0$, то $C(x) = 0$ при всех значениях x из нашего интервала.

Доказательство этой леммы будет дано в n° 2.

Есть, однако, возможность получить требуемое условие другим путем, освобождаясь в уравнении

$$\int_{x_0}^{x_1} (\eta F_u + \eta' F_{u'})\, dx = 0$$

от члена с η. Для этой цели надо применить правило интегрирования произведения к первому члену подынтегральной функции; введя для краткости обозначения $F_{u'} = A$, $F_u = B'$ и приняв во внимание граничное условие для η: $\eta(x_0) = \eta(x_1) = 0$, получим

$$\int_{x_0}^{x_1} \eta F_u\, dx = \int_{x_0}^{x_1} \eta B'\, dx = \eta B \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta' B\, dx = - \int_{x_0}^{x_1} \eta' B\, dx.$$

Если положить $\zeta = \eta'$, то наше условие примет следующий вид:

$$\int_{x_0}^{x_1} \zeta (A - B)\, dx = 0.$$

При этом методе не требуется делать дальнейших допущений о вторых производных от η и от u. Напротив, достаточно предположить, что функции φ (или u) и η непрерывны и имеют кусочно непрерывные первые производные. Правда, теперь наше интегральное условие должно выполняться не при любой кусочно непрерывной функции ζ, но только лишь при таких функциях ζ, которые являются производными от функций $\eta(x)$, удовлетворяющих нашим условиям. Однако если $\zeta(x)$ есть любая кусочно непрерывная функция, удовлетворяющая соотношению

$$\int_{x_0}^{x_1} \zeta(x)\, dx = 0,$$

а в остальном произвольная, то можно положить

$$\eta = \int_{x_0}^{x} \zeta(t)\, dt,$$

и тем самым уже построена допустимая функция $\eta(x)$, ибо $\eta' = \zeta$ и $\eta(x_0) = \eta(x_1) = 0$. Таким образом, получен следующий результат:

Необходимым условием стационарности функционала является

$$\int_{x_0}^{x_1} \zeta (A - B)\, dx = 0,$$

где ζ есть произвольная кусочно непрерывная функция, удовлетворяющая условию $\int_{x_0}^{x_1} \zeta\, dx = 0$.

Но теперь понадобится другая

Лемма II. *Если кусочно непрерывная функция $S(x)$ удовлетворяет условию*

$$\int_{x_0}^{x_1} \zeta S\, dx = 0$$

для всех функций $\zeta(x)$, *кусочно непрерывных в интервале* $[x_0, x_1]$ *и связанных соотношением*

$$\int_{x_0}^{x_1} \zeta\, dx = 0,$$

то $S(x)$ *сводится к постоянной* c.

Эта лемма тоже будет доказана в следующем n° 2. Но уже теперь можно с ее помощью вывести из условия стационарности (подставляя вместо A и B их выражения), что

$$\int_{x_0}^{x} F_u\, dx + c = F_{u'}.$$

Левая часть дифференцируема по x, и ее производная равна F_u; следовательно, и правая часть дифференцируема по x и, стало быть, выражение $\frac{d}{dx} F_{u'}$ существует для предполагаемого решения u, причем это решение удовлетворяет уравнению

$$F_u = \frac{d}{dx} F_{u'}.$$

Таким образом, уравнение Эйлера остается условием стационарности функционала и в том случае, если с самого начала расширить класс допустимых функций $\varphi(x)$, требуя только лишь кусочной непрерывности первой производной $\varphi'(x)$.

Уравнение Эйлера является *обыкновенным дифференциальным уравнением второго порядка*. Его решения (точнее, его интегральные кривые) называются *экстремалями* задачи нахождения экстремумов функционала. Для решения этой задачи надо прежде всего выделить из всех экстремалей такие, которые удовлетворяют заданным краевым условиям. Только на этих экстремалях может осуществиться экстремум функционала.

Если выполняется «условие Лежандра»

$$F_{u'u'} \neq 0,$$

то дифференциальное уравнение Эйлера можно привести к виду $u'' = f(x, u, u')$, где правая часть является известным выражением, содержащим x, u, u'.

2. Доказательства обеих лемм. Докажем теперь те две леммы, которыми мы воспользовались выше.

Для доказательства леммы I предположим, что в некоторой точке $x = \xi$ функция $C(x)$ не обращается в нуль и имеет, скажем, положительное значение. Тогда, в силу непрерывности этой функции, можно выделить в интервале $[x_0, x_1]$ такую окрестность точки ξ:

$$\xi - a \leqslant x \leqslant \xi + a,$$

что $C(x)$ остается положительной во всей этой окрестности. Функцию $\eta(x)$ мы определим теперь в интервале $[x_0, x_1]$ так:

$$\eta(x)=\begin{cases} (x-\xi+a)^4(x-\xi-a)^4=\{(x-\xi)^2-a^2\}^4 & \text{в указанной окрестности,} \\ 0 & \text{вне этой окрестности.} \end{cases}$$

Тогда интеграл $\int\limits_{x_0}^{x_1}\eta C(x)\,dx$ не может равняться нулю. (Интеграл от непрерывной неотрицательной функции вообще положителен и обращается в нуль лишь в том случае, если подынтегральная функция тождественно равна нулю; это вытекает непосредственно из определения интеграла как предела суммы.) Так как это противоречит условию леммы, то $C(\xi)\leqslant 0$. Аналогичным рассуждением устанавливаем, что $C(\xi)\geqslant 0$. Следовательно, $C(\xi)=0$ при всех значениях ξ внутри нашего интервала, что лемма I и утверждает. [$C(x)$ обращается в нуль и на границах интервала; для доказательства потребуется лишь небольшое изменение в построении функции $\eta(x)$.]

Для доказательства леммы II заметим, что наше допущение относительно $\zeta(x)$ сразу приводит к соотношениям

$$\int\limits_{x_0}^{x_1}\zeta(x)\,dx=0 \quad \text{и} \quad \int\limits_{x_0}^{x_1}\zeta(x)\,\{S(x)-c\}\,dx=0, \tag{*}$$

где c — произвольная постоянная. Выберем теперь c таким образом, чтобы $S(x)-c$ принадлежала к классу допустимых функций $\zeta(x)$, т. е. определим c равенством

$$\int\limits_{x_0}^{x_1}\{S(x)-c\}\,dx=\int\limits_{x_0}^{x_1}S(x)\,dx-c\,(x_1-x_0)=0.$$

Подставляя это значение c во второе уравнение (*) и беря в нем $\zeta=S(x)-c$, получим сразу

$$\int\limits_{x_0}^{x_1}\{S(x)-c\}^2\,dx=0.$$

Так как, согласно условию леммы II, подынтегральная функция непрерывна или по крайней мере кусочно непрерывна, то обязательно

$$S(x)-c=0,$$

что и требовалось доказать.

3. Замечания по поводу интегрирования дифференциального уравнения Эйлера. Примеры. Для того чтобы решить проблему экстремума функционала, надо найти частное решение u дифференциального уравнения Эйлера для интервала $x_0\leqslant x\leqslant x_1$, принимающее в его конечных точках заданные краевые значения y_0 и y_1. Общее решение уравнения Эйлера как дифференциального уравнения второго порядка содержит две постоянных интегрирования, а краевые условия дают два уравнения, которым эти постоянные должны удовлетворять; поэтому вообще должно быть возможно найти частное решение, удовлетворяющее краевым условиям.

Однако выразить общее решение уравнения Эйлера через элементарные функции или с помощью квадратур, как правило, невозможно. В общем случае приходится довольствоваться установлением того факта, что вариационная задача приводится к интегрированию дифференциального уравнения. С другой стороны, в важных частных случаях и фактически в большинстве классических примеров уравнение Эйлера решается в квадратурах.

Первый случай. Подынтегральная функция F не содержит явно производной $y'=\varphi'$, так что $F=F(\varphi, x)$. Тогда уравнение Эйлера есть просто $F_u(u, x)=0$, т. е. не является вовсе дифференциальным уравнением, а лишь неявным определением решения $y=u(x)$. Здесь нет постоянных интегрирования, и не может быть вопроса о таком их выборе, чтобы удовлетворялись краевые условия. [В этом случае при произвольно заданных краевых условиях вариационная задача может не иметь решения; она имеет решения при том и только при том условии, если *краевые значения* удовлетворяют уравнению $F_u(u, x)=0$.]

Второй случай. F не содержит явно функции $y=\varphi(x)$, так что $F=F(y', x)$. Тогда дифференциальное уравнение Эйлера будет $\frac{d}{dx}F_{u'}(u', x)=0$, откуда сразу получается так называемый «первый интеграл»

$$F_{u'}(u', x)=c,$$

где c — постоянная интегрирования. Если можно из этого уравнения выразить u' через x и c в явном виде:

$$u'=f(x, c),$$

то простым интегрированием (квадратурой) получим

$$u=\int_{x_0}^{x} f(\xi, c)\,d\xi+a,$$

т. е. u выражена как функция от x, причем к прежней произвольной постоянной c присоединилась еще одна постоянная интегрирования a. Стало быть, в этом случае общее решение дифференциального уравнения Эйлера получено с помощью квадратур.

Третий случай — самый важный для конкретных задач и для приложений. Подынтегральная функция F не содержит явно независимой переменной x, так что $F=F(y, y')$. В этом случае $F_{xy'}=0$, и уравнение Эйлера будет

$$L[u]=F_u-u'F_{uu'}-u''F_{u'u'}=0.$$

Помножив левую часть на u', получим

$$u'L[u]=u'F_u-u'^2F_{uu'}-u'u''F_{u'u'}=\frac{d}{dx}\{F(u,\ u')-u'F_{u'}(u,\ u')\},$$

что нетрудно проверить. Итак, после умножения на u', уравнение Эйлера принимает вид

$$\frac{d}{dx}\{F(u,\ u')-u'F_{u'}(u,\ u')\}=0,$$

откуда интегрированием получаем «первый интеграл»

$$F(u,\ u')-u'F_{u'}(u,\ u')=c,$$

где c — произвольная постоянная. Если из этого уравнения u' выражается явно: $u'=f(u,\ c)$, то переменные отделяются:

$$dx=\frac{du}{f(u,\ c)},$$

и интегрирование дает $x=g(u,\ c)+\alpha$, где α — вторая постоянная интегрирования, т. е. x выражен как функция от u, c и α. Это уравнение дает общее решение дифференциального уравнения Эйлера в неявном виде, и получено оно с помощью двух квадратур, из которых одна выполнена, а другая намечена.

Применим теперь эти методы к решению нескольких задач-примеров.

1) Задачи типа $F=\psi(y)\sqrt{1+y'^2}$. Общий класс таких задач относится к третьему случаю. Для экстремалей $y=u$ имеем первый интеграл

$$\psi(u)\sqrt{1+u'^2}-\psi(u)\frac{u'^2}{\sqrt{1+u'^2}}=c \quad \text{или} \quad \frac{\psi(u)}{\sqrt{1+u'^2}}=c,$$

откуда $u'^2=\left(\frac{\psi(u)}{c}\right)^2-1$. Отделение переменных дает

$$dx=\frac{du}{\pm\sqrt{\left(\frac{\psi(u)}{c}\right)^2-1}},$$

откуда интегрированием получим

$$x-b=\pm\int\frac{du}{\sqrt{\left(\frac{\psi(u)}{c}\right)^2-1}}.$$

Это общее решение уравнения Эйлера (в неявном виде) с постоянными интегрирования c и b.

2) Поверхность вращения наименьшей площади. Это задача рассмотренного только что типа, причем $\psi(y)=y$. Полученный выше интеграл дает

$$x-b=\pm\int\frac{du}{\sqrt{\left(\frac{u}{c}\right)^2-1}}=c\operatorname{arch}\frac{u}{c}.$$

Итак, общее решение есть

$$y=u=c\operatorname{ch}\frac{x-b}{c}.$$

Следовательно, решение задачи о нахождении кривой, при вращении которой вокруг оси x получается поверхность вращения стационарной площади, может дать только *цепная линия*. Однако необходимым условием существования такой кривой, дающей поверхность вращения стационарной площади, является возможность соединить две заданные точки A и B дугой цепной линии из полученного семейства, в которой $y>0$. Вопрос о том, действительно ли осуществляет эта цепная линия минимум площади поверхности вращения, мы не имеем возможности рассматривать здесь.

3) Брахистохрона (§ 1, n° 1). В этой задаче $\psi=\frac{1}{\sqrt{y}}$. Общее решение будет

$$x-b=\int\frac{du}{\sqrt{\frac{1}{c^2u}-1}}.$$

Полагая $1/c^2=k$ и совершая последовательно замены переменных $u=k\tau$ и $\tau=\sin^2\frac{\theta}{2}$, приводим это общее решение к виду

$$x-b=k\int\sqrt{\frac{\tau}{1-\tau}}\,d\tau=\frac{k}{2}\int(1-\cos\theta)\,d\theta,$$

откуда

$$x-b=\frac{k}{2}(\theta-\sin\theta),$$

$$y=u=k\sin^2\frac{\theta}{2}=\frac{k}{2}(1-\cos\theta).$$

Следовательно, брахистохроной является обыкновенная циклоида, точки возврата которой лежат на оси x (ср. т. I, стр. 304—305), только теперь положительная ось y направлена вниз).

Упражнения

Найти экстремали функционалов вида $\int\limits_{x_0}^{x_1} F(x, y, y')\,dx$ со следующими подынтегральными функциями:

1. $F=\sqrt{y(1+y'^2)}$. 2. $F=\frac{1}{y}\sqrt{1+y'^2}$. 3. $F=y\sqrt{1-y'^2}$.

4. Найти экстремали для случая $F=x^n y'^2$ и доказать, что при $n \geqslant 1$ две точки, лежащие по разные стороны от оси y, не могут быть соединены экстремалью.

5. Найти экстремали для случая $F=y^n(y')^m$, где n и m — четные целые числа.

6. Найти экстремали для случая $F=ay'^2+2byy'+cy^2$, где a, b, c — заданные непрерывно дифференцируемые функции от x. Доказать, что уравнение Эйлера будет линейным дифференциальным уравнением второго порядка. Выяснить: почему в случае постоянного b этот постоянный коэффициент не входит в дифференциальное уравнение?

7. Показать, что при $F=e^x\sqrt{1+y'^2}$ экстремали даются уравнениями $\sin(y-b)=e^{-(x-a)}$ и $y=b$, где a и b — постоянные. Исследовать форму этих кривых и выяснить, как должны быть расположены две точки A и B, если их можно соединить дугой экстремали вида $y=f(x)$.

8. Подынтегральная функция F не содержит явно производной y' (п° 3, первый случай). Вывести уравнение Эйлера $F_y=0$ элементарным способом.

9. Найти функцию $y(x)$, дающую наименьшее значение функционала

$$I\{y\}=\int\limits_0^1 y'^2\,dx$$

при краевых условиях: а) $y(0)=y(1)=0$, б) $y(0)=0$, $y(1)=1$.

4. Случай, когда уравнение Эйлера обращается в тождество. Дифференциальное уравнение Эйлера для функционала

$$I\{y\}=\int\limits_{x_0}^{x_1} F(x, y, y')\,dx$$

может выродиться в ничего не говорящее тождество, т. е. в соотношение, которому удовлетворяет всякая допустимая функция $y=\varphi(x)$. В этом вырожденном случае выражение Эйлера

$$L[y]=F_y-F_{xy'}-F_{yy'}y'-F_{y'y'}y''$$

должно обращаться в нуль в любой точке x интервала, какую бы функцию $y=\varphi(x)$ из класса допустимых функций в него ни подставить. Но ведь всегда можно найти кривую $y=\varphi(x)$, у которой $y=\varphi$, $y'=\varphi'$ и $y''=\varphi''$ имеют произвольно заданные значения в определенной точке. Поэтому выражение Эйлера должно обращаться в нуль при любой четверке значений x, y, y', y''. Отсюда вытекает, что выражение $F_{y'y'}$, служащее коэффициентом при y'', должно равняться нулю тождественно. Поэтому F должна быть линейной функцией от y', $F=ay'+b$, где a и b зависят только от x и y. Если подставить это выражение в уравнение Эйлера, которое имеет теперь вид тождества

$$F_{yy'}y' + F_{xy'} - F_y \equiv 0,$$

то получим сразу

$$a_y y' + a_x - a_y y' - b_y = 0 \quad \text{или} \quad a_x - b_y = 0$$

тождественно относительно x и y. Другими словами, уравнение Эйлера обращается в тождество в том и только в том случае, если функционал $I\{y\}$ имеет вид

$$I\{y\} = \int_{x_0}^{x_1} \{a(x, y)y' + b(x, y)\}\, dx = \int_{x_0}^{x_1} (a\, dy + b\, dx),$$

в котором a и b удовлетворяют условию интегрируемости (гл. V, § 1, n° 7), т. е. подынтегральное выражение $a\,dy + b\,dx$ является полным дифференциалом.

§ 3. Обобщения

1. Функционалы, зависящие от многих функциональных аргументов. Задача о нахождении экстремумов функционала может быть обобщена на тот случай, когда этот функционал зависит не от одного, а от многих функциональных аргументов (аргументных функций) $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$. Типичную задачу такого рода можно формулировать так:

Дана функция $F(x, \varphi_1, \ldots, \varphi_n, \varphi_1', \ldots, \varphi_n')$, зависящая от $(2n+1)$ символов $x, \varphi_1, \ldots, \varphi_n, \varphi_1', \ldots, \varphi_n'$, непрерывная относительно этих символов и имеющая по ним непрерывные производные первого и второго порядка в области, соответствующей рассматриваемому интервалу $x_0 \leqslant x \leqslant x_1$. Если заменить φ_i функцией $\varphi_i(x)$, имеющей непрерывные производные первого и второго порядка, а φ_i' — ее производной, то F становится функцией от одной переменной x и интеграл

$$I\{\varphi_1, \ldots, \varphi_n\} = \int_{x_0}^{x_1} F(x, \varphi_1, \ldots, \varphi_n, \varphi_1', \ldots, \varphi_n')\, dx$$

по заданному интервалу $x_0 \leqslant x \leqslant x_1$ имеет определенное значение, определяемое выбором функций $\varphi_i(x)$ [т. е. $I\{\varphi_1, \ldots, \varphi_n\}$ есть функционал, зависящий от аргументных функций $\varphi_1(x), \ldots, \varphi_n(x)$].

При этом мы считаем допустимыми все функции $\varphi_i(x)$, обладающие перечисленными выше свойствами непрерывности и принимающие заранее заданные краевые значения $\varphi_i(x_0)$ и $\varphi_i(x_1)$. Иными словами, мы рассматриваем кривые $y_i = \varphi_i(x)$, соединяющие две заданные точки A и B в пространстве $(n+1)$ измерений переменных $x, y_1, y_2, \ldots, y_n$. И вот вариационная задача требует выделения из всех этих систем функций $\varphi_i(x)$ такой системы $y_i = \varphi_i(x) = u_i(x)$, для которой функционал $I\{\varphi_1, \ldots, \varphi_n\}$ имеет экстремум (максимум или минимум).

И здесь мы не можем заняться фактической природой экстремума и ограничимся нахождением таких систем аргументных функций $\varphi_i(x) = u_i(x)$, для которых функционал имеет стационарное значение.

Само определение понятия стационарного значения функционала вводится точно таким же путем, как в § 1, n° 2. Включают систему функций $u_i(x)$ в однопараметрическое семейство функций с параметром ε. Делают это так. Выбирают произвольно n функций $\eta_1(x), \ldots, \eta_n(x)$, обращающихся в нуль при $x = x_0$ и при $x = x_1$, непрерывных в интервале $[x_0, x_1]$ и имеющих в нем непрерывные первые и вторые производные. Затем рассматривают семейство функций $y_i = \varphi_i(x) = u_i(x) + \varepsilon\eta_i(x)$.

Член $\varepsilon\eta_i(x) = \delta u_i$ называется вариацией функции $u_i(x)$. Если подставить выражения для функций $\varphi_i(x)$ построенного семейства в $I\{\varphi_1, \ldots, \varphi_n\}$, то этот функционал обратится в обычную функцию от параметра ε:

$$\Phi(\varepsilon) = \int_{x_0}^{x_1} F(x, u_1 + \varepsilon\eta_1, \ldots, u_n + \varepsilon\eta_n, u_1' + \varepsilon\eta_1', \ldots, u_n' + \varepsilon\eta_n')\,dx.$$

Необходимым условием того, чтобы для $\varphi_i = u_i$, т. е. при $\varepsilon = 0$, этот интеграл имел экстремум, является

$$\Phi'(0) = 0.$$

Точно так же, как и в § 1, n° 2, если $\Phi'(0) = 0$ или, другими словами, если вариация

$$\delta I = \varepsilon\Phi'(0) = 0$$

при любом выборе функций η_i, подчиненных перечисленным выше условиям, то говорят, что функционал I имеет стационарное значение при $\varphi_i = u_i$. Другими словами, стационарный характер интеграла I и обращение в нуль вариации δI означают один и тот же факт.

Теперь желательно установить такие условия стационарности интеграла, которые не содержали бы произвольных функций η_i. Для этого не понадобится никаких новых идей, и цель будет достигнута с помощью следующего рассуждения. Если выбрать $\eta_2, \eta_3, \ldots, \eta_n$ тождественно равными нулю, т. е. если не подвергать изменению функции $u_2, \ldots, u_n$, и, значит, рассматривать лишь первую функцию $\varphi_1(x)$ как подлежащую изменению, то условие $\Phi'(0) = 0$ (согласно § 2, n° 1) эквивалентно дифференциальному уравнению Эйлера

$$F_{u_1} - \frac{d}{dx} F_{u_1'} = 0.$$

Так как функции $u_i(x)$ равноправны, и любую из них можно выделить таким же путем, то получается следующий результат:

Для того чтобы n функций $u_i(x)$ сообщали функционалу $I = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ стационарное значение, необходимо и достаточно, чтобы эти функции удовлетворяли системе уравнений Эйлера

$$F_{u_i} - \frac{d}{dx} F_{u_i'} = 0 \qquad (i = 1, 2, \ldots, n).$$

Это система дифференциальных уравнений второго порядка для n функций $u_i(x)$, причем число уравнений равно числу искомых функций. Всякое решение этой системы дифференциальных уравнений называется *экстремалью вариационной задачи*. Таким образом, задача нахождения стационарных значений функционала приводится к нахождению общего решения этой системы дифференциальных уравнений и к выделению из него такого частного решения, которое удовлетворяет заданным краевым условиям.

Пользуясь леммой II (§ 2, n° 1, конец), можно доказать, что условие стационарности приводится к этой системе уравнений и при более широком классе допустимых функций, от которых требуется только, чтобы они имели кусочно непрерывные первые производные. Однако для начинающего желательно сосредоточить внимание на существенном механизме проблемы, а с этой целью удобнее включить в условия допустимости функций $\varphi_i(x)$ непрерывность вторых производных. Тогда можно записать в развернутом виде

$$\frac{d}{dx} F_{u_i'} = \sum_{k=1}^{n} F_{u_k' u_i'} u_k'' + \sum_{k=1}^{n} F_{u_k u_i'} u_k' + F_{x u_i'}.$$

2. Важный частный случай. Примеры. Найти общее решение системы дифференциальных уравнений Эйлера возможно еще реже, чем в случае одного функционального аргумента из § 2. Лишь в весьма частных случаях удается определить все экстремали в явном виде. Однако, если подынтегральная функция F не содержит явно независимой переменной x, так что $F = F(\varphi_1, \ldots, \varphi_n; \varphi_1', \ldots, \varphi_n')$, то способ, примененный в задаче с одним функциональным аргументом (§ 2, n° 3, третий случай), наводит на мысль, как получить из системы уравнений Эйлера «интегрируемую комбинацию». В этом случае, так как $F_{x u_i'} = 0$, эта система принимает следующий вид:

$$F_{u_i} - \sum_{k=1}^{n} F_{u_k u_i'} u_k' - \sum_{k=1}^{n} F_{u_k' u_i'} u_k'' \qquad (i = 1, 2, \ldots, n).$$

Помножим каждое из этих уравнений на u_i' и просуммируем по i от 1 до n:

$$\sum_{i=1}^{n} F_{u_i} u_i' - \sum_{i=1}^{n} \sum_{k=1}^{n} F_{u_k u_i'} u_k' u_i' - \sum_{i=1}^{n} \sum_{k=1}^{n} F_{u_k' u_i'} u_i' u_k'' = 0.$$

Левая часть этого уравнения является полной производной по x от выражения $F(u_1, \ldots, u_n; u_1', \ldots, u_n') - \sum\limits_{i=1}^{n} u_i' F_{u_i'}$, что нетрудно проверить дифференцированием. Таким образом, последнее уравнение можно записать так:

$$\frac{d}{dx}\left\{F(u_1, \ldots, u_n; u_1', \ldots, u_n') - \sum_{i=1}^{n} u_i' F_{u_i'}\right\} = 0.$$

Интегрируя его по x, получаем первый интеграл системы уравнений Эйлера

$$F(u_1, \ldots, u_n; u_1', \ldots, u_n') - \sum_{i=1}^{n} u_i' F_{u_i'} = c,$$

где c — произвольная постоянная. При подстановке в этот первый интеграл конкретной экстремали, удовлетворяющей краевым условиям, эта постоянная приобретает определенное значение.

Простой пример, результат которого заранее известен, дает задача о нахождении кратчайшего расстояния между двумя точками в трехмерном пространстве. В этой задаче требуется найти такие две функции $y = y(x)$ и $z = z(x)$, которые сообщают интегралу

$$\int_{x_0}^{x_1} \sqrt{1 + y'^2 + z'^2}\, dx$$

наименьшее возможное значение, причем эти функции $y(x)$ и $z(x)$ должны принимать в конечных точках промежутка интегрирования заранее заданные значения. Получаем систему двух дифференциальных уравнений Эйлера

$$\frac{d}{dx}\frac{y'}{\sqrt{1 + y'^2 + z'^2}} = 0, \qquad \frac{d}{dx}\frac{z'}{\sqrt{1 + y'^2 + z'^2}} = 0,$$

из которых после несложных выкладок вытекает, что производные $y'(x)$ и $z'(x)$ должны быть постоянны. Следовательно, экстремали являются прямыми линиями.

Менее тривиальной является задача о *брахистохроне в трехмерном пространстве*. (Положительное направление оси y мы и здесь выбираем по вертикали, в сторону действия силы тяжести.) Здесь требуется найти две функции $y = y(x)$ и $z = z(x)$, делающие стационарным функционал

$$\int_{x_0}^{x_1} \sqrt{\frac{1 + y'^2 + z'^2}{y}}\, dx = \int_{x_0}^{x_1} F(y, y', z')\, dx.$$

Этой задаче соответствует система двух дифференциальных уравнений Эйлера

$$-\frac{1}{2}\sqrt{\frac{1 + y'^2 + z'^2}{y^3}} - \frac{d}{dx}\frac{y'}{\sqrt{y(1 + y'^2 + z'^2)}} = 0,$$

$$\frac{d}{dx}\frac{z'}{\sqrt{y(1 + y'^2 + z'^2)}} = 0.$$

Второе из этих уравнений и первый интеграл, найденный выше в этом же номере в общем виде, дают

$$\frac{z'}{\sqrt{y}}\frac{1}{\sqrt{1+y'^2+z'^2}}=a,$$

$$F-y'F_{y'}-z'F_{z'}\equiv\frac{1}{\sqrt{y}}\frac{1}{\sqrt{1+y'^2+z'^2}}=b,$$

где a и b — произвольные постоянные. Из последних двух уравнений вытекает (в результате почленного деления), что $z'=\frac{a}{b}=k$ тоже постоянно. Поэтому кривая, для которой функционал стационарен, должна лежать в плоскости $z=kx+c$, но тогда из примера 3 на стр. 527 очевидно, что эта кривая есть циклоида. Этот же вывод можно получить из данного выше первого интеграла

$$\frac{1}{\sqrt{y}}\frac{1}{\sqrt{1+k^2+y'^2}}=b.$$

Упражнение

Пользуясь сферической системой координат (r, θ, φ), написать дифференциальные уравнения пути (луча) света для того случая, когда скорость света есть функция только полярного радиуса r (ср. § 1, стр. 520, упр. 2). Показать, что лучи света являются плоскими кривыми.

3. Принцип Гамильтона. Уравнения Лагранжа. Система дифференциальных уравнений Эйлера играет очень важную роль во многих ветвях прикладной математики, особенно в динамике. Дело в том, что движение механической системы, состоящей из конечного числа материальных точек (частиц), можно описать в виде условия, что известный функционал, так называемый *интеграл Гамильтона*, должен иметь стационарное значение. Это положение мы здесь вкратце поясним.

Если положение и конфигурация механической системы определяются n независимыми (обобщенными) координатами $(q_1, q_2, \ldots, q_n)$, то говорят, что система имеет n степеней свободы. Пусть, например, система состоит из одной материальной точки; тогда $n=3$, ибо в качестве координат q_1, q_2, q_3 можно взять три координаты точки в прямоугольной или в сферической системе координат. Пусть теперь система состоит из двух частиц, удерживаемых жесткой связью на расстоянии 1 одна от другой (предполагается, что сама связь не имеет массы); тогда $n=5$, ибо за координаты q_i можно принять три прямоугольные координаты одной из частиц и два числа, определяющих направление прямой, соединяющей обе частицы.

Динамическую систему можно с достаточной общностью описать с помощью двух функций: *кинетической энергии* и *потенциальной энергии*. Если представить себе, что система каким-либо образом движется, то координаты положения q_i будут функциями $q_i(t)$ от времени t, а «обобщенные скорости» будут $\dot{q}_i=\dot{q}_i(t)=\frac{dq_i}{dt}$. Тогда для динамической системы существует функция T, называемая *кинетической энергией*, имеющая следующий вид:

$$T(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) = \sum_{i,k=1}^{n} a_{ik}\dot{q}_i\dot{q}_k \qquad (a_{ik} = a_{ki}).$$

Стало быть, кинетическая энергия есть однородное квадратичное выражение от обобщенных скоростей, причем коэффициенты a_{ik} являются для данной системы известными функциями от самих координат положения $q_1, q_2, \ldots, q_n$ и не содержат явно времени t.

Это выражение для кинетической энергии T можно получить следующим путем. Прямоугольные координаты (x_k, y_k, z_k) материальных точек нашей системы выражаем через обобщенные координаты $q_1, q_2, \ldots, q_n$. Тогда прямоугольные координаты скоростей всех материальных точек выразятся в виде линейных однородных функций от обобщенных скоростей $\dot{q}_i$. Затем надо вычислить известное выражение для кинетической энергии системы, равное полусумме произведений масс всех частиц системы на квадраты их скоростей, и требуемое выражение для T будет получено.

Помимо кинетической энергии динамическая система характеризуется еще одной функцией, *потенциальной энергией* $U(q_1, q_2, \ldots, q_n)$, зависящей только от обобщенных координат q_i и не содержащей явно времени t.

В курсах динамики показывается, что эта потенциальная энергия определяет внешние силы, действующие на систему. При переводе системы из одного положения в другое производится работа, равная разности соответствующих значений потенциальной энергии U и не зависящая от пути, по которому совершается переход от начального до конечного положения.

Теперь мы уже можем формулировать *принцип Гамильтона*: действительное движение динамической системы за время $t_0 \leqslant t \leqslant t_1$ от заданного начального до заданного конечного положения совершается таким образом, что для этого движения функционал Гамильтона

$$H\{q_1, q_2, \ldots, q_n\} = \int_{t_0}^{t_1} (T - U)\,dt$$

имеет стационарное значение, если к сравнению допускаются все непрерывные функции $q_i(t)$, имеющие непрерывные производные первого и второго порядка и принимающие при $t = t_0$ и при $t = t_1$ заданные краевые значения.

Этот принцип Гамильтона является основным принципом динамики. Его сила и польза заключаются в том, что он представляет собой краткий суммарный сгусток законов динамики. Если написать, в качестве условий стационарности функционала Гамильтона, соответствующие уравнения Эйлера, то получим *уравнения движения Лагранжа*

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = -\frac{\partial U}{\partial q_i} \qquad (i = 1, 2, \ldots, n),$$

которые являются основными уравнениями аналитической динамики.

Здесь мы выведем из этих уравнений только один важный результат, а именно *закон сохранения энергии*. Так как подынтегральная функция интеграла Гамильтона не содержит явно независимой переменной t, то можно сразу написать первый интеграл из n° 2:

$$T-U-\sum \dot{q}_i \frac{\partial (T-U)}{\partial \dot{q}_i}=c=\text{const}.$$

Так как U не зависит от обобщенных скоростей $\dot{q}_i$, а T является однородной квадратичной функцией от них (см. Дополнения к гл. II, § 3), то

$$\sum \dot{q}_i \frac{\partial (T-U)}{\partial \dot{q}_i}=\sum \dot{q}_i \frac{\partial T}{\partial \dot{q}_i}=2T.$$

Следовательно,

$$T+U=\text{const},$$

т. е. при движении системы материальных точек сумма кинетической энергии и потенциальной энергии не изменяется со временем.

4. Функционалы, содержащие производные выше первого порядка. Аналогичные методы применимы к задачам о стационарных значениях функционалов, у которых подынтегральная функция F содержит не только искомую функцию $y=\varphi$ и ее производную φ', но и производные более высокого, например второго, порядка. Пусть, например, требуется найти экстремумы функционала вида

$$I\{\varphi\}=\int_{x_0}^{x_1} F(x, \varphi, \varphi', \varphi'')\,dx,$$

причем к сравнению допускаются те функции $y=\varphi(x)$, которые принимают вместе со своими первыми производными заданные значения на границах интервала и которые имеют непрерывные производные до четвертого порядка включительно.

Для того чтобы вывести необходимые условия экстремума, предположим, что $y=u(x)$ есть искомое решение. Включим это решение в семейство функций $y=\varphi(x)=u(x)+\varepsilon\eta(x)$, где ε—произвольный параметр, а $\eta(x)$—произвольно выбранная функция, имеющая непрерывные производные до четвертого порядка включительно и в граничных точках интервала обращающаяся в нуль вместе со своей первой производной. На кривых этого семейства наш функционал становится функцией $\Phi(\varepsilon)$ параметра ε, имеющей экстремум при $\varepsilon=0$. Поэтому необходимое условие экстремума

$$\Phi'(0)=0$$

должно удовлетворяться для всех функций $\eta(x)$ описанного выше класса. Действуем тем же путем, что и в § 2, n° 1, и выполним дифференцирование под знаком интеграла. Тогда наше условие примет следующий вид:

$$\Phi'(0) = \int_{x_0}^{x_1} (\eta F_u + \eta' F_{u'} + \eta'' F_{u''})\, dx = 0,$$

и оно должно удовлетворяться, если вместо $\varphi(x)$ подставить u. Интегрируем второй член по правилу интегрирования произведения, а третий член—по обобщенному правилу (т. I, стр. 256—257), беря в проинтегрированной части два члена. В силу условий на границах интервала, члены с $\eta'(x)$ и $\eta''(x)$ превратятся в члены, имеющие множитель $\eta(x)$, и в результате получится

$$\int_{x_0}^{x_1} \eta \left(F_u - \frac{d}{dx} F_{u'} + \frac{d^2}{dx^2} F_{u''}\right) dx = 0.$$

Следовательно, необходимым условием экстремума, т. е. условием стационарности функционала, является дифференциальное уравнение Эйлера

$$L[u] \equiv F_u - \frac{d}{dx} F_{u'} + \frac{d^2}{dx^2} F_{u''} = 0.$$

Читатель может сам проверить, что это — дифференциальное уравнение четвертого порядка.

Пример. Рассмотрим функционал

$$I\{\varphi\} = \int_{x_0}^{x_1} (\varphi''^2 - 2f\varphi)\, dx,$$

где $f = f(x)$ — заданная функция. Для него дифференциальное уравнение Эйлера будет

$$u^{\mathrm{IV}} - f(x) = 0.$$

5. Функционал, имеющий вид кратного интеграла. Общий метод нахождения необходимых условий экстремума можно применить с тем же успехом, когда функционал выражен в виде кратного интеграла, т. е. когда функциональный аргумент зависит от нескольких независимых переменных. Пусть, например, дана область D плоскости xy, ограниченная кусочно гладкой кривой Γ. Пусть дана функция $F(x, y, \varphi, \varphi_x, \varphi_y)$, непрерывная и дважды непрерывно дифференцируемая по всем своим пяти аргументам. Если в эту функцию F подставить вместо φ функцию $\varphi(x, y)$, имеющую в области D непрерывные производные первого и второго порядка и принимающую на граничной кривой Γ заданные краевые значения, а вместо φ_x и φ_y — частные производные этой функции $\varphi(x, y)$, то F обращается в функцию от x и y, и значение двойного интеграла

$$I\{\varphi\} = \iint\limits_D F(x, y, \varphi, \varphi_x, \varphi_y)\, dx\, dy$$

будет зависеть от выбора функции φ, т. е. будет функционалом, зависящим от функционального аргумента $\varphi(x, y)$. Задача состоит в том, чтобы найти такую функцию $\varphi = u(x, y)$, которая сообщает функционалу экстремум.

Для нахождения необходимых условий мы применим прежний метод. Выберем функцию $\eta(x, y)$, обращающуюся в нуль на граничной кривой Γ и имеющую непрерывные производные первого и второго порядка, а в остальном произвольную. Допустим, что u есть искомая функция, и подставим под знаком двойного интеграла $\varphi = u + \varepsilon\eta(x, y)$, где ε — произвольный параметр. Тогда этот интеграл становится функцией $\Phi(\varepsilon)$, и необходимое условие экстремума будет

$$\Phi'(0) = 0.$$

Как и в прежних задачах, это условие принимает вид

$$\iint\limits_D (\eta F_u + \eta_x F_{u_x} + \eta_y F_{u_y})\, dx\, dy = 0.$$

Для того чтобы избавиться от членов с η_x и η_y под знаком интеграла, рассматриваем двойной интеграл как повторный и интегрируем по частям — второе слагаемое по x, а третье слагаемое по y. Так как $\eta(x, y)$ обращается в нуль на граничной кривой Γ, то значения проинтегрированных членов на Γ исчезают и наше условие приводится к виду

$$\iint\limits_D \eta\left(F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y}\right) dx\, dy = 0.$$

Лемму I из § 2 нетрудно обобщить на многомерные области, и с помощью этой обобщенной леммы сразу получается *дифференциальное уравнение Эйлера второго порядка с частными производными*

$$F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} = 0.$$

[Сам Эйлер получил это уравнение для прямоугольной области, а для произвольной области D оно было впервые выведено М. В. Остроградским в 1834 г.]

Примеры.

1) $F = \varphi_x^2 + \varphi_y^2$. После сокращения на (-2) получается дифференциальное уравнение Эйлера

$$u_{xx} + u_{yy} = 0.$$

Таким образом, дифференциальное уравнение Лапласа $\nabla^2 u = 0$ получено здесь из вариационной задачи.

2) *Минимальные поверхности. Задача Плато.* Среди поверхностей $z = \varphi(x, y)$, расположенных над областью D плоскости xy и проходящих через

заданную замкнутую пространственную кривую, имеющую своей проекцией граничную кривую Γ области D, найти такую, площадь которой

$$\iint_D \sqrt{1+\varphi_x^2+\varphi_y^2}\,dx\,dy$$

есть минимум.

Для этой задачи дифференциальное уравнение Эйлера есть

$$\frac{\partial}{\partial x}\frac{u_x}{\sqrt{1+u_x^2+u_y^2}}+\frac{\partial}{\partial y}\frac{u_y}{\sqrt{1+u_x^2+u_y^2}}=0$$

или, в развернутом виде,

$$u_{xx}(1+u_y^2)-2u_{xy}u_xu_y+u_{yy}(1+u_x^2)=0.$$

Это известное дифференциальное уравнение минимальных поверхностей.

[Физическое осуществление минимальной поверхности дает жидкостная (например, мыльная) пленка, натянутая на проволочную петлю.]

6. Задачи с дополнительными условиями. Множитель Эйлера. В теории обычных экстремумов функций многих переменных мы изучили также случай, когда эти переменные подчинены некоторым дополнительным условиям (гл. III, § 6, n° 4). В этом случае метод неопределенных множителей привел к особенно ясной формулировке условий стационарности функции. В вариационном исчислении аналогичный метод имеет еще большее значение. Мы здесь рассмотрим вкратце лишь простейшие случаи.

а) Обыкновенные дополнительные условия. В качестве типичного случая рассмотрим задачу о нахождении в трехмерном пространстве кривой

$$x=x(t),\quad y=y(t),\quad z=z(t)\quad (t_0\leqslant t\leqslant t_1),$$

подчиненной следующему дополнительному условию: кривая должна лежать на заданной поверхности $G(x, y, z)=0$ и должна проходить через две данные точки A и B этой поверхности. В этой задаче требуется сообщить функционалу вида

$$I=\int_{t_0}^{t_1} F(x, y, z, \dot x, \dot y, \dot z)\,dt$$

стационарное значение путем подходящего выбора функций $x(t)$, $y(t)$, $z(t)$, связанных дополнительным условием $G(x, y, z)=0$, обычными краевыми условиями и условиями непрерывности.

Эту задачу можно сразу привести к задаче, рассмотренной в n° 1. Предположим, что функции $x(t)$, $y(t)$, $z(t)$ дают решение задачи. Допустим также, что на том куске поверхности, где должна лежать искомая кривая, z может быть выражен явно в виде $z=g(x, y)$. Это будет обеспечено, если $G_z\neq 0$ на этом куске поверхности. Мы будем предполагать, что на рассматриваемой поверхности производные G_x, G_y и G_z не обращаются одновременно в нуль. Тогда, если ограничиться достаточно малым куском поверхности, мы вправе допустить без утраты общности, что $G_z \neq 0$. Подставим теперь под знаком интеграла $z = g(x, y)$ и $\dot{z} = g_x \dot{x} + g_y \dot{y}$, и в нашей задаче уже можно будет рассматривать $x(t)$ и $y(t)$ как независимые друг от друга функции. Стало быть, можно применить непосредственно условия стационарности, сформулированные в конце n° 1, к функционалу I, представив его подынтегральную функцию в виде

$$F(x, y, g(x, y), \dot{x}, \dot{y}, \dot{x} g_x + \dot{y} g_y) = H(x, y, \dot{x}, \dot{y}).$$

Мы получим систему двух уравнений Эйлера

$$\frac{d}{dt} H_{\dot{x}} - H_x = \frac{d}{dt} F_{\dot{x}} - F_x + \frac{d}{dt}(F_{\dot{z}}\, g_x) - F_z g_x - F_{\dot{z}} \frac{\partial \dot{z}}{\partial x} = 0,$$

$$\frac{d}{dt} H_{\dot{y}} - H_y = \frac{d}{dt} F_{\dot{y}} - F_y + \frac{d}{dt}(F_{\dot{z}} g_y) - F_z g_y - F_{\dot{z}} \frac{\partial \dot{z}}{\partial y} = 0.$$

Так как

$$\frac{d}{dt} g_x = \frac{\partial \dot{z}}{\partial x}, \qquad \frac{d}{dt} g_y = \frac{\partial \dot{z}}{\partial y},$$

то эти уравнения принимают следующий вид:

$$\left.\begin{aligned} \frac{d}{dt} F_{\dot{x}} - F_x + g_x \left(\frac{d}{dt} F_{\dot{z}} - F_z\right) = 0, \\ \frac{d}{dt} F_{\dot{y}} - F_y + g_y \left(\frac{d}{dt} F_{\dot{z}} - F_z\right) = 0. \end{aligned}\right\} \qquad (A)$$

Положим

$$\frac{d}{dt} F_{\dot{z}} - F_z = \lambda G_z,$$

т. е. вводим вспомогательный множитель $\lambda(t)$, определяемый этим равенством. Так как $g_x = -\frac{G_x}{G_z}$ и $g_y = -\frac{G_y}{G_z}$, то система двух уравнений (A) превращается в систему трех уравнений

$$\frac{d}{dt} F_{\dot{x}} - F_x = \lambda G_x, \qquad \frac{d}{dt} F_{\dot{y}} - F_y = \lambda G_y, \qquad \frac{d}{dt} F_{\dot{z}} - F_z = \lambda G_z. \qquad (B)$$

Следовательно, условие стационарности функционала I при заданном выше дополнительном условии можно формулировать так:

Если G_x, G_y и G_z не обращаются одновременно в нуль на поверхности $G(x, y, z) = 0$, то необходимым условием экстремума является существование такого множителя $\lambda(t)$, который вместе с функциями $x(t)$, $y(t)$, $z(t)$ удовлетворяет системе трех уравнений (B) и дополнительному условию $G(x, y, z) = 0$. Таким образом, имеется симметричная система четырех уравнений для определения функций $x(t)$, $y(t)$, $z(t)$ и множителя $\lambda(t)$.

Самой важной конкретной задачей этого типа является задача о нахождении на поверхности $G(x, y, z)=0$ кратчайшей линии, соединяющей две точки A и B этой поверхности, причем предполагается, что на ней $\operatorname{grad} G \neq 0$. В данном случае

$$F=\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}$$

и дифференциальные уравнения Эйлера (B) будут

$$\frac{d}{dt}\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}=\lambda G_x, \quad \frac{d}{dt}\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}=\lambda G_y,$$

$$\frac{d}{dt}\frac{\dot{z}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}=\lambda G_z.$$

Эти уравнения инвариантны относительно выбора параметра. Это значит, как читатель легко сам проверит, что они сохраняют свой вид, если заменить t другим параметром $\tau=\tau(t)$, если только это преобразование взаимно однозначно, обратимо и непрерывно дифференцируемо. Введем в качестве нового параметра длину дуги s искомой кривой. Тогда после введения нового параметра будет $\dot{x}^2+\dot{y}^2+\dot{z}^2=1$, и наши дифференциальные уравнения примут следующий вид:

$$\frac{d^2x}{ds^2}=\lambda G_x, \quad \frac{d^2y}{ds^2}=\lambda G_y, \quad \frac{d^2z}{ds^2}=\lambda G_z.$$

Геометрический смысл этих дифференциальных уравнений состоит в том, что соприкасающиеся плоскости [1]) экстремалей нашей задачи ортогональны поверхности $G(x, y, z)=0$. Эти экстремали называются *геодезическими линиями* данной поверхности. Следовательно, кратчайшее расстояние между двумя точками на поверхности может быть дано только дугой геодезической линии.

Упражнение

Показать, что те же геодезические линии получаются как траектории материальной точки, которая может двигаться только по данной поверхности $G(x, y, z)=0$ и не подвергается действию внешних сил. (Применить принцип Гамильтона из n° 3, приняв во внимание, что в этом случае потенциальная энергия $U=0$.)

б) Другие виды дополнительных условий. В задаче, рассмотренной в пункте а), удалось исключить дополнительное условие, разрешая уравнение, выражающее это условие, относительно одного из аргументов, и в результате задача сразу привелась к виду, рассмотренному ранее. Однако часто встречаются такие виды дополнительных условий, для которых это невыполнимо. Самыми важными задачами этого типа являются задачи с «изопериметрическими» дополнительными условиями. Приведем типичный пример.

---

[1]) То есть плоскости, проходящие через точки кривой параллельно векторам $\dot{r}=\{\dot{x}, \dot{y}, \dot{z}\}$ и $\ddot{r}=\{\ddot{x}, \ddot{y}, \ddot{z}\}$ (ср. упр. 1, 2 и 6 на стр. 114—115).

При обычных краевых условиях и условиях непрерывности найти критерий стационарности функционала

$$I\{\varphi\}=\int_{x_0}^{x_1} F(x,\ \varphi,\ \varphi')\,dx,$$

когда функциональный аргумент $\varphi(x)$ подчинен дополнительному условию

$$H\,\{\varphi\}=\int_{x_0}^{x_1} G(x,\ \varphi,\ \varphi')\,dx=c,$$

где c — заданная постоянная. Классическая изопериметрическая задача получается отсюда как частный случай при $F=\varphi$ и $G=\sqrt{1+\varphi'^2}$.

К задаче такого типа невозможно подойти прежним методом построения «варьированной» функции $\varphi=u+\varepsilon\eta$ с помощью произвольной функции $\eta(x)$, обращающейся в нуль на границах интервала, ибо, как правило, построенные таким образом функции не удовлетворяют дополнительному условию при ε, близком к нулю, но отличном от нуля. Однако можно добиться успеха, если немного изменить этот метод и ввести, вместо одной функции $\eta(x)$ и одного параметра ε, две функции $\eta_1(x)$ и $\eta_2(x)$, обращающиеся в нуль на границах, и два параметра ε_1 и ε_2. Предполагая, что $\varphi=u$ есть функция, решающая задачу, строим варьированную функцию

$$\varphi=u+\varepsilon_1\eta_1+\varepsilon_2\eta_2.$$

Подставив эту функцию в оба интеграла, мы приведем функционал I к виду

$$I=\int_{x_0}^{x_1} F(x,\ u+\varepsilon_1\eta_1+\varepsilon_2\eta_2,\ u'+\varepsilon_1\eta_1'+\varepsilon_2\eta_2')\,dx=\Phi(\varepsilon_1,\ \varepsilon_2)$$

и дополнительное условие к виду

$$H=\int_{x_0}^{x_1} G(x,\ u+\varepsilon_1\eta_1+\varepsilon_2\eta_2,\ u'+\varepsilon_1\eta_1'+\varepsilon_2\eta_2')\,dx=\Psi(\varepsilon_1,\ \varepsilon_2)=c.$$

Теперь задача сводится к тому, что функция $\Phi(\varepsilon_1,\ \varepsilon_2)$ должна иметь стационарное значение при $\varepsilon_1=\varepsilon_2=0$, когда ε_1 и ε_2 удовлетворяют дополнительному условию

$$\Psi(\varepsilon_1,\ \varepsilon_2)=c.$$

Пользуясь известными результатами из теории обычных экстремумов функции двух переменных (гл. III, § 6, n° 4), а в дальнейшем следуя тем же путем, что и в § 2 настоящей главы, придем к следующему результату:

Стационарность функционала I при добавочном условии $H=c$ равносильна существованию такого постоянного множителя λ, при котором искомая функция и удовлетворяет уравнению $H\{u\}=c$ и дифференциальному уравнению Эйлера

$$\frac{d}{dx}(F_{u'}+\lambda G_{u'})-(F_u+\lambda G_u)=0.$$

Исключение из этого правила может возникнуть лишь в том случае, если функция $u(x)$ удовлетворяет уравнению

$$\frac{d}{dx}G_{u'}-G_u=0.$$

Подробное выполнение доказательства предоставляем читателю, который может консультироваться с литературой вопроса. [Например, М. А. Лаврентьев и Л. А. Люстерник, Основы вариационного исчисления, том первый, часть II, гл. IX.]

Упражнения

1. Пользуясь методом множителя Эйлера, доказать, что решением классической изопериметрической задачи является окружность.

2. Нить постоянной плотности и заданной длины l натянута между двумя точками A и B. Известно, что равновесным будет такое положение нити, в котором центр тяжести занимает наинизшее возможное положение. Пусть ось y направлена вертикально вверх. Тогда задача заключается в нахождении минимума интеграла

$$I=\int_{x_0}^{x_1} y\sqrt{1+y'^2}\,dx$$

при добавочном условии

$$\int_{x_0}^{x_1}\sqrt{1+y'^2}\,dx=l=\text{const}.$$

Доказать, что равновесная форма нити — цепная линия.

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ VII

1. Показать, что геодезическими линиями на цилиндрической поверхности являются винтовые линии.

2. Составить уравнения Эйлера для интегралов со следующими подынтегральными функциями:

а) $F=\sqrt{1+y'^2}+yg(x)$; б) $F=\dfrac{y''^2}{(1+y'^2)^3}+yg(x)$;

в) $F=y''^2-y'^2+y^2$; г) $F=\sqrt[4]{1+y'^2}$.

3. Составить уравнения Эйлера для двойных интегралов со следующими подынтегральными функциями, где $\varphi=\varphi(x, y)$:

а) $F=a\varphi_x^2+2b\varphi_x\varphi_y+c\varphi_y^2+m\varphi^2$; б) $F=(\varphi_{xx}+\varphi_{yy})^2=(\nabla^2\varphi)^2$;

в) $F=(\nabla^2\varphi)^2+(\varphi_{xx}\varphi_{yy}-\varphi_{xy}^2)$.

4. Составить уравнение Эйлера для задачи о стационарном значении интеграла

$$\int_{x_0}^{x_1} (au'^2 + 2buu' + cu^2)\, dx$$

при добавочном условии

$$\int_{x_0}^{x_1} u^2 dx = 1$$

(коэффициенты a, b и c — функции от x).

5. Задана функция $f(x)$. Найти функцию $\varphi = u(x)$, сообщающую максимум интегралу

$$I\{\varphi\} = \int_0^1 f(x)\, \varphi(x)\, dx$$

при добавочном условии

$$H\{\varphi\} = \int_0^1 \varphi^2\, dx = k^2,$$

где k — заданная постоянная.

Указание. а) Найти решение $u(x)$ из уравнения Эйлера; б) доказать с помощью интегрального неравенства [Буняковского —] Шварца (примечание 1 на стр. 352), что найденное в а) решение сообщает функционалу I его наибольшее значение.

ГЛАВА VIII

ФУНКЦИИ КОМПЛЕКСНОЙ ПЕРЕМЕННОЙ

Теории функций комплексной переменной мы коснулись вскользь в первом томе, в гл. VIII, § 7 и убедились, что эта теория проливает новый свет на строение функций действительной переменной. Здесь мы дадим краткий, но более систематический очерк элементов этой теории.

§ 1. Введение

1. Пределы и бесконечные ряды с комплексными членами. Мы будем исходить из элементарного понятия комплексного числа $z = x + iy$ (т. I, стр. 95), построенного из действительной единицы 1 и мнимой единицы i с помощью любых двух действительных чисел x, y. Над этими комплексными числами совершают действия по тем же правилам, что и над обыкновенными числами, но с добавлением правила, что i^2 можно всегда заменить числом (-1). Действительную часть x и мнимую часть y комплексного числа z рассматривают как декартовы прямоугольные координаты в плоскости xy или, как принято говорить, в плоскости комплексной переменной z. Число $\bar{z} = x - iy$ называется комплексным числом, *сопряженным* числу z. Если ввести полярные координаты (r, θ) с помощью соотношений $x = r\cos\theta$, $y = r\sin\theta$, то комплексное число $z = x + iy$ запишется так: $z = r(\cos\theta + i\sin\theta)$; это тригонометрическая форма комплексного числа. Полярный угол θ называется *аркусом* или *аргументом* комплексного числа z, а $r = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} = |z|$ — его *модулем*.

[Значение θ_0 полярного угла θ, удовлетворяющее неравенствам $-\pi < \theta_0 \leqslant \pi$, называется главным значением аркуса и обозначается $\operatorname{arc} z$, а все множество значений аркуса комплексного числа z обозначается символом $\operatorname{Arc} z$. Итак, если $z \neq 0$, то

$-\pi < \operatorname{arc} z \leqslant \pi$, $\quad \operatorname{Arc} z = \operatorname{arc} z + 2n\pi$, где n — любое целое число.]

Комплексные числа z_1, z_2 и $z_1 + z_2$ удовлетворяют так называемому *неравенству треугольника*

$$|z_1 + z_2| \leqslant |z_1| + |z_2|,$$

которое нетрудно доказать, а из этого неравенства, полагая $z_1 = u_1 - u_2$ и $z_2 = u_2$, выводим еще одно неравенство:

$$\big||u_1| - |u_2|\big| \leqslant |u_1 - u_2|.$$

Неравенство треугольника можно истолковать геометрически следующим образом. Всякое комплексное число $z = x + iy$ можно изобразить на плоскости xy вектором $\{x, y\}$. Тогда число z_1 изобразится вектором $\{x_1, y_1\}$, а число z_2 — вектором $\{x_2, y_2\}$; вектор же, изображающий сумму $z_1 + z_2$, будет равен сумме векторов $\{x_1, y_1\} + \{x_2, y_2\}$. Эти три вектора образуют треугольник, стороны которого имеют длины $|z_1|$, $|z_2|$ и $|z_1 + z_2|$. Ясно, что «неравенство треугольника» выражает тот факт, что любая сторона треугольника меньше суммы двух других сторон.

Теперь мы рассмотрим существенно новое понятие — *предела последовательности комплексных чисел*. Установим следующее определение: последовательность комплексных чисел z_n называется сходящейся или стремящейся к пределу z, если $|z_n - z|$ стремится к нулю при $n \to \infty$. Это означает, конечно, что как действительная, так и мнимая часть разности $z_n - z$ стремятся к нулю. А из критерия сходимости Коши для действительных последовательностей вытекает следующий вывод: необходимым и достаточным условием существования у последовательности z_n предела z является выполнение соотношения

$$\lim_{\substack{n\to\infty \\ m\to\infty}} |z_n - z_m| = 0.$$

Это критерий Коши для комплексной последовательности.

Особенно важный класс пределов возникает при изучении *бесконечных рядов с комплексными членами*. Бесконечный ряд с комплексными членами $\sum_{k=0}^{\infty} c_k$ называется сходящимся и имеющим сумму S, если последовательность частичных сумм $S_n = \sum_{k=0}^{\infty} c_k$ стремится к (конечному) пределу S.

Если ряд $\sum_{k=0}^{\infty} |c_k|$, составленный из модулей членов данного ряда сходится, то (как и в гл. VIII первого тома, стр. 431) можно доказать, что исходный ряд с комплексными членами тоже сходится; в этом случае этот исходный комплексный ряд называется *абсолютно сходящимся*.

Если члены ряда не являются постоянными, а зависят от комплексной переменной $z = x + iy$, стало быть, зависят от точки (x, y), пробегающей некоторую замкнутую область G, то уместно введение понятия *равномерной сходимости*. Ряд называется равномерно сходящимся в G, если по любому наперед заданному сколь угодно малому положительному числу ε можно найти такое число N, зависящее только от ε, что при всяком $n \geqslant N$ выполняется неравенство $|S_n - S| < \varepsilon$, независимо от того, где в области G находится точка $z = x + iy$. Ясно, что точно таким же путем можно определить *равномерную сходимость последовательности* комплексных функций

$f_n(z)$, зависящих от точки z области G. Все эти определения, теоремы и их доказательства полностью соответствуют тому, что нам уже знакомо из теории действительных переменных.

Самый простой пример функционального ряда представляет геометрический ряд

$$1+z+z^2+z^3+\ldots.$$

Как и в случае действительной переменной, имеем

$$S_n=\frac{1-z^{n+1}}{1-z} \quad \text{и} \quad \lim_{n\to\infty} S_n=\frac{1}{1-z} \quad \text{при } |z|<1.$$

Последнее утверждение часто записывают так:

$$1+z+z^2+\ldots=\frac{1}{1-z} \quad \text{при } |z|<1.$$

Ясно, что геометрический ряд *сходится абсолютно*, если $|z|<1$, а также, что эта сходимость *равномерна*, если только $|z|\leqslant q$, где q — любое постоянное положительное число, лежащее между 0 и 1. Другими словами, *геометрический ряд сходится абсолютно при всех значениях z, лежащих внутри единичной окружности, и сходится равномерно на любом замкнутом круге, имеющем общий центр с единичной окружностью и радиус, меньший единицы.*

Принцип сравнения рядов применим и к рядам с комплексными членами в следующем виде: *если $|c_k|\leqslant p_k$, где все p_k — действительные и неотрицательные числа, и если ряд $\sum\limits_{k=0}^{\infty} p_k$ — сходится, то комплексный ряд $\sum\limits_{k=0}^{\infty} c_k$ сходится абсолютно.*

Если члены c_k комплексного ряда зависят от точки z, изменяющейся в замкнутой области G: $c_k=f_k(z)$, а все числа p_k не зависят от z, причем $|c_k|\leqslant p_k$ во всей области G, то ряд $\sum\limits_{k=0}^{\infty} c_k=\sum\limits_{k=0}^{\infty} f_k(z)$ сходится *равномерно* в G.

Доказательства ничем не отличаются от соответствующих доказательств в действительной области (т. I, гл. VIII), поэтому нет надобности повторять их здесь.

Если M — произвольная положительная постоянная, а q — положительное число, лежащее между 0 и 1, то ряды с положительными членами $\sum\limits_{k=0}^{\infty} Mq^k$, $\sum\limits_{k=1}^{\infty} kMq^{k-1}$ и $\sum\limits_{k=0}^{\infty} \frac{M}{k+1}q^{k+1}$ все сходятся (т. I, гл. VIII, стр. 462). Этими рядами мы воспользуемся как рядами для сравнения в следующем же номере.

2. Степенной ряд. Среди бесконечных рядов с комплексными членами самыми важными являются *степенные ряды*, т. е. ряды с членами вида $c_k = a_k z^k$ или $a_k(z - z_0)^k$, где z_0 — постоянное комплексное число. Стало быть, степенной ряд может быть записан так

$$P(z) = \sum_{k=0}^{\infty} a_k z^k,$$

или в следующем, несколько более общем виде:

$$\mathfrak{P}(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k.$$

Так как последний ряд всегда приводится подстановкой $z - z_0 = Z$ к предыдущему виду, то достаточно будет исследовать тот случай, когда $z_0 = 0$.

Основная теорема о сходимости степенного ряда повторяет почти слово в слово соответствующую теорему для рядов с действительными членами (т. I, гл. VIII, стр. 460). *Если степенной ряд сходится при $z = z_0$, то он сходится абсолютно при всяком значении z, для которого $|z| < |z_0|$; к тому же ряд сходится равномерно на круге $|z| \leqslant q|z_0|$, где q — любое положительное число, меньшее единицы.*

Из того же предположения, что степенной ряд $P(z)$ сходится при $z = z_0$, вытекает и следующая теорема:

Оба ряда

$$D(z) = \sum_{k=1}^{\infty} k a_k z^{k-1} \quad \text{и} \quad I(z) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} z^{k+1}$$

тоже сходятся абсолютно и равномерно на круге $|z| \leqslant q|z_0|$.

Доказательство точно такое же, как в действительной области. Так как ряд $P(z)$ сходится при $z = z_0$, то его общий член $a_n z_0^n$ стремится к нулю при $n \to \infty$. Следовательно, существует такое положительное число M, что $|a_n z_0^n| < M$ при всех значениях n. Пусть теперь, $|z| = q|z_0|$, где $0 < q < 1$; тогда

$$|a_n z^n| < M q^n, \quad |n a_n z^{n-1}| < \frac{M}{|z_0|} n q^{n-1}, \quad \left| \frac{a_n}{n+1} z^{n+1} \right| < \frac{M \cdot |z_0|}{n+1} q^{n+1}.$$

Стало быть, мы получили действительные знакоположительные ряды сравнения, которые, как мы отметили в конце предшествующего n° 1, сходятся. Теорема о рядах $D(z)$ и $I(z)$ доказана.

Вернемся к рассмотрению произвольного степенного ряда вида $P(z) = \sum_{k=0}^{\infty} a_k z^k$. Существуют две возможности: либо он сходится при всех значениях z, либо существует такое значение $z = Z$, при котором ряд расходится. Во втором случае ряд расходится и при всех значениях z, для которых $|z| > |Z|$ (доказывается от противного). Несложное рассуждение (такое же, как в действительной области) приводит к выводу, что существует такое положительное действительное число ρ (оно называется *радиусом сходимости* ряда), что ряд $P(z)$ сходится при $|z| < \rho$ и расходится при $|z| > \rho$. Этот вывод относится и к рядам $D(z)$ и $I(z)$, причем их радиус сходимости ρ тот же самый, что и для исходного ряда $P(z)$. Круг $|z| \leqslant \rho$ называется *кругом сходимости* степенного ряда. По вопросу о сходимости или расходимости степенного ряда на самой окружности $|z| = \rho$ круга сходимости нет возможности делать какие-либо общие утверждения. Если степенной ряд сходится при всех значениях z, то принято говорить, что радиус сходимости $\rho = \infty$. При $z = 0$ сходятся все степенные ряды рассматриваемого вида; если ряд $P(z)$ сходится только при $z = 0$, то радиус сходимости $\rho = 0$.

3. Дифференцирование и интегрирование степенного ряда. Выражение вида

$$P_n(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n$$

с постоянными (комплексными) коэффициентами a_k естественно называть *функцией* от z, а более конкретно *многочленом степени n* от z. Аналогично степенной ряд

$$P(z) = \sum_{k=0}^{\infty} a_k z^k$$

(точнее говоря, его сумма) рассматривается внутри своего круга сходимости как функция комплексной переменной z. Внутри круга сходимости степенного ряда его сумма есть предел, к которому стремится многочлен

$$P_n(z) = \sum_{k=0}^{n} a_k z^k,$$

когда $n \to \infty$.

Для многочлена $P_n(z)$ вводится понятие *производной* по независимой переменной z с помощью того же определения, что и для действительной переменной; обозначается производная теми же знакомыми символами. Прежде всего замечаем, что алгебраическое тождество

$$\frac{z_1^n - z^n}{z_1 - z} = z_1^{n-1} + z_1^{n-2} z + \ldots + z^{n-1}$$

справедливо и при любом комплексном z. Устремим теперь z_1 к z, и мы получим[1]):

$$\frac{d}{dz}(z^n) = (z^n)' = \lim_{z_1 \to z} \frac{z_1^n - z^n}{z_1 - z} = n z^{n-1}.$$

---

[1]) Понятие предела непрерывно изменяющейся комплексной переменной ($z_1 \to z$) можно ввести таким же точно путем, как и для действительной переменной.

Точно так же сразу получаем

$$P'_n(z) = \frac{d}{dz} P_n(z) = \lim_{z_1 \to z} \frac{P_n(z_1) - P_n(z)}{z_1 - z} = \sum_{k=1}^{n} k a_k z^{k-1} = D_n(z).$$

Многочлен $P'_n(z)$ и есть производная от комплексного многочлена $P_n(z)$. Процесс нахождения производной в комплексной области тоже называется дифференцированием.

В теории степенных рядов играет фундаментальную роль следующая теорема:

Степенной ряд

$$P(z) = \sum_{k=0}^{\infty} a_k z^k,$$

радиус сходимости которого не равен нулю, можно дифференцировать почленно внутри его круга сходимости. Это значит, что предел

$$P'(z) = \lim_{z_1 \to z} \frac{P(z_1) - P(z)}{z_1 - z}$$

существует, и притом

$$P'(z) = \sum_{k=1}^{\infty} k a_k z^{k-1} = \lim_{n \to \infty} P'_n(z) = \lim_{n \to \infty} D_n(z) = D(z).$$

Из этой теоремы сразу ясно, что степенной ряд

$$I(z) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} z^{k+1}$$

можно рассматривать как *первообразную функцию* от исходного степенного ряда внутри его круга сходимости, т. е. что $I'(z) = P(z)$.

Теорема о почленном дифференцировании степенного ряда доказывается следующим образом. Из n°2 известно, что внутри круга сходимости $D(z) = \lim\limits_{n \to \infty} D_n(z)$. Для нашей цели достаточно доказать, что $\left| \frac{P(z_1) - P(z)}{z_1 - z} - D(z) \right|$ может быть сделана меньше любого наперед заданного положительного числа ε, если только выбрать z_1 внутри круга сходимости достаточно близко к z. Построим требуемое отношение приращений

$$\frac{\Delta P(z)}{\Delta z} = \frac{P(z_1) - P(z)}{z_1 - z} = \frac{P_n(z_1) - P_n(z)}{z_1 - z} + \sum_{k=n+1}^{\infty} a_k \lambda_k,$$

где для краткости положено

$$\lambda_k = \frac{z_1^k - z^k}{z_1 - z} = z_1^{k-1} + z_1^{k-2} z + \ldots + z^{k-1}.$$

Теперь, пользуясь обозначениями n°2 и считая $|z|<q|z_0|$, выберем z_1 так, что и $|z_1|<q|z_0|$; этим обеспечивается, что

$$|\lambda_k|<kq^{k-1}\cdot|z_0|^{k-1}.$$

Следовательно,

$$|R_n|=\left|\sum_{k=n+1}^{\infty} a_k\lambda_k\right|<\sum_{k=n+1}^{\infty}|a_k|\,kq^{k-1}\,|z_0|^{k-1}<\frac{M}{|z_0|}\sum_{k=n+1}^{\infty}kq^{k-1}.$$

В силу сходимости ряда $\sum\limits_{k=1}^{\infty}kq^{k-1}$, состоящего из положительных членов, выражение $|R_n|$ может быть сделано сколь угодно малым, если сделать n достаточно большим. Мы и выберем n столь большим, чтобы было $|R_n|<\frac{\varepsilon}{3}$, а также настолько большим (для этого мы, если понадобится, еще увеличим n), что $|D(z)-D_n(z)|<\frac{\varepsilon}{3}$. А теперь выберем z_1 так близко от z, что $\left|\frac{P_n(z_1)-P_n(z)}{z_1-z}-D_n(z)\right|<\frac{\varepsilon}{3}$. Тогда

$$\left|\frac{\Delta P(z)}{\Delta z}-D(z)\right|\leqslant\left|\frac{P_n(z_1)-P_n(z)}{z_1-z}-D_n(z)\right|+|D_n(z)-D(z)|+$$
$$+|R_n|<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon,$$

а это неравенство и доказывает нашу теорему.

Так как производная $P'(z)$ тоже представлена в виде степенного ряда с тем же радиусом сходимости, то полученный ряд можно снова дифференцировать почленно, и этот процесс можно повторить сколько угодно раз. Стало быть, *степенной ряд можно дифференцировать внутри его круга сходимости столько раз, сколько понадобится.*

Степенной ряд является рядом Тэйлора для той функции $P(z)$, которую он представляет; это значит, что его коэффициенты a_k могут быть выражены формулами

$$a_0=P(0),\qquad a_k=\frac{1}{k!}P^{(k)}(0)\quad \textit{при}\quad k=1,\,2,\,3,\,\ldots$$

Для степенного ряда вида $\mathfrak{P}(z)=\sum\limits_{k=0}^{\infty}a_k(z-z_0)^k$ будет

$$a_0=\mathfrak{P}(z_0),\qquad a_k=\frac{1}{k!}\mathfrak{P}^{(k)}(z_0)\quad\text{при}\quad k=1,\,2,\,3,\,\ldots$$

Доказательство точно такое же, как для действительной переменной (т. I, стр. 464—465).

4. Определение показательной функции, тригонометрических и гиперболических функций с помощью степенных рядов. Как уже упоминалось в гл. VIII, § 7 первого тома, степенные ряды для элементарных функций можно непосредственно распространить на комплексные переменные; другими словами, в степенном ряде, полученном для функции действительного аргумента, приписывают аргументу комплексные значения и рассматривают этот ряд как *определение* соответствующей функции комплексной переменной. Например, степенные ряды

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}, \quad \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!}, \quad \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}, \quad \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!}$$

сходятся при всех значениях z. (Это можно доказать, применяя признак сходимости Даламбера к ряду, составленному из модулей членов рассматриваемого ряда.) Так вот, функции, представленные этими рядами, т. е. суммы этих степенных рядов, вновь обозначаются соответственно символами e^z, $\cos z$, $\sin z$, $\operatorname{ch} z$, $\operatorname{sh} z$ — теми же символами, что и в действительной области. Стало быть, эти степенные ряды

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}, \qquad \cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}, \qquad \sin z = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!},$$

$$\operatorname{ch} z = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}, \qquad \operatorname{sh} z = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!},$$

которые для действительного аргумента были выведены из определений перечисленных здесь функций, данных ранее другим путем, теперь сами служат *определениями* новых, впервые здесь вводимых, функций комплексной переменной.

Из этих определений сразу вытекают следующие тождества:

$$e^{iz} = \cos z + i \sin z, \qquad e^z = \operatorname{ch} z + \operatorname{sh} z,$$

$$\cos iz = \operatorname{ch} z, \qquad \sin iz = i \operatorname{sh} z, \qquad \operatorname{ch} iz = \cos z, \qquad \operatorname{sh} iz = i \sin z.$$

Подставив в наши степенные ряды $z = 0$, имеем

$$e^0 = 1, \quad \cos 0 = 1, \quad \sin 0 = 0, \quad \operatorname{ch} 0 = 1, \quad \operatorname{sh} 0 = 0.$$

Почленное дифференцирование степенных рядов, дающих определение показательной функции, тригонометрических и гиперболических функций комплексной переменной, приводит к следующим формулам для производных от этих функций:

$$(e^z)' = e^z, \qquad (\cos z)' = -\sin z, \qquad (\sin z)' = \cos z, \qquad (\operatorname{ch} z)' = \operatorname{sh} z, \qquad (\operatorname{sh} z)' = \operatorname{ch} z.$$

Таким же точно путем вводим новые для нас функции комплексного аргумента: логарифмическую функцию $\ln(1+z)$ и функцию $\operatorname{arctg} z$, определяемые как суммы следующих степенных рядов:

$$\ln(1+z) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{z^k}{k} \qquad (\text{радиус сходимости } \rho = 1),$$

$$\operatorname{arctg} z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{2k+1} = \frac{1}{2i}\left[\ln(1+iz) - \ln(1-iz)\right]$$

$$(\text{радиус сходимости } \rho = 1).$$

Дифференцируя эти ряды почленно, получаем формулы для производных

$$\frac{d}{dz}\ln(1+z)=\frac{1}{1+z}, \qquad \frac{d}{dz}(\operatorname{arctg} z)=\frac{1}{1+z^2}.$$

Упражнения

1. При каких значениях $z=x+iy$ будет $\left|\frac{z-1}{z+1}\right|\leqslant 1$?

2. Доказать, что если ряд $\sum a_n z^n$ сходится *абсолютно* при $z=\zeta$, то он сходится равномерно в области $|z|\leqslant|\zeta|$.

3. Пользуясь степенными рядами для $\cos z$ и $\sin z$, показать, что

$$\cos^2 z+\sin^2 z=1.$$

4*. Выяснить, при каких значениях z сходится ряд

$$\sum_{k=1}^{\infty}\frac{z^k}{1-z^k}.$$

§ 2. Основные понятия теории функций комплексной переменной

1. Требование дифференцируемости. Как мы видели выше, все функции, представленные степенными рядами, имеют как производную, так и первообразную (интеграл). Этот факт можно сделать исходным пунктом для построения общей теории функций комплексной переменной. Цель такой теории — распространить дифференциальное и интегральное исчисление на функции комплексной переменной. Важно, в частности, чтобы обобщение понятия функции на случай комплексной независимой переменной осуществлялось таким путем, чтобы функция была дифференцируема в комплексной области.

Можно было бы, конечно, ограничиться с самого начала рассмотрением функций, представленных степенными рядами, и тем самым требование дифференцируемости было бы выполнено. Но против такого подхода существуют два возражения. Во-первых, мы не можем сказать a priori, вытекает ли из требования дифференцируемости комплексной функции, что она может быть разложена в степенной ряд. (Ведь в случае действительного аргумента мы видели (т. I, стр. 385—386), что существуют даже такие функции, которые имеют производные любого порядка и все же не могут быть разложены в степенной ряд.) Во-вторых, даже для простых функциональных выражений представляющий их степенной ряд может не охватывать все поведение функции; это видно уже на примере простой функции $\frac{1}{1-z}$, разложение которой в степенной ряд (бесконечная геометрическая прогрессия) сходится лишь в круге радиуса 1, между тем как эта функция определена при всех значениях z, кроме $z=1$.

Правда, существует метод Вейерштрасса, позволяющий избежать этих затруднений и развить теорию функций комплексной переменной на основе теории степенных рядов. Желательно, однако, подчеркнуть другую точку зрения, принадлежащую Коши и Риману. В их подходе функции характеризуются не своими *явными выражениями*, а своими основными *свойствами*. Точнее, для отграничения области определения функции надлежит пользоваться не возможностью ее представления степенным рядом, а требованием дифференцируемости функции.

Можно было бы исходить с самого начала из следующего общего понятия комплексной функции $\zeta = f(z)$ комплексной переменной z. Если G есть область плоскости z и если каждой точке $z = x + iy$ в G приводится в соответствие комплексное число $\zeta = u + iv$, с помощью любого правила, то ζ называется комплексной функцией от z в области G. Но это определение выражает просто тот факт, что всякой паре действительных чисел x и y, таких, что точка (x, y) лежит в G, соответствует пара действительных чисел u, v; это значит, что u и v суть две любые действительные функции $u(x, y)$ и $v(x, y)$, определенные в области G, от двух действительных переменных x и y.

Однако такое понятие функции было бы чрезмерно широким. Мы его сузим прежде всего условием, что $u(x, y)$ и $v(x, y)$ должны быть функциями, непрерывными в области G и имеющими в ней непрерывные частные производные первого порядка u_x, u_y, v_x, v_y. Затем мы потребуем, чтобы наша функция $\zeta = u + iv = f(z) = f(x + iy)$ была дифференцируема в области G по комплексной независимой переменной z, т. е. чтобы, при всех значениях z в G, существовал предел

$$\lim_{z_1 \to z} \frac{f(z_1) - f(z)}{z_1 - z} = \lim_{h \to 0} \frac{f(z + h) - f(z)}{h} = f'(z).$$

Этот-то предел и называется *производной* от $f(z)$.

Того факта, что функции u и v имеют непрерывные производные по x и по y, отнюдь не достаточно для того, чтобы функция $f(z)$ была дифференцируемой, т. е. имела производную по z. Наше требование дифференцируемости в комплексной области более богато содержанием, чем дифференцируемость по действительной переменной. Дело в том, что $h = h_1 + ih_2$ может стремиться к нулю как по действительным значениям ($h_2 = 0$), так и по чисто мнимым значениям ($h_1 = 0$) и вообще по любому другому пути, и по всем этим путям должен получаться *один и тот же* предел, и только тогда можно будет считать, что производная по z существует и функция $f(z)$ дифференцируема.

Положим, например, $u = x$, $v = 0$, т. е. $f(z) = f(x + iy) = x$. Очевидно, что $u(x, y)$ и $v(x, y)$ имеют непрерывные производные по x и по y. Но если, желая найти производную $f'(z)$, положим один раз $h=h_1$, а в другой раз $h=ih_2$, то получим

$$\lim_{h_1\to 0}\frac{f(z+h_1)-f(z)}{h_1}=\lim_{h_1\to 0}\frac{x+h_1-x}{h_1}=1,$$

$$\lim_{h_2\to 0}\frac{f(z+ih_2)-f(z)}{ih_2}=\lim_{h_2\to 0}\frac{x-x}{ih_2}=\lim_{h_2\to 0}\frac{0}{ih_2}=0,$$

т. е. совершенно различные пределы. Пример такого же типа представляет функция $f(z)=u+iv=x+2iy$, у которой тоже отношение приращений стремится к различным пределам, когда h стремится к нулю разными путями.

Стало быть, для того чтобы обеспечить дифференцируемость функции $f(z)$, придется наложить на нее новое ограничение. Этот основной факт в теории функций комплексной переменной выражается следующей теоремой:

Если $\zeta=u(x, y)+iv(x, y)=f(z)=f(x+iy)$, где $u(x,y)$ и $v(x,y)$ имеют непрерывные производные по x и по y, то для того, чтобы функция $f(z)$ имела производную по комплексной переменной z, необходимо и достаточно, чтобы выполнялись следующие условия:

$$u_x=v_y,\qquad u_y=-v_x,$$

называемые условиями Коши — Римана[1].

Во всякой области G, в которой u и v удовлетворяют этим условиям, $f(z)$ называется аналитической или регулярной функцией комплексной переменной z и производная от $f(z)$ дается формулой

$$f'(z)=u_x+iv_x=v_y-iu_y=\frac{1}{i}(u_y+iv_y)=u_x-iu_y=v_y+iv_x.$$

Сначала покажем, что дифференциальные уравнения Коши — Римана являются *необходимым* условием дифференцируемости. Для этого предположим, что производная $f'(z)$ существует; тогда предел отношения приращений должен равняться $f'(z)$, безразлично, возьмем ли мы действительное $h=h_1$ или чисто мнимое $h=ih_2$:

$$f'(z)=\lim_{h_1\to 0}\left[\frac{u(x+h_1,y)-u(x,y)}{h_1}+i\frac{v(x+h_1,y)-v(x,y)}{h_1}\right]=u_x+iv_x,$$

$$f'(z)=\lim_{h_2\to 0}\left[\frac{u(x,y+h_2)-u(x,y)}{ih_2}+i\frac{v(x,y+h_2)-v(x,y)}{ih_2}\right]=$$

$$=\frac{1}{i}(u_y+iv_y).$$

[1]) А. И. Маркушевич считает это название исторически несправедливым, так как эти дифференциальные уравнения изучались уже раньше Даламбером и Эйлером. Однако, в ТФКП они выражают условия — Коши-Римана. *(Прим. перев.)*

Следовательно,

$$u_x + iv_x = \frac{1}{i}(u_y + iv_y).$$

Приравнивая действительные, а также мнимые части обоих выражений, получим условия Коши — Римана.

Теперь докажем, что эти уравнения составляют также и достаточное условие дифференцируемости функции $f(z)$. Для этого напишем отношение приращений

$$\frac{f(z+h)-f(z)}{h} =$$
$$= \frac{u(x+h_1, y+h_2) - u(x, y) + i[v(x+h_1, y+h_2) - v(x, y)]}{h_1 + ih_2} =$$
$$= \frac{h_1 u_x + h_2 u_y + i(h_1 v_x + h_2 v_y) + \varepsilon_1 |h| + i\varepsilon_2 |h|}{h_1 + ih_2},$$

где ε_1 и ε_2 — действительные величины, стремящиеся к нулю вместе с $|h| = \sqrt{h_1^2 + h_2^2}$. Если выполняются условия Коши — Римана, то полученная дробь приводится к виду

$$u_x + iv_x + \varepsilon_1 \frac{|h|}{h} + i\varepsilon_2 \frac{|h|}{h},$$

и сразу видно, что при $h \to 0$ это выражение стремится к пределу $u_x + iv_x$ совершенно независимо от того, по какому пути совершается предельный переход $h \to 0$.

Отныне мы будем пользоваться условиями Коши — Римана, или эквивалентным им свойством дифференцируемости, как определением аналитической функции, из которого мы выведем все свойства такой функции.

2. Правила дифференцирования. Основные свойства показательной функции. Все многочлены и все степенные ряды (внутри своего круга сходимости) являются аналитическими функциями (согласно § 1, n° 3). Мы знаем, что все действия над величинами, приводящие к элементарным правилам дифференциального исчисления, могут быть выполнены в комплексной области точно тем же путем, как и в действительной. Стало быть, сохраняют силу и правила дифференцирования; в частности, сумма, разность, произведение (а при условии, что знаменатель не обращается в нуль) и частное аналитических функций дифференцируются по тем же элементарным правилам дифференциального исчисления и, следовательно, тоже являются аналитическими функциями. Далее, аналитическую функцию от аналитической функции можно дифференцировать по правилу цепочки, а посему она тоже является аналитической функцией.

Обратим внимание на следующую теорему: *если производная аналитической функции $\zeta = f(z)$ равна нулю всюду в области G, то функция постоянна.*

Доказательство. Имеем $f'(z)=u_x-iu_y=0$ всюду в G. Следовательно, $u_x=0$, $u_y=0$ и, в силу условий Коши — Римана, $v_x=0$ и $v_y=0$. Отсюда вытекает, что u и v постоянны, а стало быть, и функция $\zeta=f(z)$ постоянна.

Приложение к показательной функции. Показательная функция определена нами как сумма степенного ряда $e^z=\sum\limits_{k=0}^{\infty}\frac{z^k}{k!}$, сходящегося во всей плоскости. С другой стороны, в действительной области мы сумели дать и другое определение показательной функции с помощью ее дифференциального уравнения (т. I, стр. 207). Доказанная только что теорема позволяет дать и для показательной функции комплексного аргумента определение с помощью ее дифференциального свойства:

Если комплексная функция $f(z)$ удовлетворяет дифференциальному уравнению

$$f'(z)=f(z), \tag{1}$$

то $f(z)=Ce^z$, где C — произвольная (комплексная) постоянная.

Доказательство. Прежде всего убедимся, что показательная функция $g(z)=e^z$ нигде не обращается в нуль. В самом деле, предположим противное, что она обращается в нуль при некотором значении $z=z_0$, т. е. $g(z_0)=e^{z_0}=0$, и разложим $g(z)$ в ряд Тэйлора в окрестности точки z_0:

$$g(z)=e^z=\sum_{k=0}^{\infty}\frac{(z-z_0)^k}{k!}g^{(k)}(z_0).$$

Так как $g^{(k)}(z)=e^z$ при любом $k=0, 1, 2, \ldots$, то, согласно сделанному предположению, $g^{(k)}(z_0)=e^{z_0}=0$, и следовательно, $e^z=0$ при всех значениях z, что невозможно, ибо $e^0=1$ (см. § 1, n° 4). Итак, функция e^z не обращается в нуль ни при каком значении z.

Приступим теперь к доказательству самой теоремы. Мы уже знаем (§ 1, n° 4), что функция $g(z)=e^z$ удовлетворяет дифференциальному уравнению (1). Обозначим через $f(z)$ любую функцию, удовлетворяющую уравнению (1), так что $f'(z)=f(z)$. Введем вспомогательную функцию $\varphi(z)=\frac{f(z)}{e^z}$. Тогда $\varphi'(z)=\frac{e^zf'(z)-f(z)\cdot e^z}{e^{2z}}=$ $=\frac{f'(z)-f(z)}{e^z}=0$. Стало быть, $\frac{f(z)}{e^z}=C=\text{const}$, откуда $f(z)=Ce^z$, что и требовалось доказать.

Отсюда вытекает *функциональное уравнение* показательной функции:

$$e^ze^{z_1}=e^{z+z_1}.$$

(Когда показательную функцию определяют с помощью степенного ряда, это функциональное уравнение отнюдь не является тривиальным утверждением.) Для доказательства рассмотрим функцию $g(z)=e^{z+z_1}$ при фиксированном значении z_1. В силу правила цепочки, $g(z)$ удовлетворяет дифференциальному уравнению $g'(z)=g(z)$. Стало быть, по только что доказанному, $g(z)=Ce^z$. Для определения постоянной C полагаем $z=0$; так как из степенного ряда, определяющего показательную функцию, вытекает, что $e^0=1$, то

$$g(0)=e^{z_1}=C,$$

откуда и получается подлежавшее доказательству функциональное уравнение.

В § 3, n°3(стр. 563) мы разовьем более удовлетворительный метод исследования показательной функции, не опирающийся на степенной ряд. Здесь мы только отметим одно следствие, вытекающее из функционального уравнения. При $z=x$ и $z_1=iy$, где x и y — действительные числа, получается

$$e^{x+iy}=e^xe^{iy}=e^x(\cos y+i\sin y).$$

Из функционального уравнения еще раз видно, что показательная функция нигде не обращается в нуль. Действительно, если бы существовало такое значение z_1, что $e^{z_1}=0$, то $e^z=e^{z_1}e^{z-z_1}$ обращалось бы в нуль при *всех* значениях z, что заведомо неверно.

Так как $\cos 2\pi=1$ и $\sin 2\pi=0$, то

$$e^{2\pi i}=\cos 2\pi+i\sin 2\pi=1.$$

Поэтому

$$e^{z+2\pi i}=e^z,$$

а стало быть, показательная функция e^z является *периодической* функцией с периодом $2\pi i$.

Упражнение

Доказать, пользуясь условиями Коши — Римана, что произведение и частное аналитических функций, а также аналитическая функция от аналитической функции сами являются аналитическими функциями (в случае частного — в такой области, в которой знаменатель не обращается в нуль).

3. Конформные отображения. Обратные функции. Функции $u(x, y)$ и $v(x, y)$ приводят в соответствие точкам плоскости z (плоскости xy) точки плоскости ζ (плоскости uv). Таким образом, перед нами преобразование или отображение (гл. III, § 3, n° 1) некоторой области плоскости xy на область плоскости uv. Якобиан преобразования есть

$$D=\frac{\partial(u,v)}{\partial(x,y)}=u_xv_y-u_yv_x=u_x^2+u_y^2=|f'(z)|^2.$$

Стало быть, во всякой точке, в которой $f'(z)\neq 0$, якобиан отличен от нуля и к тому же имеет положительное значение. Поэтому если предположить, что $f'(z)\neq 0$ в точке z_0, то результаты, полученные в гл. III, § 3, n° 6, показывают, что достаточно малая окрестность точки z_0 в плоскости z отображается взаимно однозначно и непрерывно на некоторую окрестность точки $\zeta_0=f(z_0)$ в плоскости ζ. Это отображение *конформно*, т. е. не изменяет углов. Это следует из того, что, как мы видели в гл. III, стр. 183, уравнения Коши — Римана представляют собой необходимые и достаточные условия конформности отображения, причем сохраняется не только величина, но и направление углов. Итак, мы пришли к следующему результату:

Конформность отображения, устанавливаемого функциями $u(x, y)$ и $v(x, y)$, и аналитический характер функции $f(z)=u+iv$ означают совершенно одно и то же, если только избегать точек z_0, в которых $f'(z_0)=0$.

Пусть читатель вновь рассмотрит примеры конформного преобразования, приведенные в гл. III, § 3, стр. 155—157, и докажет, что все эти преобразования могут быть выражены несложными аналитическими функциями.

Так как в случае взаимно однозначного конформного отображения окрестности точки z_0 плоскости z на окрестность точки ζ_0 плоскости ζ обратное отображение тоже конформно, то отсюда вытекает, что величину $z=x+iy$ можно тоже рассматривать как аналитическую функцию $\varphi(\zeta)$ комплексной переменной $\zeta=u+iv$. Эта функция $z=\varphi(\zeta)$ называется *обратной* по отношению к функции $\zeta=f(z)$.

Помимо этого геометрического рассуждения, аналитический характер обратной функции можно установить, проверив, что она удовлетворяет условиям Коши — Римана. Вычисляем производные от функций $x(u, v)$ и $y(u, v)$ по формулам дифференцирования функций, осуществляющих обратное преобразование (гл. III, § 3, n° 4):

$$x_u=\frac{v_y}{D},\quad x_v=-\frac{u_y}{D},\quad y_u=-\frac{v_x}{D},\quad y_v=\frac{u_x}{D},$$

и сразу убеждаемся, что обратная функция действительно удовлетворяет уравнениям $x_u=y_v$ и $x_v=-y_u$. Нетрудно установить, что производная от обратной функции $z=\varphi(\zeta)$ связана с производной от исходной функции $\zeta=f(z)$ соотношением

$$\frac{dz}{d\zeta}\frac{d\zeta}{dz}=1.$$

Упражнения

1. Выяснить, в каких точках следующие функции непрерывны: а) $\bar{z}$; б) $|z|$; в) $\frac{z+\bar{z}}{1+|z|}$; г) $\frac{z^2+\bar{z}^2}{|z|^2}$.

2. Какие из функций упр. 1 к тому же и дифференцируемы?

3*. Доказать, что функция вида

$$\zeta=\frac{\alpha z+\bar{\beta}}{\beta z+\bar{\alpha}},$$

где α и β — любые комплексные числа, удовлетворяющие соотношению

$$\alpha\bar{\alpha}-\beta\bar{\beta}=1,$$

преобразует единичную окружность $|z|=1$ самое в себя и ее внутреннюю область тоже самое в себя. Доказать, что, если

$$\beta\bar{\beta}-\alpha\bar{\alpha}=1,$$

то внутренняя область преобразуется во внешнюю.

4. Доказать, что преобразование $\zeta=\frac{1}{2}\left(z+\frac{1}{z}\right)$ переводит окружности плоскости z с центром в начале координат в софокусные эллипсы плоскости ζ, а прямые плоскости z, проходящие через начало, — в гиперболы плоскости ζ (ср. упр. 5, стр. 176).

5. Доказать, что функция $\zeta=\frac{\alpha z+\beta}{\gamma z+\delta}$ оставляет неизменным сложное (ангармоническое) отношение $\frac{z_1-z_3}{z_2-z_3}:\frac{z_1-z_4}{z_2-z_4}$ четырех точек z_1, z_2, z_3, z_4.

6*. Доказать, что любой круг можно отобразить на верхнюю полуплоскость, ограниченную вещественной осью, с помощью функции $w=\frac{\alpha z+\beta}{\gamma z+\delta}$. (Воспользоваться упр. 1, стр. 552.)

7. Доказать следующее свойство общего линейного преобразования:

$$\zeta=\frac{az+b}{cz+d},$$

где a b, c, d — постоянные и $ad-bc\neq 0$.

Это преобразование переводит все окружности и прямые плоскости z в совокупность всех прямых и окружностей плоскости ζ.

Если считать плоскость z и плоскость ζ, а также одноименные оси этих плоскостей совпадающими, то точки z, для которых $\zeta=z$, называются *неподвижными точками* линейного преобразования. Вообще говоря, существуют две различные неподвижные точки. Показать, что в этом случае семейство окружностей, проходящих через обе неподвижные точки, и семейство окружностей, ортогональных к ним, преобразуются каждое само в себя.

8. Функция, обратная степенной функции $\zeta=z^n$ (n — натуральное число), однозначна в окрестности любой точки z_0, если только $z_0\neq 0$, ибо при этом условии производная $\zeta'=nz^{n-1}$ не обращается в нуль. Однако точка $z_0=0$, в которой производная обращается в нуль, представляет исключение; отсюда проистекает многозначность функции $z=\sqrt[n]{\zeta}$. Мы исследуем эти обстоятельства более подробно в § 6.

§ 3. Интегрирование аналитических функций

1. Определение интеграла. Центральным фактом дифференциального и интегрального исчисления функций одной переменной является теорема о том, что определенный интеграл с переменным верхним пределом является первообразной функцией для своей подынтегральной функции (т. I, стр. 137 и 140). Соответствующее свойство образует ядро теории аналитических функций комплексной переменной.

Мы начнем с того, что обобщим понятие определенного интеграла на функцию $f(z)$ комплексной переменной z, т. е. дадим определение этого понятия. Для этого удобно изменить обозначение независимой переменной и вместо z писать $t=r+is$, так что t будет у нас переменной интегрирования. Возьмем функцию $f(t)$, аналитическую в области G, и пусть в этой области даны две точки $t=t_0$ и $t=z$, соединенные ориентированной кусочно гладкой криволинейной дугой C, лежащей целиком внутри G (рис. 107). Точками деления $t_0, t_1, \ldots, t_n=z$ разобьем дугу C на n частей-дужек. На каждой такой дужке выберем произвольную точку t'_k, лежащей на C между ее точками t_{k-1} и t_k, и построим сумму

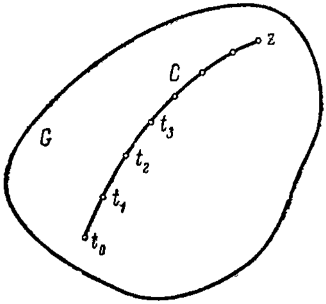


Рис. 107.

$$S_n=\sum_{k=1}^{n} f(t'_k)(t_k-t_{k-1}).$$

Будем теперь изменять наше разбиение, делая его все более и более частым, неограниченно увеличивая число n точек деления таким образом, что наибольшая из хорд $|t_k-t_{k-1}|$ будет стремиться к нулю. Тогда сумма S_n будет стремиться к некоторому пределу, не зависящему от выбора точек деления t_k и промежуточных точек t'_k.

Это можно доказать прямым путем, применяя метод, аналогичный методу доказательства соответствующей теоремы существования определенного интеграла от функции действительной переменной. Однако, для нашей цели будет более удобно привести нашу теорему к тому, что мы уже знаем о криволинейных интегралах в действительной области (гл. V, § 1). Положим $f(t)=u(r, s)+iv(r, s)$, $t_k=r_k+is_k$, $t'_k=r'_k+is'_k$, $\Delta t_k=t_k-t_{k-1}=\Delta r_k+i\Delta s_k$. Тогда

$$S_n=\sum_{k=1}^{n}[u(r'_k, s'_k)\Delta r_k-v(r'_k, s'_k)\Delta s_k]+$$
$$+i\left\{\sum_{k=1}^{n}[v(r'_k, s'_k)\Delta r_k+u(r'_k, s'_k)\Delta s_k]\right\}.$$

При неограниченном возрастании n суммы в правой части имеют своими пределами действительные криволинейные интегралы

$$\int_C (u\,dx-v\,dy) \quad \text{и} \quad \int_C (v\,dx+u\,dy);$$

стало быть, S_n стремится к пределу, что мы и утверждали. Этот предел и называется *определенным интегралом функции* $f(t)$ *вдоль кривой* C от t_0 до z; он обозначается символом

$$\int_{t_0}^{z} f(t)\,dt \quad \text{или} \quad \int_C f(t)\,dt.$$

Таким образом,

$$\int_C f(t)\,dt = \int_C (u\,dx - v\,dy) + i\int_C (v\,dx + u\,dy).$$

Из этого определения сразу вытекает (ср. гл. V, § 1, конец n° 3) следующая важная оценка определенного интеграла: если L есть длина дуги пути интегрирования (дуги C), а $|f(t)| \leqslant M$ на этом пути, где M — постоянная, то

$$\left|\int_C f(t)\,dt\right| \leqslant ML.$$

В дополнение к этому отметим, что действия над комплексными интегралами (в частности, в отношении соединения различных путей интегрирования) подчиняются всем соответствующим правилам, установленным для криволинейных интегралов в гл. V, § 1, n° 3.

[Полезно заметить, что в доказательстве существования интеграла $\int_C f(z)\,dz$ фактически использовано только свойство непрерывности функции $f(z)$. Из свойств криволинейного интеграла в действительной области вытекает, что если $f(z)$ — кусочно непрерывная функция комплексной переменной z, а путь интегрирования C — кусочно гладкая кривая, то $\int_C f(z)\,dz$ существует.]

2. Теорема Коши. Существенным фактом теории функций комплексной переменной является следующая теорема Коши:

Если функция $f(z)$ является аналитической в односвязной области G, то интеграл $\int_{z_0}^{z} f(t)\,dt = \int_C f(t)\,dt$ не зависит от выбора пути интегрирования C, лежащего в области G и содержащего точки z_0 и z; этот интеграл является аналитической функцией $F(z)$, причем

$$\frac{d}{dz}F(z) = \frac{d}{dz}\int_{z_0}^{z} f(t)\,dt = f(z).$$

Это значит, что $F(z)$ является первообразной для подынтегральной функции $f(z)$.

Теорему Коши можно изложить и в следующей, эквивалентной формулировке:

Если функция $f(z)$ является аналитической в односвязной области G, то ее интеграл вдоль любой замкнутой кривой, лежащей в этой области, равен нулю.

Доказательство того, что интеграл не зависит от пути, непосредственно вытекает из основной теоремы о криволинейном интеграле (гл. V, § 1, n° 6 и 7). В самом деле, в силу условий Коши — Римана, оба выражения $u\,dx - v\,dy$ и $v\,dx + u\,dy$, являющиеся подынтегральными функциями в действительной и мнимой части, удовлетворяют условию интегрируемости (стр. 378). Поэтому наш интеграл является функцией от $z = x + iy$, которую можно обозначить через $F(z) = U(x, y) + iV(x, y)$, а из выведенных ранее свойств криволинейного интеграла вытекают следующие соотношения:

$$U_x = u, \qquad U_y = -v, \qquad V_x = v, \qquad V_y = u;$$

это значит, что

$$U_x = V_y, \qquad U_y = -V_x, \qquad U_x + iV_x = u + iv,$$

т. е. $F(z)$ действительно является аналитической функцией в области G, и ее производная $F'(z) = f(z)$.

Односвязность области G является *существенным* условием для справедливости теоремы Коши.

Рассмотрим, например, функцию $\frac{1}{z}$, которая является аналитической всюду в плоскости z, кроме точки $z = 0$. Однако из теоремы Коши мы не в праве делать вывод, что интеграл $\int \frac{dt}{t}$, взятый вдоль замкнутой кривой, содержащей внутри себя начало координат, равен нулю; не в праве потому, что эту кривую невозможно включить в односвязную область, в которой функция $\frac{1}{z}$ была бы аналитической. Односвязность области нарушается исключительной ролью точки $z = 0$. Это можно проверить прямым вычислением. Например, если взять интеграл вдоль окружности K, заданной уравнением $|z| = a$ против часовой стрелки, то на этой окружности можно положить $z = t = ae^{i\theta}$ и принять θ за переменную интегрирования; тогда $dt = aie^{i\theta}\,d\theta$, и мы получим

$$\oint_K \frac{dt}{t} = \int_0^{2\pi} \frac{aie^{i\theta}}{ae^{i\theta}}\,d\theta = 2\pi i,$$

т. е. интеграл равен не нулю, а $2\pi i$.

Однако теорему Коши можно распространить и на многосвязные области следующим образом:

Если граница области G состоит из конечного числа кусочно гладких замкнутых кривых $C_1, C_2, \ldots$ и если функция $f(z)$ является аналитической внутри этой области, а также на ее границе[1]*, то сумма интегралов от функции $f(z)$ вдоль всех граничных кривых равна нулю при условии, что все эти граничные кривые описываются в одном и том же направлении относительно области G, т. е. при обходе всех граничных кривых об-*

[1] Функция называется *аналитической на кривой*, если она является аналитической в некоторой сколь угодно малой окрестности этой кривой.

ласть G расположена всегда по одну и ту же сторону, например по левую сторону. Так, на рис. 108 *имеем*

$$\oint_{C_1}+\oint_{C_2}+\oint_{C_3}=0.$$

Доказательство получается сразу, по образцу соответствующего доказательства для криволинейного интеграла: мы разбиваем многосвязную область G на конечное число односвязных областей (рис. 109)

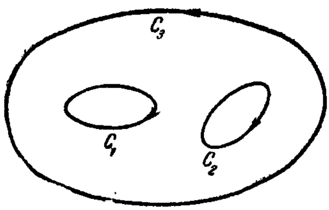


Рис. 108.

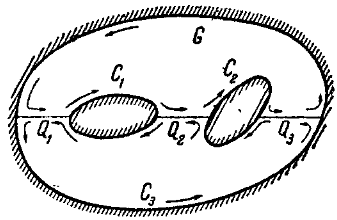


Рис. 109.

с помощью конечного числа разрезов $Q_1, Q_2, \ldots$, применяем теорему Коши к каждой из односвязных областей и все результаты складываем.

Эту теорему можно также изложить и в следующей формулировке: *Если область G образована из внутренней области замкнутой кривой C путем удаления из нее внутренних областей других кривых $C_1, C_2, \ldots$, то*

$$\oint_C f(t)\,dt=\sum_k \oint_{C_k} f(t)\,dt,$$

причем интегралы (по внешней границе C и по внутренним границам $C_1, C_2, \ldots$) должны быть взяты либо все против часовой стрелки, либо все по часовой стрелке (рис. 110).

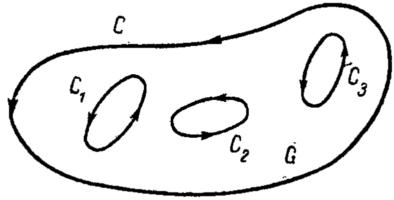


Рис. 110.

3. Приложения. Логарифм, показательная функция и общая степенная функция. На базе теоремы Коши можно теперь построить удовлетворительную теорию логарифма, показательной функции, а затем и других элементарных функций, следуя процедуре, подобной той, которую мы применили для действительной переменной (т. I, гл. III, § 6).

Начнем с того, что дадим определение логарифма как интеграла от функции $\frac{1}{t}$. Разрежем плоскость комплексной переменной вдоль отрицательной действительной полуоси и ограничим сперва путь интегрирования условием, чтобы он лежал в полученной односвязной области, т. е. не будем допускать такого пути интегрирования, который пересекает отрицательную действительную полуось. Точнее, полагая $t = |t|(\cos\theta + i\sin\theta)$, мы ограничим θ неравенством $-\pi < \theta \leqslant \pi$. В разрезанной плоскости комплексной переменной t мы соединим точку $t=1$ с произвольной точкой z любой криволинейной дугой C и будем интегрировать функцию $\frac{1}{t}$ от точки $t=1$ до точки $t=z$; по теореме Коши этот интеграл не зависит от выбора кривой C и представляет собой аналитическую функцию, которую мы обозначим символом $\ln z$:

$$\zeta = \ln z = \int_1^z \frac{dt}{t}.$$

Ясно, что $\ln 1 = 0$. Логарифмическая функция обладает тем свойством, что

$$\frac{d}{dz}\ln z = \frac{1}{z}.$$

Так как эта производная нигде не обращается в нуль, то существует обратная функция от логарифма, $z = g(\zeta)$, для которой $g(0) = 1$, и по формуле для производной от обратной функции

$$g'(\zeta) = \frac{1}{(\ln z)'} = z = g(\zeta).$$

Согласно § 4 (n° 2 и конец n° 3), эта обратная функция совпадает с показательной функцией, определение которой было дано ранее: $g(\zeta) = e^{\zeta}$.

Функция $\ln z$ определена выше однозначно. Мы видели, что $(\ln z)' = \frac{1}{z}$, т. е. $f(z) = \ln z$ удовлетворяет дифференциальному уравнению $f'(z) = \frac{1}{z}$. Нетрудно убедиться, что само это уравнение определяет функцию $f(z)$ и в комплексной области с точностью до произвольной аддитивной постоянной. Действительно, для любого решения $f(z)$ этого уравнения имеем $(f(z) - \ln z)' = \frac{1}{z} - \frac{1}{z} = 0$, откуда (см. § 2, n° 2) $f(z) - \ln z = C = \text{const}$, а стало быть, $f(z) = \ln z + C$.

Интеграл $\ln z = \int_1^z \frac{dt}{t}$ легко вычислить в явном виде, если взять в качестве пути интегрирования линию, состоящую из отрезка действительной оси от точки $t=1$ до точки $t=|z|$ и дуги L окружности $|t| = |z|$ от точки $|z|$ до точки $t=z$ (рис. 111). Тогда интеграл вдоль отрезка оси x будет $\int_1^{|z|} \frac{dt}{t} = \ln|z|$, а на дуге L можно положить $t=|z|e^{i\varphi}$, так что $dt=i|z|e^{i\varphi}d\varphi$ и соответствующий интеграл будет $\int\limits_L \frac{dt}{t}=\int\limits_0^\theta \frac{i|z|e^{i\varphi}}{|z|e^{i\varphi}}d\varphi=i\int\limits_0^\theta d\varphi=i\theta$, где θ есть главное значение аркуса комплексного числа z. В результате имеем

$$\ln z=\ln|z|+i\theta=\ln|z|+i\operatorname{arc} z, \qquad -\pi<\operatorname{arc} z=\theta\leqslant\pi.$$

Полученное таким путем значение логарифма любого комплексного числа называется *главным значением* логарифма. Это название оправдывается тем фактом, что для логарифма можно получить и

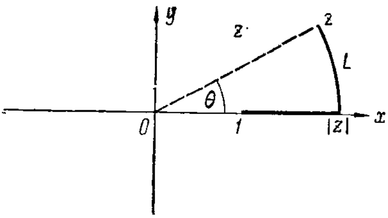


Рис. 111.

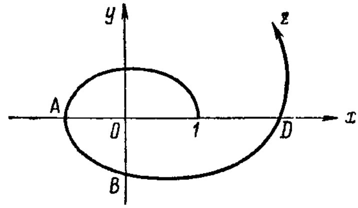


Рис. 112.

другие значения, если отказаться от условия, что путь интегрирования не должен пересекать отрицательной действительной полуоси. Тогда можно путь интегрирования от точки 1 до точки z провести таким образом, что он обойдет вокруг начала $t=0$. Возьмем, например, путь $1ABDz$ (рис. 112). Для этого пути будет

$$\int\limits_{1ABDz}\frac{dt}{t}=\int\limits_{1ABD1}\frac{dt}{t}+\int\limits_{1Dz}\frac{dt}{t}.$$

Первый интеграл в правой части, согласно теореме в конце n° 2, равен интегралу $\int\frac{dt}{t}$ вдоль окружности малого радиуса a с центром в начале, т. е. равен $2\pi i$ (см. пример, набранный петитом в n° 2). Второй интеграл нам уже тоже известен:

$$\int\limits_{1Dz}\frac{dt}{t}=\ln z=\ln|z|+i\theta.$$

В итоге получаем: вдоль пути, обходящего начало координат один раз против часовой стрелки,

$$\int\limits_1^z\frac{dt}{t}=\ln|z|+i\theta+2\pi i=\ln z+2\pi i.$$

Можно, очевидно, взять путь интегрирования, обходящий сколько угодно раз вокруг точки $z=0$ в положительном или отрицательном направлении. Тогда получится

$$\operatorname{Ln} z=\ln|z|+i\theta+2n\pi i=\ln z+2n\pi i,$$

где n — любое целое число, положительное, отрицательное или нуль. Эта формула и выражает тот факт, что *логарифм является многозначной функцией*, и для обозначения множества всех его значений мы будем пользоваться символом $\operatorname{Ln} z$ (с прописной буквой L).

Очевидно, функция $f(z) = \operatorname{Ln} z$, как и $\ln z$, удовлетворяет дифференциальному уравнению $f'(z) = \frac{1}{z}$. Пользуясь этим результатом легко вывести формулу для логарифма произведения

$$\operatorname{Ln}(az) = \ln z + \operatorname{Ln} a.$$

Для этого заметим, что по правилу цепочки $\frac{d}{dz}[\operatorname{Ln}(az)] = \frac{1}{az} \cdot a = \frac{1}{z}$. Следовательно, функция $\operatorname{Ln}(az)$ является решением дифференциального уравнения $f'(z) = \frac{1}{z}$, и поэтому (ср. стр. 564) $\operatorname{Ln}(az) = \ln z + C$. Постоянную C определяем, подставляя сюда значение $z = 1$, так что $\ln z = \ln 1 = 0$, откуда получается $\operatorname{Ln} a = C$, и окончательно: $\operatorname{Ln}(az) = \ln z + \operatorname{Ln} a$, что и требовалось доказать.

[Ясно, что можно также писать

$$\operatorname{Ln}(az) = \operatorname{Ln} z + \ln a.$$

Обе эти записи говорят, что все множество значений логарифма произведения можно получить, складывая главное значение логарифма одного из сомножителей с каждым значением логарифма другого сомножителя. Верно, очевидно, и равенство

$$\operatorname{Ln}(az) = \operatorname{Ln} z + \operatorname{Ln} a,$$

понимаемое в том смысле, что, складывая любое значение $\operatorname{Ln} z$ с любым значением $\operatorname{Ln} a$, получим все множество значений $\operatorname{Ln}(az)$. Однако нельзя утверждать, что

$$\ln(az) = \ln z + \ln a;$$

сумма в правой части равна, конечно, одному из значений $\operatorname{Ln}(az)$, но не обязательно главному. Последняя формула верна лишь в том случае, если сумма $\operatorname{arc} z + \operatorname{arc} a$ удовлетворяет условию

$$-\pi < \operatorname{arc} z + \operatorname{arc} a \leqslant \pi.$$

В частности, если a есть положительное действительное число, то она верна при всяком комплексном z, отличном от нуля.]

Эта многозначность логарифма приводит к важному свойству показательной функции, выражаемому равенством $e^{2n\pi i} = 1$ (в частности, $e^{2\pi i} = 1$). Дело в том, что одно и то же значение z соответствует всем различным значениям $\zeta = \operatorname{Ln} z$, которые разнятся друг от друга лишь слагаемым, равным целому кратному числа $2\pi i$. Поэтому функция, обратная логарифму, т. е. показательная функция $g(\zeta) = e^{\zeta}$, не изменяется, когда ее аргумент увеличивается или уменьшается на $2\pi i$: $g(\zeta + 2\pi i) = g(\zeta)$ или $e^{\zeta + 2\pi i} = e^{\zeta}$. При $\zeta = 0$ получается $e^{2\pi i} = 1$. Итак, показательная функция комплексного аргумента является периодической с периодом $2\pi i$. Введем теперь тригонометрические функции комплексной переменной, $\cos z$ и $\sin z$, с помощью равенств

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i},$$

которые теперь служат определением этих функций. Из этого определения сразу следует, что тригонометрические функции являются периодическими и имеют период 2π. Таким образом, мы вывели периодичность тригонометрических функций, не опираясь на их элементарные геометрические определения. Кстати, из этих формул вытекает тождество $e^{iz} = \cos z + i \sin z$.

Располагая определениями логарифмической и показательной функции, легко теперь ввести общую показательную функцию a^z и общую степенную функцию z^α, где a и α — любые комплексные постоянные (ср. соответствующее построение для действительной области в т. I, гл. III, § 6, n° 5). Функцию a^z мы определим равенством

$$a^z = e^{z \ln a},$$

в котором берется главное значение логарифма постоянной a. Аналогичное определение мы дадим функции z^α:

$$z^\alpha = e^{\alpha \operatorname{Ln} z}.$$

Между тем как функция a^z определена однозначно, если в ее определении берется главное значение $\ln a$, функция z^α является многозначной функцией. В силу многозначности функции $\operatorname{Ln} z$, ясно, что наряду с каким-либо одним значением функции z^α мы получим все другие ее значения, умножая это исходное значение на $e^{2n\pi i\alpha}$, где n — любое положительное или отрицательное целое число. Если постоянное число α рационально, скажем, $\alpha = \frac{p}{q}$, где p и q — взаимно простые целые числа, то среди значений множителя $e^{2n\pi i\alpha}$ окажется лишь конечное число различных — это все значения корня q-й степени из единицы. Если же α — число иррациональное, то количество различных значений этого множителя бесконечно. Многозначность функции z^α будет исследована подробнее в § 6.

Пользуясь правилом цепочки, легко вывести формулы для производных от функций a^z и z^α:

$$\frac{d(a^z)}{dz} = a^z \ln a, \qquad \frac{d(z^\alpha)}{dz} = \alpha z^{\alpha-1}.$$

Упражнения

1. *Гамма-функция.* Доказать, что интеграл $\Gamma(z) = \int\limits_0^\infty t^{z-1} e^{-t}\, dt$ (с главным значением функции t^{z-1}), распространенный на все положительные действительные значения переменной интегрирования t, является аналитической

функцией параметра $z=x+iy$, если $x>0$. (Показать прямо, что функцию $\Gamma(z)$ можно дифференцировать по z.) Доказать, что гамма-функция, определенная таким образом для комплексного аргумента z, удовлетворяет функциональному уравнению $\Gamma(z+1)=z\Gamma(z)$.

2*. *Дзета-функция Римана.* Составить бесконечный ряд

$$\sum_{n=1}^{\infty}\frac{1}{n^z}=\zeta(z),$$

беря главное значение функции n^z. Доказать, что этот ряд сходится, если $\mathrm{Re}\,z=x>1$, и представляет собой дифференцируемую функцию. (Эта функция $\zeta(z)$ называется *дзета-функцией* Римана). Доказательство можно выполнить прямым путем, наподобие доказательства для степенного ряда (ср. т. I, стр. 443).)

§ 4. Интегральная формула Коши и ее приложения

1. Формула Коши. Теорема Коши для многосвязной области приводит к замечательной формуле, тоже принадлежащей Коши. Эта формула выражает значение аналитической функции $f(z)$ в любой точке z_0, лежащей внутри замкнутой области G, в которой эта функция является аналитической, через значения, принимаемые функцией на границе области.

Пусть функция $f(z)$ является аналитической в односвязной области G и на ее границе C. Тогда функция

$$g(z)=\frac{f(z)}{z-z_0}$$

является аналитической во всех точках области G и ее границы C, за исключением точки $z=z_0$. Вырежем из области G круг малого радиуса ρ с центром в точке $z=z_0$, лежащий полностью внутри G (рис. 113), а затем применим к функции $g(z)$ теорему Коши в той ее формулировке для многосвязной области, которая дана в конце n° 2:

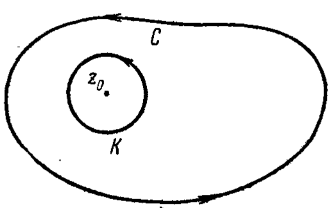


Рис. 113.

$$\oint_C g(z)\,dz=\oint_K g(z)\,dz,$$

причем оба интеграла будем брать, обходя границу C области G и окружность K вырезанного круга против часовой стрелки. На окружности K имеем $z=z_0+\rho e^{i\theta}$, где угол θ определяет положение переменной точки на окружности. Поэтому на окружности $dz=i\rho e^{i\theta}\,d\theta$, откуда

$$\oint_K g(z)\,dz=i\int_0^{2\pi}f(z_0+\rho e^{i\theta})\,d\theta.$$

Так как функция $f(z)$ непрерывна в точке z_0, то при достаточно малом радиусе ρ

$$f(z_0+\rho e^{i\theta})=f(z_0)+\eta,$$

где $|\eta|$ меньше любого наперед заданного положительного числа ε. Стало быть,

$$\int_0^{2\pi} f(z_0+\rho e^{i\theta})\,d\theta=\int_0^{2\pi} f(z_0)\,d\theta+\int_0^{2\pi}\eta\,d\theta=2\pi f(z_0)+\int_0^{2\pi}\eta\,d\theta.$$

Но

$$\left|\int_0^{2\pi}\eta\,d\theta\right|\leqslant\int_0^{2\pi}\varepsilon\,d\theta=2\pi\varepsilon.$$

Поэтому

$$\int_0^{2\pi} f(z_0+\rho e^{i\theta})\,d\theta=2\pi f(z_0)+\alpha,\quad \text{где}\quad |\alpha|\leqslant 2\pi\varepsilon.$$

Таким образом, если радиус ρ достаточно мал, то

$$\oint_C g(z)\,dz=2\pi i f(z_0)+\alpha i,\quad \text{где}\quad |\alpha i|\leqslant 2\pi\varepsilon.$$

Заставим теперь ε стремиться к нулю, безгранично уменьшая ρ; тогда правая часть последнего равенства стремится к пределу $2\pi i f(z_0)$, между тем как левая часть, а именно $\int_C g(z)\,dz$, остается неизменной:

$$\oint_C g(z)\,dz=\oint_C \frac{f(z)}{z-z_0}\,dz=2\pi i f(z_0).$$

Следовательно,

$$f(z_0)=\frac{1}{2\pi i}\oint_C \frac{f(z)}{z-z_0}\,dz.$$

Так как z_0 есть любая внутренняя точка области G, то мы индекс 0 опустим и будем писать z вместо z_0, а переменную интегрирования обозначим вместо z через ζ; тогда формула Коши примет следующий вид:

$$f(z)=\frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta-z}\,d\zeta.$$

Это и есть *интегральная формула Коши*, имеющая фундаментальное значение.

Эта формула выражает значения функции внутри замкнутой области, в которой эта функция является аналитической, через значения, принимаемые ею на границе области.

Если, в частности, граничная кривая C есть окружность $\zeta = z + re^{i\theta}$ с центром в точке z, то $d\zeta = ire^{i\theta}\,d\theta$, и мы получим

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} f(z + re^{i\theta})\,d\theta.$$

Стало быть, *если функция является аналитической в некотором круге и на его границе, то значение функции в центре круга равно среднему значению функции на его окружности.*

2. Разложение аналитической функции в степенной ряд. Интегральная формула Коши имеет много важных теоретических приложений. Главным из этих приложений является доказательство теоремы, что *всякая аналитическая функция может быть разложена в степенной ряд*; эта теорема устанавливает связь между излагаемой нами теперь теорией и материалом, данным в § 1. Точная формулировка теоремы такова:

Если функция $f(z)$ является аналитической на круге $|z - z_0| \leqslant R$ (включая его границу), то она может быть разложена в степенной ряд по степеням $z - z_0$, сходящийся внутри этого круга.

При доказательстве можно, не теряя общности, положить $z_0 = 0$, так как общий случай приводится подстановкой $z - z_0 = Z$ к случаю $Z_0 = 0$. Применим теперь интегральную формулу Коши к граничной окружности $C: |z| = R$ и запишем подынтегральную функцию (пользуясь геометрической прогрессией) в следующем виде:

$$\frac{f(\zeta)}{\zeta - z} = \frac{f(\zeta)}{\zeta}\,\frac{1}{1 - \frac{z}{\zeta}} = \frac{f(\zeta)}{\zeta}\left(1 + \frac{z}{\zeta} + \ldots + \frac{z^n}{\zeta^n}\right) + \frac{f(\zeta)}{\zeta}\left(\frac{z}{\zeta}\right)^{n+1}\frac{1}{1 - \frac{z}{\zeta}}.$$

Так как точка z лежит внутри окружности, а точка ζ — на ней самой, то $\left|\frac{z}{\zeta}\right| = q < 1$, и для остаточного члена геометрического ряда $r_n = \frac{1}{\zeta}\left(\frac{z}{\zeta}\right)^{n+1}\frac{1}{1 - \frac{z}{\zeta}}$ получается такая оценка:

$$|r_n| \leqslant \frac{1}{R}\,q^{n+1}\,\frac{1}{1-q}.$$

Внеся выражение для подынтегральной функции в формулу Коши и интегрируя почленно, имеем

$$f(z) = c_0 + c_1 z + c_2 z^2 + \ldots + c_n z^n + R_n,$$

где

$$c_k = \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta^{k+1}}\,d\zeta,$$

$$R_n = \frac{1}{2\pi i}\oint_C f(\zeta)\,r_n\,d\zeta.$$

Пусть M — верхняя граница значений $|f(\zeta)|$ на окружности круга; тогда формула оценки комплексного интеграла (§ 3, конец n° 1) дает сразу следующую оценку модуля остаточного члена:

$$|R_n| \leqslant \frac{1}{2\pi R}\frac{q^{n+1}}{1-q}2\pi RM = \frac{q^{n+1}}{1-q}M.$$

Так как q есть положительное число, меньшее единицы, то $R_n \to 0$ при $n \to \infty$, а стало быть, функция $f(z)$ разложена в бесконечный степенной ряд

$$f(z) = \sum_{k=0}^{\infty} c_k z^k$$

с коэффициентами

$$c_k = \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta^{n+1}}\,d\zeta.$$

Тем самым теорема доказана.

Эта теорема имеет важные следствия. Во-первых, из § 1, n° 3 мы знаем, что всякий степенной ряд можно дифференцировать почленно внутри его круга сходимости сколько угодно раз, а так как всякую аналитическую функцию можно представить в виде степенного ряда, то отсюда вытекает, что *внутри области, в которой функция является аналитической, ее производная* в свою очередь дифференцируема, т. е. также является *аналитической функцией*. Другими словами, *операция дифференцирования не выводит из класса аналитических функций*. Во-вторых, мы уже знаем, что степенной ряд можно также интегрировать внутри его круга сходимости. Следовательно, *как дифференцирование, так и интегрирование аналитических функций всегда выполнимо без каких-либо ограничений*. Это, конечно, приятная ситуация, которая не имеет места у функций действительной переменной.

В-третьих, в § 1 (в конце n° 3) мы видели, что всякий степенной ряд является рядом Тэйлора для функции, которую он представляет. Отсюда вытекает общее положение, что всякая функция, аналитическая в некоторой области, может быть разложена в окрестности любой внутренней точки z_0 этой области в ряд Тэйлора

$$f(z) = f(z_0) + \sum_{k=1}^{\infty}\frac{f^{(k)}(z_0)}{k!}(z-z_0)^k,$$

коэффициенты которого выражаются формулами

$$c_k = \frac{f^{(k)}(z_0)}{k!} = \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{(\zeta-z_0)^{k+1}}\,d\zeta,$$

где интеграл берется по окружности C, внутри которой функция $f(z)$ является аналитической.

Из полученных результатов можно также вывести правило для суждения о радиусе сходимости ряда Тэйлора. А именно, ряд Тэйлора, представляющий функцию $f(z)$ в окрестности точки $z=z_0$, обязательно сходится внутри наибольшего круга с центром в z_0, лежащего полностью в той области, в которой функция определена и является аналитической.

В силу теорем о дифференцировании и интегрировании степенного ряда, справедливость которых мы теперь установили также и для комплексной области, все те элементарные функции, которые мы разложили в ряд Тейлора в действительной области, имеют точно тот же самый ряд Тэйлора и в комплексной области. Для большинства этих функций мы уже ранее убедились в справедливости этого положения.

Здесь можно, например, отметить, что разложение в биномиальный ряд

$$(1+z)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k$$

сохраняет силу и для комплексного аргумента z, если $|z|<1$, при условии что функция

$$(1+z)^\alpha = e^{\alpha \ln(1+z)}$$

построена с помощью главного значения логарифма $\ln(1+z)$.

Тот факт, что радиус сходимости этого ряда равен единице, непосредственно вытекает из сказанного выше, если учесть, что функция $(1+z)^\alpha$ уже не является аналитической в точке $z=-1$. Действительно, если бы эта функция была аналитической при $z=-1$, то она имела бы в этой точке производные любого порядка, что явно не имеет места. Поэтому круг радиуса 1 с центром в точке $z=0$ есть наибольший круг, внутри которого функция еще является аналитической.

Как мы уже указывали в гл. VIII первого тома, стр. 474, поведение степенного ряда в отношении сходимости становится вполне понятным в свете только что доказанного факта о круге сходимости.

Например, то обстоятельство, что геометрический ряд, представляющий функцию $\frac{1}{1+z^2}$, перестает сходиться на окружности $|z|=1$, есть простое следствие того факта, что эта функция уже не является аналитической в точках $z=i$ и $z=-i$. Теперь нам столь же ясно, что степенной ряд

$$\frac{z}{e^z-1} = \sum_{k=0}^{\infty} \frac{B_k z^k}{k!},$$

определяющий числа Бернулли (ср. т. I, стр. 478), должен иметь круг сходимости, ограниченный окружностью $|z|=2\pi$, ибо знаменатель функции обращается в нуль в точке $z=2\pi i$, но нигде не обращается в нуль внутри окружности $|z|=2\pi$ (если не считать начала координат).

Упражнение

Доказать, что аналитическая функция имеет производные любого порядка, опираясь на интегральную формулу Коши (стало быть, без прямого использования степенного ряда).

3. Теория аналитических функций и теория потенциала. Из того факта, что аналитическую функцию можно дифференцировать сколько угодно раз, вытекает, что ее действительная часть $u(x, y)$ и мнимая часть $v(x, y)$ имеют непрерывные частные производные любого порядка. Поэтому уравнения Коши — Римана

$$u_x = v_y, \qquad u_y = -v_x$$

можно дифференцировать. Продифференцируем первое уравнение по x, а второе — по y и сложим результаты; тогда получится уравнение Лапласа

$$\nabla^2 u \equiv u_{xx} + u_{yy} = 0.$$

Аналогичным путем обнаруживается, что мнимая часть $v(x, y)$ удовлетворяет тому же уравнению:

$$\nabla^2 v \equiv v_{xx} + v_{yy} = 0.$$

Другими словами, *как действительная часть, так и мнимая часть аналитической функции удовлетворяет уравнению Лапласа*, т. е. обе они являются потенциальными или гармоническими функциями. Если две гармонические функции u и v удовлетворяют уравнениям Коши — Римана, то они называются сопряженными гармоническими функциями.

Отсюда видно, что теория функций комплексной переменной и теория двумерного потенциала по существу эквивалентны.

Упражнение

Показать, что для всякой функции $u(x, y)$, удовлетворяющей уравнению Лапласа, можно построить сопряженную ей функцию $v(x, y)$, и притом однозначным образом, если не считать аддитивной постоянной.

4. Теорема, обратная теореме Коши. *Если функция $w = f(z)$ непрерывна в односвязной области G и обладает тем свойством, что ее интеграл вдоль любой замкнутой кривой, лежащей в G, равен нулю, то функция $f(z)$ является аналитической в области G.*

Для доказательства достаточно заметить следующее. Из условия теоремы сразу вытекает, что интеграл $\int\limits_{z_0}^{z} f(\zeta)\, d\zeta$, взятый вдоль любого пути, идущего от фиксированной точки z_0 до переменной точки z области G, является дифференцируемой функцией $F(z)$, причем $F'(z) = f(z)$.

[Что производная $F'(z) = \frac{d}{dz}\int\limits_{z_0}^{z} f(\zeta)\, d\zeta$ существует и равна $f(z)$, доказывается тем же путем, что и соответствующая теорема для криволинейного интеграла в действительной области. См., например, Лаврентьев М. А. и Шабат Б. В., Методы теории функций комплексного переменного, изд. 2, М., 1958.] Стало быть, $F(z)$ есть аналитическая функция, и ее производная $F'(z)=f(z)$, согласно доказанному выше, тоже является аналитической функцией.

Эта теорема, обратная теореме Коши, показывает, что при построении теории аналитических функций можно было бы заменить требование дифференцируемости требованием интегрируемости. Эквивалентность этих двух требований представляет собой весьма характерную черту теории функций комплексной переменной.

5. Нули, полюсы и вычеты аналитической функции. Если аналитическая функция $f(z)$ обращается в нуль в точке $z=z_0$, то постоянный член $f(z_0)$ в разложении этой функции в ряд Тэйлора по степеням $z-z_0$ равен нулю; может случиться, что, кроме него, еще несколько следующих по порядку коэффициентов ряда Тэйлора обращаются в нуль. Предположим, что первым не исчезающим членом ряда будет член, содержащий $(z-z_0)^n$. Тогда можно вынести $(z-z_0)^n$ за скобку и написать

$$f(z)=(z-z_0)^n g(z),$$

причем $g(z_0)\neq 0$. В точке z_0, очевидно,

$$f(z_0)=f'(z_0)=f''(z_0)=\ldots=f^{(n-1)}(z_0)=0, \quad \text{а} \quad f^{(n)}(z_0)\neq 0.$$

Точка z_0, обладающая таким свойством, называется *нулем n-го порядка функции* $f(z)$.

Функция $\frac{1}{f(z)}=\varphi(z)$, как мы знаем, тоже является аналитической функцией в области G, но уже за исключением тех точек этой области, в которых $f(z)$ обращается в нуль. Если z_0 есть нуль n-го порядка функции $f(z)$, то функция $\varphi(z)$ может быть представлена в окрестности точки z_0 в следующем виде:

$$\varphi(z)=\frac{1}{(z-z_0)^n g(z)}=\frac{1}{(z-z_0)^n}\psi(z),$$

где $\psi(z)$ — аналитическая функция в окрестности точки z_0; функция же $\varphi(z)$ перестает быть аналитической в точке $z=z_0$. Эта точка называется *полюсом n-го порядка* функции $\varphi(z)$. Полюс принадлежит к числу так называемых *особенностей* или *особых точек* функции. Если функцию $\psi(z)$ разложить по степеням $z-z_0$ и полученный ряд разделить почленно на $(z-z_0)^n$, то для $\varphi(z)$ получится в окрестности полюса z_0 разложение в ряд вида

$$\varphi(z)=c_{-n}(z-z_0)^{-n}+\ldots+c_{-1}(z-z_0)^{-1}+c_0+$$
$$+c_1(z-z_0)+c_2(z-z_0)^2+\ldots, \qquad (A)$$

где коэффициенты при различных степенях $(z-z_0)$ обозначены через $c_{-n}, \ldots, c_{-1}, c_0, c_1, c_2, \ldots$ [Ряд (A) называется *рядом Лорана.*]

Если мы имеем дело с полюсом первого порядка, т. е. $n=1$, то коэффициент c_{-1} получается сразу из соотношения

$$c_{-1}=\lim_{z\to z_0}(z-z_0)\varphi(z).$$

Так как в этом случае

$$\frac{1}{(z-z_0)\varphi(z)}=\frac{f(z)}{z-z_0}=\frac{f(z)-f(z_0)}{z-z_0},$$

то

$$c_{-1}=\frac{1}{f'(z_0)}.$$

Подобным же образом, если $\varphi(z)=\frac{r(z)}{f(z)}$, а $f(z)$ имеет нуль первого порядка в точке z_0, между тем как $r(z_0)\neq 0$, то имеем

$$c_{-1}=\frac{r(z_0)}{f'(z_0)}.$$

Если функция $\varphi(z)$ определена и является аналитической в некоторой окрестности точки z_0, но не в самой этой точке, то ее интеграл вдоль замкнутой кривой, содержащей эту точку внутри себя, вообще говоря, не будет равен нулю. Однако, по теореме Коши, этот интеграл не зависит от выбора кривой и значение его одинаково для всех замкнутых кривых достаточно малого поперечника, окружающих точку z_0. Значение интеграла от такой функции $\varphi(z)$ вдоль замкнутой кривой C, окружающей точку z_0, взятого в положительном направлении [т. е. против часовой стрелки], называется *вычетом* функции в точке z_0, т. е. в полюсе (говорят также: относительно полюса z_0). [Здесь автор отступает от установившейся традиции: обычно вычетом функции $\varphi(z)$ относительно точки z_0 называют не само значение интеграла $\oint\limits_C \varphi(z)\,dz$, а его значение, деленное на $2\pi i$.]

Пусть точка z_0 есть полюс n-го порядка функции $f(z)$. Тогда разложение функции $f(z)$ по степеням $z-z_0$ будет иметь вид (A). Интеграл от функции $f(z)$ вдоль замкнутой кривой C будет состоять из трех слагаемых: 1) интеграла от ряда вида $\sum\limits_{k=0}^{\infty} c_k(z-z_0)^k$, 2) интеграла от члена $c_{-1}(z-z_0)^{-1}$ и 3) интеграла от членов $c_{-n}(z-z_0)^{-n}+\ldots+c_{-2}(z-z_0)^{-2}$ с отрицательными показателями при $(z-z_0)$. Первое слагаемое равно нулю, так как подлежащий интегрированию степенной ряд представляет функцию аналитическую также и в точке z_0. Интеграл от члена $c_{-1}(z-z_0)^{-1}$ равен, как известно, $2\pi i c_{-1}$. Третье же слагаемое есть сумма интегралов от членов вида $c_{-k}(z-z_0)^{-k}$ при $k>1$ и равно нулю по следующей причине. Первообразная от $(z-z_0)^{-k}$ есть $\frac{(z-z_0)^{-k+1}}{1-k}$, как и в действительной области, так что при $k>1$ интеграл вдоль замкнутой кривой равен нулю.

Итак, *вычет функции в ее полюсе равен* $2\pi i c_{-1}$.

В следующем параграфе мы познакомимся с предложениями понятия вычета и убедимся в его плодотворности. Эти приложения основываются на следующей теореме:

Теорема вычетов. *Если функция $f(z)$ является аналитической в области G и на ее границе C, за исключением конечного числа полюсов, лежащих внутри области, то интеграл от этой функции вдоль граничной кривой C в положительном ее направлении равен сумме вычетов функции $f(z)$ во всех ее полюсах, лежащих внутри области G.*

Упражнения

1*. Показать, что функция

$$f(z)-\frac{1}{2\pi i}\oint\frac{f(\zeta)}{\zeta-z}\frac{z^n}{\zeta^n}\,d\zeta,$$

где интеграл берется вдоль простого контура, окружающего точки $\zeta=0$ и $\zeta=z$, есть многочлен $g(z)$ степени $n-1$, для которого $g^{m}(0)=f^{(m)}(0)$ при $m=0, 1, \ldots, n-1$.

2. Пусть $f(z)$ — аналитическая функция при $|z|\leqslant\rho$. Доказать, что если M есть наибольшее значение модуля функции $f(z)$ на окружности $|z|=\rho$, то коэффициенты степенного ряда для этой функции,

$$f(z)=\sum_{k=0}^{\infty}a_kz^k,$$

удовлетворяют неравенству $|a_k|\leqslant\frac{M}{\rho^k}$.

3*. Доказать, что если область ограничена одной-единственной замкнутой кривой C и если функция $f(z)$ — аналитическая внутри C и на самой кривой C и не обращается в нуль на C, то выражение

$$\frac{1}{2\pi i}\oint_C\frac{f'(z)}{f(z)}\,dz$$

равно числу нулей функции $f(z)$ внутри C.

4. а) Два многочлена $P(z)$ и $Q(z)$ таковы, что во всякой точке некоторой замкнутой кривой C

$$|Q(z)|<|P(z)|.$$

Доказать, что уравнения $P(z)=0$ и $P(z)+Q(z)=0$ имеют одинаковое число корней внутри кривой C. (Рассмотреть семейство функций $P(z)+\theta Q(z)$, где параметр θ изменяется от 0 до 1.)

б) Доказать, что все корни уравнения

$$z^5+az+1=0$$

лежит внутри окружности $|z|=r$, если $|a|<r^4-\frac{1}{r}$.

5. Пусть уравнение $f(z)=0$ имеет простой корень a внутри замкнутой кривой C. Доказать, что этот корень дается формулой

$$a=\frac{1}{2\pi i}\oint_C z\frac{f'(z)}{f(z)}\,dz.$$

§ 5. Приложение к вычислению действительных определенных интегралов

Интегральная теорема Коши и теорема вычетов дают часто возможность применить комплексное интегрирование по контуру (т. е. вдоль замкнутой кривой) к вычислению действительных определенных интегралов. Такой интеграл рассматривается при этом как интеграл от функции комплексного аргумента, взятый вдоль действительной оси. Таким путем порою получают поразительно изящное вычисление сложных на вид определенных интегралов, не нуждаясь в соответствующей первообразной функции, которую в таких случаях обычно и найти-то не удается. Мы рассмотрим здесь несколько типичных примеров.

1. Вывод формулы $\int\limits_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$. Мы дадим здесь поучительное доказательство этой важной формулы; напомним, что этот интеграл мы уже исследовали в т. I, стр. 292 и 485—486, и вычислили другими способами (т. 1, стр. 527; т. II, стр. 337—338). Будем интегрировать функцию $\frac{e^{iz}}{z}$ в плоскости комплексной переменной z вдоль замкнутого пути C, обходя последовательно (рис. 114) отрезок I_1 оси x, полуокружность H_r радиуса r, отрезок I_2 оси x и полуокружность H_R радиуса R в направлении, показанном на рисунке стрелками. Обе полуокружности имеют своим центром начало координат. Соединяя вместе интегралы вдоль отрезков I_1 и I_2, имеем по теореме Коши

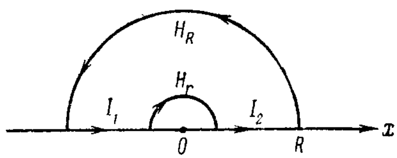


Рис. 114.

$$\int\limits_{H_R} \frac{e^{iz}}{z}\,dz + \int\limits_{H_r} \frac{e^{iz}}{z}\,dz + 2i\int\limits_r^R \frac{\sin x}{x}\,dx = 0.$$

Заставим теперь R стремиться к бесконечности, а r — к нулю. Тогда интеграл по полуокружности H_R стремится к нулю. В самом деле, положим на этой полуокружности $z = Re^{i\theta} = R(\cos\theta + i\sin\theta)$; тогда $e^{iz} = e^{iR\cos\theta}e^{-R\sin\theta}$ и $dz = iRe^{i\theta}\,d\theta$. Интеграл вдоль H_R обратится в определенный интеграл $i\int\limits_0^\pi e^{iR\cos\theta}e^{-R\sin\theta}\,d\theta$. Модуль множителя $e^{iR\cos\theta}$ равен единице, а модуль множителя $e^{-R\sin\theta}$ меньше единицы во всяком интервале $\varepsilon \leqslant \theta \leqslant \pi - \varepsilon$ $(\varepsilon > 0)$; более того, при таких значениях θ $|e^{-R\sin\theta}|$ стремится равномерно к нулю при $R \to \infty$. Отсюда сразу видно, что интеграл вдоль H_R стремится к нулю при $R \to \infty$. Интеграл вдоль другой полуокружности H_r имеет своим пределом $-\pi i$ при $r \to 0$, что читатель сам легко докажет. Третий интеграл в последнем равенстве, при $R \to \infty$ и $r \to 0$, имеет своим пределом $2i \int\limits_0^\infty \frac{\sin x}{x}\, dx$. Стало быть, получаем в пределе равенство

$$-\pi i + 2i \int\limits_0^\infty \frac{\sin x}{x}\, dx = 0, \text{ откуда } \int\limits_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

2. Доказательство формулы $\int\limits_0^\infty e^{-x^2} \cos ax\, dx = \frac{1}{2}\sqrt{\pi}\, e^{-\frac{1}{4}a^2}$

Проинтегрируем комплексную функцию e^{-z^2} по периметру прямоугольника $ABB'A'$ (рис. 115) с вершинами $A(-b,\ 0)$, $B(b,\ 0)$, $B'\left(b,\ \frac{a}{2}\right)$, $A'\left(-b,\ \frac{a}{2}\right)$, так что вертикальные стороны $AA' = BB' = \frac{a}{2}$, а горизонтальные стороны $AB = A'B' = 2b$. По теореме Коши этот интеграл равен нулю. На вертикальных сторонах имеем

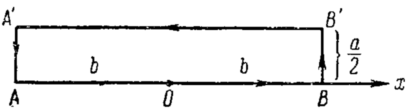


Рис. 115.

$$|e^{-z^2}| = |e^{-(x+iy)^2}| = |e^{-(x^2-y^2)} e^{-2ixy}| = e^{-b^2} e^{y^2} < e^{-b^2} e^{\frac{1}{4}a^2},$$

а последнее выражение стремится равномерно к нулю при $b \to \infty$. Поэтому $\int\limits_{BB'} e^{-z^2} dz \to 0$ и $\int\limits_{A'A} e^{-z^2} dz \to 0$. Остаются интегралы $\int\limits_{AB}$ и $\int\limits_{B'A'}$. Если выполнить переход к пределу $b \to \infty$, то $\int\limits_{AB} e^{-z^2} dz \to \int\limits_{-\infty}^{\infty} e^{-x^2} dx$, а интеграл вдоль $B'A'$, поскольку на этой стороне $z = x + \frac{1}{2} ia$ и $dz = dx$, будет $\int\limits_b^{-b} e^{-\left(x+\frac{1}{2}ia\right)^2} dx \to \int\limits_{\infty}^{-\infty} e^{-\left(x+\frac{1}{2}ia\right)^2} dx$.

Поэтому результат применения теоремы Коши можно записать так:

$$\int\limits_{-\infty}^{\infty} e^{-x^2} dx + \int\limits_{\infty}^{-\infty} e^{-\left(x+\frac{1}{2}ia\right)^2} dx = 0$$

или

$$\int\limits_{-\infty}^{\infty} e^{-\left(x+\frac{1}{2}ia\right)^2} dx = \int\limits_{-\infty}^{\infty} e^{-x^2} dx.$$

Это значит, что путь интегрирования (ось x) можно сместить параллельно самому себе, не изменяя этим результата. Но $\int\limits_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

(см. гл. IV, § 5, n° 4 и т. I, стр. 592). С другой стороны,

$$\int_{-\infty}^{\infty} e^{-\left(x+\frac{1}{2}ia\right)^2} dx = e^{\frac{1}{4}a^2} \int_{-\infty}^{\infty} e^{-x^2}(\cos ax - i\sin ax)\,dx =$$

$$= 2e^{\frac{1}{4}a^2} \int_{0}^{\infty} e^{-x^2} \cos ax\, dx$$

(в силу того, что $\cos ax$ — четная функция, а $\sin ax$ — нечетная функция). Равенство этих двух результатов и доказывает требуемую формулу. (Ср. стр. 340, упр. 4а.)

3. Приложение теоремы вычетов к интегрированию рациональных функций. Пусть дана несократимая дробно-рациональная функция

$$Q(z) = \frac{a_0 + a_1 z + \ldots + a_m z^m}{b_0 + b_1 z + \ldots + b_n z^n},$$

знаменатель которой не имеет действительных корней, а его степень превосходит степень числителя не менее чем на две единицы, т. е. $n \geqslant m+2$. Тогда интеграл

$$I = \int_{-\infty}^{\infty} Q(x)\,dx$$

можно вычислить следующим путем.

Начнем с того, что возьмем интеграл от функции $Q(z)$ по контуру, состоящему из полуокружности H с параметрическим уравнением $z = Re^{i\theta}$, $0 \leqslant \theta \leqslant \pi$, и ее диаметра, лежащего на действительной оси от $x = -R$ до $x = R$. Радиус R выберем столь большим, что ни один из полюсов дроби $Q(z)$ не лежит ни на окружности, ни вне ее. Тогда наш интеграл по контуру равен

$$\int_{-R}^{R} Q(x)\,dx + \int_{H} Q(z)\,dz,$$

но, с другой стороны, он должен быть равен сумме вычетов функции $Q(z)$ относительно всех ее полюсов, лежащих внутри контура. Легко получить оценку для модуля интеграла вдоль полуокружности H. Из свойств функции $Q(z)$ следует, что $Q(z) = \frac{\varphi(z)}{z^2}$, где при $z \to \infty$ $\varphi(z) \to 0$, если $n > m+2$, и $\varphi(z) \to \frac{a_m}{b_n}$, если $n = m+2$. Отсюда вытекает, что существует такое постоянное положительное число M, что

$$|Q(z)| < \frac{M}{R^2}.$$

Длина полуокружности H равна πR. Поэтому, в силу формулы оценки интеграла (§ 3, в конце n° 1),

$$\left|\int\limits_H Q(z)\,dz\right| < \frac{M}{R^2}\pi R = \frac{\pi M}{R}$$

и, следовательно, стремится к нулю при $R \to \infty$. Стало быть, при $R \to \infty$ интеграл по нашему контуру имеет своим пределом *искомый интеграл*

$$I = \int\limits_{-\infty}^{\infty} Q(x)\,dx,$$

и этот искомый интеграл должен быть равен сумме вычетов функции $Q(z)$ относительно всех ее полюсов, лежащих в верхней полуплоскости.

Этот результат мы применим к нескольким интересным частным случаям.

1) Пусть $Q(z) = \dfrac{1}{az^2 + bz + c} = \dfrac{1}{f(z)}$, где коэффициенты a, b, c — действительные числа, удовлетворяющие условиям $a > 0$ и $b^2 - 4ac < 0$. Тогда функция $Q(z)$ имеет в верхней полуплоскости только один простой полюс (т. е. полюс первого порядка)

$$z_1 = \frac{1}{2a}\left(-b + i\sqrt{4ac - b^2}\right)$$

с положительным значением квадратного корня. Поэтому, в силу общего правила (§ 4, n° 5), вычет относительно этого полюса равен $2\pi i\,\dfrac{1}{f'(z_1)}$. Так как

$$f'(z_1) = 2az_1 + b = i\sqrt{4ac - b^2},$$

то

$$\int\limits_{-\infty}^{\infty} \frac{dx}{ax^2 + bx + c} = \frac{2\pi}{\sqrt{4ac - b^2}}.$$

2) В качестве второго примера докажем формулу (ср. т. I, стр. 274)

$$\int\limits_{-\infty}^{\infty} \frac{dx}{1 + x^4} = \frac{1}{2}\pi\sqrt{2}.$$

И здесь можно сразу применить наш общий результат. В верхней полуплоскости функция $Q(z) = \dfrac{1}{1 + z^4} = \dfrac{1}{f(z)}$ имеет два простых полюса $z_1 = e^{\frac{1}{4}\pi i}$ и $z_2 = e^{\frac{3}{4}\pi i} = \dfrac{e^{\pi i}}{e^{\frac{1}{4}\pi i}} = -e^{-\frac{1}{4}\pi i}$. Это те два значения корня четвертой степени из -1, которые имеют положительную мнимую часть. Сумма вычетов равна

$$2\pi i\left(\frac{1}{f'(z_1)}+\frac{1}{f'(z_2)}\right)=2\pi i\cdot\frac{1}{4}\left(\frac{1}{z_1^3}+\frac{1}{z_2^3}\right)=\frac{\pi i}{2}\left(\frac{z_1}{z_1^4}+\frac{z_2}{z_2^4}\right)=$$

$$=-\frac{\pi i}{2}(z_1+z_2)=-\frac{\pi i}{2}\left(e^{\frac{1}{4}\pi i}-e^{-\frac{1}{4}\pi i}\right)=-\frac{\pi i}{2}\,2i\sin\frac{\pi}{4}=\pi\sin\frac{\pi}{4}=\frac{\pi\sqrt{2}}{2},$$

и наша формула доказана.

Упражнения

1. Тем же путем доказать формулу

$$\int_{-\infty}^{\infty}\frac{x^2}{1+x^4}\,dx=\frac{\pi}{2}\sqrt{2}.$$

2. Доказать, что вообще, если n и m — положительные целые числа и $n>m$, то

$$\int_{-\infty}^{\infty}\frac{x^{2m}}{1+x^{2n}}\,dx=\frac{\pi}{n\sin\left(\frac{2m+1}{2n}\pi\right)}.$$

В приведенных примерах и упражнениях нам приходилось вычислять вычеты только относительно простых полюсов. Теперь мы докажем формулу

$$\int_{-\infty}^{\infty}\frac{dx}{(1+x^2)^{n+1}}=\frac{\pi}{4^n}\frac{(2n)!}{(n!)^2},$$

для чего потребуется найти вычет относительно полюса более высокого порядка.

После замены x на z, знаменатель подынтегральной функции можно привести к виду $(z+i)^{n+1}(z-i)^{n+1}$. Следовательно, подынтегральная функция имеет в верхней полуплоскости лишь один полюс $(n+1)$-го порядка в точке $z=i$. Для того чтобы найти соответствующий вычет, напишем

$$\frac{1}{(z^2+1)^{n+1}}=\frac{1}{f(z)}=\frac{1}{(z-i)^{n+1}}\frac{1}{(2i+z-i)^{n+1}}=$$

$$=\frac{1}{(z-i)^{n+1}}\frac{1}{(2i)^{n+1}}\left(1+\frac{z-i}{2i}\right)^{-n-1}.$$

Если разложим последний множитель в биномиальный ряд, то член этого ряда, содержащий $(z-i)^n$, имеет коэффициент

$$\frac{1}{(2i)^n}\binom{-n-1}{n}=\frac{1}{(2i)^n}(-1)^n\frac{(n+1)\ldots 2n}{1\cdot 2\ldots n}=\frac{i^n}{2^n}\frac{(2n)!}{(n!)^2}.$$

Стало быть, коэффициент c_{-1} в разложении подынтегральной функции по степеням $z-i$ равен $\frac{1}{2^{2n+1}}\frac{1}{i}\frac{(2n)!}{(n!)^2}$. Поэтому вычет относительно полюса i будет $2\pi i c_{-1}=\frac{\pi}{2^{2n}}\frac{(2n)!}{(n!)^2}$, что и доказывает нашу формулу.

Упражнения

1. С помощью теории вычетов доказать формулу

$$\int_{-\infty}^{\infty}\frac{x\sin x}{x^2+c^2}\,dx=\frac{\pi}{2}\,e^{-|c|}.$$

2. Дан многочлен $f(x)$ степени n, имеющий простые корни $\alpha_1, \alpha_2, \ldots, \alpha_n$. Доказать, что

$$\sum_{k=1}^{n}\frac{\alpha_k^m}{f'(\alpha_k)}=0 \qquad (m=0,\ 1,\ 2,\ \ldots,\ n-2).$$

$\left(\text{Рассмотреть } \oint \frac{z^m}{f(z)}\,dz\right.$ вдоль замкнутой кривой, содержащей все корни α_k внутри себя.$\Big)$

4. Теорема вычетов и линейные дифференциальные уравнения с постоянными коэффициентами. Дан многочлен степени n

$$P(z)=a_0+a_1z+a_2z^2+\ldots+a_nz^n,$$

и пусть $f(z)$ — любой многочлен, степень которого меньше чем n (в частности, постоянная). Если t есть действительный параметр, то интеграл

$$u(t)=\oint_C\frac{e^{tz}f(z)}{P(z)}\,dz,$$

взятый вдоль любой замкнутой кривой C плоскости z, не проходящей через нули знаменателя $P(z)$, является функцией $u(t)$ параметра t. По правилам дифференцирования под знаком интеграла, которые действуют без изменения и в комплексной области, функцию $u(t)$ можно дифференцировать требуемое число раз по t. Это дифференцирование по t под знаком интеграла равносильно умножению подынтегральной функции на $z, z^2, z^3, \ldots$, в зависимости от порядка вычисляемой производной. Поэтому, если составить дифференциальное выражение

$$L[u]=a_0u+a_1u'+a_2u''+\ldots+a_nu^{(n)},$$

или, в символическом обозначении, $P(D)u$, где D есть символ дифференцирования: $D=\frac{d}{dt}$, то

$$P(D)u=L[u]=\oint_C e^{tz}f(z)\,dz.$$

По теореме Коши этот комплексный интеграл равен нулю; это значит, что функция $u(t)$ является решением дифференциального уравнения $L[u]=0$. Если $f(z)$ есть произвольный многочлен степени $n-1$, то это решение содержит n произвольных постоянных. Поэтому можно ожидать, что таким путем получено самое общее решение линейного дифференциального уравнения с постоянными коэффициентами $L[u]=0$.

И действительно, это решение можно привести к уже известному нам виду (гл. VI, § 4); для этого надо вычислить полученный для $u(t)$ интеграл по теореме вычетов, предполагая, что замкнутая кривая C содержит внутри себя все нули $z_1, z_2, \ldots, z_n$ знаменателя

$$P(z)=a_n(z-z_0)(z-z_1)\ldots(z-z_n).$$

Предположим сначала, что все эти нули простые; тогда они являются простыми полюсами подынтегральной функции, и вычет относительно полюса z_k будет $2\pi i\frac{f(z_k)}{P'(z_k)}e^{z_k t}$. В силу произвольности многочлена $f(z)$, выражения $\frac{f(z_k)}{P'(z_k)}$ можно сделать *произвольными* постоянными, и, стало быть, решение $u(t)$ принимает обычный вид общего решения

$$u(t)=\sum_{k=1}^{n} c_k e^{z_k t},$$

в полном согласии с прежними результатами.

Если же нуль z_k многочлена $P(z)$ является кратным, скажем точнее, r-кратным, то соответствующий полюс подынтегральной функции будет полюсом r-го порядка. Тогда вычет относительно полюса z_k придется определить, разлагая числитель подынтегральной функции, а именно $e^{tz}f(z)=e^{tz_k}e^{t(z-z_k)}f(z)$, в ряд по степеням $z-z_k$. Предоставляем читателю показать, что вычет относительно полюса z_k дает наряду с решением $e^{z_k t}$ еще и $r-1$ других частных решений $te^{z_k t}$, $t^2e^{z_k t}, \ldots, t^{r-1}e^{z_k t}$.

5. Доказательство формулы $\int\limits_{-\infty}^{\infty} e^{-x^2}dx=\sqrt{\pi}$ с помощью теории вычетов. При вычислении интеграла в n° 2 мы пользовались этой формулой, считая ее известной из теории действительных переменных. Однако этот интеграл можно вычислить с помощью интегрирования в комплексной области, пользуясь теоремой вычетов. Так как это доказательство весьма поучительно, то мы его здесь приведем, хотя с нашей элементарной точки зрения его исходный пункт может показаться искусственным. Мы будем исходить из комплексного интеграла, который возникает в других математических дисциплинах (например, в теории чисел).

Будем обозначать символом l_a прямую $z=a+se^{\frac{i\pi}{4}}$ $(-\infty<s<\infty)$ в плоскости z, т. е. прямую, пересекающую ось x в точке $x=a$ и образующую с этой осью угол в 45°. Пусть u — действительный параметр. Рассмотрим интеграл

$$f(u)=\int\limits_{l_{1/2}}\frac{e^{\pi iz^2+2\pi iuz}}{e^{2\pi iz}-1}\,dz.$$

Этот интеграл следует рассматривать как несобственный, т. е. сперва интегрировать от $s=-R$ до $s=R$, а затем заставить R стремиться к бесконечности. Читатель может установить сходимость этого интеграла по образцу аналогичного доказательства для несобственного интеграла в действительной области. Тогда

$$f(u+1)-f(u)=\int\limits_{l_{1/2}}\frac{e^{\pi iz^2}}{e^{2\pi iz}-1}\,e^{2\pi iuz}\left(e^{2\pi iz}-1\right)dz=$$
$$=\int\limits_{l_{1/2}}e^{\pi iz^2+2\pi iuz}\,dz=e^{-\pi iu^2}\int\limits_{l_{1/2}}e^{\pi i(z+u)^2}\,dz.$$

Так как подынтегральная функция в правой части всюду аналитическая, то можно произвести параллельное смещение пути интегрирования на любую величину (это доказывается с помощью теоремы Коши, как в n° 2, стр. 578) и написать, например,

$$f(u+1)-f(u)=e^{-\pi iu^2}\int\limits_{l_0}e^{\pi iz^2}\,dz=e^{-\pi iu^2}I,$$

где $z=se^{\frac{i\pi}{4}}$ вдоль прямой l_0, биссектрисы первой и четвертой четверти, и следовательно,

$$I=e^{\frac{i\pi}{4}}\int\limits_{-\infty}^{\infty}e^{-\pi s^2}\,ds.$$

Замена переменной $s\sqrt{\pi}=t$, $\sqrt{\pi}\,ds=dt$ дает

$$I=e^{\frac{i\pi}{4}}\frac{1}{\sqrt{\pi}}\int\limits_{-\infty}^{\infty}e^{-t^2}\,dt.$$

С другой стороны, введя в интеграл, определяющий $f(u)$, вместо z новую переменную интегрирования λ с помощью формулы $z=\lambda+1$ и учитывая, что $e^{\pi i}=-1$, получим

$$f(u)=-\int\limits_{l_{-1/2}} \frac{e^{\pi i\lambda^2+2\pi i\lambda u}}{e^{2\pi i\lambda}-1} e^{2\pi i u} e^{2\pi i\lambda} d\lambda,$$

или

$$-e^{-2\pi i u} f(u)=\int\limits_{l_{-1/2}} e^{\pi i\lambda^2+2\pi i\lambda u} d\lambda+\int\limits_{l_{-1/2}} \frac{e^{\pi i\lambda^2+2\pi i\lambda u}}{e^{2\pi i\lambda}-1} d\lambda.$$

В первом интеграле правой части можно опять сместить путь интегрирования параллельно самому себе на 1 вправо, и тогда будет видно, что

$$\int\limits_{l_{-1/2}} e^{\pi i\lambda^2+2\pi i\lambda u} d\lambda=\int\limits_{l_{1/2}} e^{\pi i z^2+2\pi i u z} dz=e^{-\pi i u^2} I.$$

Если во втором интеграле правой части сместить путь интегрирования вправо на 1, то обнаружим, что он равен интегралу, определяющему $f(u)$, причем полюс $\lambda=0$ лежит между обоими путями интегрирования, $l_{-1/2}$ и $l_{1/2}$.

Применим теперь теорему вычетов (тот факт, что пути интегрирования $l_{-1/2}$ и $l_{1/2}$ простираются в бесконечность, нас не беспокоит в силу такого же рассуждения, как на стр. 578), докажем, что вычет подынтегральной функции относительно полюса $\lambda=0$ имеет значение 1, и тогда из нашего уравнения сразу получится следующий результат:

$$-f(u) e^{-2\pi i u}=e^{-\pi i u^2} I+f(u)-1.$$

Правда, здесь неизвестны в явном виде ни I, ни функция $f(u)$. Однако если подставить сюда $u=\frac{1}{2}$, то $f(u)$ исключится, и останется

$$e^{-\frac{\pi i}{4}} I=1,$$

а так как (см. выше)

$$I=e^{\frac{\pi i}{4}} \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} e^{-t^2} dt,$$

то сразу получится формула, подлежавшая доказательству.

6. Многозначные функции и аналитическое продолжение. Когда мы до сих пор давали определения действительным и комплексным функциям, мы всегда держались той точки зрения, что каждому значению независимой переменной должно соответствовать *единственное* значение функции. Даже, например, теорема Коши опирается на предположение, что функция может быть определена однозначно в рассматриваемой области. Правда, при фактическом построении функций многозначность часто возникает в силу необходимости; примером может служить нахождение обратной функции для однозначной степенной функции z^n. В таких процессах обращения, как $\sqrt{z}$ или $\sqrt[n]{z}$, получались в случае действительной переменной *различные обособленные однозначные ветви* обратной функции. Мы увидим, однако, что в комплексной области такое обособление уже невозможно, ибо различные однозначные ветви оказываются теперь взаимно связанными.

Нам придется здесь удовольствоваться очень простым рассмотрением этого вопроса на основе типичных примеров.

Рассмотрим, например, обратную функцию $\zeta = \sqrt{z}$ для функции $z = \zeta^2$. Одному значению z соответствуют два возможных решения ζ и $-\zeta$ уравнения $z = \zeta^2$. Эти две ветви обратной функции связаны между собой следующим образом. Пусть $z = re^{i\theta}$. Если мы теперь положим $\zeta = \sqrt{r}\, e^{i\frac{\theta}{2}} = f(z)$, то $\zeta = f(z)$ является аналитической функцией во всякой односвязной области G, не содержащей начала координат, где $f(z)$ уже не имеет производной. В такой области ζ определена однозначно. Если же мы заставим точку z двигаться вокруг начала по окружности с центром в начале, скажем, в положительном направлении, то $\zeta = \sqrt{r}e^{\frac{i\theta}{2}}$ будет изменяться непрерывно; однако угол θ не вернется к своему первоначальному значению, но увеличится на 2π. Стало быть, при этом непрерывном изменении, когда мы вернемся в точку z, то корень уже не вернется к своему исходному значению $\zeta = \sqrt{r}e^{\frac{i\theta}{2}}$, но придет в точку z с новым значением $\sqrt{r}e^{\frac{i(\theta+2\pi)}{2}} = \sqrt{r}\,e^{\frac{i\theta}{2}}e^{i\pi} = -\zeta$. Можно поэтому сказать, что когда функция $\zeta = f(z)$ непрерывно продолжается по замкнутой кривой C, то она перестает быть однозначной.

Функция $\sqrt[n]{z}$, где n — целое число, ведет себя точно таким же образом. Здесь каждый оборот вокруг начала приводит к умножению функции на корень n-й степени из единицы, т. е. на $\varepsilon = e^{\frac{2\pi i}{n}}$, и функция возвращается к своему исходному значению только после n оборотов.

У функции $\ln z$ мы обнаружили (стр. 565) аналогичную многозначность: при непрерывном движении точки z вокруг начала в положительном направлении значение функции $\ln z$ увеличивается после одного оборота на $2\pi i$.

Многозначна и функция z^{α}: после каждого оборота вокруг начала она *умножается* на $e^{2\pi i\alpha}$.

Все эти функции, хотя и получили первоначально однозначное определение в некоторой области G, оказываются многозначными, когда их непрерывно продолжают (как аналитические функции), и возвращаются в исходную точку по какому-либо замкнутому пути, окружающему начало координат. Это явление многозначности и связанную с ним общую теорию аналитического продолжения мы не можем исследовать подробно в рамках этой книги. Мы лишь отметим, что теоретически однозначность функции можно обеспечить проведением в плоскости z некоторых линий, которые путь, описываемый точкой z, не имеет права пересекать или, как принято говорить, с помощью разрезов вдоль надлежащих линий. Эти разрезы производятся таким образом, что замкнутые пути, вызывающие многозначность, становятся невозможными в плоскости z.

Например, функция $\ln z$ становится однозначной, если разрезать плоскость z вдоль отрицательной действительной полуоси. Такой же разрез делает однозначной функцию $\sqrt{z}$. Функция $\sqrt{1-z^2}$ станет однозначной, если сделать разрез вдоль отрезка действительной оси между точками -1 и $+1$.

Коль скоро в плоскости z произведен надлежащий разрез, делающий функцию $f(z)$ однозначной, к $f(z)$ уже можно применять теорему Коши.

В качестве простого примера, показывающего, как применять теорему Коши в случае, когда возникает многозначная функция, мы докажем формулу

$$I = \int_{-1}^{1} \frac{dx}{(x-a)\sqrt{1-x^2}} = \frac{\pi}{\sqrt{a^2-1}},$$

где a — постоянная, не лежащая на отрезке действительной оси от -1 до $+1$.

Сначала замечаем, что подынтегральная функция будет однозначной в плоскости z, если сделать в ней разрез вдоль отрезка действительной оси от -1 до $+1$. Когда точка z приблизится к этому разрезу сначала сверху, а затем снизу, то квадратный корень $\sqrt{1-z^2}$ примет значения, отличающиеся только знаком, скажем, со знаком плюс сверху и со знаком минус снизу. Возьмем теперь комплексный интеграл

$$\oint_C \frac{dz}{(z-a)\sqrt{1-z^2}}$$

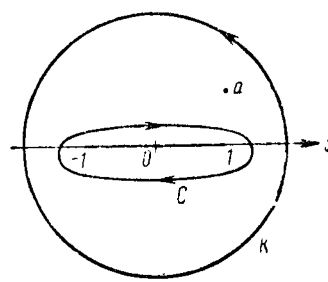


Рис. 116.

вдоль замкнутой кривой C, как показано на рис. 116. По теореме Коши эту кривую можно деформировать, стягивая ее к разрезу, не изменяя при этом значения интеграла. Поэтому это значение интеграла совпадает с его предельным значением, которое, очевидно, равно $2I$. С другой стороны, если взять интеграл от той же подынтегральной функции вдоль окружности K радиуса R с центром в начале, то полученный интеграл стремится к нулю при $R \to \infty$, что нетрудно доказать. Кстати, его значение равно нулю, так как, по теореме Коши, оно не зависит от радиуса R, коль скоро окружность K содержит внутри себя полюс a. Однако, по теореме вычетов, сумма интегралов вдоль кривых C и K, описываемых в одинаковом направлении относительно лежащей между ними области, притом так, что область остается всегда по левую руку, равна вычету подынтегральной функции относительно полюса a, заключенного между обоими контурами, C и K, а этот вычет равен

$$2\pi i \lim_{z\to a} (z-a) \frac{1}{(z-a)\sqrt{1-z^2}} = \frac{2\pi}{\sqrt{a^2-1}}.$$

Следовательно, $2I$ равно этому же числу, что и доказывает нашу формулу.

7. Пример аналитического продолжения. Гамма-функция. В заключение покажем на примере, как аналитическая функция, определенная первоначально лишь в некоторой части плоскости z, может быть продолжена за пределы первоначальной области ее определения. Мы продолжим аналитически гамма-функцию, которая была ранее определена для значений $\operatorname{Re} z = x > 0$ интегралом

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt,$$

на все остальные значения z, для которых $\operatorname{Re} z = x \leqslant 0$ (ср. стр. 567, упр. 1). Можно это выполнить, например, с помощью функционального уравнения $\Gamma(z) = \frac{1}{z}\Gamma(z+1)$, используя это уравнение для определения $\Gamma(z-1)$ по известной $\Gamma(z)$. С помощью этого уравнения можно аналитически продолжить $\Gamma(z)$ сначала на полосу $-1 < x \leqslant 0$, затем на следующую параллельную полосу $-2 < x \leqslant -1$, и т. д.

Для того чтобы аналитически продолжить гамма-функцию, можно, однако, воспользоваться другим методом, представляющим значительно больший теоретический интерес. Рассмотрим путь интегрирования C в плоскости комплексной переменной t, изображенный на рис. 117. Этот путь охватывает положительную действительную полуось плоскости t и приближается к этой полуоси асимптотически сверху и снизу. С помощью теоремы Коши нетрудно убедиться, что этот интеграл[1])

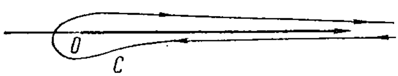


Рис. 117.

$$\int_C t^{z-1} e^{-t} dt$$

вдоль «петли» C не изменяет своего значения, когда петля C деформируется, стягиваясь к положительной полуоси x. Но при этом подынтегральная функция $t^{z-1}e^{-t}$ стремится к различным значениям, когда точки петли приближаются к одной и той же точке оси x сверху или снизу: эти значения отличаются множителем $e^{2\pi i z}$. Таким образом получается формула

$$(1 - e^{2\pi i z})\,\Gamma(z) = \int_C t^{z-1}\, e^{-t}\, dt, \qquad (*)$$

справедливая, когда $\operatorname{Re} z = x > 0$. Еще раз подчеркиваем, что эта формула выведена при допущении, что действительная часть аргумента z положительна. Однако сам интеграл вдоль петли C имеет смысл при всех значениях комплексного числа z, так как путь ин-

---

[1]) Это тоже несобственный интеграл, который получается из интеграла вдоль конечной части петли C путем перехода к пределу. Его сходимость можно доказать тем же методом, которым мы уже не раз пользовались для этой цели.

тегрирования обходит начало $t=0$. Поэтому интеграл вдоль петли C представляет собой функцию, определенную на всей плоскости z. Так вот, мы определяем функцию $\Gamma(z)$ последним равенством (*). Это равенство и дает аналитическое продолжение гамма-функции на всю плоскость z, за исключением тех точек отрицательной полуоси x, в которых множитель $(1-e^{2\pi iz})$ обращается в нуль, т. е. за исключением точек $z=0$, $z=-1$, $z=-2$ и т. д.

СМЕШАННЫЕ УПРАЖНЕНИЯ К ГЛАВЕ VIII

1. Составить условие, при котором три точки z_1, z_2, z_3 лежат на одной прямой.

2*. Составить условие, при выполнении которого четыре точки z_1, z_2, z_3, z_4 лежат на одной окружности.

3*. Четыре точки $A(z_1)$, $B(z_2)$, $C(z_3)$ и $D(z_4)$ лежат в этом порядке на одной окружности плоскости z. Пользуясь комплексными координатами этих точек, доказать, что $AB\cdot CD+BC\cdot AD=AC\cdot BD$.

4. Доказать, что уравнение $\cos z=c$ имеет решения при всех значениях c.

5. Для каких значений c уравнение $\operatorname{tg} z=c$ не имеет решений?

6. Выяснить, при каких значениях z функции а) $\cos z$, б) $\sin z$ имеют действительные значения.

7. Найти радиус сходимости степенного ряда $\sum a_n z^n$, если

а) $a_n=\frac{1}{n^s}$, где s — комплексное число с положительной действительной частью;

б) $a_n=n^n$;

в) $a_n=\ln n$.

8. Доказать формулу $e^z=\lim\limits_{n\to\infty}\left(1+\frac{z}{n}\right)^n$ для комплексных значений z.

9. С помощью комплексного интегрирования вычислить интегралы

$$\text{а) } \int_0^\infty \frac{\cos x}{1+x^4}\,dx;\quad \text{б) } \int_0^\infty \frac{x^2\cos x}{1+x^4}\,dx;\quad \text{в) } \int_0^\infty \frac{\cos x}{c^2+x^2}\,dx;$$

$$\text{г) } \int_0^\infty \frac{x^{\alpha-1}}{(x+1)(x+2)}\,dx,$$

где α — действительное число $(1<\alpha<2)$.

10. Найти полюсы и вычеты относительно этих полюсов функций

$$\text{а) } \frac{1}{\sin z};\quad \text{б) } \frac{1}{\cos z};\quad \text{в) } \Gamma(z);\quad \text{г) } \operatorname{ctg} z=\frac{\cos z}{\sin z}.$$

11*. Найти предел интеграла

$$\oint_{C_n} \frac{\operatorname{ctg}\pi\zeta}{\zeta-z}\,d\zeta$$

при $n\to\infty$, где путь интегрирования есть периметр квадрата C_n, стороны которого параллельны осям и проходят на расстоянии $n \pm \frac{1}{2}$ от начала координат. Затем получить отсюда с помощью теоремы вычетов разложение $\operatorname{ctg} \pi z$ на элементарные дроби.

12*. Пользуясь формулой

$$\ln(1+z) = \int_0^z \frac{d\zeta}{1+\zeta},$$

показать, что степенной ряд для $\ln(1+z)$ сходится на единичной окружности $|z| = 1$, за исключением точки $z = -1$. Приравнивая мнимую часть этого ряда мнимой части функции $\ln(1+e^{i\theta})$, установить справедливость следующего разложения в ряд Фурье (ср. т. I, стр. 514):

$$\frac{1}{2}\theta = \sin\theta - \frac{1}{2}\sin 2\theta + \frac{1}{3}\sin 3\theta - + \ldots, \qquad -\pi < \theta < \pi.$$

13*. а) Доказать, что ряд

$$f(z) = f(x+iy) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^z}$$

сходится при $x > 0$.

б) Доказать, что этот ряд дает возможность аналитического продолжения дзета-функции (определенной в упр. 2, стр. 568, при $x > 1$) на такие значения z, для которых $0 < x \leqslant 1$, с помощью формулы

$$f(z) = (1 - 2^{1-z})\,\zeta(z),$$

справедливой при $x > 1$.

в) Доказать, что дзета-функция имеет полюс $z = 1$ с вычетом, равным 1.

14. а) Применить теорему Коши к интегралу

$$\oint \left(z + \frac{1}{z}\right)^m z^{n-1}\,dz \qquad (n > m > 0),$$

взятому вдоль контура, состоящего из четверти единичной окружности $|z| = 1$, лежащей в первом квадранте, четверти малой окружности $|z| = \varepsilon$, лежащей в первом квадранте (для обхода начала координат), и двух отрезков осей x и y между обеими окружностями. Вывести отсюда формулу

$$\int_0^{\pi/2} \cos^m\theta \cos n\theta\,d\theta = \frac{\sin\frac{(n-m)\pi}{2}}{2^{m+1}} \frac{\Gamma(m+1)\,\Gamma\left(\frac{n-m}{2}\right)}{\Gamma\left(\frac{n+m}{2}+1\right)}.$$

б) Доказать, что если $n = m$, то последний интеграл равен $\frac{\pi}{2^{m+1}}$.

15. Доказать, что если $f(z)$ есть аналитическая функция, то $\frac{d^n}{dx^n} f(\sqrt{x})$ равно результату, который получится после подстановки в выражение $2\frac{d^n}{dy^n}\frac{yf(y)}{(y+a)^{n+1}}$ вместо y и a, каждого в отдельности, $\sqrt{x}$.

16. а) Показать, что при действительных x и y

$$|\operatorname{sh}(x+iy)| \geqslant A(x),$$

где $A(x)$ не зависит от y и стремится к ∞ при $x \to \pm\infty$.

б) Интегрированием функции $\dfrac{1}{(z-w)\operatorname{sh} z}$ вдоль подходящей последовательности контуров показать, что

$$\frac{1}{\operatorname{sh} w} = \frac{1}{w} + 2w \sum_{n=1}^{\infty} \frac{(-1)^n}{w^2+\pi^2 n^2}.$$

17*. На плоскости дана система n материальных точек; каждая из них притягивает к себе с силой, обратно пропорциональной расстоянию. Доказать, что частица, помещенная в этом поле, может иметь не более чем $n-1$ положений равновесия.

Вычислить эти положения равновесия для системы четырех притягивающих точек (a, b), $(-a, b)$, $(-a, -b)$, $(a, -b)$, где $a > b > 0$.

СВОДКА ВАЖНЕЙШИХ ТЕОРЕМ И ФОРМУЛ

1. Дифференцирование

Дифференцирование сложной функции — правило цепочки для функции многих переменных.

Если $u = f(\xi, \eta, \zeta, \ldots)$, где $\xi = \xi(x, y)$, $\eta = \eta(x, y), \ldots$, то

$$u_x = f_\xi \xi_x + f_\eta \eta_x + f_\zeta \zeta_x + \cdots, \qquad u_y = f_\xi \xi_y + f_\eta \eta_y + f_\zeta \zeta_y + \cdots;$$

$$\begin{aligned} u_{xx} = f_{\xi\xi}\xi_x^2 + f_{\eta\eta}\eta_x^2 + f_{\zeta\zeta}\zeta_x^2 + \cdots \\ + 2f_{\xi\eta}\xi_x\eta_x + 2f_{\xi\zeta}\xi_x\zeta_x + \cdots \\ \cdots\cdots\cdots\cdots\cdots \\ + f_\xi \xi_{xx} + f_\eta \eta_{xx} + f_\zeta \zeta_{xx} + \cdots \end{aligned}$$

и соответствующие формулы для u_{xy} и u_{yy} (стр. 89).

Неявные функции. Если $F(x, y) = 0$, то

$$\frac{dy}{dx} = -\frac{F_x}{F_y},$$

$$\frac{d^2y}{dx^2} = -\frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_y^3} \quad \text{(стр. 138)}.$$

Якобианы. Если $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$, то

$$\frac{\partial x}{\partial \xi} = \frac{\psi_y}{D}, \qquad \frac{\partial x}{\partial \eta} = -\frac{\varphi_y}{D}, \qquad \frac{\partial y}{\partial \xi} = -\frac{\psi_x}{D}, \qquad \frac{\partial y}{\partial \eta} = \frac{\varphi_x}{D},$$

где

$$D = \frac{\partial(\xi, \eta)}{\partial(x, y)} = \begin{vmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{vmatrix} = \varphi_x\psi_y - \varphi_y\psi_x$$

есть якобиан или функциональный определитель (стр. 163—164).

Свойства якобианов:

а) $$\frac{\partial (x, y)}{\partial (\xi, \eta)} = \frac{1}{\dfrac{\partial (\xi, \eta)}{\partial (x, y)}} \quad \text{(стр. 165).}$$

б) Если $u = u(\xi, \eta)$, $v = v(\xi, \eta)$ и $\xi = \xi(x, y)$, $\eta = \eta(x, y)$, то

$$\frac{\partial (u, v)}{\partial (x, y)} = \frac{\partial (u, v)}{\partial (\xi, \eta)} \frac{\partial (\xi, \eta)}{\partial (x, y)} \quad \text{(стр. 166).}$$

2. Сходимость двойных последовательностей

Критерий сходимости Коши для двойных последовательностей (стр. 121—125). Последовательность a_{nm} сходится в том и только в том случае, если для любого $\varepsilon > 0$ можно найти такое число N, что

$$|a_{nm} - a_{n'm'}| < \varepsilon,$$

коль скоро $n > N$, $m > N$, $n' > N$, $m' > N$. Тогда существует такое число a, что

$$\lim_{\substack{n \to \infty \\ m \to \infty}} a_{nm} = a.$$

Если при этом существует $\lim\limits_{m \to \infty} a_{nm}$ или $\lim\limits_{n \to \infty} a_{nm}$, то

$$\lim_{\substack{n \to \infty \\ m \to \infty}} a_{nm} = \lim_{n \to \infty} \left(\lim_{m \to \infty} a_{nm} \right) = \lim_{m \to \infty} \left(\lim_{n \to \infty} a_{nm} \right).$$

3. Равномерная сходимость и изменение порядка предельных операций

Теорема Дини. Если ряд, члены которого — *положительные* непрерывные функции, сходится в некоторой замкнутой области к непрерывной же предельной функции, то эта сходимость *равномерна* (стр. 126).

Изменение порядка дифференцирования и интегрирования. (Дифференцирование интеграла по параметру.)

$$\frac{d}{dx} \int_a^b f(x, y)\, dy = \int_a^b f_x(x, y)\, dy,$$

если $f(x, y)$ и $f_x(x, y)$ непрерывны на отрезке $[a, b]$ (стр. 241).

Изменение порядка дифференцирования и интегрирования в несобственных интегралах:

$$\frac{d}{dx} \int_0^\infty f(x, y)\, dy = \int_0^\infty f_x(x, y)\, dy,$$

если $f_x(x, y)$ непрерывна в рассматриваемом бесконечном промежутке, а интегралы $\int\limits_0^\infty f(x, y)\, dy$ и $\int\limits_0^\infty f_x(x, y)\, dy$ — равномерно сходящиеся (стр. 334).

Изменение порядка двух интегрирований. Если $f(x, y)$ непрерывна и a, b, α, β — постоянные, то

$$\int_a^b dx \int_\alpha^\beta f(x, y)\, dy = \int_\alpha^\beta dy \int_a^b f(x, y)\, dx \quad \text{(стр. 263).}$$

Порядок интегрирования можно переставить и в том случае, когда пределы не являются постоянными, при условии, что оба раза интегрирование производится по *всей* рассматриваемой области и в соответствии с этим переходят к новым пределам интегрирования (стр. 266).

Изменение порядка двух интегрирований в несобственном интеграле (стр. 332).

$$\int_\alpha^\beta dx \int_0^\infty f(x, y)\, dy = \int_0^\infty dy \int_\alpha^\beta f(x, y)\, dx,$$

если интеграл $\int_0^\infty f(x, y)\, dy$ сходится равномерно в интервале $\alpha \leqslant x \leqslant \beta$.

4. Некоторые определенные интегралы

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \quad \text{(стр. 284, 583; см. также т. I, стр. 592).}$$

Интеграл Дирихле:

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2} \quad \text{(стр. 338, 577; см. также т. I, стр. 292, 293, 486, 527).}$$

Интегралы Френеля:

$$\int_{-\infty}^{+\infty} \sin(\tau^2)\, d\tau = \int_{-\infty}^{+\infty} \cos(\tau^2)\, d\tau = \sqrt{\frac{\pi}{2}} \quad \text{(стр. 339; см. также т. I, стр. 294).}$$

Интеграл Фурье (стр. 341—345). Если 1) $f(x)$ есть кусочно гладкая функция в любом конечном интервале, 2) $\int_{-\infty}^{+\infty} |f(x)|\, dx$ сходится и 3) $f(x+0) + f(x-0) = 2f(x)$, то

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(\tau)\, e^{i\tau x}\, d\tau,$$

где

$$g(\tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{-i\tau x}\, dx \quad \text{(стр. 342).}$$

Гамма-функция (стр. 346—361). Если $x > 0$, то гамма-функция $\Gamma(x)$ определяется равенством

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt = 2\int_0^\infty e^{-t^2} t^{2x-1} dt.$$

Она удовлетворяет функциональному уравнению

$$\Gamma(x+1) = x\Gamma(x).$$

Следовательно, при любом натуральном n

$$\Gamma(n) = (n-1)!$$

При всяком $x \neq 0, -1, -2, \ldots$

$$\Gamma(x) = \lim_{n\to\infty} \frac{(n-1)!}{x(x+1)\ldots(x+n-1)} n^x = \frac{1}{x}\prod_{\nu=1}^{\infty} \frac{\left(1+\frac{1}{\nu}\right)^x}{1+\frac{x}{\nu}},$$

$$\frac{1}{\Gamma(x)} = xe^{\gamma x} \prod_{\nu=1}^{\infty}\left(1+\frac{x}{\nu}\right)e^{-\frac{x}{\nu}},$$

где $\gamma = \lim\limits_{n\to\infty}\left(\sum\limits_{k=1}^{n}\frac{1}{k} - \ln n\right)$ есть постоянная Эйлера (т. I, стр. 444, 548).

При всяком целом $m \geqslant 2$

$$\sum_{k=0}^{\infty} \frac{1}{(x+k)^m} = \frac{(-1)^m}{(m-1)!}\frac{d^m}{dx^m}\ln\Gamma(x).$$

Формула дополнения:

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}.$$

Производная от логарифма гамма-функции:

$$\frac{\Gamma'(x)}{\Gamma(x)} = \frac{d}{dx}\ln\Gamma(x) = -\frac{1}{x} - \gamma - \sum_{k=1}^{\infty}\left(\frac{1}{x+k} - \frac{1}{k}\right) \quad \text{(стр. 356).}$$

При любом комплексном z (за исключением точек $z=0, -1, -2, \ldots$) гамма-функция определяется равенством

$$(1-e^{2\pi i z})\Gamma(z) = \int_C t^{z-1}e^{-t}\,dt,$$

где путь интегрирования C по комплексной переменной t окружает действительную положительную полуось и приближается к ней асимптотически сверху и снизу (см. рис. 117, стр. 588).

Бета-функция $B(x, y)$ (стр. 358—361) определяется при $x>0$ и $y>0$ равенством

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt = \int_{-1/2}^{+1/2}\left(\frac{1}{2}+t\right)^{x-1}\left(\frac{1}{2}-t\right)^{y-1}dt =$$

$$= 2\int_0^{\frac{\pi}{2}} \sin^{2x-1}\varphi \cos^{2y-1}\varphi\, d\varphi.$$

Связь между функциями бета и гамма:

$$\mathrm{B}(x, y)=\frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}.$$

Функциональное уравнение бета-функции:

$$\mathrm{B}(x+1, y)=\frac{x}{x+y}\,\mathrm{B}(x, y).$$

5. Теоремы о среднем значении. Формула и ряд Тэйлора

Теорема о среднем значении для функции двух переменных (стр. 97):

$$f(x+h, y+k)-f(x, y)=hf_x(x+\theta h, y+\theta k)+kf_y(x+\theta h, y+\theta k),$$
$$0<\theta<1.$$

Формула Тэйлора с остаточным членом для $f(x, y)$ (стр. 98):

$$f(x+h, y+k)-f(x, y)=hf_x+kf_y+\frac{1}{2!}\{h^2f_{xx}+2hkf_{xy}+k^2f_{yy}\}+$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$+\frac{1}{n!}\left\{h^nf_{x^n}+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}+\ldots+k^nf_{y^n}\right\}+R_n,$$

где остаточный член R_n (в символической записи стр. 97) имеет следующий вид:

$$R_n=\frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{(n+1)}f(x+\theta h, y+\theta k), \qquad 0<\theta<1.$$

Если $R_n\to 0$ при $n\to\infty$, то функция разлагается в бесконечный *ряд Тэйлора*:

$$f(x+h, y+k)=$$
$$=f(x, y)+\frac{1}{1!}\{hf_x+kf_y\}+\frac{1}{2!}\{h^2f_{xx}+2hkf_{xy}+k^2f_{yy}\}+\ldots$$
$$\ldots+\frac{1}{n!}\left\{h^nf_{x^n}+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}+\ldots+k^nf_{y^n}\right\}+\ldots$$

Теоремы о среднем значении для двойного интеграла (стр. 255—256):

$$\iint\limits_G f(x, y)\,dS=\mu S,$$

где S есть площадь области G, а μ есть некоторое число, промежуточное между наибольшим и наименьшим значениями функции $f(x, y)$ в области G.

Аналогично, если непрерывная функция $p(x, y)$ не изменяет своего знака в области G, то

$$\iint\limits_G f(x, y)\,p(x, y)\,dS=\mu\iint\limits_G p(x, y)\,dS.$$

6. Векторы

Определение вектора см. на стр. 17.

Пусть пространственный вектор задан своими координатами:

$$\boldsymbol{a} = \{a_1,\ a_2,\ a_3\}.$$

Модуль (длина) вектора: $a = |\boldsymbol{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$.

При повороте координатных осей координаты вектора преобразуются по тем же формулам, что и координаты точки.

Сумма векторов $\boldsymbol{a} = \{a_1,\ a_2,\ a_3\}$ и $\boldsymbol{b} = \{b_1,\ b_2,\ b_3\}$. Вектор

$$\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} = \{a_1 + b_1,\ \ a_2 + b_2,\ \ a_3 + b_3\}.$$

Произведение вектора на число (стр. 21):

$$\lambda\boldsymbol{a} = \boldsymbol{a}\lambda = \{\lambda a_1,\ \lambda a_2,\ \lambda a_3\}.$$

Скалярное произведение (стр. 21):

$$\boldsymbol{ab} = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cdot \cos\delta = a_1 b_1 + a_2 b_2 + a_3 b_3,$$

где δ есть угол между векторами $\boldsymbol{a}$ и $\boldsymbol{b}$.

Векторное произведение (стр. 28): $\boldsymbol{c} = [\boldsymbol{ab}]$ есть вектор $\boldsymbol{c} = \{c_1,\ c_2,\ c_3\}$ с координатами

$$c_1 = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \qquad c_2 = -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \qquad c_3 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Смешанное произведение трех векторов (стр. 31, 51):

$$\boldsymbol{abc} = \boldsymbol{a}\,[\boldsymbol{bc}] = [\boldsymbol{ab}]\,\boldsymbol{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}, \qquad \boldsymbol{abc} = \boldsymbol{bca} = \boldsymbol{cab}, \qquad \boldsymbol{bac} = -\,\boldsymbol{abc}.$$

Двойное векторное произведение (стр. 51):

$$[[\boldsymbol{ab}]\,\boldsymbol{c}] = \boldsymbol{b}\,(\boldsymbol{ac}) - \boldsymbol{a}\,(\boldsymbol{bc}), \qquad [\boldsymbol{a}[\boldsymbol{bc}]] = \boldsymbol{b}(\boldsymbol{ac}) - \boldsymbol{c}\,(\boldsymbol{ab}).$$

Дифференцирование вектор-функции скалярного аргумента $\boldsymbol{u} = \boldsymbol{u}\,(t)$ (стр. 103). Определение производной:

$$\dot{\boldsymbol{u}} = \frac{d\boldsymbol{u}}{dt} = \lim_{\Delta t \to 0} \frac{\boldsymbol{u}\,(t + \Delta t) - \boldsymbol{u}\,(t)}{\Delta t}.$$

$$\frac{d}{dt}(\boldsymbol{u} \pm \boldsymbol{v}) = \frac{d\boldsymbol{u}}{dt} \pm \frac{d\boldsymbol{v}}{dt}; \qquad \frac{d\,(\boldsymbol{uv})}{dt} = \frac{d\boldsymbol{u}}{dt}\,\boldsymbol{v} + \boldsymbol{u}\,\frac{d\boldsymbol{v}}{dt};$$

$$\frac{d}{dt}(\varphi\boldsymbol{u}) = \frac{d\varphi}{dt}\,\boldsymbol{u} + \varphi\,\frac{d\boldsymbol{u}}{dt}; \qquad \frac{d}{dt}[\boldsymbol{uv}] = \left[\frac{d\boldsymbol{u}}{dt}\,\boldsymbol{v}\right] + \left[\boldsymbol{u}\,\frac{d\boldsymbol{v}}{dt}\right].$$

Производная от скалярной функции точки $f\,(P) = f\,(\boldsymbol{r}) = f\,(x,\ y,\ z)$ *по направлению вектора* $\boldsymbol{c}$. Определение:

$$D_{\boldsymbol{c}} f = \frac{\partial f}{\partial \boldsymbol{c}} = \lim_{s \to 0} \frac{f\,(\boldsymbol{r} + s\boldsymbol{c}^\circ) - f\,(\boldsymbol{r})}{s} =$$

$$= \lim_{s \to 0} \frac{f\,(x + s\cos\alpha,\ y + s\cos\beta,\ z + s\cos\gamma) - f\,(x,\ y,\ z)}{s}$$

где c° — единичный вектор направления дифференцирования (вектора c), а α, β, γ — углы, образуемые вектором c с осями координат.

Отсюда

$$D_c = \frac{\partial}{\partial c} = \cos\alpha \frac{\partial}{\partial x} + \cos\beta \frac{\partial}{\partial y} + \cos\gamma \frac{\partial}{\partial z},$$

$$D_{e_1} f = \frac{\partial f}{\partial x}, \qquad D_{e_2} f = \frac{\partial f}{\partial y}, \qquad D_{e_3} f = \frac{\partial f}{\partial z},$$

$$D_{-e_1} f = -\frac{\partial f}{\partial x}, \qquad D_{-e_2} f = -\frac{\partial f}{\partial y}, \qquad D_{-e_3} f = -\frac{\partial f}{\partial z},$$

где e_1, e_2, e_3 — орты осей координат.

На плоскости xy

$$D_c = \frac{\partial}{\partial c} = \cos\alpha \frac{\partial}{\partial x} + \sin\alpha \frac{\partial}{\partial y},$$

где α — угол от оси x до вектора c.

Дифференциальные операции в скалярном и векторном поле.

Скалярной функции точки $f(P) = f(\boldsymbol{r}) = f(x, y, z)$ относят вектор, называемый ее *градиентом*

$$\operatorname{grad} f = \left\{\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right\} \text{ (стр. 109)}.$$

Производная от функции $f(P)$ по направлению вектора c равна проекции градиента этой функции на направление c:

$$D_c f = \frac{\partial f}{\partial c} = \boldsymbol{c}^\circ \operatorname{grad} f.$$

Векторному полю $\boldsymbol{u}(P) = \boldsymbol{u}(\boldsymbol{r}) = \boldsymbol{u}(x, y, z)$ относят скалярную функцию точки, называемую *дивергенцией* поля:

$$\operatorname{div} \boldsymbol{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \qquad \text{(стр. 112)},$$

и векторную функцию точки, *ротор* поля:

$$\operatorname{rot} \boldsymbol{u} = \left\{\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}, \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}, \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right\} \qquad \text{(стр. 112)}.$$

Пользуясь символическим вектором *набла* $\nabla = \left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\}$, имеем

$$\operatorname{grad} f = \nabla f, \qquad \operatorname{div} \boldsymbol{u} = \nabla \boldsymbol{u}, \qquad \operatorname{rot} \boldsymbol{u} = [\nabla \boldsymbol{u}] \qquad \text{(стр. 113)}.$$

Дифференциальные операции второго порядка:

$$\operatorname{rot} \operatorname{grad} f = 0, \qquad \operatorname{div} \operatorname{grad} f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2},$$

$$\operatorname{div} \operatorname{rot} \boldsymbol{u} = 0, \qquad \operatorname{rot} \operatorname{rot} \boldsymbol{u} = \operatorname{grad} \operatorname{div} \boldsymbol{u} - \nabla^2 \boldsymbol{u}.$$

7. Преобразование кратных интегралов

Если ориентированная область G плоскости xy отображается на соответствующим образом ориентированную область G' плоскости uv с помощью взаимно однозначного преобразования, якобиан которого $D = \frac{\partial(x, y)}{\partial(u, v)}$ нигде не обращается в нуль, то

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_{G'} f(x, y)\,D\,du\,dv \qquad \text{(стр. 275, 401).}$$

Аналогичная формула для любого числа измерений дана на стр. 276.

В частности, преобразование к *полярным координатам* на плоскости $x = r\cos\theta$, $y = r\sin\theta$ приводит к формуле

$$\iint\limits_G f(x, y)\,dx\,dy = \iint\limits_{G'} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta \qquad \text{(стр. 276).}$$

Преобразование к *сферическим координатам* (полярным координатам в пространстве):

$$x = r\sin\theta\cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad z = r\cos\theta,$$

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz =$$

$$= \iiint\limits_{G'} f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)\,r^2\sin\theta\,dr\,d\theta\,d\varphi \qquad \text{(стр. 277).}$$

8. Интегральные теоремы Гаусса — Остроградского, Грина и Стокса

Определение криволинейного интеграла см. стр. 368 и след.

А. На плоскости

Если область G односвязна, то криволинейный интеграл

$$\int\limits_{AB} (F_1\,dx + F_2\,dy) = \int\limits_{AB} \boldsymbol{F}\,d\boldsymbol{r}$$

не зависит от пути, соединяющего в области G любые две точки A и B этой области, в том и только в том случае, если *критерий интегрируемости*

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$$

выполняется во всех точках области G. При этом условии, если фиксировать начальную точку пути, то криволинейный интеграл является функцией $U(\xi, \eta)$ конечной точки и вектор $\boldsymbol{F} = \{F_1, F_2\}$ есть градиент этой функции:

$$\boldsymbol{F} = \operatorname{grad} U \qquad \text{(стр. 377—378).}$$

Теорема Гаусса [*теорема Остроградского на плоскости*]. Пусть G есть односвязная область, а C — ее контур. Тогда

$$\iint\limits_G \{f_x(x, y) + g_y(x, y)\}\,dx\,dy = \oint\limits_C \{f(x, y)\,dy - g(x, y)\,dx\} \qquad \text{(стр. 384)}$$

или, в векторной записи,

$$\iint\limits_G \operatorname{div} \boldsymbol{F}\,dS = \oint\limits_C \boldsymbol{F}\boldsymbol{n}^\circ\,ds = \oint\limits_C F_n\,ds, \qquad \text{(стр. 388)}$$

где $\boldsymbol{n}^\circ$ — единичный вектор внешней нормали граничной кривой C, F_n — проекция вектора поля $\boldsymbol{F} = \{f, g\}$ на направление внешней нормали, dS — элемент площади области G, а ds — дифференциал дуги контура C.

Теоремы Грина (стр. 390):

$$\iint_G (u_x v_x + u_y v_y)\,dx\,dy = -\iint_G u\nabla^2 v\,dx\,dy + \oint_C (-uv_y\,dx + uv_x\,dy) =$$

$$= -\iint_G v\nabla^2 u\,dx\,dy + \oint_C (-vu_y\,dx + vu_x\,dy),$$

$$\iint_G (u\nabla^2 v - v\nabla^2 u)\,dx\,dy = \oint_C \{(vu_y - uv_y)\,dx - (vu_x - uv_x)\,dy\} =$$

$$= \oint_C (uD_n v - vD_n u)\,ds.$$

Векторная запись первой формулы Грина:

$$\iint_G \operatorname{grad} u \operatorname{grad} v\,dS + \iint_G u \operatorname{div} \operatorname{grad} v\,dS = \oint_C uD_n v\,ds,$$

где $\nabla^2 v = \operatorname{div} \operatorname{grad} v = v_{xx} + v_{yy}$, а D_n — оператор дифференцирования по направлению внешней нормали.

Б. В трехмерном пространстве

Если область G односвязна, то криволинейный интеграл

$$\int_{AB} (F_1\,dx + F_2\,dy + F_3\,dz) = \int_{AB} \boldsymbol{F}\,d\boldsymbol{r}$$

не зависит от пути интегрирования, соединяющего в G любые две точки A и B этой области, в том и только в том случае, если выполняется критерий интегрируемости

$$\operatorname{rot} F = 0$$

или

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}, \quad \frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}, \quad \frac{\partial F_3}{\partial x} = \frac{\partial F_1}{\partial z} \qquad \text{(стр. 382).}$$

Интеграл по поверхности (стр. 406) имеет следующий вид:

$$\iint_\Sigma \{F_1(x, y, z)\,dy\,dz + F_2(x, y, z)\,dz\,dx + F_3(x, y, z)\,dx\,dy\} =$$

$$= \iint_\Sigma \boldsymbol{F}(P)\,\boldsymbol{n}^\circ\,d\sigma = \iint_\Sigma F_n\,d\sigma,$$

где вектор $\boldsymbol{F} = \{F_1, F_2, F_3\}$, $\boldsymbol{n}^\circ$ — единичный нормальный вектор поверхности Σ, F_n — проекция вектора $\boldsymbol{F}$ на направление $\boldsymbol{n}^\circ$, а $d\sigma$ — элемент поверхности Σ.

Если поверхность Σ задана параметрическими уравнениями $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, то поверхностный интеграл записывается так:

$$\iint_\Sigma = \iint_B \left\{F_1(x, y, z)\frac{\partial(y, z)}{\partial(u, v)} + F_2(x, y, z)\frac{\partial(z, x)}{\partial(u, v)} + \right.$$

$$\left. + F_3(x, y, z)\frac{\partial(x, y)}{\partial(u, v)}\right\} du\,dv,$$

причем ориентация области B плоскости uv должна быть согласована с ориентацией поверхности Σ.

Теорема Гаусса — Остроградского (стр. 408—411). Пусть $n^\circ = \{\cos\alpha, \cos\beta, \cos\gamma\}$ — единичный вектор внешней нормали замкнутой поверхности Σ, и пусть вектор поля есть $F = \{F_1, F_2, F_3\}$. Тогда

$$\iiint\limits_G \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\right) dx\,dy\,dz = \oiint\limits_\Sigma (F_1\cos\alpha + F_2\cos\beta + F_3\cos\gamma)\,d\sigma$$

или, в векторной записи,

$$\iiint\limits_G \operatorname{div} F\,dV = \oiint\limits_\Sigma Fn^\circ\,d\sigma = \oiint\limits_\Sigma F_n\,d\sigma,$$

где тройной интеграл берется по области G, а двойной интеграл — по замкнутой поверхности Σ, ограничивающей эту область.

Теоремы Грина (стр. 414):

$$1)\quad \iiint\limits_G (u_x v_x + u_y v_y + u_z v_z)\,dx\,dy\,dz + \iiint\limits_G u\nabla^2 v\,dx\,dy\,dz = \oiint\limits_\Sigma u\,D_n v\,d\sigma.$$

$$2)\quad \iiint\limits_G (u\nabla^2 v - v\nabla^2 u)\,dx\,dy\,dz = \oiint\limits_\Sigma \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) d\sigma,$$

где $\dfrac{\partial}{\partial n} = D_n$ есть оператор дифференцирования по направлению внешней нормали, а $\nabla^2 = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}$ есть оператор Лапласа.

Теорема Стокса (стр. 417). Пусть кусок поверхности Σ ограничен замкнутой кривой C, ориентированной соответствующим образом. Тогда

$$\iint\limits_\Sigma \left\{\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) dy\,dz + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) dz\,dx + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) dx\,dy\right\} = \oint\limits_C (F_1\,dx + F_2\,dy + F_3\,dz)$$

или, в векторной записи,

$$\iint\limits_\Sigma n^\circ \operatorname{rot} F\,d\sigma = \iint\limits_\Sigma (\operatorname{rot} F)_n\,d\sigma = \oint\limits_C F_t\,ds,$$

где вектор поля $F = \{F_1, F_2, F_3\}$, n° — единичный вектор нормали к Σ, $(\operatorname{rot} F)_n$ — проекция вектора $\operatorname{rot} F$ на направление n°, F_t — проекция вектора поля на направление касательной к граничной кривой C. Направление обхода контура C согласовано с нормальным вектором поверхности по правилу правого винта.

9. Максимумы и минимумы функции многих переменных

Свободные экстремумы функции двух переменных $u = f(x, y)$. *Необходимые* условия экстремума:

$$f_x = 0, \qquad f_y = 0 \qquad \text{(стр. 202)},$$

или f_x, f_y *не существуют* (стр. 205).

Достаточные условия экстремума. Если в некоторой точке (x_0, y_0) выполнены условия $f_x = 0$, $f_y = 0$, и при этом

$$f_{xx}f_{yy} - f_{xy}^2 > 0,$$

то в этой точке функция $u = f(x, y)$ имеет экстремум. Этот экстремум будет максимумом или минимумом, смотря по тому, каков знак второй производной f_{xx} в точке (x_0, y_0): отрицательный или положительный. Отличить максимум от минимума можно и по знаку f_{yy}, ибо в этом случае f_{xx} и f_{yy} имеют одинаковые знаки.

Если в точке (x_0, y_0)

$$f_{xx}f_{yy} - f_{xy}^2 < 0,$$

то в этой точке нет экстремума, — она является седловиной (стр. 225).

Если в точке (x_0, y_0)

$$f_{xx}f_{yy} - f_{xy}^2 = 0,$$

то вопрос о наличии в ней экстремума требует дополнительного исследования.

Условные или относительные экстремумы. Метод неопределенных множителей (стр. 206—216). Если n аргументов функции $u = f(x_1, x_2, \ldots, x_n)$ связаны между собой с помощью m дополнительных условий $(m < n)$

$$\varphi_1(x_1, \ldots, x_n) = 0, \quad \varphi_2(x_1, \ldots, x_n) = 0, \ \ldots, \quad \varphi_m(x_1, \ldots, x_n) = 0, \qquad (*)$$

то надо ввести m неопределенных множителей $\lambda_1, \lambda_2, \ldots, \lambda_m$ и построить вспомогательную функцию

$$F(x_1, \ldots, x_n) = f + \lambda_1\varphi_1 + \lambda_2\varphi_2 + \ldots + \lambda_m\varphi_m.$$

Тогда m условий связи (*) и n уравнений

$$\frac{\partial F}{\partial x_1} = 0, \quad \frac{\partial F}{\partial x_2} = 0, \ \ldots, \ \frac{\partial F}{\partial x_n} = 0$$

дают систему $(m + n)$ *необходимых* условий экстремума.

10. Кривые и поверхности

(Сводка формул для плоских кривых дана в т. I, стр. 657—658. Формулы для кривизны кривой $F(x, y) = 0$ (с учетом знака) даны в т. II, стр. 147.)

А. Пространственные кривые

Уравнение кривой

$$\boldsymbol{r} = \boldsymbol{r}(t)$$

или

$$x = x(t), \qquad y = y(t), \qquad z = z(t).$$

Касательный вектор

$$\boldsymbol{T} = \dot{\boldsymbol{r}} = \frac{d\boldsymbol{r}}{dt} = \{\dot{x}, \dot{y}, \dot{z}\}$$

направлен в сторону возрастания параметра t.

Единичный касательный вектор:

$$\boldsymbol{T}^\circ = \frac{\dot{\boldsymbol{r}}}{|\dot{\boldsymbol{r}}|} = \boldsymbol{t}$$

Уравнение касательной плоскости (стр. 82, 151):

а) $\zeta - z = (\xi - x) f_x + (\eta - y) f_y$;

б) $(\xi - x) F_x + (\eta - y) F_y + (\zeta - z) F_z = 0$;

в) $(\xi - x) \begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix} - (\eta - y) \begin{vmatrix} x_u & z_u \\ x_v & z_v \end{vmatrix} + (\zeta - z) \begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix} = 0.$

В. Огибающие (стр. 188—199)

а) Для того чтобы получить огибающую кривую однопараметрического семейства плоских кривых

$$f(x, y, c) = 0$$

или огибающую поверхность однопараметрического семейства поверхностей

$$f(x, y, z, c) = 0,$$

надо найти дискриминантную кривую или дискриминантную поверхность, исключая параметр c из уравнений

$$f = 0, \qquad f_c = 0.$$

Дискриминантная кривая или поверхность содержит огибающую, а также и геометрическое место особых точек.

Если однопараметрическое семейство кривых задано в параметрическом виде

$$x = \varphi(t, c), \qquad y = \psi(t, c),$$

где t — параметр на кривой, c — параметр семейства, то дискриминантная кривая получается исключением параметров c и t из уравнений

$$x = \varphi(t, c), \qquad y = \psi(t, c), \qquad \varphi_t\psi_c - \varphi_c\psi_t = 0 \qquad (\text{стр. } 191).$$

б) Огибающая двупараметрического семейства поверхностей

$$f(x, y, z, c_1, c_2) = 0$$

содержится в уравнении, полученном путем исключения двух параметров c_1 и c_2 из системы уравнений

$$f = 0, \qquad f_{c_1} = 0, \qquad f_{c_2} = 0.$$

11. Длина дуги, площадь, объем

Формулы для длины дуги *плоской кривой* и площади *плоской* фигуры даны в т. I, стр. 658—659.

Длина дуги пространственной кривой

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

есть

$$s = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt \qquad (\text{стр. } 105).$$

Гауссовы коэффициенты на кривой поверхности.

Если поверхность задана в параметрическом виде

$$\mathbf{r} = \mathbf{r}(u, v) \text{ или } x = x(u, v), \quad y = y(u, v), \quad z = z(u, v),$$

(направлен в сторону возрастания параметра t) или

$$\mathbf{t}=\frac{d\boldsymbol{r}}{ds}=\left\{\frac{dx}{ds},\ \frac{dy}{ds},\ \frac{dz}{ds}\right\}.$$

(направлен в сторону возрастания длины дуги s).

Главный нормальный вектор:

$$\frac{d\mathbf{t}}{ds}=\frac{d^2\boldsymbol{r}}{ds^2}=\left\{\frac{d^2x}{ds^2},\ \frac{d^2y}{ds^2},\ \frac{d^2z}{ds^2}\right\}\perp\mathbf{t}\qquad(\text{стр. } 106).$$

Единичный главный нормальный вектор:

$$\boldsymbol{n}=\frac{\dfrac{d\mathbf{t}}{ds}}{\left|\dfrac{d\mathbf{t}}{ds}\right|}.$$

Бинормальный вектор (единичный):

$$\boldsymbol{b}=[\mathbf{t}\boldsymbol{n}]\qquad(\text{стр. } 115).$$

Кривизна $k>0$ (стр. 107, 115):

$$k=\boldsymbol{n}\frac{d\mathbf{t}}{ds}=\left|\frac{d\mathbf{t}}{ds}\right|=\left|\frac{d^2\boldsymbol{r}}{ds^2}\right|=\sqrt{\left(\frac{d^2x}{ds^2}\right)^2+\left(\frac{d^2y}{ds^2}\right)^2+\left(\frac{d^2z}{ds^2}\right)^2},$$

$$k=\frac{\left|\left[\dfrac{d\boldsymbol{r}}{dt}\,\dfrac{d^2\boldsymbol{r}}{dt^2}\right]\right|}{\left|\dfrac{d\boldsymbol{r}}{dt}\right|^3}\qquad(\text{стр. } 132).$$

Кручение (стр. 115):

$$\varkappa=-\boldsymbol{n}\frac{d\boldsymbol{b}}{ds}\qquad(\text{стр. 615, решение упр. 7}).$$

Формулы Френэ (стр. 115):

$$1)\ \frac{d\mathbf{t}}{ds}=k\boldsymbol{n},\qquad 2)\ \frac{d\boldsymbol{n}}{ds}=-k\mathbf{t}+\varkappa\boldsymbol{b},\qquad 3)\ \frac{d\boldsymbol{b}}{ds}=-\varkappa\boldsymbol{n}.$$

Уравнение соприкасающейся плоскости (стр. 114, 615):

$$\begin{vmatrix} X-x & \dot{x} & \ddot{x} \\ Y-y & \dot{y} & \ddot{y} \\ Z-z & \dot{z} & \ddot{z} \end{vmatrix}=0,$$

где точками обозначены производные по параметру t.

Б. Поверхности

Уравнение поверхности:

а) $z=f(x,\ y)$; б) $F(x,\ y,\ z)=0$;

в) $\boldsymbol{r}=\boldsymbol{r}(u,\ v)$ или $x=x(u,\ v)$, $y=y(u,\ v)$, $z=z(u,\ v)$.

Нормальный вектор:

а) $\boldsymbol{n}=\{f_x,\ f_y,\ -1\}$; б) $\boldsymbol{n}=\{F_x,\ F_y,\ F_z\}$;

в) $\boldsymbol{n}=[\boldsymbol{r}_u\boldsymbol{r}_v]=\left\{\begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix},\ -\begin{vmatrix} x_u & z_u \\ x_v & z_v \end{vmatrix},\ \begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix}\right\}.$

то гауссовыми коэффициентами называются величины

$$\left.\begin{aligned} E &= (\mathbf{r}_u)^2 = x_u^2 + y_u^2 + z_u^2, \\ F &= \mathbf{r}_u\mathbf{r}_v = x_u x_v + y_u y_v + z_u z_v, \\ G &= (\mathbf{r}_v)^2 = x_v^2 + y_v^2 + z_v^2 \end{aligned}\right\} \quad \text{(стр. 180).}$$

Тогда

$$EG - F^2 = [\mathbf{r}_u\mathbf{r}_v]^2 = \begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix}^2 + \begin{vmatrix} x_u & z_u \\ x_v & z_v \end{vmatrix}^2 + \begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix}^2 \quad \text{(стр. 181).}$$

Длина дуги кривой, лежащей на этой поверхности, есть

$$s = \int_{t_0}^{t_1} \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}\, dt \quad \text{(стр. 180).}$$

Площадь куска Σ кривой поверхности, проекция которого на плоскость xy есть область G этой плоскости, выражается двойным интегралом

$$\Sigma = \iint_{\Sigma} d\sigma.$$

Если уравнение поверхности $z = f(x, y)$, то

$$\Sigma = \iint_{G} \sqrt{1 + f_x^2 + f_y^2}\, dx\, dy.$$

Если уравнение поверхности $F(x, y, z) = 0$, то

$$\Sigma = \iint_{G} \frac{1}{F_z} \sqrt{F_x^2 + F_y^2 + F_z^2}\, dx\, dy.$$

Если уравнение поверхности дано в параметрическом виде $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, то

$$\Sigma = \iint_{B} \sqrt{EG - F^2}\, du\, dv,$$

причем последний интеграл берется по области B плоскости uv, соответствующей области G плоскости xy.

Площадь поверхности вращения

$$x = u\cos v, \qquad y = u\sin v, \qquad z = \varphi(u),$$

полученной при вращении кривой $z = \varphi(x)$ плоскости zx вокруг оси z, есть

$$\Sigma = 2\pi \int_{u_0}^{u_1} u\sqrt{1 + \{\varphi'(u)\}^2}\, du = 2\pi \int_{s_0}^{s_1} u\, ds,$$

где s — длина дуги меридианной кривой $z = \varphi(x)$ (стр. 296; см. также т. I, стр. 329).

Площадь ω_n единичной сферы $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$ в n-мерном пространстве дается формулой (стр. 325).

$$\omega_n = \frac{2(\sqrt{\pi})^n}{\Gamma\left(\frac{n}{2}\right)}.$$

Объемы. Объем вертикального цилиндрического бруса, ограниченного областью G плоскости xy и куском поверхности S с уравнением $z = f(x, y)$, выражается формулой

$$V = \iint_{G} f(x, y)\, dx\, dy.$$

Если пространственная область (V) полностью ограничена замкнутой поверхностью Σ, то объем V этой области есть

$$V=\iiint_{(V)} dx\, dy\, dz=\oiint_{\Sigma} z\, dx\, dy=-\oiint_{\Sigma} x\, dy\, dz=-\oiint_{\Sigma} y\, dz\, dx$$

(стр. 411)

В *сферических координатах* этот же объем выражается так:

$$V=\iiint_{B} r^2 \sin\theta\, dr\, d\theta\, d\varphi \qquad \text{(стр. 277, 289)},$$

где B есть область пространства $r\theta\varphi$, соответствующая области (V).

Объем тела, ограниченного *поверхностью вращения*

$$x=u\cos v, \qquad y=u\sin v, \qquad z=\varphi(u),$$

образованной вращением кривой $z=\varphi(x)$ плоскости zx вокруг оси z, есть

$$V=\pi\int_{z_0}^{z_1} u^2\, dz \text{ (стр. 289; см. также т. I, стр. 329).}$$

Объем v_n, ограниченный единичной сферой $x_1^2+x_2^2+\ldots+x_n^2=1$ в n-мерном пространстве есть $v_n=\dfrac{(\sqrt{\pi})^n}{\Gamma\left(\dfrac{n+2}{2}\right)}$ (стр. 326).

12. Вариационное исчисление

Необходимое и достаточное условие стационарности интеграла

$$I(u)=\int_{x_0}^{x_1} F(x, u, u')\, dx$$

выражается *уравнением Эйлера*

$$F_u-\frac{d}{dx}F_{u'}=0$$

или

$$F_{u'u'}u''+F_{uu'}u'+F_{xu'}-F_u=0 \qquad \text{(стр. 520).}$$

Если подынтегральная функция F содержит несколько функциональных аргументов $u_1(x), u_2(x), \ldots, u_n(x)$ и их производные, то необходимое и достаточное условие стационарности функционала

$$I(u)=\int_{x_0}^{x_1} F(x; u_1, \ldots, u_n; u_1', \ldots, u_n')\, dx$$

выражается системой уравнений Эйлера

$$F_{u_i}-\frac{d}{dx}F_{u_i'}=0 \qquad (i=1, 2, \ldots, n) \quad \text{(стр. 531).}$$

Если подынтегральная функция F зависит от x, $u(x)$, $u'(x)$ и $u''(x)$, то уравнение Эйлера будет

$$F_u-\frac{d}{dx}F_{u'}+\frac{d^2}{dx^2}F_{u''}=0 \quad \text{(стр. 536).}$$

Если надо найти стационарное значение функционала

$$\int_{t_0}^{t_1} F(x, y, z; \dot{x}, \dot{y}, \dot{z})\, dt$$

при добавочном условии $G(x, y, z) = 0$, то условия стационарности выражаются системой трех уравнений

$$\frac{d}{dt} F_{\dot{x}} - F_x = \lambda G_x,$$

$$\frac{d}{dt} F_{\dot{y}} - F_y = \lambda G_y,$$

$$\frac{d}{dt} F_{\dot{z}} - F_z = \lambda G_z,$$

где λ — множитель Лагранжа (стр. 539).

13. Аналитические функции

Определение см. на стр. 554.

Необходимое и достаточное условие того, чтобы функция

$$f(z) = f(x + iy) = u(x, y) + iv(x, y)$$

была аналитической в области G, выражается *условиями Коши — Римана*

$$u_x = v_y, \qquad u_y = -v_x \qquad \text{(стр. 554)}$$

(предполагается, что эти производные существуют и непрерывны).

Теорема Коши. Если функция $f(z)$ является аналитической в односвязной области G, то

$$\int_{z_0}^{z} f(\zeta)\, d\zeta$$

не зависит от пути интегрирования, соединяющего в области G две точки z_0 и z этой области или, что то же самое,

$$\oint_C f(\zeta)\, d\zeta = 0$$

вдоль любой замкнутой кривой C, лежащей в области G (стр. 561).

Формула Коши. Если $f(z)$ является аналитической функцией в односвязной области G и на его контуре C, то

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{\zeta - z}\, d\zeta,$$

где z есть любая точка, лежащая внутри контура C (стр. 569).

Если функция $f(z)$ является аналитической внутри и на границе круга $|z - z_0| \leqslant R$, то она разлагается в степенной ряд Тэйлора

$$f(z) = f(z_0) + \sum_{k=1}^{\infty} c_k (z - z_0)^k,$$

который сходится внутри этого круга. При этом

$$c_k = \frac{f^{(k)}(z_0)}{k!} = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{(\zeta - z_0)^{k+1}}\, d\zeta \qquad \text{(стр. 570—571).}$$

ОТВЕТЫ И УКАЗАНИЯ

ГЛАВА I

§ 1, стр. 26.

1. Обозначим орты (единичные векторы) осей Ox, Oy, Oz через $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ и орты осей Ox', Oy', Oz' через $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$, $\boldsymbol{e}'_3$. Тогда $\boldsymbol{e}'_i = \{\alpha_i, \beta_i, \gamma_i\}$. Следовательно, $\boldsymbol{e}'^2_i = \boldsymbol{e}'_i\boldsymbol{e}'_i = \alpha_i^2 + \beta_i^2 + \gamma_i^2 = 1$. Скалярные произведения $\boldsymbol{e}'_i\boldsymbol{e}'_k = \alpha_i\alpha_k + \beta_i\beta_k + \gamma_i\gamma_k = 0$ при $i \neq k$, ибо тогда $\boldsymbol{e}'_i \perp \boldsymbol{e}'_k$.

3. Обозначим радиус-векторы, идущие от начала O к вершинам треугольника A, B, C, через $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$. Тогда радиус-вектор середины D стороны AB будет $\overline{OD} = \frac{1}{2}(\boldsymbol{a} + \boldsymbol{b})$, а радиус-вектор центра масссы M вершин треугольника (ср. упр. 2)

$$\overline{OM} = \frac{\boldsymbol{c} + 2 \cdot \dfrac{\boldsymbol{a}+\boldsymbol{b}}{2}}{1+2} = \frac{1}{3}(\boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c}).$$

Это выражение не зависит от того, в каком порядке брать вершины треугольника.

4. Обозначим радиус-векторы, идущие от начала O к вершинам P, Q, R, S, через $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$, $\boldsymbol{s}$. Тогда радиус-вектор центра масс M вершин треугольника PQR есть $\overline{OM} = \frac{1}{3}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r})$ (ср. упр. 3), а радиус-вектор центра масс всех вершин тетраэдра будет (ср. упр. 2)

$$\overline{ON} = \frac{\boldsymbol{s} + 3 \cdot \dfrac{1}{3}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r})}{1+3} = \frac{1}{4}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r} + \boldsymbol{s}).$$

Это выражение не зависит от того, в каком порядке брать вершины тетраэдра.

5. Радиус-векторы вершин тетраэдра обозначим через $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{r}$, $\boldsymbol{s}$. Тогда радиус-векторы середин ребер, т. е. точек A, A', B, B', C, C', равны $\frac{1}{2}(\boldsymbol{p} + \boldsymbol{q})$, $\frac{1}{2}(\boldsymbol{r} + \boldsymbol{s})$, ..., $\frac{1}{2}(\boldsymbol{q} + \boldsymbol{r})$ и три отрезка AA', BB' и CC' имеют общую середину в точке с радиус-вектором $\frac{1}{4}(\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r} + \boldsymbol{s})$, а эта точка и есть центр массы вершин тетраэдра.

§ 2, стр. 33.

1. Обозначим через s направляющий вектор прямой l, $s = \{a, c, e\}$ и проведем вектор $\overline{AP}$ от точки $A(b, d, f)$ прямой к данной точке P. Тогда искомое расстояние равно

$$\frac{|[\overline{AP}s]|}{|s|} = \frac{1}{\sqrt{a^2 + c^2 + e^2}} \times$$
$$\times \sqrt{\begin{vmatrix} d - y_0 & f - z_0 \\ c & e \end{vmatrix}^2 + \begin{vmatrix} b - x_0 & f - z_0 \\ a & e \end{vmatrix}^2 + \begin{vmatrix} b - x_0 & d - y_0 \\ a & c \end{vmatrix}^2}.$$

2. Три вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ параллельны одной плоскости в том и только в том случае, если (после их приведения к общему началу) построенный на них параллелепипед имеет объем, равный нулю. Отсюда искомое условие:

$$[\boldsymbol{ab}]\boldsymbol{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0.$$

3. Задача сводится к тому, чтобы вывести условие, при котором две данные прямые лежат в одной плоскости. На первой прямой известна точка $A(x_1, y_1, z_1)$, на второй — точка $B(x_2, y_2, z_2)$. Данные две прямые лежат в одной плоскости в том и только в том случае, если направляющие векторы обеих прямых $\boldsymbol{a} = \{a_1, a_2, a_3\}$ и $\boldsymbol{b} = \{b_1, b_2, b_3\}$, а также вектор $\overline{AB}$ параллельны одной плоскости; отсюда искомое условие (см. упр. 2):

$$[\boldsymbol{ab}]\,\overline{AB} = 0, \text{ т. е. } \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{vmatrix} = 0.$$

4. 1 случай $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$. Прямые l и l' параллельны, и для них можно взять общий направляющий вектор $s = \{a, c, e\}$. На l известна точка $A(b, d, f)$, на l' — точка $A'(b', d', f')$.

$$\text{пр}_l\overline{AA'} = \frac{\overline{AA'}s}{|s|} = \frac{(b'-b)a + (d'-d)c + (f'-f)e}{\sqrt{a^2+c^2+e^2}}.$$

Расстояние между прямыми постоянно и есть

$$h = \sqrt{(AA')^2 - (\text{пр}_l\overline{AA'})^2},$$

где $(AA')^2 = (b'-b)^2 + (d'-d)^2 + (f'-f)^2$.

2 случай. Прямые l и l' не параллельны. Как правило, они и не пересекаются. Направляющие векторы прямых l и l' суть $s = \{a, c, e\}$ и $s' = \{a', c', e'\}$. На l известна точка $A(b, d, f)$, на l' — точка $A'(b', d', f')$. Сперва доказать, что прямые l и l' имеют общий перпендикуляр, т. е. существует прямая (единственная), пересекающая обе данные прямые и перпендикулярная к каждой из них. За направляющий вектор этого общего перпендикуляра можно принять $[ss']$. Искомое кратчайшее расстояние h равно абсолютной величине проекции вектора $\overline{AA'}$ на направление $[ss']$, т. е. $h = \frac{|[ss']\overline{AA'}|}{|[ss']|}$ где

$$[ss'] = \left\{\begin{vmatrix} c & e \\ c' & e' \end{vmatrix}, -\begin{vmatrix} a & e \\ a' & e' \end{vmatrix}, \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}\right\} \quad \text{и} \quad \overline{AA'} = \{b'-b,\ d'-d,\ f'-f\}.$$

Следовательно,

$$h = \text{абс. вел.}\ \frac{(b'-b)\begin{vmatrix} c & e \\ c' & e' \end{vmatrix} - (d'-d)\begin{vmatrix} a & e \\ a' & e' \end{vmatrix} + (f'-f)\begin{vmatrix} a & c \\ a' & c' \end{vmatrix}}{|[ss']|} =$$

$$= \text{абс. вел.}\ \frac{\begin{vmatrix} a & c & e \\ a' & c' & e' \\ b'-b & d'-d & f'-f \end{vmatrix}}{\sqrt{(ce'-c'e)^2 + (ae'-a'e)^2 + (ac'-a'c)^2}}.$$

Если прямые l и l' пересекаются, то числитель обратится в нуль (см. упр. 3)

5. Левую часть уравнения можно истолковать как объем тетраэдра, построенного на векторах $\overline{PP}_1$, $\overline{PP}_2$ и $\overline{PP}_3$, где $P(x, y, z)$ — переменная точка плоскости.

6. Построим вектор $\boldsymbol{\omega} = \omega\{\alpha, \beta, \gamma\}$. Вектор-скорость точки P есть $\boldsymbol{v} = [\boldsymbol{\omega r}]$, где $\boldsymbol{r}$ — радиус-вектор точки P. Величина скорости равна $\omega\sqrt{(\beta z - \gamma y)^2 + (\alpha z - \gamma x)^2 + (\alpha y - \beta x)^2}$.

8. Определители второго порядка равны последовательно удвоенным площадям (с надлежащими знаками) треугольников OP_1P_2, OP_2P_3, ..., $OP_{n-1}P_n$, OP_nP_1, где O — начало координат. Если O лежит внутри многоугольника, то все слагаемые площади одинакового знака. Если же O лежит вне многоугольника, то часть треугольников обходится в одном направлении, а другая — в противоположном направлении (сделать чертеж!). В обоих случаях алгебраическая сумма даст площадь многоугольника (с тем или другим знаком).

9. а) Ввести три вектора $\boldsymbol{x} = \{a, b, c\}$, $\boldsymbol{y} = \{a_1, b_1, c_1\}$, $\boldsymbol{z} = \{a_2, b_2, c_2\}$. Тогда $D = [\boldsymbol{xy}]\boldsymbol{z}$. Так как для любых двух векторов $\boldsymbol{a}$ и $\boldsymbol{b}$

$$|[\boldsymbol{ab}]| \leqslant |\boldsymbol{a}|\cdot|\boldsymbol{b}| \quad \text{и} \quad |\boldsymbol{ab}| \leqslant |\boldsymbol{a}|\cdot|\boldsymbol{b}|,$$

то

$$D \leqslant |\boldsymbol{x}|\cdot|\boldsymbol{y}|\cdot|\boldsymbol{z}|.$$

б) В том и только в том случае, если векторы $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ попарно перпендикулярны.

§ 3, стр. 40.

2. Существование решения у заданной системы равносильно наличию у однородной системы уравнений

$$\begin{aligned} a_1x + a_2y + Az &= 0, \\ b_1x + b_2y + Bz &= 0, \\ c_1x + c_2y + Cz &= 0 \end{aligned}$$

такого решения, что $z = -1$. Поэтому необходимым условием существования решения является $D = 0$. Но если $D = 0$ и, например, $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0$, то третье уравнение заданной системы есть следствие первых двух, а система первых двух уравнений с двумя неизвестными x и y имеет решение, так как ее определитель не равен нулю.

3. Данные две прямые пересекаются, если система трех уравнений

$$\begin{aligned} a_1t + x_1 &= b_1\tau + x_2, \\ a_2t + y_1 &= b_2\tau + y_2, \\ a_3t + z_1 &= b_3\tau + z_2 \end{aligned}$$

с двумя неизвестными t и τ имеет решение, но для этого должно выполняться условие (см. упр. 2):

$$\begin{vmatrix} a_1 & b_1 & x_2 - x_1 \\ a_2 & b_2 & y_2 - y_1 \\ a_3 & b_3 & z_2 - z_1 \end{vmatrix} = 0.$$

См. геометрическое (векторное) решение этой задачи на стр. 609.

5. Вычесть элементы первой строки этого определителя из соответствующих элементов второй и третьей строки.

6. Вычесть элементы первой строки определителя из соответствующих элементов остальных трех строк.

7. а) 0; б) 2; в) 12; г) $(x - y)(y - z)(z - x)(x + y + z)$.

8. $a + c = 2b$.

9. $x = 3$, $y = 2$, $z = 1$.

§ 4, стр. 50.

1. $ab + a_1b_1 = 0$, $a^2 + a_1^2 = b^2 + b_1^2 = 1$.

3. При аффинном преобразовании координаты любого вектора $\boldsymbol{v}$ преобразуются по тем же формулам, что и координаты точки (см. § 4, n° 1, теорема 5). Достаточно поэтому показать, что существует такой не равный нулю вектор $\boldsymbol{u} = \{u_1, u_2, u_3\}$, который преобразуется в вектор $\boldsymbol{u}' = \lambda\boldsymbol{u}$, так что

$$\lambda u_1 = a_1u_1 + b_1u_2 + c_1u_3,$$
$$\lambda u_2 = a_2u_1 + b_2u_2 + c_2u_3,$$
$$\lambda u_3 = a_3u_1 + b_3u_2 + c_3u_3.$$

Теперь надо только найти такое значение λ, чтобы определитель этой однородной системы уравнений для неизвестных u_1, u_2, u_3 обратился в нуль; это даст для определения λ уравнение третьей степени с действительными коэффициентами, которое всегда имеет по крайней мере один действительный корень.

4.
$$x' = \frac{1}{2}(1 + \cos\varphi)\,x - \frac{\sqrt{2}}{2}\sin\varphi\cdot y - \frac{1}{2}(1 - \cos\varphi)\,z,$$
$$y' = \frac{\sqrt{2}}{2}\sin\varphi\cdot x + \cos\varphi\cdot y + \frac{\sqrt{2}}{2}\sin\varphi\cdot z,$$
$$z' = \frac{1}{2}(\cos\varphi - 1)\,x - \frac{\sqrt{2}}{2}\sin\varphi\cdot y + \frac{1}{2}(\cos\varphi + 1)\,z.$$

6. Согласно правилу умножения определителей, с учетом формул упр. 1, стр. 26, квадрат определителя равен $+1$.

См. на стр. 608 в решении задачи 1, стр. 26, обозначения ортов обеих систем координат. В исходной системе $[\boldsymbol{e}_1\boldsymbol{e}_2] = \boldsymbol{e}_3$ и $[\boldsymbol{e}_1\boldsymbol{e}_2]\,\boldsymbol{e}_3 = \boldsymbol{e}_3^2 = +1$. Если новая система координат имеет ту же ориентацию, что и старая, то $[\boldsymbol{e}_1'\boldsymbol{e}_2'] = \boldsymbol{e}_3'$ и $\Delta = [\boldsymbol{e}_1'\boldsymbol{e}_2']\,\boldsymbol{e}_3' = \boldsymbol{e}_3'\boldsymbol{e}_3' = \boldsymbol{e}_3'^2 = +1$. Если же ориентации обеих систем различны, то $[\boldsymbol{e}_1'\boldsymbol{e}_2'] = -\boldsymbol{e}_3'$ и $\Delta = [\boldsymbol{e}_1'\boldsymbol{e}_2']\,\boldsymbol{e}_3' = -\boldsymbol{e}_3'^2 = -1$.

Смешанные упражнения к гл. I, стр. 50.

1. а) Если бы существовало линейное соотношение $\alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c} = 0$, в котором, например, $\alpha \neq 0$, то, помножив это соотношение скалярно на $\boldsymbol{a}$, мы получили бы $\alpha\boldsymbol{a}\boldsymbol{a} + \beta\boldsymbol{a}\boldsymbol{b} + \gamma\boldsymbol{a}\boldsymbol{c} = 0$ или, так как $\boldsymbol{a} \perp \boldsymbol{b}$ и $\boldsymbol{a} \perp \boldsymbol{c}$, $\alpha\boldsymbol{a}\boldsymbol{a} = \alpha\,|\boldsymbol{a}|^2 = 0$; но $|\boldsymbol{a}| \neq 0$, следовательно, $\alpha = 0$, что противоречит условию.

б) Соотношение $\alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c} = 0$ равносильно системе линейных уравнений для α, β, γ:

$$\alpha a_1 + \beta b_1 + \gamma c_1 = 0,$$
$$\alpha a_2 + \beta b_2 + \gamma c_2 = 0,$$
$$\alpha a_3 + \beta b_3 + \gamma c_3 = 0.$$

Эта система уравнений имеет единственное решение $\alpha = \beta = \gamma = 0$ в том и только в том случае, если ее определитель не равен нулю.

в) Векторное уравнение $\boldsymbol{v} = \alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c}$ равносильно системе трех скалярных линейных уравнений для α, β, γ; а эта система непременно имеет решение, так как, согласно б), определитель системы не равен нулю.

2. Выберем какую-либо прямоугольную систему координат Ox, Oy, Oz и отнесем к ней все рассматриваемые векторы. Тогда тождество а) сведется к теореме умножения определителей. Тождество б) можно проверить, вычислив координаты вектора, стоящего в левой части, и вектора в правой части.

в) 1-й способ. Правая часть равна определителю

$$\begin{vmatrix} a_1x_1+a_2x_2+a_3x_3 & a_1y_1+a_2y_2+a_3y_3 \\ b_1x_1+b_2x_2+b_3x_3 & b_1y_1+b_2y_2+b_3y_3 \end{vmatrix};$$

его можно представить в виде суммы девяти определителей, из которых три обращаются в нуль, а остающуюся сумму нетрудно привести к виду

$$\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \cdot \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix},$$

а это и есть координатная запись левой стороны.

2-й способ. Левую часть можно рассматривать как смешанное произведение векторов $\boldsymbol{a}$, $\boldsymbol{b}$ и $[\boldsymbol{xy}]$, а в смешанном произведении можно заключить в квадратные скобки по желанию любую пару стоящих рядом сомножителей; поэтому

$$[\boldsymbol{ab}][\boldsymbol{xy}] = \boldsymbol{a}\,[\boldsymbol{b}\,[\boldsymbol{xy}]] = \boldsymbol{a}\,\{\boldsymbol{x}(\boldsymbol{by}) - \boldsymbol{y}\,(\boldsymbol{bx})\} = (\boldsymbol{ax})\,(\boldsymbol{by}) - (\boldsymbol{ay})\,(\boldsymbol{bx}),$$

что и требовалось. Двойное векторное произведение $[\boldsymbol{b}\,[\boldsymbol{xy}]]$ мы раскрыли по б).

г) 1-й способ. Сначала доказывают тождество

$$[\boldsymbol{a}\,[\boldsymbol{bc}]] + [\boldsymbol{b}\,[\boldsymbol{ca}]] + [\boldsymbol{c}\,[\boldsymbol{ab}]] = 0,$$

раскрывая каждый его член по б). Это значит, что три вектора, сумма которых записана в левой части последнего тождества, линейно зависимы; но тогда, согласно 1б), их смешанное произведение равно нулю.

2-й способ (прямым вычислением). Согласно б),

$$[\boldsymbol{a}\,[\boldsymbol{bc}]] = \boldsymbol{b}\,(\boldsymbol{ac}) - \boldsymbol{c}\,(\boldsymbol{ab}),$$
$$[\boldsymbol{b}\,[\boldsymbol{ca}]] = \boldsymbol{c}\,(\boldsymbol{ab}) - \boldsymbol{a}\,(\boldsymbol{bc}),$$
$$[\boldsymbol{c}\,[\boldsymbol{ab}]] = \boldsymbol{a}\,(\boldsymbol{bc}) - \boldsymbol{b}\,(\boldsymbol{ac}).$$

Векторы, записанные в первых двух строчках, перемножаем векторно и в итоге получим вектор, выраженный следующей суммой:

$$[[\boldsymbol{a}\,[\boldsymbol{bc}]]\,[\boldsymbol{b}\,[\boldsymbol{ca}]]] = [\boldsymbol{bc}]\,(\boldsymbol{ac})\,(\boldsymbol{ab}) + [\boldsymbol{ab}]\,(\boldsymbol{ac})\,(\boldsymbol{bc}) + [\boldsymbol{ca}]\,(\boldsymbol{ab})\,(\boldsymbol{bc}).$$

Помножив этот вектор скалярно на вектор, записанный в третьей строчке (в правой части перемножаем почленно, причем из шести членов четыре обращаются в нуль), получим

$$[\boldsymbol{a}\,[\boldsymbol{bc}]]\,[\boldsymbol{b}\,[\boldsymbol{ca}]]\,[\boldsymbol{c}\,[\boldsymbol{ab}]] = (\boldsymbol{abc})\,(\boldsymbol{ab})\,(\boldsymbol{bc})\,(\boldsymbol{ac}) - (\boldsymbol{cab})\,(\boldsymbol{ab})\,(\boldsymbol{bc})\,(\boldsymbol{ac}) = 0.$$

Наконец, следствие: векторы $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ после приведения к общему началу лежат соответственно на трех прямых, имеющих общую точку; плоскость, проходящая через вектор $\boldsymbol{a}$ перпендикулярно к векторам $\boldsymbol{b}$ и $\boldsymbol{c}$, содержит также и вектор $[\boldsymbol{bc}]$, а стало быть, вектор $[\boldsymbol{a}\,[\boldsymbol{bc}]]$ является ее нормальным вектором. Аналогично находим нормальные векторы двух других плоскостей, о которых идет речь. Согласно г) и 1б), эти три нормальных вектора лежат в одной плоскости; следовательно, три рассматриваемые плоскости (имея к тому же общую точку) проходят через общую прямую.

3. Обозначим орты осей Ox, Oy через $\boldsymbol{e}_1$ и $\boldsymbol{e}_2$, а орты осей Ox', Oy' через $\boldsymbol{i}_1$ и $\boldsymbol{i}_2$. Приравнивая друг другу разложения радиус-вектора произвольной точки плоскости по ортам обеих систем координат, имеем

$$x\boldsymbol{e}_1 + y\boldsymbol{e}_2 = x'\boldsymbol{i}_1 + y'\boldsymbol{i}_2.$$

Умножив это равенство скалярно на $\boldsymbol{e}_1$, получим

$$x = x'\boldsymbol{i}_1\boldsymbol{e}_1 + y'\boldsymbol{i}_2\boldsymbol{e}_1 = x'\cos\alpha + y'\cos\left(\frac{\pi}{2}+\alpha\right) = x'\cos\alpha - y'\sin\alpha,$$

так как $e_1e_1=1$, $e_1e_2=0$. Аналогично получается формула для y — скалярным умножением на e_2. Выражения для x' и y' через x и y получаются скалярным умножением того же равенства на l_1 и l_2.

4. Поворот системы $x'Oy'$ на угол β приводит к новой системе координат $x''Oy''$. Сравнение последовательного перехода от первой системы ко второй, а затем от второй к третьей, с прямым переходом от x, y к x'', y'' приводит к требуемому результату.

5. а) В системе координат Ox, Oy, Oz взять три вектора: $\{\alpha_1, \beta_1, \gamma_1\}$, $\{\alpha_2, \beta_2, \gamma_2\}$ и $\{\alpha_3, \beta_3, \gamma_3\}$. Если определитель Δ ортогональный, то эти векторы определяют новую прямоугольную координатную систему Ox', Oy', Oz'.

б) Переход от системы Ox', Oy', Oz' к системе Ox, Oy, Oz определяется таблицей косинусов:

	x'	y'	z'
x	α_1	α_2	α_3
y	β_1	β_2	β_3
z	γ_1	γ_2	γ_3

определитель которой

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix}$$

тоже будет ортогональным.

6. Переходим от системы Ox, Oy, Oz к системе Ox', Oy', Oz' с помощью следующих трех поворотов: 1) Вращаем систему Ox, Oy, Oz на угол φ вокруг оси Oz и получаем тем самым новую систему Ox_1, Oy_1, Oz_1 $(Oz=Oz_1)$. 2) Вращаем Ox_1, Oy_1, Oz_1 на угол θ вокруг оси Ox_1 $(Ox_1=Ox_2,\ Oz_2=Oz')$. 3) Систему Ox_2, Oy_2, Oz_2 повернем на угол ψ вокруг оси Oz_2, и в результате получим систему Ox', Oy', Oz'. В каждом из этих поворотов преобразование переменных совершается по формулам упр. 3. Затем надо из полученных трех преобразований исключить промежуточные переменные x_1, y_1, z_1, x_2, y_2, z_2; это лучше всего достигается перемножением (в правильном порядке) трех определителей составляющих преобразований.

7. Воспользоваться формулой $\cos(\widehat{xOx'})=\cos\varphi\cos\psi-\sin\varphi\sin\psi\cos\theta$.

8. Нормальный вектор плоскости $N=\{A, B, C\}$, а направляющий вектор прямой $b=\{b_1, b_2, b_3\}$. Обозначим угол между этими векторами через ψ, $\psi=(\widehat{N, b})$. По определению, углом φ между прямой и плоскостью называется острый угол между прямой и ее проекцией на эту плоскость; поэтому $\varphi=\frac{\pi}{2}-\psi$ (если ψ — острый угол), либо $\varphi=\psi-\frac{\pi}{2}$ (если ψ — тупой угол). В обоих случаях

$$\sin\varphi=|\cos\psi|=\frac{|Nb|}{|N|\,|b|},\quad \text{т. е.}\quad \varphi=\arcsin\frac{|Ab_1+Bb_2+Cb_3|}{\sqrt{A^2+B^2+C^2}\cdot\sqrt{b_1^2+b_2^2+b_3^2}}.$$

10.
$$D=\begin{vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}\cdot\begin{vmatrix} \cos\alpha & \cos\beta & \cos\gamma \\ \sin\alpha & \sin\beta & \sin\gamma \\ \sin(\beta-\gamma) & \sin(\gamma-\alpha) & \sin(\alpha-\beta) \end{vmatrix},$$

а первый множитель равен единице.

11. К первому столбцу прибавить второй и третий столбец и после этого вынести множитель $A+2B$ за знак определителя. Затем вычесть первую строку из второй и из третьей строки. В результате получим

$$D=(A+2B)(B-A)^2=\{(x+y+z)(x^2+y^2+z^2-xy-xz-yz)\}^2.$$

12. Вычесть первый столбец из всех остальных столбцов; тогда станет ясно, что определитель является линейной функцией от x. Коэффициенты A и B этой функции можно определить, подставляя в определитель один раз $x=-a$, другой раз $x=-b$.

13. Так как $uv=1$, то по правилу Лейбница (т. I, стр. 228—229) имеем

$$\begin{aligned} u'v+uv' &= 0, \\ u''v+2u'v'+uv'' &= 0, \\ u'''v+3u''v'+3u'v'' &= -uv'''. \end{aligned}$$

Эти уравнения можно рассматривать как систему линейных уравнений для v, v' и v'' с определителем, равным $(-D)$. Если найти из этой системы v (т. е. первую из неизвестных) по правилу, данному на стр. 39, то получим

$$v=-\frac{1}{D}\begin{vmatrix} 0 & u & 0 \\ 0 & 2u' & u \\ -uv''' & 3u'' & 3u' \end{vmatrix}=\frac{u^3v'''}{D},$$

откуда

$$v'''=\frac{Dv}{u^3}=\frac{D}{u^4}.$$

ГЛАВА II

§§ 1, 2, стр. 64.

2. $\frac{1}{2}(n+1)(n+2)$. 4. $\begin{vmatrix} a & b \\ b & c \end{vmatrix}>0$. 5. а) Нет. б) Нет. в) Нет. г) Нет. д) Да. е) Нет. ж) Да. з) Нет.

§ 3, стр. 73.

1. $\frac{1}{2}(n+1)(n+2)$. Ср. упр. 2, стр. 64.

4. а) и б) для определителя четвертого порядка воспользоваться результатом упр. 4, стр. 40.

в) Умножить определитель Δ на искомый определитель по правилу умножения определителей.

§§ 4—5, стр. 95.

1. $ax_0x+by_0y+cz_0z=1$.

2. Если поместить начало координат в вершину конической поверхности, то ее уравнение можно записать в виде $u=x\varphi\left(\frac{y}{x}\right)$.

5. $6x+2(a+e+k)$. Воспользоваться теоремой упр. 4.

6. Воспользоваться теоремой упр. 4.

7. а) $f_{xx}+f_{yy}+f_{zz}=g_{rr}+\frac{2}{r}g_r$.

б) Найти общее решение обыкновенного дифференциального уравнения $g_{rr}+\frac{2}{r}g_r=0$.

8. $g_{rr}+\frac{n-1}{r}g_r$.

10. $f_{xx}+f_{yy}+f_{zz}=\frac{1}{r^2\sin\theta}\left\{\frac{\partial}{\partial r}(r^2f_r\sin\theta)+\frac{\partial}{\partial\varphi}\left(\frac{f_\varphi}{\sin\theta}\right)+\frac{\partial}{\partial\theta}(f_\theta\sin\theta)\right\}$.

§ 6, стр. 99.

1. xy.

3. а) $\sum\limits_{n=0}^{\infty}\sum\limits_{m=0}^{\infty}\binom{m+n}{n}x^m y^n$ при $|x|+|y|<1$.

б) $\sum\limits_{n=0}^{\infty}\sum\limits_{m=0}^{\infty}\frac{x^m y^n}{m!n!}$ при всех значениях x и y.

§ 7, стр. 114.

1. Выразить по формуле Тэйлора координаты точки кривой через значения функций f, φ, ψ и их первых и вторых производных в t_0; затем написать уравнение плоскости, проходящей через три данные точки. Уравнение соприкасающейся плоскости:

$$\begin{vmatrix} x-f(t_0) & \dot f(t_0) & \ddot f(t_0) \\ y-\varphi(t_0) & \dot\varphi(t_0) & \ddot\varphi(t_0) \\ z-\psi(t_0) & \dot\psi(t_0) & \ddot\psi(t_0) \end{vmatrix}=0,$$

или $(\boldsymbol{r}-\boldsymbol{r}_0)[\dot{\boldsymbol{r}}(t_0)\,\ddot{\boldsymbol{r}}(t_0)]=0$; отсюда видно, что $[\dot{\boldsymbol{r}}(t_0)\,\ddot{\boldsymbol{r}}(t_0)]$ есть нормальный вектор соприкасающейся плоскости в точке t_0 кривой.

3. Пусть $\boldsymbol{R}$ — радиус-вектор центра C искомой сферы; тогда скалярный квадрат $f(s)=\{\boldsymbol{r}(s)-\boldsymbol{R}\}^2$, равный квадрату расстояния от C до переменной точки кривой, должен иметь при $s=s_0$ наибольший возможный порядок стационарности, т. е. в этой точке должны обратиться в нуль $f'(s)$, $f''(s)$ и $f'''(s)$. Дифференцируя $f(s)$ трижды по s и принимая во внимание $(\boldsymbol{r}')^2=1$ и $\boldsymbol{r}'\boldsymbol{r}''=0$, получим три уравнения $(\boldsymbol{r}-\boldsymbol{R})\boldsymbol{r}'=0$, $(\boldsymbol{r}-\boldsymbol{R})\boldsymbol{r}''+1=0$ и $(\boldsymbol{r}-\boldsymbol{R})\boldsymbol{r}'''=0$. Из первого и третьего уравнения видно, что вектор $\boldsymbol{r}-\boldsymbol{R}=\lambda[\boldsymbol{r}'\boldsymbol{r}''']$, где число λ подлежит определению из второго уравнения.

Отв. $\boldsymbol{R}=\boldsymbol{r}-\frac{[\boldsymbol{r}'\boldsymbol{r}''']}{[\boldsymbol{r}'\boldsymbol{r}'']\boldsymbol{r}'''}$.

5. Если $\boldsymbol{R}$ — радиус-вектор центра сферы, то $\{\boldsymbol{r}(s)-\boldsymbol{R}\}^2=1$ при всех значениях s. Дифференцируя три раза последовательно это тождество, получим те же три уравнения, что и в упр. 3, с той лишь разницей, что здесь они удовлетворяются при всех значениях s. Принимая во внимание исходное тождество $(\boldsymbol{r}-\boldsymbol{R})^2=1$ и тождество Лагранжа (упр. 7 на стр. 33), получим требуемый результат.

7. Из определения векторов $\mathfrak{t}$, $\boldsymbol{n}$, $\boldsymbol{b}$ имеем $\mathfrak{t}=\boldsymbol{r}'$, $(\boldsymbol{r}')^2=1$, $\boldsymbol{n}=\frac{1}{k}\boldsymbol{r}''$, $\boldsymbol{b}=[\mathfrak{t}\boldsymbol{n}]$, $\varkappa=\pm|\boldsymbol{b}'|$. Очевидно, $\mathfrak{t}'=k\boldsymbol{n}$, и первая формула доказана. Для определения производных по s от $\boldsymbol{b}$ и $\boldsymbol{n}$ выразим $\boldsymbol{b}'$ и $\boldsymbol{n}'$ через единичные векторы $\mathfrak{t}$, $\boldsymbol{n}$ и $\boldsymbol{b}$. Из соотношений $\boldsymbol{b}^2=1$ и $\mathfrak{t}\boldsymbol{b}=0$ получаем дифференцированием $\boldsymbol{b}\boldsymbol{b}'=0$ и $\boldsymbol{b}'\mathfrak{t}=-\boldsymbol{b}\mathfrak{t}'$, а так как $\mathfrak{t}'=k\boldsymbol{n}$, то $\boldsymbol{b}'\mathfrak{t}=-k\boldsymbol{b}\boldsymbol{n}=0$; стало быть, $\boldsymbol{b}'$ перпендикулярен к обоим векторам $\mathfrak{t}$ и $\boldsymbol{b}$, а потому $\boldsymbol{b}'=\pm|\boldsymbol{b}'|\boldsymbol{n}=\pm\varkappa\boldsymbol{n}$. Оставшийся неопределенным знак кручения в формуле $\varkappa=\pm|\boldsymbol{b}'|$ выберем так, чтобы было $\boldsymbol{b}'=-\varkappa\boldsymbol{n}$, и третья формула доказана. Геометрическое рассмотрение этой формулы показывает, что при таком выборе знака движение точки вдоль кривой по направлению вектора $\mathfrak{t}$, т. е. в сторону возрастания s, в сочетании с вращением бинормального вектора $\boldsymbol{b}$ образует правый винт.

Осталось доказать вторую формулу. Для этой цели продифференцируем по s формулы $\boldsymbol{n}^2=1$, $\mathfrak{t}\boldsymbol{n}=0$ и $\boldsymbol{n}\boldsymbol{b}=0$ и получим $\boldsymbol{n}\boldsymbol{n}'=0$, $\mathfrak{t}\boldsymbol{n}'=-\mathfrak{t}'\boldsymbol{n}=$

$= -kn^2 = -k$ и $n'b = -nb' = \varkappa n^2 = \varkappa$. Следовательно, проекции вектора n' на орты $\mathbf{t}$, n и b равны соответственно $(-k)$, 0, $\varkappa$. Стало быть, $n' = -k\mathbf{t} + \varkappa b$.

Добавим, что, на основании упр. 2, бинормальный вектор перпендикулярен к соприкасающейся плоскости.

8. Воспользоваться упр. 7 и упр. 3.

Отв. а) $r'''(s) = -k^2\mathbf{t} + k'n + k\varkappa b$; б) $R - r = \frac{1}{k}n - \frac{k'b}{k^2\varkappa}$.

9. $|b'| = |\varkappa| = 0$, следовательно, b есть постоянный вектор. Далее, $r'b = \mathbf{t}b = 0$, откуда $rb = \alpha$, где α — постоянное число. Мы получили уравнение плоскости, в которой лежит наша кривая.

10. б) Если кривая задана параметрическими уравнениями $x = f(t)$, $y = \varphi(t)$, $z = \psi(t)$, то поверхность имеет параметрические уравнения (с двумя параметрами s и t):

$$x = f(t) + sf'(t), \qquad y = \varphi(t) + s\varphi'(t), \qquad z = \psi(t) + s\psi'(t);$$

из этих уравнений надо выразить $\frac{\partial^2 z}{\partial x^2}$, $\frac{\partial^2 z}{\partial x\,\partial y}$, $\frac{\partial^2 z}{\partial y^2}$ через производные по t и по s.

Дополнения к гл. II

§ 1, стр. 121.

1. а) Так как область G замкнута, то в ней существует такая точка Q, которая удалена от P на меньшее расстояние, чем любая другая точка области G. Пусть n есть перпендикуляр к PQ в точке Q. Тогда ни одна точка R в G не лежит на той же стороне от n, что и P, ибо, в противном случае, области G принадлежали бы не только точки Q и R, но и весь отрезок QR, и на этом отрезке нашлись бы точки, более близкие к P, чем точка Q. Следовательно, прямая, проходящая через точку P параллельно прямой n, не может иметь общие точки с областью G.

б) Существует последовательность точек $P_1, P_2, \ldots$, не принадлежащих области G и сходящихся к точке P. Обозначим через $l_1, l_2, \ldots$ прямые линии, проходящие через точки $P_1, P_2, \ldots$ и делящие плоскость на две полуплоскости, одна из которых не содержит ни одной точки из G; согласно а) такие прямые существуют. Из этой последовательности можно выделить подпоследовательность прямых, направления которых тоже сходятся. Предельная прямая и будет опорной прямой, проходящей через точку P.

в) Если бы точка M не лежала в G, то, согласно доказательству пункта а), существовала бы опорная прямая n, отделяющая точку M от области G.

г) Примем одну (любую) из опорных прямых за ось x. Тогда ординаты y всех точек области G имеют одинаковый знак. Согласно определению центра массы, его ордината имеет тот же знак. Следовательно, область G и ее центр массы лежат на одной и той же стороне от любой опорной прямой; теперь применить в).

е) Кривизна $k = \frac{d\varphi}{ds}$, где φ есть угол, образуемый касательной с осью x, а s — длина дуги кривой; φ является непрерывной функцией от s. Так как $k = \frac{d\varphi}{ds} > 0$, то угол φ монотонно возрастает от $\varphi(0)$ до $\varphi(0) + 2\pi$, т. е. φ не может иметь одно и то же значение для двух различных точек кривой. Если бы кривая пересекала какую-либо прямую l с уравнением $ax + by - c = 0$ в трех точках s_0, s_1, s_2, то функция

$$F(s) = ax(s) + by(s) - c$$

имела бы три корня; в таком случае $F'(s)$ тоже имела бы по крайней мере три корня, т. е. существовали бы три касательных, параллельных прямой l.

К тому же две из этих касательных непременно были бы направлены в одну сторону, т. е. имели бы одинаковое значение φ, что противоречит сделанному выше выводу.

2. а) Множество, состоящее из точек, лежащих во всех выпуклых областях, содержащих S, обладает свойствами 1, 2 и 3.

б) Если точка P принадлежит множеству E, то не может быть такой прямой l, которая отделяла бы P от S; действительно, если бы такая прямая существовала, мы могли бы взять, например, большой квадрат Q с одной стороной на прямой l, содержащий S; тогда квадрат Q был бы выпуклой областью, содержащей S и не содержащей P.

Если же точка P не принадлежит множеству E, то существует по крайней мере одна выпуклая область B, содержащая S, но не содержащая точки P. Тогда (ср. упр. 1а) существует прямая, отделяющая область B от точки P; а так как B содержит S, то эта прямая отделяет также множество S от P.

в) Ср. упр. 1г.

§ 2, стр. 127.

1. а) Нет. б) Нет. в) Да (см. т. I, стр. 509).

Смешанные упражнения к гл. II, стр. 131.

1. б) Положить $z = \ln u$; для z получится уравнение $z_{xy} = 0$, т. е. z_x не зависит от y. Положим $z_x = \varphi(x)$; тогда $z = \int\limits_a^x \varphi(x)\,dx + \psi(y)$ и $u = e^z =$
$$= e^{\int\limits_a^x \varphi(x)\,dx} \cdot e^{\psi(y)} = f(x)\cdot g(y).$$

3. Дифференцировать соотношение $F(u_x, u_y) = 0$ по x и по y.

4. Функция u должна иметь вид $u = xf\left(\frac{y}{x}\right) + g\left(\frac{y}{x}\right)$ или, что то же, $u = x\varphi\left(\frac{x}{y}\right) + \psi\left(\frac{x}{y}\right)$.

5. а)
$$|f(x+h, y+k) - f(x, y)| =$$
$$= \left|\frac{2hx + 4ky + h^2 + 2k^2}{\sqrt{1+(x+h)^2+2(y+k)^2} + \sqrt{1+x^2+2y^2}}\right| \leqslant$$
$$\leqslant |2hx + 4ky + h^2 + 2k^2| \leqslant 2\,|h^2 + k^2 + 2hx + 2ky| \leqslant$$
$$\leqslant 2(h^2 + k^2 + 2\sqrt{h^2+k^2}\sqrt{x^2+y^2}) \leqslant 2\sqrt{h^2+k^2}\,(1 + 2\sqrt{x^2+y^2}),$$
если предположить, что $h^2 + k^2 < 1$. Теперь видно, что
$$|f(x+h, y+k) - f(x, y)| < \varepsilon, \text{ когда } \sqrt{h^2+k^2} < \frac{\varepsilon}{2 + 4\sqrt{x^2+y^2}}.$$

6. Положим $x = at$, $y = bt$; тогда
$$\lim_{t\to 0} f(x, y) = \lim_{t\to 0}\frac{a^4b^4t^2}{(a^2+b^4t^2)^3} = 0, \qquad \lim_{t\to 0} g(x, y) = \lim\frac{a^2t}{(a^2+b^2)\,t - a} = 0;$$
однако если точка (x, y) стремится к началу, двигаясь вдоль параболы $y^2 = x$, то $f(x, y) = \frac{1}{8}$, а $g(x, y) = 1$.

7. Зададим кривую уравнениями $x = x(t)$, $y = y(t)$, где $x(t)$ и $y(t)$ имеют непрерывные производные. Пусть точки P и Q кривой соответствуют значениям t_1 и t_2 параметра. По теореме о среднем значении интегрального исчисления

$$l = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = (t_2 - t_1) \sqrt{[\dot{x}(\tau_1)]^2 + [\dot{y}(\tau_1)]^2},$$

а по теореме о среднем значении дифференциального исчисления

$$d = \sqrt{[x(t_2) - x(t_1)]^2 + [y(t_2) - y(t_1)]^2} = (t_2 - t_1) \sqrt{[\dot{x}(\tau_2)]^2 + [\dot{y}(\tau_3)]^2},$$

где τ_1, τ_2, τ_3 — некоторые промежуточные значения между t_1 и t_2. Теперь ясно, что $l - d = o(t_2 - t_1)$, так как

$$\sqrt{[\dot{x}(\tau_1)]^2 + [\dot{y}(\tau_1)]^2} - \sqrt{[\dot{x}(\tau_2)]^2 + [\dot{y}(\tau_3)]^2} \to 0, \qquad \text{когда} \qquad t_2 - t_1 \to 0,$$

а d имеет при этом тот же порядок малости, что и $t_2 - t_1$.

8. Так как все члены ряда положительны, то достаточно доказать его сходимость и вычислить его сумму при любом порядке членов. Положим $a + b = n$. Тогда

$$S = \sum_{n=0}^{\infty} \sum_{a=0}^{n} \binom{n}{a} \frac{a}{x^a y^{n-a}} = \sum_{n=0}^{\infty} \frac{1}{y^n} \sum_{a=0}^{n} \binom{n}{a} a \left(\frac{y}{x}\right)^a.$$

Внутреннюю сумму легко вычислить по формуле

$$\sum_{a=0}^{n} \binom{n}{a} a z^a = nz(1+z)^{n-1},$$

которая доказывается дифференцированием тождества

$$\sum_{a=0}^{n} \binom{n}{a} z^a = (1+z)^n.$$

Поэтому

$$S = \sum_{n=0}^{\infty} \frac{1}{y^n} \frac{y}{x} n \left(1 + \frac{y}{x}\right)^{n-1} = \frac{1}{x} \sum_{n=0}^{\infty} n \left(\frac{1}{y} + \frac{1}{x}\right)^{n-1} = \frac{1}{x} \frac{1}{\left(1 - \frac{1}{x} - \frac{1}{y}\right)^2}.$$

10. Будем обозначать точками производные по t, штрихами — по s. Согласно стр. 107, $k = |\boldsymbol{r}''| = \sqrt{(\boldsymbol{r}'')^2}$ (Сокращать показатель корня с показателем степени нельзя, так как под корнем стоит не квадрат числа, а скалярный квадрат вектора.) Выразим $\boldsymbol{r}''$, вторую производную радиус-вектора по s, через его производные по t. Имеем

$$\boldsymbol{r}' = \dot{\boldsymbol{r}} \frac{dt}{ds} = \frac{\dot{\boldsymbol{r}}}{\frac{ds}{dt}};$$

$$\boldsymbol{r}'' = \frac{\frac{ds}{dt} \cdot \ddot{\boldsymbol{r}} \frac{dt}{ds} - \dot{\boldsymbol{r}} \frac{d^2 s}{dt^2} \frac{dt}{ds}}{\left(\frac{ds}{dt}\right)^2} = \frac{\ddot{\boldsymbol{r}} \left(\frac{ds}{dt}\right)^2 - \dot{\boldsymbol{r}} \frac{ds}{dt} \frac{d^2 s}{dt^2}}{\left(\frac{ds}{dt}\right)^4}.$$

Так как $\left(\frac{ds}{dt}\right)^2=\dot{r}^2$ (стр. 105), а отсюда дифференцированием получается $\frac{ds}{dt}\frac{d^2s}{dt^2}=\dot{r}\ddot{r}$, то $r''=\frac{\ddot{r}\dot{r}^2-\dot{r}(\dot{r}\ddot{r})}{(\dot{r}^2)^2}$, а скалярный квадрат этого вектора

$$(r'')^2=\frac{\ddot{r}^2(\dot{r}^2)^2-\dot{r}^2(\dot{r}\ddot{r})^2}{(\dot{r}^2)^4}=\frac{\dot{r}^2\ddot{r}^2-(\dot{r}\ddot{r})^2}{(\dot{r}^2)^3}.$$

Числитель равен $[\dot{r}\ddot{r}]^2=|[\dot{r}\ddot{r}]|^2$, а знаменатель равен $|\dot{r}|^6$.

11. Кривая C лежит на цилиндрической поверхности $x=x(s)$, $y=y(s)$, образующие которой параллельны оси z. Параметр s есть длина дуги проекции кривой C на плоскость xOy, но не является длиной дуги самой кривой C; поэтому будем считать s «произвольным» параметром и производные по s обозначать точками.

По формуле для производной от длины дуги по параметру имеем

$$\dot{x}^2+\dot{y}^2=\left(\frac{ds}{ds}\right)^2=1, \tag{1}$$

откуда дифференцированием получаем

$$\dot{x}\ddot{x}+\dot{y}\ddot{y}=0. \tag{2}$$

Введем обозначение:

$$\dot{x}\ddot{y}-\dot{y}\ddot{x}=\gamma. \tag{3}$$

Решая эту систему уравнений, получим

$$\ddot{x}=-\gamma\dot{y}, \qquad \ddot{y}=\gamma\dot{x}, \tag{4}$$

откуда

$$\ddot{x}^2+\ddot{y}^2=\gamma^2. \tag{5}$$

Обозначим через σ длину дуги кривой C, отсчитываемую от некоторой начальной точки на C, причем пусть σ возрастает с возрастанием s. Тогда, так как $\dot{z}=a$, в силу (1), имеем

$$\frac{d\sigma}{ds}=\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}=\sqrt{1+a^2} \quad \text{и} \quad \frac{ds}{d\sigma}=\frac{1}{\sqrt{1+a^2}}.$$

Единичный касательный вектор

$$\mathbf{t}=\frac{dr}{d\sigma}=\frac{dr}{ds}\frac{ds}{d\sigma}=\{\dot{x},\ \dot{y},\ a\}\frac{1}{\sqrt{1+a^2}}.$$

Далее,

$$\frac{d\mathbf{t}}{d\sigma}=\frac{d^2r}{d\sigma^2}=\frac{d\mathbf{t}}{ds}\frac{ds}{d\sigma}=\{\ddot{x},\ \ddot{y},\ 0\}\frac{1}{1+a^2}.$$

Согласно определению (стр. 107), *кривизна*

$$k=\left|\frac{d^2r}{d\sigma^2}\right|=\left|\frac{d\mathbf{t}}{d\sigma}\right|=\frac{\sqrt{\ddot{x}^2+\ddot{y}^2}}{1+a^2}=\frac{|\gamma|}{1+a^2}=\frac{|\dot{x}\ddot{y}-\dot{y}\ddot{x}|}{1+a^2},$$

см. (5) и (3). Впрочем, кривизну можно вычислить и по формуле упр. 10, но только точки должны теперь означать дифференцирование по s.

По первой формуле Френэ (упр. 7, стр. 115 и его решение стр. 615), единичный главный нормальный вектор

$$\mathbf{n}=\frac{1}{k}\frac{d\mathbf{t}}{d\sigma}=\frac{1}{|\gamma|}\{\ddot{x},\ \ddot{y},\ 0\}=\frac{\gamma}{|\gamma|}\{-\dot{y},\ \dot{x},\ 0\}=\operatorname{sgn}\gamma\,\{-\dot{y},\ \dot{x},\ 0\} \quad [\text{см. (4)}].$$

Единичный бинормальный вектор

$$b = [\boldsymbol{t}n] = \frac{\operatorname{sgn}\gamma}{\sqrt{1+a^2}}\{-a\dot{x},\ -a\dot{y},\ 1\} \qquad [\text{см. } (1)];$$

$$\frac{db}{d\sigma} = \frac{db}{ds}\frac{ds}{d\sigma} = \frac{\operatorname{sgn}\gamma}{1+a^2}\{-a\ddot{x},\ -a\ddot{y},\ 0\}.$$

Из третьей формулы Френэ (упр. 7, стр. 115) скалярным умножением на единичный вектор $(-n)$ находим *кручение*

$$\varkappa = -n\frac{db}{d\sigma} = \frac{1}{1+a^2}(-a\ddot{x}\dot{y} + a\dot{x}\ddot{y}) = \frac{a(\dot{x}\ddot{y} - \dot{y}\ddot{x})}{1+a^2}.$$

За нормальный вектор *соприкасающейся плоскости* можно принять вектор $\{-a\dot{x}, -a\dot{y}, 1\}$, параллельный бинормальному вектору b; поэтому уравнение соприкасающейся плоскости в точке (x, y, z)

$$-a\dot{x}(\xi - x) - a\dot{y}(\eta - y) + (\zeta - z) = 0.$$

Уравнения нормали к цилиндрической поверхности в той же точке:

$$\dot{x}(\xi - x) + \dot{y}(\eta - y) = 0, \qquad \zeta = z.$$

Нетрудно убедиться, что эта нормаль является линией сечения соприкасающейся плоскости с плоскостью $\zeta = z$.

12. Ср. упр. 1, стр. 114. Уравнение соприкасающейся плоскости

$$(\ddot{f}\cos\theta + \dot{f}\sin\theta)\,\xi + (\ddot{f}\sin\theta - \dot{f}\cos\theta)\,\eta + \zeta - (f + \ddot{f}) = 0.$$

Расстояние от начала координат до этой плоскости равно абсолютной величине дроби

$$\frac{f + \ddot{f}}{\sqrt{1 + \dot{f}^2 + \ddot{f}^2}}.$$

В частном случае, когда $f(\theta) = \frac{1}{A}\operatorname{ch} A\theta$, это расстояние оказывается постоянным и равным $\sqrt{1 + \frac{1}{A^2}}$.

13. См. упр. 10.

14. а) Согласно упр. 5, стр. 33, уравнение искомой плоскости есть

$$\begin{vmatrix} \frac{1}{3}at_1^3 - x & \frac{1}{2}bt_1^2 - y & ct_1 - z \\ \frac{1}{3}at_2^3 - x & \frac{1}{2}bt_2^2 - y & ct_2 - z \\ \frac{1}{3}at_3^3 - x & \frac{1}{2}bt_3^2 - y & ct_3 - z \end{vmatrix} = 0.$$

б) Согласно упр. 1, стр. 114, соприкасающаяся плоскость в точке t кривой есть предельное положение плоскости, проходящей через три точки этой кривой, когда все эти точки стремятся к общей предельной точке t; поэтому ее уравнение есть

$$\frac{3x}{a} - \frac{6ty}{b} + \frac{3t^2z}{c} - t^3 = 0 \qquad \text{или} \qquad t^3 - \frac{3t^2z}{c} + \frac{6ty}{b} - \frac{3x}{a} = 0. \qquad (*)$$

Если сюда подставить вместо t последовательно t_1, t_2, t_3, то получим уравнения соприкасающихся плоскостей в каждой из этих точек. В точке пересечения (x, y, z) этих трех плоскостей уравнению (*) должны поэтому удовлетворять значения $t=t_1$, $t=t_2$, $t=t_3$. Это значит, что t_1, t_2, t_3 суть три корня кубического уравнения для t. Следовательно,

$$\frac{3z}{c}=t_1+t_2+t_3, \qquad \frac{6y}{b}=t_1t_2+t_1t_3+t_2t_3, \qquad \frac{3x}{a}=t_1t_2t_3.$$

Но эти три выражения для x, y, z удовлетворяют уравнению плоскости пункта а).

15. Если, сохраняя неизменными b и c, изменять только a, то из формулы $S=\frac{1}{2}bc\sin A$ вытекает $d_aS=\frac{1}{2}bc\cos A\,dA$. Дифференцируя равенство $a^2=b^2+c^2-2bc\cos A$, получим $a\,da=bc\sin A\,dA$, откуда $bc\,dA=\frac{a}{\sin A}da=2R\,da$; следовательно, $d_aS=R\cos A\,da$. Аналогично $d_bS=R\cos B\,db$ и $d_cS=R\cos C\,dc$. Но $dS=d_aS+d_bS+d_cS$.

16. Полярный радиус $r=\sqrt{\mathbf{r}^2}$. Дифференцируя по t, получим

$$\frac{dr}{dt}=\frac{1}{2\sqrt{\mathbf{r}^2}}\cdot 2\mathbf{r}\frac{d\mathbf{r}}{dt}=\frac{\mathbf{r}}{r}\frac{d\mathbf{r}}{dt}=\mathbf{r}^\circ\frac{d\mathbf{r}}{dt}.$$

Геометрическая формулировка: производная от полярного радиуса по скалярному параметру равна проекции производной от радиус-вектора по тому же параметру на направление самого радиус-вектора.

ГЛАВА III

§ 1, стр. 144.

2. а) $-\frac{5}{4}$, б) $-\frac{\pi}{2}$, в) -1; г) -1.

3. а) $-\frac{21}{32}$, б) π, в) 2; г) $-\frac{19}{3}$.

4. При $x=-3$ максимум $y=6$; при $x=3$ минимум $y=-6$. [Максимум и минимум принадлежат различным ветвям двузначной функции $y=f(x)$, но через каждую из этих точек проходит лишь по одной ее ветви.]

5. $z_x=\frac{yz-x^2}{z^2-xy}$, $z_y=\frac{zx-y^2}{z^2-xy}$.

6. $u_x(0, 0)=-1$, $u_y(0, 0)=-1$.

§ 2, стр. 153.

1. а) $5x+7y-21z+9=0$. б) $20x+13y+3z-36=0$.

в) $x-y-z+\frac{\pi}{6}=0$.

2. 1.

4. $\xi(\eta+\zeta)=a\eta$, где ξ, η, ζ — текущие координаты на искомой поверхности.

[**5.** Две точки пересечения $(0, a)$ и $(a, 0)$. В первой точке $F_y=0$, а следовательно, $x'=x_y=0$ (вертикальная касательная), во второй точке $F_x=0$, а значит, $y'=y_x=0$ (горизонтальная касательная). В первой точке $x''=0$, но $x'''\neq 0$; во второй точке $y''=0$, но $y'''\neq 0$.]

6. $a=1$, $b=-\frac{1}{2}$.

8. Пусть окружности K, K_1, K_2 заданы уравнениями

$$K \equiv x^2 + y^2 + ax + by + c = 0,$$
$$K_1 \equiv x^2 + y^2 + a_1x + b_1y + c_1 = 0,$$
$$K_2 \equiv x^2 + y^2 + a_2x + b_2y + c_2 = 0.$$

Тогда $K_1 + \lambda K_2 = 0$ есть уравнение пучка всех окружностей, проходящих через точки A и B. Условия ортогональности окружности K к окружностям K_1 и K_2 приводятся к виду

$$aa_1 + bb_1 - 2(c + c_1) = 0, \qquad aa_2 + bb_2 - 2(c + c_2) = 0.$$

Из этих условий сразу вытекает как их следствие условие ортогональности окружностей $K = 0$ и $K_1 + \lambda K_2 = 0$.

§ 3, стр. 175.

2. в) $\dfrac{\partial(\xi, \eta)}{\partial(x, y)} = \dfrac{-1}{(x^2 + y^2)^2}$.

3. Принять точку O за начало координат и произвести инверсию; тогда криволинейный треугольник преобразуется в прямолинейный с теми же углами.

5. б) Обозначим левую часть уравнения, определяющего t_1 и t_2, через $F(x, y, t)$; тогда кривые $t_1 = \text{const}$ и $t_2 = \text{const}$ задаются неявными уравнениями $F = (x, y, t_1) = 1$ и $F(x, y, t_2) = 1$. Условие их ортогональности запишется так:

$$0 = F_x(x, y, t_1)\, F_x(x, y, t_2) + F_y(x, y, t_1)\, F_y(x, y, t_2) =$$
$$= \frac{4x^2}{(a - t_1)(a - t_2)} + \frac{4y^2}{(b - t_1)(b - t_2)};$$

но это соотношение является прямым следствием тождества

$$F(x, y, t_1) - F(x, y, t_2) = 0.$$

в) $x = \pm\sqrt{\dfrac{(a - t_1)(a - t_2)}{a - b}}, \qquad y = \pm\sqrt{\dfrac{(b - t_1)(b - t_2)}{b - a}}$.

г) $\dfrac{\partial(t_1, t_2)}{\partial(x, y)} = \dfrac{4(a - b)xy}{\sqrt{(a + b)^2 - 2(a - b)(x^2 - y^2) + (x^2 + y^2)^2}}$.

д) $\dfrac{f_1' g_1'}{(a - t_1)(b - t_1)} = \dfrac{f_2' g_2'}{(a - t_2)(b - t_2)}$.

6. а) Обозначим через $F(t)$ левую часть уравнения, определяющего t. Функция $F(t)$ непрерывна в промежутке $-\infty < t < c$, причем $F(-\infty) = 0$, $F(c - 0) = +\infty$; следовательно, $F(t_1) = 1$ по крайней мере в одной точке t_1 этого промежутка. Аналогичные рассуждения применить к остальным двум промежуткам.

б) Ср. упр. 5б).

в) $x = \pm\sqrt{\dfrac{(a - t_1)(a - t_2)(a - t_3)}{(a - b)(a - c)}}$ и аналогичные формулы для y и z. Ср. упр. 5в).

7. б) Пусть $x = r\cos\theta$, $y = r\sin\theta$. Тогда прямая $\theta = \text{const}$ преобразуется в кривую второго порядка $t_1 = \dfrac{1}{2} - \cos^2\theta$, а окружность $r = \text{const}$ — в кривую второго порядка $t_2 = -\dfrac{1}{4}\left(r^2 + \dfrac{1}{r^2}\right)$.

9. Введем функцию $v = x^2 + y^2$. Тогда

$$\frac{\partial(u, v)}{\partial(x, y)} = 2\begin{vmatrix} u_x & u_y \\ x & y \end{vmatrix} = 0.$$

§ 4, стр. 185.

1. б) Окружность на сфере задается присоединением к уравнению сферы линейного уравнения относительно x, y, z.

г) $ds^2 = \dfrac{4r^4(du^2+dv^2)}{(u^2+v^2+r^2)^2}$.

2. а) $ds^2 = a^2(d\theta^2 + \sin^2\theta\, d\varphi^2)$.

б) $ds^2 = (a^2 \sin^2 u + b^2 \cos^2 u) \operatorname{ch}^2 v\, du^2 - \dfrac{a^2-b^2}{2} \sin 2u \operatorname{sh} 2v\, du\, dv +$

$$+ [(a^2\cos^2 u + b^2 \sin^2 u) \operatorname{sh}^2 v + c^2 \operatorname{ch}^2 v]\, dv^2.$$

в) Параметрические уравнения поверхности вращения:

$$x = f(z)\cos\theta, \qquad y = f(z)\sin\theta, \qquad z = z;$$

$$ds^2 = (1 + f'^2)\, dz^2 + f^2\, d\theta^2.$$

г) $ds^2 = \dfrac{(t_1-t_2)(t_1-t_3)}{4(a-t_1)(b-t_1)(c-t_1)}\, dt_1^2 + \dfrac{(t_2-t_1)(t_2-t_3)}{4(a-t_2)(b-t_2)(c-t_2)}\, dt_2^2$.

3. $EG - F^2 = [r_u r_v]^2 = \begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix}^2 + \begin{vmatrix} z_u & x_u \\ z_v & x_v \end{vmatrix}^2 + \begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix}^2$;

воспользоваться формулой преобразования якобианов.

4. Примем точку P за начало прямоугольной системы координат x, y, z. Касательную t примем за ось x, а касательную плоскость к поверхности S в точке P — за плоскость xy. Уравнение поверхности S будет тогда $z = f(x, y)$, где $f(0, 0) = f_x(0, 0) = f_y(0, 0) = 0$. Пучок плоскостей Σ, проходящих через t, выражается уравнением $z = \alpha y$. Введем теперь $\rho = \sqrt{y^2+z^2}$ и x в качестве координат на плоскости Σ. Тогда на этой плоскости $y = \dfrac{\rho}{\sqrt{1+\alpha^2}}$ и $z = \dfrac{\alpha\rho}{\sqrt{1+\alpha^2}}$, и линия пересечения плоскости Σ с поверхностью S будет иметь в координатах ρ и x неявное уравнение

$$\frac{\alpha\rho}{\sqrt{1+\alpha^2}} = f\left(x, \frac{\rho}{\sqrt{1+\alpha^2}}\right).$$

Кривизна этой линии пересечения в точке $x = 0$, $\rho = 0$ имеет поэтому (ср. стр. 147) значение

$$k = \frac{f_{xx}(0, 0)}{\alpha}\sqrt{1+\alpha^2} \qquad \text{(с точностью до знака).}$$

Следовательно, координаты центра кривизны этой линии, соответствующего точке $P(0, 0, 0)$, будут

$$x = 0, \qquad y = \frac{1}{k\sqrt{1+\alpha^2}} = \frac{\alpha}{(1+\alpha^2) f_{xx}(0, 0)},$$

$$z = \frac{\alpha}{k\sqrt{1+\alpha^2}} = \frac{\alpha^2}{(1+\alpha^2) f_{xx}(0, 0)}.$$

Исключив α из этих уравнений, обнаружим, что геометрическим местом центров кривизны является окружность

$$f_{xx}(0, 0)(y^2+z^2) - z = 0, \qquad x = 0.$$

5. Точку P примем за начало координат, а касательную плоскость в этой точке — за плоскость xy. Уравнение поверхности S запишем в виде $z = f(x, y)$, причем $f(0, 0) = f_x(0, 0) = f_y(0, 0) = 0$. Пучок нормальных плоскостей выражается уравнением $y = \alpha x$. В качестве координат в каждой нормальной плоскости примем $\rho = \sqrt{x^2+y^2}$ и z. Тогда в этой плоскости

$x=\frac{\rho}{\sqrt{1+\alpha^2}}$, $y=\frac{\alpha\rho}{\sqrt{1+\alpha^2}}$ и уравнения соответствующего нормального сечения в координатах ρ, z будут

$$z=f\left(\frac{\rho}{\sqrt{1+\alpha^2}},\ \frac{\alpha\rho}{\sqrt{1+\alpha^2}}\right),\qquad z=0,$$

а его кривизна в точке P, где $\rho=0$ и $z=0$, будет

$$k=f_{xx}(0,0)\frac{1}{1+\alpha^2}+2f_{xy}(0,0)\frac{\alpha}{1+\alpha^2}+f_{yy}(0,0)\frac{\alpha^2}{1+\alpha^2}.$$

Конечная точка Q вектора длины $\frac{1}{\sqrt{k}}$, отложенного вдоль касательной t, имеет координаты

$$x=\frac{1}{\sqrt{1+\alpha^2}}\frac{1}{\sqrt{k}},\qquad y=\frac{\alpha}{\sqrt{1+\alpha^2}}\frac{1}{\sqrt{k}},\qquad z=0.$$

Исключив из этих уравнений и из выражения для k параметр α и кривизну k, обнаружим, что точка Q лежит на кривой

$$x^2f_{xx}+2xyf_{xy}+y^2f_{yy}=1.$$

6. а) Дифференцируя оба уравнения по параметру t кривой, получим

$$x\dot{x}+y\dot{y}+z\dot{z}=0,\quad ax\dot{x}+by\dot{y}+cz\dot{z}=0. \tag{1}$$

Решая эту систему однородных уравнений относительно неизвестных $\dot{x}$, $\dot{y}$, $\dot{z}$, находим

$$\frac{\dot{x}}{yz(c-b)}=\frac{\dot{y}}{zx(a-c)}=\frac{\dot{z}}{xy(b-a)}. \tag{2}$$

Следовательно, за касательный вектор в точке (x, y, z) кривой можно принять вектор $\left\{\frac{c-b}{x},\ \frac{a-c}{y},\ \frac{b-a}{z}\right\}$. Обозначив через ξ, η, ζ текущие координаты на касательной, получим ее уравнения в следующем виде:

$$\frac{x(\xi-x)}{c-b}=\frac{y(\eta-y)}{a-c}=\frac{z(\zeta-z)}{b-a}.$$

б) Дифференцируя по t уравнения (1) из а), получим с помощью пропорций (2):

$$x\ddot{x}+y\ddot{y}+z\ddot{z}=-(\dot{x}^2+\dot{y}^2+\dot{z}^2)=\lambda\left\{\frac{(c-b)^2}{x^2}+\frac{(a-c)^2}{y^2}+\frac{(b-a)^2}{z^2}\right\}$$

и

$$ax\ddot{x}+by\ddot{y}+cz\ddot{z}=\lambda\left\{\frac{a(c-b)^2}{x^2}+\frac{b(a-c)^2}{y^2}+\frac{c(b-a)^2}{z^2}\right\},$$

где λ — коэффициент пропорциональности. Исключив λ, найдем

$$(x\ddot{x}+y\ddot{y}+z\ddot{z})\left\{\frac{a(c-b)^2}{x^2}+\frac{b(a-c)^2}{y^2}+\frac{c(b-a)^2}{z^2}\right\}=$$
$$=(ax\ddot{x}+by\ddot{y}+cz\ddot{z})\left\{\frac{(c-b)^2}{x^2}+\frac{(a-c)^2}{y^2}+\frac{(b-a)^2}{z^2}\right\}.$$

Это уравнение, линейное относительно $\ddot{x}$, $\ddot{y}$ и $\ddot{z}$, остается верным, если заменить в нем $\ddot{x}$, $\ddot{y}$, $\ddot{z}$ величинами $\dot{x}$, $\dot{y}$, $\dot{z}$. Следовательно, этому уравнению удовлетворяют и любые линейные комбинации $\alpha\dot{x}+\beta\ddot{x}$, $\alpha\dot{y}+\beta\ddot{y}$, $\alpha\dot{z}+\beta\ddot{z}$, если их подставить вместо $\dot{x}$, $\dot{y}$, $\dot{z}$ соответственно. Вместе с тем, если точка (ξ, η, ζ) лежит в соприкасающейся плоскости, то разности $\xi-x$, $\eta-y$, $\zeta-z$ равны как раз таким линейным комбинациям. Отсюда и вытекает, что уравнение соприкасающейся плоскости имеет следующий вид:

$$\frac{ax^3}{c-b}(\xi-x)+\frac{by^3}{a-c}(\eta-y)+\frac{cz^3}{b-a}(\zeta-z)=0.$$

§ 5, стр. 199.

1. Пусть $P(x, y, z)$ есть точка трубчатой поверхности Σ, а S есть та шаровая поверхность семейства, которая имеет с Σ эту общую точку P. Тогда S и Σ имеют в P общую касательную плоскость, т. е. имеют в этой точке одинаковые значения x, y, z, z_x и z_y. Поэтому достаточно доказать, что искомое соотношение справедливо для любой шаровой поверхности радиуса 1, центр которой лежит в плоскости xOy, т. е. для поверхности

$$z=u(x, y)=\sqrt{1-(x-a)^2-(y-b)^2}.$$

2. а) $\sqrt{x}+\sqrt{y}+\sqrt{z}=1$; б) $x^{2/3}+y^{2/3}+z^{2/3}=1$.

4. Можно ввести t в качестве параметра на кривой и считать ее заданной уравнениями $x=x(t)$, $y=y(t)$, $z=z(t)$; тогда ее касательная в точке со значением параметра t лежит на двух плоскостях, соответствующих тому же значению t. Это дает соотношения:

$$a\dot{x}+b\dot{y}+c\dot{z}=0, \qquad A\dot{x}+B\dot{y}+C\dot{z}=0.$$

Дифференцируя по t уравнения прямых линий семейства, мы, в силу этого, получаем

$$\dot{a}x+\dot{b}y+\dot{c}z=0, \qquad \dot{A}x+\dot{B}y+\dot{C}z=0.$$

Присоединив к этим двум третье уравнение

$$ax+by+cz=Ax+By+Cz,$$

имеем систему трех однородных уравнений для неизвестных x, y, z; отсюда и вытекает, что определитель должен обратиться в нуль.

5. Для огибающей имеем два уравнения:

$$f(x, y, z, t)\equiv x\cos t+y\sin t+z=t,$$

$$f_t(x, y, z, t)\equiv -x\sin t+y\cos t=1.$$

Эти два уравнения определяют семейство прямых с параметром t. Если существует кривая, для которой эти прямые являются касательными, то она должна также удовлетворять уравнению

$$f_{tt}\equiv -x\cos t-y\sin t=0.$$

а) Уравнение огибающей поверхности $\rho\sin\left(z+\sqrt{\rho^2-1}-\theta\right)+1=0$.

б) Уравнения искомой кривой в цилиндрических координатах $z=\theta-\frac{\pi}{2}$, $\rho=1$.

7. Применить инверсию. Так как сферы S_1, S_2, S_3 проходят через начало координат, то они преобразуются в плоскости, и задача приводится к нахождению огибающей семейства шаровых поверхностей, касающихся трех плоскостей, т. е. некоторой конической поверхности, которую мы подвергнем обратному преобразованию инверсии.

Отв. $(x^2+y^2+z^2)^2-2(x^2+y^2+z^2)(x+y+z)-3(x^2+y^2+z^2-2xy-2xz-2yz)=0$.

8. б) $\xi^2(1-a^2)+\eta^2(1-b^2)-2ab\,\xi\eta+2a\xi+2b\eta=1$; в) $a^2\xi^2+b^2\eta^2=1$.

§ 6, стр. 219.

1. $\dfrac{4+\sqrt{5}}{\sqrt{2}}$, $\dfrac{4-\sqrt{5}}{\sqrt{2}}$.

2. $\dfrac{a}{20}$, $\dfrac{a}{10}$, $\dfrac{a}{10}$.

3. Максимумы при $x=0$, $y=\pm 1$; минимум при $x=y=0$.

4. Этот максимум совпадает с максимумом функции $ax^2+2bxy+cy^2$, подчиненным дополнительному условию $ex^2+2fxy+gy^2=1$.

5. Ср. упр. 4. а) $\dfrac{14+2\sqrt{67}}{3}$; б) функция имеет *несобственный максимум* (гл. III, § 6, п° 1), равный 1,95, когда $\dfrac{y}{x}=0{,}64$.

6. Седловины: $y=0$, $x=\dfrac{\pi}{3}, \dfrac{7\pi}{3}, \dfrac{13\pi}{3}, \ldots$

Минимумы: $y=0$, $x=\dfrac{5\pi}{3}, \dfrac{11\pi}{3}, \dfrac{17\pi}{3}, \ldots$

7. Искомый эллипс, очевидно, касается окружности; поэтому после исключения y из обоих уравнений должно получиться уравнение, имеющее двойной корень x. Условие касания будет поэтому $a^2(b^2-1)=b^4$.

Отв. $a=\dfrac{3}{\sqrt{2}}$, $b=\sqrt{\dfrac{3}{2}}$.

8. В качестве переменных ввести угол между сторонами a, b и угол между сторонами c, d.

Отв. Четырехугольник, вписываемый в окружность.

9. Точка $\left(-\dfrac{1}{\sqrt{14}}, -\dfrac{2}{\sqrt{14}}, -\dfrac{3}{\sqrt{14}}\right)$.

10. Ср. аналогичное решение для треугольника в п° 3, пример 4. Точка O с наименьшей суммой расстояний должна существовать. Сперва доказать, что если искомой точкой не является одна из вершин, то ею может быть лишь точка пересечения диагоналей. Воспользоваться тем, что четыре единичных вектора, сумма которых равна нулю, образуют ромб. Затем доказать, что сумма расстояний от вершин четырехугольника меньше для точки пересечения диагоналей, чем для любой из вершин.

11. $A=\dfrac{a^2}{x}$, $B=\dfrac{b^2}{y}$, $C=\dfrac{c^2}{z}$. Из условия $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$ получить уравнение связи для A, B, C.

а) $x=\dfrac{a^{4/3}}{\sqrt{a^{2/3}+b^{2/3}+c^{2/3}}}$ и т. д.; б) $x=\dfrac{a^{3/2}}{\sqrt{a+b+c}}$ и т. д.

12. Вершины искомого параллелепипеда имеют координаты:

$$x=\pm\frac{a}{\sqrt{3}}, \quad y=\pm\frac{b}{\sqrt{3}}, \quad z=\pm\frac{c}{\sqrt{3}}.$$

13. Координаты вершин: $x=\dfrac{\pm a^2}{\sqrt{a^2+b^2}}$, $y=\dfrac{\pm b^2}{\sqrt{a^2+b^2}}$.

14. $x=1$, $y=1$.

15. Самая большая ось равна максимуму величины $\sqrt{x^2+y^2+z^2}$ при дополнительном условии, что точка (x, y, z) лежит на эллипсоиде. Для иско-

мых координат x, y, z и неопределенного множителя λ получаются следующие три уравнения:

$$\frac{x}{\sqrt{x^2+y^2+z^2}} \equiv \frac{x}{l} = \lambda(Ax+Dy+Ez),$$

$$\frac{y}{\sqrt{x^2+y^2+z^2}} = \frac{y}{l} = \lambda(Dx+By+Gz),$$

$$\frac{z}{\sqrt{x^2+y^2+z^2}} = \frac{z}{l} = \lambda(Ex+Gy+Cz).$$

Помножив первое уравнение на x, второе на y, третье на z и сложив полученные уравнения, получим $\lambda = \sqrt{x^2+y^2+z^2} = l$. С другой стороны, эти три уравнения можно рассматривать как систему линейных однородных уравнений с неизвестными x, y, z; следовательно, определитель системы должен равняться нулю.

Дополнения к главе III

§ 1, стр. 226.

Указание. $f(x)+f(y)+f(z) = 3f(a)+[(x-a)+(y-a)+(z-a)] \times \times f'(a)+\frac{1}{2}\rho^2[f''(a)+\varepsilon]$, где $\rho^2 = (x-a)^2+(y-a)^2+(z-a)^2$. С другой стороны, дополнительное условие дает

$$(x-a)+(y-a)+(z-a) = \rho^2\left(-\frac{\varphi''(a)}{2\varphi'(a)}+\varepsilon\right)-$$

$$-\frac{\varphi'(a)}{\varphi(a)}[(x-a)(y-a)+(x-a)(z-a)+(y-a)(z-a)] =$$

$$= \left[-\frac{\varphi''(a)}{2\varphi'(a)}+\frac{\varphi'(a)}{2\varphi(a)}+\varepsilon\right]\rho^2,$$

где $\lim\limits_{x,\, y,\, z \to a} \varepsilon = 0$.

§ 2, стр. 229.

1. Воспользоваться тем фактом, что касательные в начале координат к обеим ветвям даны уравнениями $y=0$ и $ax+by=0$. Кривизна $k = \frac{2c}{a}$,

$$k = \frac{2(a^3g - a^2bf + ab^2e - b^3c)}{a(a^2+b^2)^{3/2}}.$$

2. а) Двойная точка. б) Две ветви, касающиеся друг друга. в) Угловая точка. г) Точка заострения. д) Точка заострения.

3. Уравнение $F=0$ дифференцировать два раза по x и воспользоваться тем, что $F_y=0$; $\operatorname{tg}\varphi = -\frac{2\sqrt{F_{xy}^2 - F_{xx}F_{yy}}}{F_{xx}+F_{yy}}$. а) $\frac{\pi}{2}$; б) $\frac{\pi}{2}$.

Смешанные упражнения к гл. III, стр. 235.

1. Точку O примем за начало координат, и пусть $\frac{x^2}{a^2} \pm \frac{y^2}{b^2} = 1$ есть уравнение заданной кривой второго порядка. Тогда уравнение огибающей будет $(x^2+y^2)^2 = 4(a^2x^2 \pm b^2y^2)$. Интересно заметить, что если задана равнобочная гипербола, то огибающей будет лемниската $(x^2+y^2)^2 = 4a^2(x^2-y^2)$.

3. Пусть точка P описывает подэру Γ' кривой Γ; построить на отрезке OP, как диаметре, окружность в плоскости, перпендикулярной к плоскости кривой Γ. Огибающей будет поверхность, описываемая этой переменной окружностью.

5. Эллипс.

6. Плоскость, касающаяся обеих парабол, имеет уравнение вида

$$-c^2x+cy+cz=1 \qquad \text{или} \qquad -c^2x+cy-cz=1.$$

Соответствующие огибающие будут $(y+z)^2=4x$ и $(y-z)^2=4x$.

7. Доказательство аналогично частному случаю $n=2$. (Дополнения к гл. III, § 1, n° 1). Положительно определенную квадратичную форму $\sum a_{ik}x_ix_k$ можно привести с помощью надлежащего преобразования $x_i = \sum_{k=1}^{n} c_{ik}y_k$ $(i=1, 2, \ldots, n)$, с не равным нулю определителем, к виду

$$\sum a_{ik}x_ix_k = y_1^2+y_2^2+\ldots+y_n^2 > m(x_1^2+x_2^2+\ldots+x_n^2),$$

где m — подходящая положительная постоянная. Для приложений важно помнить, что необходимое и достаточное условие положительной определенности формы $\Phi=\sum a_{ik}x_ix_k$ состоит в том, чтобы ее главные миноры порядков $1, 2, \ldots, n$, отмеченные в напечатанной здесь схеме:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \vdots \\ a_{31} & a_{32} & a_{33} & \vdots \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix},$$

были все положительны. Форма Φ — отрицательно определенная, если $(-\Phi)$ является положительно определенной.

8. Построить кривую $f(x, y)=0$ и исследовать распределение знаков функции f на плоскости.

9. Пусть вершины треугольника $P_i(x_i, y_i)$ и $r_i=PP_i$; тогда

$$d^2f=\sum_1^3 d^2r_i=\sum_{i=1}^{3} r_i^{-3}[(y-y_i)\,dx-(x-x_i)\,dy]^2,$$

а эта квадратичная форма — положительно определенная.

10. В точке P_1. Заметить, что функция $f=r_1+r_2+r_3$ непрерывна на всей плоскости, но не дифференцируема в точках P_1, P_2 и P_3, в которых она имеет *конические точки* (подобно функции $u=\sqrt{(x-x_1)^2+(y-y_1)^2}$, изображаемой круговым конусом). Исследовать производную функции f в точке P_1 по всем направлениям от этой точки.

11. Согласно первому правилу, надо вычислить d^2f по формуле (3) и подставить туда $dx_1, \ldots, dx_m$ и $d^2x_1, \ldots, d^2x_m$, найденные из (1). Учесть, что из формулы (1) вытекает, что

$$d^2\varphi_\mu=\sum \varphi_{\mu x_ix_k}\,dx_idx_k+\varphi_{\mu x_1}\,d^2x_1+\ldots+\varphi_{\mu x_m}\,d^2x_m=0 \quad (\mu=1, \ldots, m);$$

помножим эту формулу на λ_μ, просуммируем по всем значениям $\mu=1, 2, \ldots, m$ и прибавим к (3); тогда получим $d^2f=d^2F=\sum F_{x_ix_k}\,dx_i\,dx_k$, так как $d^2x_1, \ldots, d^2x_m$ выпадают в силу соотношений (2).

12. Для $F=f+\lambda\varphi$ (с точностью до положительного множителя) получим

$$d^2F=\sum_{i,k=1}^{n} dx_i\,dx_k, \qquad \text{причем} \qquad d\varphi=dx_1+\ldots+dx_n=0.$$

После исключения dx_n мы должны показать, что квадратичная форма

$$-d^2F=(dx_1+\ldots+dx_{n-1})^2-\sum_{i,k=1}^{n-1}dx_i\,dx_k=\sum_{i=1}^{n-1}dx_i^2+\sum_{i,k=1}^{n-1}dx_i\,dx_k$$

положительно определенная.

14. Оси координат.

15. Явное уравнение кривой: $y=x^2(1\pm x^{1/2})$. Обе ветви кривой, образующие точку заострения в начале координат, лежат по одну сторону от их общей касательной.

16. а) Положим $\varphi=x^p+y^p+z^p-c^p$, $F=f-\lambda\varphi$. Тогда уравнения для определения стационарных точек будут

$$l=\lambda px^{p-1},\qquad m=\lambda py^{p-1},\qquad n=\lambda pz^{p-1}.\tag{A}$$

Помножим эти уравнения соответственно на x, y, z и сложим результаты

$$lx+my+nz=\lambda pc^p.\tag{B}$$

Выразив x, y, z из (A) и подставив в уравнение связи $\varphi=0$, получим

$$\lambda p=(l^q+m^q+n^q)^{1/q}c^{1-p}.$$

Подстановка этого выражения вместо λp в (B) дает искомое стационарное значение функции f.

б) См. упр. 11. Здесь мы имеем

$$d^2F=-\lambda p(p-1)(x^{p-2}\,dx^2+y^{p-2}\,dy^2+z^{p-2}\,dz^2);$$

так как $\lambda p>0$, то эта квадратичная форма положительно определенная при $p<1$ и отрицательно определенная при $p>1$.

17. Минимум при $x=1$, $y=4$; седловина при $x=-1$, $y=2$.

18. Пусть сторона AB касается кривой в точке P. Пусть $A'B'$ — другая, соседняя, касательная и P' — новая точка касания. Обозначим через $d\varphi$ угол между AB и $A'B'$. Тогда, если пренебречь членами второго порядка, разность площадей треугольников $A'B'C'$ и ABC будет

$$\Delta S=\frac{d\varphi}{2}(AP^2-BP^2).$$

Для треугольника наименьшей площади $\Delta S=0$, откуда $AP=BP$.

19. Сделать преобразование координат

$$X=x\cos\alpha-y\sin\alpha,\qquad Y=x\sin\alpha+y\cos\alpha.$$

20. Обозначим через S кривую $f(x, y)=C$, а через S' — кривую $\varphi(x, y)=C'$. Линии S и S' касаются друг друга в точке (a, b). Как правило, $f(x, y)-C$ положительно по одну сторону от S и отрицательно по другую ее сторону в некоторой окрестности точки (a, b). Аналогичное свойство имеет знак функции $\varphi(x, y)-C'$ относительно кривой S'. Теперь, если, например, $f(a, b)$ является максимумом функции f, то $f(x, y)-C\leqslant 0$ на S', т. е. S' лежит по одну сторону от S; но тогда и S лежит по одну сторону от S'. Стало быть, $\varphi(x, y)-C'$ имеет постоянный знак на S, а так как эта функция равна нулю в точке (a, b), то она имеет в этой точке либо максимум, либо минимум.

21. Уравнение производящей касательной есть

$$x\sin\theta+y\cos\theta=a(\theta\sin\theta+\cos\theta-1).$$

ГЛАВА IV

§ 1, стр. 245.

1. $F(y)=0$ при $y>0$.

2. Воспользоваться соотношением

$$\frac{1}{z}(f_x\cos\varphi+f_y\sin\varphi)=f_{xx}\sin^2\varphi-2f_{xy}\sin\varphi\cos\varphi+f_{yy}\cos^2\varphi+$$
$$+\frac{1}{z}\frac{d}{d\varphi}(f_x\sin\varphi-f_y\cos\varphi).$$

3. Преобразовать u_{xx} по обобщенной формуле интегрирования произведения, беря два члена в проинтегрированной части (специальная предосторожность необходима, если $p<\frac{5}{2}$).

4. Интеграл, выражающий $J_0'(x)$, преобразовать интегрированием по частям.

§ 2, 3, стр. 270.

1. $\pi/24$. 2. 0. 3. 0.

4. $\pi/8$, если область интегрирования ограничена условием $z>0$; в противном случае нуль.

5. 1/50 400. 6. $\pi\left(2-\frac{3}{2}\ln 3\right)$.

7. Ввести сферические координаты и интегрировать в первую очередь по φ и по θ.

Отв. $\pi\left(2+\frac{3}{2}\ln 3\right)$.

8. $8\ln(1+\sqrt{2})$.

9. Разбить промежуток интегрирования на три промежутка $-1\leqslant x\leqslant-\sqrt[4]{h}$, $-\sqrt[4]{h}\leqslant x\leqslant\sqrt[4]{h}$, $\sqrt[4]{h}\leqslant x\leqslant 1$ и найти пределы интегралов по каждому из этих промежутков.

§ 4, стр. 277.

1. $\frac{1}{4}\left(e-\frac{1}{e}\right)$. Воспользоваться преобразованием $x+y=\xi$, $x-y=\eta$.

2. Перейти к полярным координатам.

Отв. а) $\frac{\pi}{4}-\frac{1}{2}$; б) $\frac{\sqrt{3}}{2}\operatorname{arctg}\frac{1}{2}$.

3. $\frac{1}{48}a^2b^2c^2$. Сделать замену переменных

$$x=a\xi,\qquad y=b\eta,\qquad z=c\zeta.$$

7. Ввести поворотом осей новые прямоугольные координаты (ξ', η', ζ'), так что $\xi'=\frac{x\xi+y\eta+z\zeta}{r}$, где $r=\sqrt{x^2+y^2+z^2}$. Тогда $d\xi\,d\eta\,d\zeta=d\xi'\,d\eta'\,d\zeta'$ (ср. упр. 6 на стр. 50), откуда

$$I=\iiint\cos(r\xi')\,d\xi'\,d\eta'\,d\zeta'$$

по объему шара $\xi'^2+\eta'^2+\zeta'^2\leqslant 1$. Выполнив интегрирование по η' и по ζ', получим

$$I=\pi\int_{-1}^{+1}\cos(r\xi')(1-\xi'^2)\,d\xi'.$$

Отв. $\frac{4\pi}{r^2}\left(\frac{\sin r}{r}-\cos r\right)$.

8. Ввести новые переменные $\xi'=\frac{x\xi+y\eta}{r}$, $\eta'=\frac{-y\xi+x\eta}{r}$ и выполнить интегрирование по η'.

§ 6, стр. 296.

1. $2\pi^2 ab$. Воспользоваться правилом Гульдина и принять во внимание, что центр тяжести эллипса находится в его центре.

2. $\frac{\pi}{2}abh^2$.

3. $\frac{1}{3}\pi abc\left(1-\frac{p}{\sqrt{(a^2l^2+b^2m^2+c^2n^2)}}\right)^2\left(2+\frac{p}{\sqrt{(a^2l^2+b^2m^2+c^2n^2)}}\right)$.

Сделать преобразование $x=a\xi$, $y=b\eta$, $z=c\zeta$.

4. а) Сравнить соответствующие элементы площади.

б) $a^2\int\limits_0^{2\pi}\{1-\cos f(\varphi)\}\,d\varphi$. в) $2\pi\left(1-\frac{1}{2}\sqrt{2}\right)a^2$.

5. $2\pi a^2\left[1+(1-\varepsilon^2)\frac{\operatorname{arth}\varepsilon}{\varepsilon}\right]$, где $2a$ — длина большой оси.

6. Объем $=\frac{1}{3}\pi cp^2$, площадь поверхности $=\pi(a+b)p$, где a, b, c — длины сторон треугольника, а p — длина перпендикуляра, опущенного из вершины C на сторону AB.

7. Из дифференциального уравнения трубчатой поверхности $z=u(x, y)$ (упр. 1, стр. 199) находим

$$\sigma=2\iint\sqrt{1+u_x^2+u_y^2}\,dx\,dy=2\iint\frac{dx\,dy}{u}.$$

Если ввести в качестве параметров длину дуги s кривой L и расстояние t по нормали к L (ср. т. I, стр. 334, упр. 21 и т. II, стр. 199, упр. 3), то

$$\sigma=2\int\limits_L ds\int\limits_{-1}^{+1}\frac{1+kt}{\sqrt{1-t^2}}\,dt=2\pi\int\limits_L ds,$$

где k обозначает кривизну кривой L [она зависит от s, но не зависит от t].

8. а) $\frac{16}{9}a^3$; б) $8a^2$.

9. $\frac{\pi}{2}\left\{R\sqrt{R^2+h^2}-r\sqrt{r^2+h^2}+h^2\ln\frac{R+\sqrt{R^2+h^2}}{r+\sqrt{r^2+h^2}}\right\}$.

10. $\frac{\pi^2}{2}$. Ввести полярные координаты.

§ 7, стр. 308.

1.

Объем	Центр массы	Моменты инерции относительно		
		оси x	оси y	оси z
а) abc	$\left(\frac{a}{2}, \frac{b}{2}, \frac{c}{2}\right)$	$\frac{1}{3}abc\,(b^2+c^2)$	$\frac{1}{3}abc\,(c^2+a^2)$	$\frac{1}{3}abc\,(a^2+b^2)$
б) $\frac{2}{3}\pi a^3$	$\left(0, 0, \frac{3a}{8}\right)$	$\frac{4\pi a^5}{15}$	$\frac{4\pi a^5}{15}$	$\frac{4\pi a^5}{15}$
в) $\frac{1}{6}abc$	$\left(\frac{a}{4}, \frac{b}{4}, \frac{c}{4}\right)$	$\frac{abc}{60}(b^2+c^2)$	$\frac{abc}{60}(c^2+a^2)$	$\frac{abc}{60}(a^2+b^2)$

2. На оси конуса, отсекает две трети отрезка от вершины до центра основания.

3. $x=\frac{2}{3}a$, $y=z=0$. **4.** Ср. упр. 7, стр. 297 и упр. 1, стр. 199.

5. а) $\pi h(R^4-R'^4)$; б) $2\pi h(R^2-R'^2)\left\{\frac{1}{4}(R^2+R'^2)+\frac{1}{3}h^2\right\}$.

6. Например, $A+B-C=2\iiint \mu z^2\,dx\,dy\,dz$, а это величина положительная.

7. а) $\frac{4}{15}\pi abc\,(a^2+b^2)$; б) $\frac{4}{15}\pi abc\,\{(1-\alpha^2)\,a^2+(1-\beta^2)\,b^2+(1-\gamma^2)\,c^2\}$. Сделать замену переменных: $x=a\xi$, $y=b\eta$, $z=c\zeta$ и воспользоваться данными в тексте выражениями для моментов инерции, а также свойствами симметрии эллипсоида.

8. Расстояние от точки (x, y, z) до плоскости $ux+vy+wz+1=0$ равно абсолютной величине выражения $\frac{ux+vy+wz+1}{\sqrt{u^2+v^2+w^2}}$. Поэтому момент инерции эллипсоида относительно этой плоскости есть

$$I=\frac{Au^2+Bv^2+Cw^2+V}{u^2+v^2+w^2},$$

где A, B, C — моменты инерции эллипсоида относительно плоскостей координат, а V — объем эллипсоида, т. е. $A=\frac{4}{15}a^3bc$, $B=\frac{4}{15}ab^3c$, $C=\frac{4}{15}abc^3$ и $V=\frac{4\pi}{3}abc$. Требуется, стало быть, найти огибающую плоскостей, для которых $I=h$. Эта огибающая определяется уравнениями

$$(A-h)\,u=\lambda x,\qquad (B-h)\,v=\lambda y,\qquad (C-h)\,w=\lambda z, \tag{1}$$

где λ — общий множитель, который нетрудно определить из уравнения $I=h$ и уравнения плоскости, причем оказывается, что $\lambda=V$. Исключив параметры u, v и w из уравнений (1) и уравнения $I=h$, получим уравнение огибающей поверхности

$$\frac{x^2}{h-A}+\frac{y^2}{h-B}+\frac{z^2}{h-C}=\frac{1}{V}.$$

10. $a^2(x-\xi)^2+b^2(y-\eta)^2+c^2(z-\zeta)^2=\{a^2+b^2+c^2+5(\xi^2+\eta^2+\zeta^2)\}\times$ $\times\{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2\}$.

11. $\left(\frac{1}{3}, 0, 0\right)$.

12. $\xi=\frac{5a}{16}\frac{2a^2+b^2+c^2}{a^2+b^2+c^2}$.

14. $\frac{2\pi a^2 b}{\sqrt{b^2-a^2}}\ln\left(\frac{b}{a}+\sqrt{\frac{b^2}{a^2}-1}\right)$.

15. Интегрировать сначала по x и y.

Отв. $2\pi\int\limits_a^b\sqrt{z^2+\{f(z)\}^2}\,dz-\pi\,|\,b^2\mp a^2\,|$,

где надо взять верхний или нижний знак, смотря по тому, находится ли начало координат внутри тела или вне его.

Дополнения к главе IV

§ 3, стр. 329.

1. Сделать преобразование $x_1=a_1\xi_1, \ldots, x_n=a_n\xi_n$.

Отв. $\frac{(\sqrt{\pi})^n}{\Gamma\left(\frac{n+2}{2}\right)}a_1a_2\ldots a_n$.

2. Согласно определению интеграла по поверхности (стр. 323),

$$I=\int\ldots\int\frac{f(x_1)+f(-x_1)}{\sqrt{1-x_2^2-\ldots-x_n^2}}\,dx_2\ldots dx_n,$$

где интеграл берется по внутренней области $(n-1)$-мерного единичного шара в пространстве переменных $x_2\ldots x_n$. Введя полярные координаты, получим

$$I=\int\limits_0^1 dr\int\limits_{S(r)}\frac{f(\sqrt{1-r^2})+f(-\sqrt{1-r^2})}{\sqrt{1-r^2}}\,d\sigma,$$

где $S(r)$ обозначает поверхность сферы радиуса r с центром O в пространстве $x_2\ldots x_n$. Так как подынтегральная функция зависит только от r, то

$$I=\omega_{n-1}\int\limits_0^1\frac{f(\sqrt{1-r^2})+f(-\sqrt{1-r^2})}{\sqrt{1-r^2}}\,r^{n-2}\,dr.$$

Полагая $\sqrt{1-r^2}=t$, получим

$$I=\omega_{n-1}\int\limits_{-1}^1 f(t)\,(1-t^2)^{\frac{n-3}{2}}\,dt.$$

§ 4, стр. 340.

1. Способ 1. Сделать подстановку $x^2=t$, а затем применить обобщенное правило интегрирования произведения (т. I, стр. 256).

Способ 2. То же преобразование дает

$$\int_0^\infty x^n e^{-x^2}\,dx = \frac{1}{2}\,\Gamma\left(\frac{n+1}{2}\right) \qquad \text{(ср. стр. 325).}$$

Способ 3. Положить $I_n(a) = \int_0^\infty x^n e^{-ax^2}dx$; тогда $I_n(a) = -I'_{n-2}(a) = = I''_{n-4}(a) = \ldots$, где штрихи означают производные по a.

Отв. $\frac{1}{2}\left(\frac{n-1}{2}\right)!$ при нечетном n; $\sqrt{\pi}\,\frac{1\cdot 3\ldots(n-1)}{2^{\frac{n+2}{2}}}$ при четном n.

2. Ввести новые переменные $\xi = \alpha x + \beta y$, $\eta = \gamma x + \delta y$, выбрав α, β, γ, δ таким образом, чтобы было

$$\xi^2 + \eta^2 = ax^2 + 2bxy + cy^2.$$

Тогда $(\alpha\delta - \beta\gamma)^2 = ac - b^2$, и интеграл преобразуется к виду

$$\frac{1}{\sqrt{ac-b^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(\xi^2+\eta^2)}\,d\xi\,d\eta.$$

Отв. $ac - b^2 = \pi^2$, $a > 0$.

3. Сделать то же самое преобразование, что и в упр. 2, и вычислить полученные интегралы: а) используя результат упр. 2, б) введением полярных координат.

Отв. а) $\frac{\pi(aC + cA - 2bB)}{(ac-b^2)^{3/2}}$; б) $\frac{2\pi}{\sqrt{ac-b^2}}$.

4. а) Найти $K'(a)$, т. е. производную по a, а затем интегрировать по частям два раза (в качестве одного множителя взять xe^{-ax^2}). Получится дифференциальное уравнение

$$K'(a) = -\frac{K(a)}{2a} + \frac{K(a)}{4a^2},$$

откуда

$$K(a) = \frac{C}{\sqrt{a}}\,e^{-\frac{1}{4a}},$$

причем $C = \lim\limits_{a\to\infty}\sqrt{a}\,K(a) = \lim\limits_{a\to\infty}\int_0^\infty e^{-t^2}\cos\frac{t}{\sqrt{a}}\,dt = \frac{1}{2}\sqrt{\pi}.$

Отв. $K(a) = \frac{1}{2}\sqrt{\frac{\pi}{a}}\,e^{-\frac{1}{4a}}$.

б) Формулу $\frac{t}{1+t^2} = \int_0^\infty e^{-tx}\cos x\,dx\ (t>0)$ интегрировать по t от a до b.

Отв. $\frac{1}{2}\ln\frac{1+a^2}{1+b^2}$.

в) Интеграл для $I'(a)$ преобразовать подстановкой $x=\frac{a}{t}$, $a>0$. Получим дифференциальное уравнение $I'=-2I$, откуда $I=Ce^{-2a}$, где $C=\lim\limits_{a\to 0} I(a)=\frac{\sqrt{\pi}}{2}$.

Отв. $I(a)=\frac{1}{2}\sqrt{\pi}e^{-2a}$.

г) Подставить интегральное выражение для J_0 и переставить порядок интегрирования. Воспользоваться формулой $2\sin ax\cos b\,tx=\sin(a+bt)x+\sin(a-bt)x$.

Отв. $\frac{\pi}{2}$ при $a>b$, $\arcsin\frac{a}{b}$ при $a<b$; ср. выражение для $\int\limits_0^\infty \frac{\sin xy}{y}\,dy$ в гл. IV, § 4, n° 1.

6. Существует такое число $\varepsilon>0$, что при всяком A найдется некоторое число $A'>A$, для которого

$$\left|\int\limits_{A'}^{\infty} f(x, y)\,dy\right|\geqslant\varepsilon$$

для какого-либо значения x.

§ 6, стр. 361.

1. Сделать замену переменных $x^m=a^m\xi$, $y^m=b^m\eta$.

3. Начать с интегрирования по y и z.

Отв. $\xi=\frac{3}{4}a\frac{\Gamma(2n)\Gamma(3n)}{\Gamma(n)\Gamma(4n)}$.

4. $2R^4\mathrm{B}\left(\frac{3}{2}, \frac{11}{2}\right)=\frac{21\pi}{2^9}R^4$.

5. Показать, что $G_{2n}(2x)=\frac{1}{2}2^{2x}G_n(x)\,G_n\left(x+\frac{1}{2}\right)\frac{(2n)!\sqrt{n}}{2^{2n}(n!)^2}$ (определение функции $G_n(x)$ см. на стр. 353), а затем устремить $n\to\infty$ и применить формулу Валлиса (т. I, стр. 263—265).

Смешанные упражнения к гл. IV, стр. 366.

1. $f'(x)=\frac{2}{x}\int\limits_0^{\frac{\pi}{2}}\left(1-\frac{1}{1-x^2\cos^2\theta}\right)d\theta$, а так как

$$\int\frac{1}{1-x^2\cos^2\theta}\,d\theta=\frac{1}{\sqrt{1-x^2}}\operatorname{arctg}\left(\frac{\operatorname{tg}\theta}{\sqrt{1-x^2}}\right),$$

то

$$f'(x)=\frac{\pi}{x}\left(1-\frac{1}{\sqrt{1-x^2}}\right),$$

откуда $f(x)=\pi\ln\left(1+\sqrt{1-x^2}\right)-\pi\ln 2$.

2. Согласно стр. 295,

$$\Sigma=\iint\sqrt{EG-F^2}\,dr\,d\theta=\int\limits_{\theta_1}^{\theta_2}d\theta\int\limits_0^{f'(\theta)}\sqrt{r^2+f'^2}\,dr=$$

$$=[\sqrt{2}+\ln(1+\sqrt{2})]\int\limits_{\theta_1}^{\theta_2}\frac{1}{2}f'^2\,d\theta,$$

а это значит, что Σ равна произведению площади проекции, т. е. фигуры

$$\theta_1\leqslant\theta\leqslant\theta_2,\qquad 0\leqslant r\leqslant f'(\theta),$$

на число $[\sqrt{2}+\ln(1+\sqrt{2})]$.

3. Так как $A-BR^2=2{,}5$ и $A-\frac{3}{5}BR^2=5{,}5$, то $A=10$ и $B=\frac{15}{2R^2}$. Напряженность силы тяжести во внутренней точке на расстоянии r от центра равна силе притяжения массы концентрического шара радиуса r, сосредоточенной в его центре.

4. С помощью параллельного переноса можно добиться того, что треугольник будет лежать в верхней полуплоскости. Тогда его момент инерции будет

$$\varphi(x_1y_1,\ x_2y_2)+\varphi(x_2y_2,\ x_3y_3)+\varphi(x_3y_3,\ x_1y_1),$$

где $\varphi(x_1y_1,\ x_2y_2)$ обозначает момент инерции четырехугольника с вершинами $(x_1,\ 0)$, $(x_1,\ y_1)$, $(x_2,\ y_2)$ и $(x_2,\ 0)$, помноженный на $\operatorname{sgn}(x_1-x_2)$. Затем надо показать, что

$$\varphi(x_1y_1,\ x_2y_2)=\frac{1}{12}(x_1-x_2)(y_1^3+y_1^2y_2+y_1y_2^2+y_2^3).$$

6. $2-\frac{\pi}{2}$.

7. Ввести полярные координаты с полюсом (началом) в полюсе сфероида.

8. $I=\int\limits_1^2(y-4)\,dy\int\limits_{\frac{8y-20}{y-4}}^{4/y}dx=12-16\ln 2.$

9. а) $K=\int\limits_0^\beta d\theta\int\limits_0^a r\ln(r^2)\,dr=a^2\beta\left(\ln a-\frac{1}{2}\right).$

б) $K=\int\limits_0^a dx\int\limits_0^{\varphi(x)}\ln(x^2+y^2)\,dy$, где $\varphi(x)=x\operatorname{tg}\beta$ при $0\leqslant x\leqslant a\cos\beta$ и $\varphi(x)=\sqrt{a^2-x^2}$ при $a\cos\beta\leqslant x\leqslant a$.

10. $V=\frac{7}{72}\pi h^3\operatorname{tg}^2\alpha$. Пользуясь цилиндрическими координатами $(r,\ \theta,\ z)$, имеем

$$V=\iint z\,dS=\iint\left(h-\frac{r}{\operatorname{tg}\alpha}\right)dS$$

по области интегрирования

$$h\operatorname{tg}\alpha\,(1-\sqrt[3]{\sin^4\theta\cos^2\theta})\leqslant r\leqslant h\operatorname{tg}\alpha.$$

Стало быть,

$$V=\int_0^{2\pi} d\theta \int\limits_{h\operatorname{tg}\alpha\left(1-\sqrt[3]{\sin^4\theta\cos^2\theta}\right)}^{h\operatorname{tg}\alpha} \left(h-\frac{r}{\operatorname{tg}\alpha}\right) r\,dr=$$

$$=h^3\operatorname{tg}^2\alpha\int_0^{2\pi}\frac{1}{2}\sqrt[3]{\sin^8\theta\cos^4\theta}\,d\theta-h^3\operatorname{tg}^2\alpha\int_0^{2\pi}\frac{1}{3}\sin^4\theta\cos^2\theta\,d\theta=$$

$$=h^3\operatorname{tg}^2\alpha\cdot 2\int_0^{\pi/2}\sqrt[3]{\sin^8\theta\cos^4\theta}\,d\theta-\frac{2h^3\operatorname{tg}^2\alpha}{3}\cdot 2\int_0^{\pi/2}\sin^4\theta\cos^2\theta\,d\theta=$$

$$=h^3\operatorname{tg}^2\alpha\cdot \mathrm{B}\left(\frac{11}{6},\frac{7}{6}\right)-\frac{2}{3}h^3\operatorname{tg}^2\alpha\cdot\mathrm{B}\left(\frac{5}{2},\frac{3}{2}\right).$$

В последней строке мы воспользовались формулой (4), стр. 359.

Далее,

$$\mathrm{B}\left(\frac{11}{6},\frac{7}{6}\right)=\frac{\Gamma\left(\frac{11}{6}\right)\Gamma\left(\frac{7}{6}\right)}{\Gamma(3)}=\frac{5}{72}\Gamma\left(\frac{1}{6}\right)\Gamma\left(\frac{5}{6}\right)=$$

$$=\frac{5}{72}\Gamma\left(\frac{1}{6}\right)\Gamma\left(1-\frac{1}{6}\right)=\frac{5}{72}\frac{\pi}{\sin\frac{\pi}{6}}=\frac{5}{36}\pi.$$

Аналогично вычисляется $\mathrm{B}\left(\frac{5}{2},\frac{3}{2}\right)$.

11. Прямолинейными образующими нашего гиперболического параболоида являются прямые, по которым он пересекается плоскостями $x=\text{const}$ или $y=\text{const}$. Поэтому, так как $d\sigma=\sqrt{1+z_x^2+z_y^2}\,dx\,dy$, то

$$I=-\int_0^{\eta}dy\int_0^{\xi}\frac{1}{(1+x^2+y^2)^{3/2}}\,dx=-\int_0^{\eta}\frac{\xi}{(1+y^2)(1+\xi^2+y^2)^{1/2}}\,dy=$$

$$=-\operatorname{arctg}\frac{\xi\eta}{(1+\xi^2+\eta^2)^{1/2}}.$$

12. $\displaystyle\frac{d}{da}K(a)=\int_0^{\pi}\frac{d}{da}\left(\frac{\ln(1+a\cos x)}{\cos x}\right)dx=\int_0^{\pi}\frac{dx}{1+a\cos x}=\frac{\pi}{\sqrt{1-a^2}}$;

следовательно, $K(a)=\pi\arcsin a+C$; постоянная C определяется из условия $K(0)=0$.

13. Ввести новые переменные u и v с помощью преобразования $u=\frac{x^3}{y}$, $v=\frac{y^2}{x}$. Искомая площадь

$$S=\int_{a^2}^{b^2}\int_{\alpha}^{\beta}\frac{\partial(x,y)}{\partial(u,v)}\,du\,dv=\frac{1}{5}\int_{a^2}^{b^2}u^{-\frac{2}{5}}\,du\int_{\alpha}^{\beta}v^{-\frac{1}{5}}\,dv.$$

ГЛАВА V

§ 1, стр. 384.

1. $e^{\xi}\sin\eta$.

2. Пусть $F_1=\dfrac{e^x}{x^2+y^2}(x\sin y-y\cos y)$, $F_2=\dfrac{e^x}{x^2+y^2}(x\cos y+y\sin y)$. Эти функции имеют неустранимый разрыв в начале координат, во всех же остальных точках плоскости x, y они непрерывны и имеют непрерывные частные производные, причем $\dfrac{\partial F_1}{\partial y}=\dfrac{\partial F_2}{\partial x}$. Поэтому можно лишь утверждать, что искомый интеграл $\oint(F_1\,dx+F_2\,dy)$ имеет одинаковое значение вдоль любой замкнутой кривой C, обходящей один раз начало координат (см. стр. 383, конец п° 8).

Положим $F_1=u-\dfrac{y}{x^2+y^2}$, $F_2=v+\dfrac{x}{x^2+y^2}$.

Функции u и v можно доопределить в начале координат таким образом, что они станут непрерывными и будут иметь непрерывные частные производные во всей плоскости x, y (включая начало). Тогда криволинейный интеграл $\oint\limits_C(u\,dx+v\,dy)=0$ вдоль любого замкнутого пути C. Искомый интеграл

$$\oint\limits_C(F_1\,dx+F_2\,dy)=\oint\limits_C(u\,dx+v\,dy)+\oint\limits_C\left(-\frac{y}{x^2+y^2}\,dx+\frac{x}{x^2+y^2}\,dy\right)=$$
$$=\oint\limits_K\left(-\frac{y}{x^2+y^2}\,dx+\frac{x}{x^2+y^2}\,dy\right)$$

(где K есть окружность $x=\cos t$, $y=\sin t$), а этот последний интеграл равен 2π, согласно стр. 383, п° 8.

§ 5, стр. 416.

1. а) Ср. упр. 9, стр. 33. в) Пусть G — произвольная область, а v — произвольная функция, обращающаяся в нуль на границе области G. Тогда по первой формуле Грина

$$\iiint\limits_G(u_{x_1}v_{x_1}+u_{x_2}v_{x_2}+u_{x_3}v_{x_3})\,dx_1\,dx_2\,dx_3=$$
$$=-\iiint\limits_G v\nabla^2u\,dx_1\,dx_2\,dx_3=-\iiint\limits_G v\nabla^2u\sqrt{e_1e_2e_3}\,dp_1\,dp_2\,dp_3.$$

Но

$$u_{x_i}=u_{p_1}\frac{\partial p_1}{\partial x_i}+u_{p_2}\frac{\partial p_2}{\partial x_i}+u_{p_3}\frac{\partial p_3}{\partial x_i}=u_{p_1}\frac{a_{i1}}{e_1}+u_{p_2}\frac{a_{i2}}{e_2}+u_{p_3}\frac{a_{i3}}{e_3}$$

и

$$v_{x_i}=v_{p_1}\frac{a_{i1}}{e_1}+v_{p_2}\frac{a_{i2}}{e_2}+v_{p_3}\frac{a_{i3}}{e_3}.$$

Следовательно

$$\iiint\limits_G (u_{x_1}v_{x_1}+u_{x_2}v_{x_2}+u_{x_3}v_{x_3})\,dx_1\,dx_2\,dx_3=$$

$$=\iiint\limits_G \left(\frac{1}{e_1}u_{p_1}v_{p_1}+\frac{1}{e_2}u_{p_2}v_{p_2}+\frac{1}{e_3}u_{p_3}v_{p_3}\right)dx_1\,dx_2\,dx_3=$$

$$=\iiint\left(\sqrt{\frac{e_2e_3}{e_1}}u_{p_1}v_{p_1}+\sqrt{\frac{e_3e_1}{e_2}}u_{p_2}v_{p_2}+\sqrt{\frac{e_1e_2}{e_3}}u_{p_3}v_{p_3}\right)dp_1\,dp_2\,dp_3=$$

$$=\iiint (U_1v_{p_1}+U_2v_{p_2}+U_3v_{p_3})\,dp_1\,dp_2\,dp_3,$$

где $U_i=\frac{\sqrt{e_1e_2e_3}}{e_i}u_{p_i}$. Применяя теорему Гаусса к вектору $\boldsymbol{F}=\{U_1v,\ U_2v,\ U_3v\}$, получим $-\iiint\left(\frac{\partial U_1}{\partial p_1}+\frac{\partial U_2}{\partial p_2}+\frac{\partial U_3}{\partial p_3}\right)v\,dp_1\,dp_2\,dp_3$. Таким образом, имеем для произвольной функции v, обращающейся в нуль на границе области G:

$$\iiint v\nabla^2u\sqrt{e_1e_2e_3}\,dp_1\,dp_2\,dp_3=\iiint v\left(\frac{\partial U_1}{\partial p_1}+\frac{\partial U_2}{\partial p_2}+\frac{\partial U_3}{\partial p_3}\right)dp_1\,dp_2\,dp_3,$$

откуда (ср. лемму I, стр. 521)

$$\nabla^2u=\left(\frac{\partial U_1}{\partial p_1}+\frac{\partial U_2}{\partial p_2}+\frac{\partial U_3}{\partial p_3}\right)\frac{1}{\sqrt{e_1e_2e_3}}=$$

$$=\frac{1}{\sqrt{e_1e_2e_3}}\left[\frac{\partial}{\partial p_1}\left(\sqrt{\frac{e_2e_3}{e_1}}\frac{\partial u}{\partial p_1}\right)+\frac{\partial}{\partial p_2}\left(\sqrt{\frac{e_3e_1}{e_2}}\frac{\partial u}{\partial p_2}\right)+\frac{\partial}{\partial p_3}\left(\sqrt{\frac{e_1e_2}{e_3}}\frac{\partial u}{\partial p_3}\right)\right].$$

г) Воспользоваться упр. 6в), стр. 176;

$$\frac{1}{4}(t_2-t_1)(t_3-t_1)(t_3-t_2)\nabla^2u=(t_3-t_2)\sqrt{\varphi(t_1)}\frac{\partial}{\partial t_1}\left(\sqrt{\varphi(t_1)}\frac{\partial u}{\partial t_1}\right)+$$

$$+(t_3-t_1)\sqrt{-\varphi(t_2)}\frac{\partial}{\partial t_2}\left(\sqrt{-\varphi(t_2)}\frac{\partial u}{\partial t_2}\right)+(t_2-t_1)\sqrt{\varphi(t_3)}\frac{\partial}{\partial t_3}\left(\sqrt{\varphi(t_3)}\frac{\partial u}{\partial t_3}\right),$$

где $\varphi(x)=(a-x)(b-x)(c-x)$.

§ 7, стр. 424.

1. $\iint\frac{z}{p}\,dS=\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{2}{c^2}\right)\iiint z\,dx\,dy\,dz$, где тройной интеграл берется по объемной области, ограниченной верхней половиной эллипсоида и плоскостью xy. Поверхностный интеграл слева берется по всей замкнутой поверхности, ограничивающей упомянутую объемную область, но интеграл по плоскому основанию полуэллипсоида равен нулю.

Отв. $\frac{\pi}{4}\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{2}{c^2}\right)abc^2$. Наш интеграл есть $\iint \boldsymbol{F}n^\circ\,dS$, где $\boldsymbol{F}=\left\{\frac{xz}{a^2},\ \frac{yz}{b^2},\ \frac{z^2}{c^2}\right\}$.

2. Так как H есть однородная функция четвертой степени, то

$$4\iint H\,dS=\iint(xH_x+yH_y+zH_z)\,dS=\iint\frac{\partial H}{\partial n}\,dS=$$

$$=\iiint\nabla^2H\,dx\,dy\,dz=6\iiint[x^2(2a_1+a_4+a_6)+y^2(2a_2+a_4+a_5)+$$

$$+z^2(2a_3+a_5+a_6)]\,dx\,dy\,dz.$$

Отв. $\frac{4\pi}{5}(a_1+a_2+a_3+a_4+a_5+a_6)$.

Дополнения к гл. V

§ 2, стр. 429.

1. Систему друх уравнений $u=f_x$, $v=f_y$ можно разрешить относительно x и y, ибо $\frac{\partial(u, v)}{\partial(x, y)} \neq 0$. Пусть $x=\sigma(u, v)$, $y=\tau(u, v)$. Так как $u_y=v_x$, то (ср. стр. 163) имеем $x_v=y_u$, $\sigma_v=\tau_u$. Следовательно, существует такая функция $g(u, v)$, что $x=g_u(u, v)$, $y=g_v(u, v)$.

2. $$B=\left\{\frac{yz}{(x^2+y^2)\sqrt{x^2+y^2+z^2}},\ -\frac{xz}{(x^2+y^2)\sqrt{x^2+y^2+z^2}},\ 0\right\}.$$

Смешанные упражнения к гл. V, стр. 430.

2. Пусть (ξ, η) и (x, y) — прямоугольные координаты на Π и на Q соответственно. Тогда движение точки $P(x, y)$ может быть описано уравнениями $\xi=x\cos\varphi-y\sin\varphi+a$, $\eta=x\sin\varphi+y\cos\varphi+b$ (т. е. сочетание вращения с поступательным движением). Тогда

$$S(P)=A(x^2+y^2)+Bx+Cy+D.$$

а) Если $A=n\pi\neq 0$, то $S(P)=n\pi[(x-x_0)^2+(y-y_0)^2]+S(M)$, где M есть точка, в которой $x=x_0=-B/2n\pi$, $y=y_0=-C/2n\pi$; следовательно, A, B, C, D имеют значения, указанные в упр. 1.

б$_1$) Если $A=n\pi=0$, но $B^2+C^2>0$, то

$$S(P)=\sqrt{B^2+C^2}\cdot\frac{Bx+Cy+D}{\sqrt{B^2+C^2}}=\lambda d(P),$$

где $\lambda=\sqrt{B^2+C^2}$, а Δ есть прямая $Bx+Cy+D=0$.

б$_2$) Если $A=B=C=0$, то $S(P)=D=\text{const}$.

3. Для движения плоскости Q, жестко связанной с шатуном AB, имеем $n=0$, $S(A)=0$, $S(B)=\pi CB^2=\pi\gamma^2$. Следовательно, прямая Δ проходит через точку A и, в силу симметрии, перпендикулярна к AB. Поэтому $S(K)=\frac{\pi}{l}\gamma^2 d(K)$, где l есть длина AB.

4. Для движения плоскости Q, жестко связанной с хордой AB, имеем $n=1$, $S(A)=S(B)=S=$ площади, ограниченной контуром Γ. Поэтому точка M теоремы Штейнера (упр. 2) лежит на равных расстояниях от A и B, и $S(A)=\pi MA^2+S(M)$, $S(K)=\pi MK^2+S(M)$. Следовательно, $S(A)-S(K)=$ площади Γ — площадь $\Gamma'=\pi(MA^2-MK^2)=\pi ab$.

5. Пусть l — длина кривой Γ. По формулам Френэ (стр. 115, упр. 7)

$$\oint_{\Gamma}\frac{\boldsymbol{n}}{\rho}\,ds=\int_0^l\dot{\mathbf{t}}\,ds=\int_0^l\frac{d^2\boldsymbol{r}}{ds^2}\,ds=0;$$

$$\oint_{\Gamma}\frac{[\boldsymbol{rn}]}{\rho}\,ds=\int_0^l[\boldsymbol{r}\dot{\mathbf{t}}]\,ds=[\boldsymbol{r}\mathbf{t}]\Big|_0^l-\int_0^l[\dot{\boldsymbol{r}}\mathbf{t}]\,ds=-\int_0^l[\mathbf{tt}]\,ds=0.$$

6. Пусть $\boldsymbol{n}_i=\{\cos\alpha,\ \cos\beta,\ \cos\gamma\}$ и $\boldsymbol{r}=\{x, y, z\}$. Если в координатную запись формулы Гаусса

$$\oint\!\!\oint_{\Sigma}(F_1\cos\alpha+F_2\cos\beta+F_3\cos\gamma)\,dS=-\iiint_{G}\left(\frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}+\frac{\partial F_3}{\partial z}\right)dx\,dy\,dz$$

подставим $F_1=1$, $F_2=F_3=0$, а затем $F_1=0$, $F_2=-z$, $F_3=y$, то получим

$$\oiint \cos\alpha\, dS=0 \quad \text{и} \quad \oiint (y\cos\gamma - z\cos\beta)\, dS=0.$$

[Это проекции доказываемых векторных равенств на ось x. Аналогично получаются проекции этих векторных равенств на оси y и z.]

7. Поместим начало O системы координат (x, y, z) на свободной горизонтальной поверхности жидкости и направим ось Oz вертикально вниз, так что на свободной поверхности $z=0$. Давление жидкости на площадку dS поверхности тела с единичным нормальным вектором $\boldsymbol{n}$, направленным внутрь тела, равно $\boldsymbol{n}z\,dS$, где z — глубина погружения площадки; момент этой элементарной силы относительно начала равен $[\boldsymbol{rn}]\,z\,dS$. Применяя координатную запись формулы Гаусса в пространстве три раза и выбирая каждый раз надлежащим образом функции F_1, F_2 и F_3, получим для проекций результирующей силы на оси координат следующие формулы:

$$\oiint z\cos\alpha\, dS=0, \qquad \oiint z\cos\beta\, dS=0,$$

$$\oiint z\cos\gamma\, dS=-\iiint dx\,dy\,dz=-V.$$

Для координат результирующего момента относительно начала O получим аналогичным путем, с помощью теоремы Гаусса:

$$\oiint (yz\cos\gamma - z^2\cos\beta)\, dS=-\iiint y\,dx\,dy\,dz=-Vy_0,$$

$$\oiint (z^2\cos\alpha - xz\cos\gamma)\, dS=\iiint x\,dx\,dy\,dz=Vx_0,$$

$$\oiint (xz\cos\beta - yz\cos\alpha)\, dS=0,$$

где x_0, y_0, z_0 — координаты центра массы C объемной области, занимаемой телом. Заметим теперь, что результирующая сила $\boldsymbol{F}=\{0, 0, -V\}$, а ее момент относительно начала координат есть

$$[\boldsymbol{rF}]=\{-Vy_0, Vx_0, 0\}.$$

8. Из параметрических уравнений эллипсоида

$$x=a\cos u\cos v, \qquad y=b\sin u\cos v, \qquad z=c\sin v$$

$$\left(0\leqslant u<2\pi, \quad -\frac{\pi}{2}\leqslant v\leqslant\frac{\pi}{2}\right)$$

вытекают следующие формулы:

$$p\,dS=abc\cos v\,du\,dv, \qquad \frac{dS}{p}=\frac{D^2\,du\,dv}{abc\cos v},$$

где

$$D^2=b^2c^2\cos^2 u\cos^2 v+a^2c^2\sin^2 u\cos^2 v+a^2b^2\sin^2 v\cos^2 v.$$

10. Этот интеграл представляет «расправленный» телесный угол, под которым вся плоскость $z=0$ видна из точки $M(0, 0, 1)$. Для прямого вычисления интеграла ввести полярные координаты на плоскости xy.

12. Установить тождество

$$\frac{\partial}{\partial x}\left(\frac{a-x}{R^3}\right)+\frac{\partial}{\partial y}\left(\frac{b-y}{R^3}\right)+\frac{\partial}{\partial z}\left(\frac{c-z}{R^3}\right)=0, \quad R^2=(x-a)^2+(y-b)^2+(z-c)^2,$$

для всех точек (x, y, z), отличных от (a, b, c). Из формулы Гаусса для пространства вытекает тогда, что если Σ есть замкнутая поверхность, то: а) $\Omega = 0$, если точка $A(a, b, c)$ лежит вне Σ и б) если же точка A лежит внутри Σ, то значение поверхностного интеграла не зависит от формы поверхности Σ. Выбрав в роли этой поверхности сферу с центром A, нетрудно получить $\Omega = 4\pi$.

13. Поверхностный интеграл

$$\frac{\partial\Omega}{\partial a} = \iint\limits_{\Sigma} \frac{\partial}{\partial a}\left(\frac{a-x}{R^3}\right) dy\,dz + \frac{\partial}{\partial a}\left(\frac{b-y}{R^3}\right) dz\,dx + \frac{\partial}{\partial a}\left(\frac{c-z}{R^3}\right) dx\,dy$$

не зависит от выбора поверхности Σ и зависит только от ее граничной кривой Γ, ибо из тождества, данного в решении упр. 12, вытекает, что

$$\frac{\partial}{\partial x}\left[\frac{\partial}{\partial a}\left(\frac{a-x}{R^3}\right)\right] + \frac{\partial}{\partial y}\left[\frac{\partial}{\partial a}\left(\frac{b-y}{R^3}\right)\right] + \frac{\partial}{\partial z}\left[\frac{\partial}{\partial a}\left(\frac{c-z}{R^3}\right)\right] = 0.$$

По теореме Стокса и в силу Дополнений к гл. V, § 2, поверхностный интеграл в выражении для $\frac{\partial\Omega}{\partial a}$ можно представить в виде криволинейного интеграла $\oint (u\,dz + v\,dy + w\,dz)$ вдоль Γ. Проверить, что функции $u = 0$, $v = \frac{z-c}{R^3}$, $w = -\frac{y-b}{R^3}$ удовлетворяют тождествам

$$\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} = \frac{\partial}{\partial a}\left(\frac{a-x}{R^3}\right), \qquad \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} = \frac{\partial}{\partial a}\left(\frac{b-y}{R^3}\right), \qquad \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{\partial}{\partial a}\left(\frac{c-z}{R^3}\right).$$

14. Установить следующие факты: 1) значение криволинейного интеграла Θ остается неизменным, если кривая Γ деформируется таким образом, что она при своей деформации не проходит ни через точку $(-1; 0)$, ни через точку $(1; 0)$; 2) $\Theta = 2\pi$, если Γ есть малая окружность с центром $(1; 0)$, ориентированная против часовой стрелки; 3) $\Theta = 2\pi$, если Γ есть малая окружность с центром $(-1; 0)$, ориентированная по часовой стрелке.

15. Представим себе C в виде жесткого проволочного кругового кольца, а Γ в виде нити. Деформируем нить Γ в новое положение Γ', лежащее целиком в плоскости $y = 0$. В процессе этой деформации числа p и n не изменяются и первая формула получается сразу, если применить упр. 14 к отрезку $-1 < x < 1$, $y = 0$, $z = 0$ плоскости $y = 0$ и кривой Γ', лежащей в этой плоскости. Множитель 4π (вместо участвующего в упр. 14 множителя 2π) получается вследствие того, что телесный угол Ω возрастает на 4π вдоль замкнутого пути, для которого $p = 1$, $n = 0$.

Упомянутую выше деформацию кривой Γ в Γ' можно, например, выполнить аналитически следующим образом. Предположим, что кривая Γ не встречает оси z и задана параметрическими уравнениями

$$z = \gamma(t)\cos\varphi(t), \qquad y = \gamma(t)\sin\varphi(t), \qquad z = z(t) \qquad (0 \leqslant t \leqslant 2\pi).$$

Рассмотрим семейство кривых $\Gamma(\tau)$:

$$x = \gamma(t)\cos(\tau\varphi(t)), \quad y = \gamma(t)\sin(\tau\varphi(t)), \quad z = z(t),$$

зависящее от параметра τ, который убывает от $\tau = 1$ до $\tau = 0$. Заметим, что $\Gamma(1)$ есть кривая Γ, а $\Gamma(0)$ есть замкнутая кривая, лежащая в плоскости $y = 0$, и ее можно принять за Γ'. Заметим также, что при фиксированном значении t каждая точка P кривой $\Gamma(\tau)$, имея постоянный z, вращается вокруг оси z при изменении τ; следовательно, телесный угол Ω с вершиной P, стягиваемый кривой C, не изменяется при этом изменении τ. Отсюда вытекает, что $\Omega_1 - \Omega_0$ имеет одинаковое значение для кривых $\Gamma' = \Gamma(0)$ и $\Gamma = \Gamma(1)$. Для доказательства второй формулы надо заметить, что

$$\Omega_1 - \Omega_0 = \int\limits_\Gamma d\Omega = \int\limits_\Gamma \operatorname{grad}\Omega \cdot dP = -\int\limits_\Gamma dP \cdot \int\limits_C \frac{[\overline{PP'}\,dP']}{|\overline{PP'}|^3} =$$

$$= -\int\limits_\Gamma\int\limits_C \frac{dP \cdot [\overline{PP'}\,dP']}{|\overline{PP'}|^3} = \int\limits_\Gamma\int\limits_C \frac{\overline{PP'} \cdot [dP \cdot dP']}{|\overline{PP'}|^3}.$$

16. Возьмем координатную систему Ox, Oy, Oz и обозначим радиус-вектор переменной точки кривой Γ через $\boldsymbol{r}$. Тогда вектор

$$\boldsymbol{a} = \frac{1}{2}\oint\limits_\Gamma [\boldsymbol{r}\, d\boldsymbol{r}]$$

обладает требуемым свойством, так как

$$a_3 = \frac{1}{2}\oint\limits_\Gamma (x\,dy - y\,dx)$$

есть площадь проекции кривой Γ на плоскость xOy.

ГЛАВА VI

§ 2, стр. 450.

1. Воспользоваться уравнением, выражающим закон сохранения энергии, и доказать, что $r \to \infty$ при $t \to \infty$.

2. Пусть (ξ, η) — прямоугольные координаты планеты относительно главных осей эллипса, а (x, y) — ее координаты относительно системы параллельных осей с началом в центре Солнца. Тогда параметрические уравнения эллипса будут

$$\xi = x + \varepsilon a = a\cos\omega, \qquad \eta = y = b\sin\omega.$$

Согласно закону площадей,

$$k(t - t_0) = \int\limits_0^\omega \left(x\frac{dy}{d\omega} - y\frac{dx}{d\omega}\right) d\omega = ab\int\limits_0^\omega (1 - \varepsilon\cos\omega)\, d\omega.$$

3, 4. Воспользоваться уравнением, выражающим закон сохранения энергии, и законом площадей.

6. В любом центральном силовом поле движение происходит в одной плоскости (в п° 3, стр. 445—446 это было доказано). Имеем

$$\ddot{x} = -\frac{x}{r}f, \qquad \ddot{y} = -\frac{y}{r}f.$$

Отсюда

$$x\dot{y} - y\dot{x} = k = \text{const}, \qquad \dot{x}\ddot{x} + \dot{y}\ddot{y} = \frac{-x\dot{x} - y\dot{y}}{r}f = -\dot{r}f.$$

Следовательно,

$$\frac{1}{2}\frac{d}{dt}(\dot{x}^2 + \dot{y}^2) = -\dot{r}f.$$

Расстояние от начала до касательной есть

$$q=\frac{|x\dot{y}-y\dot{x}|}{\sqrt{\dot{x}^2+\dot{y}^2}}=\frac{|k|}{\sqrt{\dot{x}^2+\dot{y}^2}}.$$

Поэтому

$$\frac{1}{2}\frac{d}{dt}\frac{k^2}{q^2}=-f\frac{dr}{dt},$$

откуда

$$\frac{1}{2}\frac{d}{dr}\frac{k^2}{q^2}=-f$$

или

$$f=\frac{k^2}{q^3}\frac{dq}{dr}.$$

Для кардиоиды $q=\frac{r^2}{\sqrt{2ar}}$.

§ 3, стр. 465.

1. а) Воспользоваться тем, что криволинейный интеграл

$$\int(3x^2+6xy^2)\,dx+(6x^2y+4y^3)\,dy$$

не зависит от пути. Выбрав в качестве пути интегрирования ломаную OAP, где $O\,(0,\,0)$, $A\,(x,\,0)$, $P\,(x,\,y)$, получим общее решение

$$\int\limits_{(0,\,0)}^{(x,\,y)}(3x^2+6xy^2)\,dx+(6x^2y+4y^3)\,dy=x^3+3x^2y^2+y^4=c.$$

б) Общее решение $\sqrt{1+x^2+y^2}-\operatorname{arctg}\frac{y}{x}=c$.

2. $x^2y-2xy^2-2cy-2=0$; интегрирующий множитель $\mu=\frac{1}{y^2}$.

3. $xy+\operatorname{arctg}x=c$; интегрирующий множитель $\mu=\frac{1}{1+x^2}$.

4. Дифференциальное уравнение является линейным, если считать x искомой функцией; его общее решение есть

$$(xy^2+1)^2=cy.$$

Из тождества

$$d\left(\frac{(xy^2+1)^2}{y}\right)=\frac{xy^2+1}{y^2}\left[2y^3\,dx+(3xy^2-1)\,dy\right]$$

выясняется интегрирующий множитель данного уравнения.

5. а) $x^2+y^2+cx+1=0$ $(-\infty<c<\infty)$ и прямая $x=0$.

б) $x^2+2y^2=c^2$.

в) Дифференциальное уравнение этого семейства эллипсов и гипербол (ср. стр. 176, упр. 5) есть

$$(y')^2+\frac{x^2-y^2-a^2+b^2}{xy}y'-1=0;$$

оно не изменяется при замене y' на $(-1/y')$. Семейство эллипсов $(-b^2<c<\infty)$ ортогонально семейству гипербол $(-a^2<c<-b^2)$.

г) $y=\ln|\operatorname{tg}(x/2)|+c$ и вертикальные прямые $x=k\pi$ (k — целое).

д) Семейство кривых (трактрисс)

$$x-c=\pm\left(\sqrt{a^2-y^2}-a\operatorname{arch}\frac{a}{y}\right)$$

и семейство, ему симметричное относительно оси x.

6. а) Семейство парабол $y=cx^2$.
б) Семейство гипербол $xy=c$.
7. а) $y=x^2$; б) $y=-x+x\ln(-x)$ $(-\infty<x<0)$.
8. $y=xp+a\sqrt{1+p^2}-ap\,\operatorname{arsh}p$.
9. $x=ce^{-p/a}+\frac{1}{2}p$, $\quad y=c(p+a)e^{-p/a}+\frac{1}{2}p(p+a)-\frac{1}{4}(p+a)^2$.
Обратите внимание, что при $c=0$ получается парабола

$$y=x^2-\frac{a^2}{4}.$$

Каков геометрический смысл этого результата?

10. а) $y=\sin(x+c)$, особые решения $y=\pm 1$;

б) $x=\pm\frac{1}{2}(\arcsin y+y\sqrt{1-y^2})+c$;

в) $x=\mp\left(\sqrt{(2a-y)y}-2a\operatorname{arctg}\sqrt{\frac{y}{2a-y}}\right)+c$; это семейство циклоид, которое может быть представлено в параметрическом виде:

$$x=c+a(t-\sin t),\qquad y=a(1-\cos t).$$

Особое решение $y=2a$;

г) $x=\pm\int\limits_0^y\sqrt{\frac{1+y^2}{1-y^2}}\,dy+c$, $-1\leqslant y\leqslant 1$; особые решения $y=\pm 1$.

Доказать, что эти кривые не являются синусоидами.

11. $MN=y\sqrt{1+(y')^2}$, $MC=-\frac{[1+(y')^2]^{3/2}}{y''}$, и дифференциальное уравнение искомых кривых будет $(1+y'^2)^2y+ky''=0$. Это — неполное дифференциальное уравнение второго порядка, не содержащее явно независимой переменной. По методу т. I, стр. 607—609 (§ 2, п° 2б) отсюда получается

$$\left(\frac{dy}{dx}\right)^2=\frac{k+c-y^2}{y^2-c},$$

где c — произвольная постоянная.

Возможны следующие различные случаи (все они имеют значение в дифференциальной геометрии):

1) $k=\lambda^2\ (>0)$, $c=-\gamma^2\ (<0)$, $\gamma^2<\lambda^2$. Кривая — всюду гладкая, осциллирующая (т. е. имеет колебательный характер); попеременно касается прямых

$$y=\pm\sqrt{\lambda^2-\gamma^2}.$$

Она напоминает синусоиду, но не является синусоидой.

2) $k=\lambda^2$, $c=0$. Кривая является окружностью радиуса λ с центром на оси x.

3) $k=\lambda^2$, $c=\gamma^2\ (>0)$. Кривая состоит из последовательности конгруэнтных дуг, соединенных между собой в точках возврата, лежащих на прямых $y=\gamma$; все эти дуги касаются прямой

$$y=\sqrt{\lambda^2+\gamma^2}.$$

Она похожа на циклоиду, но не является циклоидой.

4) $k=-\lambda^2\ (<0)$, $c=\gamma^2>\lambda^2$. Кривая состоит из последовательности тождественных дуг, опрокинутых по сравнению со случаем 3), точки возврата на прямой $y=\gamma$; все дуги касаются прямой

$$y=\sqrt{\gamma^2-\lambda^2}.$$

5) $k=-\lambda^2$, $c=\gamma^2=\lambda^2$. Кривая является трактриссой.

6) $k=-\lambda^2$, $c=\gamma^2<\lambda^2$. Кривая имеет бесконечное множество точек возврата, лежащих попеременно на прямых $y=\gamma$ и $y=-\gamma$; касательные в точках возврата перпендикулярны к этим прямым.

12. Уравнение семейства дифференцировать последовательно три раза; параметры c и a исключаются автоматически. После исключения b получится искомое дифференциальное уравнение

$$(1+y'^2)\,y'''-3y'\,(y'')^2=0.$$

13. $y=x\sin ax$; особые решения

$$y=x \text{ и } y=-x.$$

14. Положим

$$y(x)=\sum_{k=0}^{\infty} c_k x^k.$$

Тогда

$$c_{k+2}=-\frac{c_k}{(k+2)^2} \quad \text{и} \quad c_0=1,\ c_1=0,$$

откуда

$$y(x)=\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{2k}(k!)^2}x^{2k}.$$

Возьмем выражение для $J_0(x)$ из упр. 4, стр. 245, подставим в него вместо $\cos xt$ его разложение в степенной ряд и переставим порядок суммирования и интегрирования (почему это допустимо?). Тогда получится

$$J_0(x)=\frac{1}{\pi}\sum_{k=0}^{\infty}\frac{x^{2k}}{(2k)!}(-1)^k\int_{-1}^{+1}\frac{t^{2k}}{\sqrt{1-t^2}}\,dt.$$

Подстановка $t=\sin\tau$ преобразует интеграл в правой части в интеграл, вычисленный в т. I, гл. IV, § 4, n° 7, так что

$$\int_{-1}^{+1}\frac{t^{2k}}{\sqrt{1-t^2}}\,dt=\frac{(2k)!\,\pi}{(k!)^2\,2^{2k}},$$

и степенные ряды для $y(x)$ и $J_0(x)$ оказываются тождественными.

15. Кривая удовлетворяет дифференциальному уравнению

$$n\left(x\frac{dy}{dx}-y\right)=r,$$

где

$$r=\sqrt{x^2+y^2}.$$

Это однородное уравнение первого порядка. Его общее решение есть

$$x=a\left(\frac{y}{x}+\sqrt{1+\frac{y^2}{x^2}}\right)^n$$

или (в полярных координатах)

$$r\cos\theta=a\left(\operatorname{tg}\theta+\frac{1}{\cos\theta}\right)^n,$$

откуда

$$r=\frac{a(1+\sin\theta)^n}{\cos^{n+1}\theta}.$$

Другой способ. Преобразовать дифференциальное уравнение к полярным координатам:

$$\frac{nr^2}{\cos\theta\,\dfrac{dr}{d\theta}-r\sin\theta}=r$$

или

$$\frac{1}{r}\,dr=\left(\frac{n}{\cos\theta}+\frac{\sin\theta}{\cos\theta}\right)d\theta,$$

откуда

$$r=a\,\frac{\left[\operatorname{tg}\left(\frac{\theta}{2}+\frac{\pi}{4}\right)\right]^n}{\cos\theta}=a\,\frac{(1+\sin\theta)^n}{\cos^{n+1}\theta}.$$

§ 4, стр. 479.

1. Воспользоваться полной индукцией. Сначала убедиться, что *две* такие функции линейно независимы, а затем доказать, что из предположения линейной независимости $k-1$ функций φ_i вытекает линейная независимость k таких функций. Последнее доказывается от противного. Пусть существует тождество

$$c_1\varphi_1+c_2\varphi_2+\ldots+c_k\varphi_k=0, \qquad (1)$$

между тем как любые $k-1$ таких функций линейно независимы; тогда можно считать, что все $c_k\neq 0$. Разделим тождество (1) на e^{a_kx} и затем продифференцируем (n_k+1) раз, где n_k есть степень многочлена $P_k(x)$. Степени многочленов, которые будут множителями при остальных e^{a_ix}, останутся неизменными, так что эти многочлены тоже будут отличны от нуля.

2. Помножить обе части уравнения Бернулли на $(1-n)\,y^{-n}$.

а) $y^{-1}=cx+\ln x+1$; б) $y^3=cx^{-3}+\dfrac{3a^2}{2x}$; в) $(y^{-1}+a)^2=c\,(x^2-1)$.

3. Преобразование искомой функции $y=y_1+z$ переводит уравнение Риккати в уравнение Бернулли

$$z'+Pz^2+(2Py_1+Q)\,z=0,$$

а затем замена $z=u^{-1}$ (см. упр. 2) преобразует последнее уравнение в линейное

$$u'-(2Py_1+Q)\,u=P.$$

Оба последовательных преобразования можно заменить одним: $y=y_1+u^{-1}$.

Общее решение предложенного уравнения:

$$y=x-\frac{e^{\frac{1}{2}x^4}}{c+\int\limits_0^x x^2e^{\frac{1}{2}x^4}\,dx}.$$

4. Приравнивая друг другу правые части обоих уравнений, находим их совпадающее решение $y=x^2$.

5. Общее решение

$$y=x^2-\frac{e^{\frac{2}{3}x^3}}{c+\int\limits_{-\infty}^x e^{\frac{2}{3}x^3}\,dx}\quad(=f(x,\,c)).]$$

Для построения эскиза семейства интегральных кривых начертим сначала обе ветви кривой

$$y^2+2x-x^4=0,$$

т. е.

$$y=\pm\sqrt{x(x^3-2)},$$

которые разбивают плоскость на две области, где $y'<0$, и одну область, где $y'>0$. Две бесконечные ветви этой кривой приближаются асимптотически к двум параболам $y=\pm x^2$. Показать, что интегральные кривые приближаются асимптотически к этим параболам; для этого надо доказать два соотношения:

а) $f(x, c)=-x^2+o(1)$ при $x\to+\infty$ $(-\infty<c<\infty)$

и

б) $f(x, c)=x^2+o(1)$ при $x\to-\infty$ $(c\neq 0)$, где $o(1)$ обозначает функцию, стремящуюся к нулю.

6. Положим

$$y_1-y_3=a,\quad y_1-y_4=b,\quad y_2-y_3=c,\quad y_2-y_4=d.$$

Тогда

$$a'+Pa(y_1+y_3)+Qa=0,$$

так что

$$P(y_1+y_3)=-Q-\frac{a'}{a},\qquad P(y_1-y_3)=aP,$$

или

$$2Py_1=aP-Q-\frac{a'}{a}.$$

Аналогично

$$2Py_1=bP-Q-\frac{b'}{b}.$$

Следовательно,

$$\frac{d\ln(a/b)}{dx}=P(a-b)=-P(y_3-y_4)$$

и аналогично

$$\frac{d\ln(c/d)}{dx}=-P(y_3-y_4);$$

вычитая из первого равенства второе и интегрируя, получим

$$\ln\left(\frac{a}{b}:\frac{c}{d}\right)=\text{const}.$$

7. Воспользоваться соотношением

$$\frac{d\ln(a/b)}{dx}=P(y_4-y_3),$$

доказанным в предыдущем упражнении. Предложенное конкретное уравнение имеет частные решения

$$y_1=\frac{1}{\cos x}\ \text{и}\ y_2=-\frac{1}{\cos x};$$

его общее решение

$$y=\frac{1+ce^{2x}}{(1-ce^{2x})\cos x}.$$

8. Исключая y'' из обоих уравнений, найдем их совпадающее решение $y_1 = e^x$. Их общие решения:

$$\text{а) } y = c_1 e^x + c_2 x; \qquad \text{б) } y = c_1 e^x + c_2 \sqrt{x}.$$

§ 4, стр. 483.

1. а) $y = c_1 e^x + c_2 e^{-\frac{1}{2}x} \cos\frac{\sqrt{3}}{2}x + c_3 e^{-\frac{1}{2}x} \sin\frac{\sqrt{3}}{2}x$;

б) $y = c_1 e^x + c_2 x e^x + c_3 e^{2x}$; в) $y = c_1 e^x + c_2 x e^x + c_3 x^2 e^x$;

г) $y = c_1 e^x + c_2 e^{-x} + c_3 e^{\sqrt{2}\,x} + c_4 e^{-\sqrt{2}\,x}$.

2. Ввести новую независимую переменную $t = \ln|x|$, откуда $x = e^t$ при $x > 0$ и $x = -e^t$ при $x < 0$.

Отв. а) $y = c_1 x^n + c_2 x^{-n}$; б) $y = c_1 x + c_2 x^2 + c_3/x^2$. Уравнение Эйлера можно также решить, если сразу искать частные решения вида x^α.

3. Составляем уравнения движения

$$\ddot{x} = -\lambda^2 x - 2\mu\dot{y}, \quad \ddot{y} = -\lambda^2 y + 2\mu\dot{x}, \quad \ddot{z} = -\lambda^2 z. \tag{α}$$

Так как в начальный момент $z = \dot{z} = 0$, то при любом t будет $z = 0$, и движение происходит в плоскости xy. Дифференцируя первые два уравнения два раза по t и исключая из шести уравнений пять величин y, $\dot{y}$, $\ddot{y}$, $\dddot{y}$ и $\ddddot{y}$, получим л. д. у. для одного лишь x:

$$\ddddot{x} + (2\lambda^2 + 4\mu^2)\ddot{x} + \lambda^4 x = 0.$$

Аналогично, исключая x и его производные, получим уравнение для одного только y:

$$\ddddot{y} + (2\lambda^2 + 4\mu^2)\ddot{y} + \lambda^4 y = 0.$$

Следовательно, x и y являются линейными комбинациями четырех частных решений

$$e^{\pm i(\sqrt{\lambda^2+\mu^2} \pm \mu)t}$$

или четырех частных решений $\cos(\sqrt{\lambda^2+\mu^2}+\mu)t$, $\cos(\sqrt{\lambda^2+\mu^2}-\mu)t$, $\sin(\sqrt{\lambda^2+\mu^2}+\mu)t$ и $\sin(\sqrt{\lambda^2+\mu^2}-\mu)t$ с постоянными коэффициентами A, B, C, D и A_1, B_1, C_1, D_1. (Это вытекает из формул Эйлера, ср. т. I, стр. 615.) Подставив полученные выражения для x и y в уравнения (α), найдем, что $A_1 = -C$, $B_1 = D$, $C_1 = A$ и $D_1 = -B$, а затем из начальных условий $x(0) = y(0) = \dot{y}(0) = 0$ и $\dot{x}(0) = v_0$ определятся A, B, C, D.

4. б) После умножения на x^3 уравнение примет тот вид, который рассмотрен в а). Оно имеет частные решения $u = x^3$ и $v = x^5$; следовательно, в силу а), оно имеет также частное решение $w = x^2 + 1$. Эти три решения линейно независимы. Поэтому общее решение есть

$$y = A(x^2+1) + Bx^3 + Cx^5.$$

§ 4, стр. 487.

1. а) Подставив проверяемое решение в левую часть дифференциального уравнения, получим

$$(a_0 b_0 - 1) P(x) + (a_0 b_1 + a_1 b_0) P'(x) + (a_0 b_2 + a_1 b_1 + a_2 b_0) P''(x) + \ldots,$$

но это выражение равно нулю в силу тождеств

$$a_0 b_0 = 1, \quad a_0 b_1 + a_1 b_0 = 0, \quad a_0 b_2 + a_1 b_1 + a_2 b_0 = 0, \ \ldots,$$

вытекающих из разложения дроби в степенной ряд.

б) Вытекает из а), если положить $y' = z$.

2. а) $\frac{1}{1+t^2} = 1 - t^2 + t^4 - + \ldots$; следовательно,

$$y = P(x) - P''(x) = 3x^2 - 5x - 6;$$

б) $\frac{1}{t+t^2} = \frac{1}{t} - 1 + t - t^2 + - \ldots$, откуда

$$y = \int P(x)\,dx - P(x) + P'(x) - P''(x) = -1 + x + \frac{1}{3}x^3.$$

3. а) $y = \frac{3}{8}e^x$; б) $y = \frac{1}{6}x^3e^x$.

4. $y = e^x\left(\frac{x^2}{2} + \frac{3}{2}x + \frac{1}{4}\right) + c_1e^{3x} + c_2e^{2x}$.

§ 5, стр. 504.

1. Формула Пуассона (при $R = 1$) дает потенциальную функцию $u(r, \theta)$ внутри единичного круга, принимающую на его окружности краевые значения $f(\theta)$. Но $u\left(\frac{1}{r}, \theta\right)$ тоже является потенциальной функцией (ср. т. I, гл. X, стр. 574, упр. 3), принимает она те же самые краевые значения и является ограниченной в области, лежащей вне единичной окружности. Поэтому выражение

$$\frac{r^2-1}{2\pi}\int_0^{2\pi} \frac{f(\alpha)\,d\alpha}{1 - 2r\cos(\theta - \alpha) + r^2}$$

является решением внешней краевой задачи уравнения Лапласа для единичного круга.

2. Потенциал заряженного отрезка равен

$$\mu \ln \frac{z + l + \sqrt{(z+l)^2 + x^2 + y^2}}{z - l + \sqrt{(z-l)^2 + x^2 + y^2}}.$$

а) Так как на эллипсоиде $z = l\alpha\cos\varphi$, $\sqrt{x^2+y^2} = l\sqrt{\alpha^2 - 1}\sin\varphi$, то потенциал есть $\mu\ln\frac{\alpha+1}{\alpha-1}$, а эквипотенциальными поверхностями являются софокусные эллипсоиды

$$\frac{z^2}{l^2\alpha^2} + \frac{x^2+y^2}{l^2(\alpha^2-1)} = 1 \qquad (1 \leqslant \alpha < \infty).$$

б) Силовые линии (как ортогональные траектории эквипотенциальных поверхностей) являются (ср. стр. 466, упр. 5в)) софокусными гиперболами, выражаемыми тем же уравнением, но при $0 \leqslant \alpha \leqslant 1$ и при постоянном отношении x к y.

3. Обозначим через Σ сферу радиуса ρ с центром (x, y, z), лежащую внутри S. Так как $\nabla^2\left(\frac{1}{r}\right) = 0$ и $\nabla^2 u = 0$ в области, лежащей между Σ и S, то по теореме Грина (стр. 414)

$$0 = \oint\!\!\oint_S \left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial(1/r)}{\partial n}\right)d\sigma - \oint\!\!\oint_\Sigma \left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial(1/r)}{\partial n}\right)d\sigma,$$

где в первом интеграле производная берется по внешней нормали к S, а во втором — по внешней нормали к Σ. Но на сфере Σ имеем

$$\frac{\partial(1/r)}{\partial n}=\frac{\partial(1/r)}{\partial r}=-\frac{1}{\rho^2},\quad r=\text{const}=\rho.$$

Поэтому

$$\oiint_{\Sigma}\frac{1}{r}\frac{\partial u}{\partial n}d\sigma=\frac{1}{\rho}\oiint_{\Sigma}\frac{\partial u}{\partial n}d\sigma=0,$$

так как u есть потенциальная функция (см. сноску стр. 498); кроме того,

$$-\frac{1}{4\pi}\oiint_{\Sigma}u\frac{\partial(1/r)}{\partial n}d\sigma=\frac{1}{4\pi\rho^2}\oiint_{\Sigma}u\,d\sigma,$$

а при $\rho\to 0$ это выражение имеет, очевидно, своим пределом $u(x, y, z)$, так как оно является средним значением потенциальной функции u на сфере Σ.

§ 6, стр. 512.

1. а) $u=f(x)+g(y)$, где f и g — произвольные функции;
б) $u=f(x, y)+g(x, z)+h(y, z)$, где f, g, h — произвольные функции.

2. Воспользоваться линейным преобразованием

$$x=\xi+\eta,\qquad y=3\xi+2\eta.$$

Отв. $u=f(y-2x)+g(3x-y)+\frac{1}{12}e^{x+y}$.

3. $z^2(z_x^2+z_y^2+1)=1$.

4. $u(x, t)=f(x-at)+g(x+at)$. Отсюда при $x\geqslant 0$ (в силу начальных условий)

$$u(x, 0)=f(x)+g(x)=0,$$
$$u_t(x, 0)=-af'(x)+ag'(x)=0.$$

Дифференцируя первое уравнение и сопоставляя со вторым, имеем $f'(x)=0$ и $g'(x)=0$ при $x\geqslant 0$, откуда

$$f(x)=c\ (=\text{const}),$$
$$g(x)=-c\quad\text{при } x\geqslant 0.$$

Далее, в силу краевого условия,

$$\varphi(t)=u(0, t)=f(-at)+g(at)=f(-at)-c,$$

т. е.

$$f(\xi)=\varphi\left(\frac{\xi}{-a}\right)+c,$$

если $\xi<0$.

Так как всегда $x+at\geqslant 0$, то $g(x+at)=-c$, а следовательно

$$u(x, t)=\begin{cases}0, & \text{если } x-at\geqslant 0,\\ \varphi\left(\dfrac{x-at}{-a}\right), & \text{если } x-at\leqslant 0\end{cases}$$

коль скоро x и t не отрицательны.

5. Если положить $u(x, y)=\sum a_{nk}x^ny^k$, то

$$a_{n+1,\,k+1}=\frac{a_{nk}}{(n+1)(k+1)}.$$

Затем $a_{n0}=a_{0n}=0$ при $n \geqslant 1$ и $a_{00}=1$. Следовательно,

$$u(x, y)=\sum_{n=0}^{\infty} \frac{x^n y^n}{(n!)^2}=J_0(2i\sqrt{xy}),$$

где J_0 — символ бесселевой функции нулевого индекса (стр. 245, упр. 4).

6. а) Из дифференциального уравнения имеем

$$[f'(x)]^2+[g'(y)]^2=1,$$

откуда

$$[f'(x)]^2=1-[g'(y)]^2.$$

Так как левая часть не зависит от y, а правая не зависит от x, то обе части должны быть равны постоянной, которая должна быть $\geqslant 0$; обозначим ее через c^2. Тогда

$$[f'(x)]^2=c^2, \qquad 1-[g'(y)]^2=c^2.$$

Следовательно,

$$u=cx \pm \sqrt{1-c^2}\, y+b,$$

где c и b — произвольные постоянные и $c^2 \leqslant 1$.

б) Если $u=f(x)+g(y)$, то

$$f'(x)=\frac{1}{g'(y)}=a=\text{const},$$

так что

$$u=ax+\frac{1}{a}y+b,$$

где a и b — произвольные постоянные.

Если $u=f(x)\,g(y)$, то

$$\frac{d}{dx}[f(x)]^2=\frac{4}{\frac{d}{dy}[g(y)]^2}=2c=\text{const}.$$

В этом случае отсюда получается

$$u=\sqrt{(2cx+a)\left(\frac{2}{c}y+b\right)},$$

где a, b и c — произвольные постоянные.

7. Из двупараметрического семейства решений $z=u(x, y;\ a, b)$ выделяется однопараметрическое семейство, если представить a и b как функции одного параметра t, т. е. положить $a=f(t)$, $b=g(t)$. Тогда

$$z=u\{x, y;\ f(t), g(t)\}. \tag{1}$$

Огибающая этого однопараметрического семейства получится, если выразить t из уравнения

$$z_t \equiv u_a f'(t)+u_b g'(t)=0$$

и подставить полученное выражение вместо t в уравнение (1). Полученная функция z от x и y тоже будет решением заданного дифференциального уравнения, так как

$$z=u(x, y;\ a, b),$$
$$z_x=u_x+u_t t_x=u_x(x, y;\ a, b),$$
$$z_y=u_y+u_t t_y=u_y(x, y;\ a, b),$$

а $z=u(x, y;\ a, b)$ удовлетворяет уравнению, заданному в условии.

8. $u = x\sqrt{\frac{y}{x+k}} + y\sqrt{\frac{x+k}{y}} + k\sqrt{\frac{y}{x+k}}$.

9. Согласно § 6, n° 2, решение первого уравнения имеет вид

$$z = f(x+at) + g(x-at).$$

Подставив это выражение во второе уравнение, получим

$$f'g' = 0,$$

откуда либо $f = \text{const}$, либо $g = \text{const}$. Стало быть,

$$z = f(x+at)$$

или

$$z = f(x-at)$$

есть самое общее решение, удовлетворяющее обоим уравнениям.

[10. Пусть K есть однородная функция степени α. Дифференциальное уравнение можно записать в следующем виде:

$$K\nabla^2 u + \frac{\partial K}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial K}{\partial y}\frac{\partial u}{\partial y} + \frac{\partial K}{\partial z}\frac{\partial u}{\partial z} = 0.$$

Положим

$$u = r^n = (x^2 + y^2 + z^2)^{n/2}.$$

Тогда

$$\frac{\partial u}{\partial x} = nr^{n-2}x, \quad \frac{\partial u}{\partial y} = nr^{n-2}y,$$

$$\frac{\partial u}{\partial z} = nr^{n-2}z \text{ и } \nabla^2 u = n(n+1)r^{n/2}.$$

Подставив $u = r^n$ в дифференциальное уравнение, получим

$$K\, n(n+1)r^{n-2} + nr^{n-2}\left(x\frac{\partial K}{\partial x} + y\frac{\partial K}{\partial y} + z\frac{\partial K}{\partial z}\right) = 0,$$

а так как (Дополнения к гл. II, § 3)

$$x\frac{\partial K}{\partial x} + y\frac{\partial K}{\partial y} + z\frac{\partial K}{\partial z} = \alpha K,$$

то $u = r^n$ будет решением, если $n + 1 + \alpha = 0$ или $n = -(\alpha+1)$. Следовательно, наше уравнение имеет решение

$$u = r^{-(\alpha+1)} = (x^2 + y^2 + z^2)^{-\frac{\alpha+1}{2}}.]$$

Глава VII

§ 1, стр. 520.

1. $\frac{2}{\sqrt{2g}}\sqrt{\frac{(x_1-x_0)^2+(y_1-y_0)^2}{y_1-y_0}}$.

2. $T = \int_{\sigma_0}^{\sigma_1} f(r)\sqrt{\dot r^2 + r^2\dot\theta^2 + r^2\sin^2\theta\,\dot\varphi^2}\,d\sigma$.

§ 2, стр. 528.

1. Параболы $y = c^2 + \frac{x^2}{4c^2}$. 2. Окружности с центром на оси x.

3. $y = c \sin \frac{x-a}{c}$.

4. $y = \frac{a}{x^{n-1}} + b$ при $n > 1$ и $y = a \ln x + b$ при $n = 1$.

5. $y = a(x-b)^{\frac{n}{n+m}}$, если $n + m \neq 0$; $y = ae^{bx}$, если $n = -m$.

6. Уравнение Эйлера есть

$$ay'' + a'y' + (b' - c)y = 0.$$

При $b = \text{const}$ второй член функционала,

$$\int\limits_{x_0}^{x_1} byy'\,dx = \frac{b}{2}(y_1^2 - y_0^2),$$

зависит только от граничных точек кривой $y = y(x)$.

7. $y_1 - y_0 < \frac{\pi}{2}$.

8. Рассмотрим $F(x, y)$ при фиксированном x как функцию от y; пусть эта функция от y имеет минимум при $y = \bar{y}$. Тогда $F(x, y) \geqslant F(x, \bar{y})$ в некоторой окрестности значения $\bar{y}$ и $F_y(x, \bar{y}) = 0$. Но $\bar{y}$ зависит от x как параметра: $\bar{y} = \bar{y}(x)$. Тогда для функции $y(x)$, близкой к $\bar{y}(x)$, имеем

$$\int\limits_{x_0}^{x_1} F(x, y(x))\,dx \geqslant \int\limits_{x_0}^{x_1} F(x, \bar{y}(x))\,dx,$$

где $\bar{y}(x)$ удовлетворяет уравнению $F_y(x, \bar{y}(x)) = 0$.

9. а) $y = 0$.

б) Воспользоваться неравенством Буняковского — Шварца для интегралов (примечание 1 на стр. 352). Для любой допустимой функции $y(x)$

$$1 = y(1) - y(0) = \int\limits_0^1 y'\,dx \leqslant \sqrt{\int\limits_0^1 1^2\,dx} \cdot \sqrt{\int\limits_0^1 y'^2\,dx} = \sqrt{I},$$

причем знак равенства имеет место для функции $y = x$.

§ 3, стр. 533.

Если $v = \frac{1}{f(r)}$, то время T, в течение которого свет проходит путь от точки A до точки B, можно взять из упр. 2, стр. 520 (ответ на стр. 652), а наша подынтегральная функция есть

$$F = f(r)\sqrt{(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\, \dot{\varphi}^2)}.$$

Уравнение Эйлера для переменной φ дает вдоль луча света

$$F_{\dot{\varphi}} = \frac{\dot{\varphi} f^2 r^2 \sin^2\theta}{F} = C = \text{const}.$$

Выберем нашу систему сферических координат так, чтобы меридианная плоскость $\varphi = 0$ проходила как через начальную, так и через конечную точку. Так как в обеих этих точках $\varphi = 0$, то, по теореме о среднем значении, в некоторой промежуточной точке производная $\dot{\varphi}=0$; отсюда и $C=0$. Но тогда $\dot{\varphi}=0$ для всего светового луча. Стало быть, весь луч должен лежать в плоскости $\varphi=0$.

§ 3, стр. 540.

Согласно закону сохранения энергии,

$$T+U=T=\frac{1}{2}m\left(\frac{ds}{dt}\right)^2=\text{const}=\frac{1}{2}mC^2,$$

откуда $\frac{ds}{dt}=C=\text{const}=$ начальной скорости.

Принцип Гамильтона устанавливает стационарность функционала

$$\int_{t_0}^{t_1}(T-U)\,dt=\int_{t_0}^{t_1}T\,dt=\frac{1}{2}mC^2\int_{t_0}^{t_1}dt=\frac{1}{2}mC\int_{s_0}^{s_1}ds.$$

Таким образом, из стационарности интеграла Гамильтона вытекает стационарность длины траектории.

Смешанные упражнения к гл. VII, стр. 542.

1. Выберем ось z параллельно образующим цилиндрической поверхности; тогда уравнение этой поверхности будет $G(x, y)=0$ и не содержит z. Из уравнений геодезических линий [стр. 540, конец пункта а)] вытекает, что $\frac{dz}{ds}=\text{const}$. Следовательно, геодезическая линия на цилиндрической поверхности образует постоянный угол с осью z, т. е. пересекает все образующие под одним и тем же углом.

2. а) $g(x)-\frac{y''}{\sqrt{(1+y'^2)^3}}=0$;

   б) $g(x)-\frac{6y''(y''^2+4y'y''')}{(1+y'^2)^4}+\frac{2y^{IV}}{(1+y'^2)^3}+\frac{48y'^2y''^3}{(1+y'^2)^5}=0$;

   в) $y+y''+y^{IV}=0$;

   г) $(2-y'^2)y''=0$.

3. а) $m\varphi=(a_x+b_y)\varphi_x+(b_x+c_y)\varphi_y+a\varphi_{xx}+2b\varphi_{xy}+c\varphi_{yy}$.

   б) $(\nabla^2)^2\varphi=0$; в) $(\nabla^2)^2\varphi=0$.

4. $\frac{au''+a'u'+u(b'-c)}{u}=\lambda=\text{const}$.

5. а) Уравнение Эйлера будет $f(x)+2\lambda u(x)=0$, откуда $u=-\frac{f(x)}{2\lambda}$. Подставив это выражение в добавочное условие

$$\int_0^1\varphi^2\,dx=k^2,$$

получим

$$\lambda=\pm\frac{1}{2k}\sqrt{\int_0^1 f^2\,dx};$$

следовательно,

$$u(x) = \pm \frac{kf(x)}{\sqrt{\int\limits_0^1 f^2\,dx}}.$$

б) Для любой непрерывной допустимой функции $\varphi(x)$ имеем

$$I = \int\limits_0^1 f\varphi\,dx \leqslant \sqrt{\int\limits_0^1 f^2\,dx} \cdot \sqrt{\int\limits_0^1 \varphi^2\,dx} = k\sqrt{\int\limits_0^1 f^2\,dx},$$

причем знак равенства будет как раз, если $\varphi = u$.

Глава VIII

§ 1, стр. 552.

1. Когда $\operatorname{Re} z = x \geqslant 0$.

2. Воспользоваться принципом сравнения рядов.

3. В разложении функции $\cos^2 z + \sin^2 z$ в степенной ряд коэффициент при z^n в случае $n > 0$ будет (см. т. I, стр. 98, упр. 12 *а*)

$$(-1)^{\frac{n}{2}} \sum_{\nu=0}^{n} \frac{(-1)^\nu}{\nu!\,(n-\nu)!} = \frac{(-1)^{\frac{n}{2}}}{n!} \sum_{\nu=0}^{n} (-1)^\nu \binom{n}{\nu} = 0.$$

4. Ряд сходится в том и только в том случае, если $|z| < 1$. Действительно, если $|z| = r < 1$, то

$$\left|\frac{z^k}{1-z^k}\right| \leqslant \frac{r^k}{1-r^k} \leqslant \frac{1}{1-r}\,r^k,$$

и сходимость данного ряда вытекает из сравнения ряда, составленного из модулей его членов, со сходящимся геометрическим рядом. Если $|z| > 1$, то $\frac{z^k}{1-z^k} \to -1$ при $k \to \infty$, и ряд не может сходиться, так как общий член a_k сходящегося ряда стремится к нулю при $k \to \infty$. Если же $|z| = 1$, то либо некоторые члены ряда вовсе не будут определены, либо по крайней мере его члены не будут ограничены, так как z^k может подойти к 1 как угодно близко.

§ 2, стр. 557.

Пусть $f(z) = u + iv$, $g(z) = u_1 + iv_1$; тогда, например, $fg = p + iq$, где $p = uu_1 - vv_1$, $q = uv_1 + vu_1$. Предположив, что u и v удовлетворяют условиям Коши — Римана, доказать, что p и q тоже удовлетворяют этим условиям.

§ 2, стр. 558.

1. Функции а), б), в) непрерывны на всей плоскости z; г) разрыв в точке $z = 0$.

2. Ни одна.

3. Имеем

$$|\zeta|^2 = \zeta\bar{\zeta} = \frac{\alpha\bar{\alpha}z\bar{z} + \beta\bar{\beta} + (\alpha\bar{\beta}z + \bar{\alpha}\beta\bar{z})}{\beta\bar{\beta}z\bar{z} + \alpha\bar{\alpha} + (\alpha\bar{\beta}z + \bar{\alpha}\beta\bar{z})}.$$

Отсюда видно, что если $\alpha\bar{\alpha} - \beta\bar{\beta} = 1$, то разность между числителем и знаменателем равна $z\bar{z}-1$, так что числитель больше знаменателя при $|z|>1$ и меньше его при $|z|<1$. Если $\beta\bar{\beta}-\alpha\bar{\alpha}=1$, то будет наоборот.

4. Полагая $z=re^{i\varphi}$, $\zeta=\xi+i\eta$, имеем

$$\xi=\frac{1}{2}\left(r+\frac{1}{r}\right)\cos\varphi, \qquad \eta=\frac{1}{2}\left(r-\frac{1}{r}\right)\sin\varphi.$$

Если $r=c=\text{const}$, то

$$\frac{\xi^2}{\frac{1}{4}\left(c+\frac{1}{c}\right)^2}+\frac{\eta^2}{\frac{1}{4}\left(c-\frac{1}{c}\right)^2}=1;$$

если же $\varphi=c=\text{const}$, то

$$\frac{\xi^2}{\cos^2 c}+\frac{\eta^2}{\cos^2 c-1}=1$$

(ср. упр. 7, стр. 176).

6. Заданный (произвольный) круг отобразить сперва на единичный круг с помощью функции вида $\zeta=az+b$, а затем выполнить преобразование $w=i\,\dfrac{1+\zeta}{1-\zeta}$.

7. Уравнение окружности или прямой в плоскости ζ имеет следующий вид:

$$\alpha\zeta\bar{\zeta}+\beta\zeta+\bar{\beta}\bar{\zeta}+\gamma=0,$$

где α и γ — действительные числа (для прямой $\alpha=0$). Если в это уравнение подставить выражение ζ через z, то для z получится уравнение такого же вида.

Неподвижные точки $\zeta=z$ определятся из квадратного уравнения

$$cz^2+(d-a)z-b=0;$$

если корни этого квадратного уравнения различны, то преобразование имеет две неподвижные точки. Как мы только что видели, окружность, проходящая через эти неподвижные точки, преобразуется в окружность, которая тоже должна проходить через неподвижные точки; семейство ортогональных окружностей тоже преобразуется само в себя, потому что окружности переходят в окружности, а наше преобразование является конформным.

§ 3, стр. 567.

2. Согласно т. I, стр. 443, ряд сходится абсолютно.

§ 4, стр. 572.

В формуле Коши дифференцировать последовательно под знаком интеграла и доказать правомерность этого процесса.

§ 4, стр. 573.

Так как функция $u(x, y)$ задана, то условия Коши — Римана вполне определяют частные производные v_x и v_y функции $v(x, y)$. Функция v, имеющая такие производные, действительно существует, ибо условие интегрируемости выполнено, в силу того, что u удовлетворяет уравнению $u_{xx}+u_{yy}=0$ (ср. стр. 378). Искомая функция v определяется однозначно, если не считать аддитивной постоянной c, и дается криволинейным интегралом

$$v(x, y) = \int_{(x_0, y_0)}^{(x, y)} (v_y\,dy + v_x\,dx) + c.$$

Из условий Коши — Римана вытекает, что v также удовлетворяет уравнению Лапласа.

§ 4, стр. 576.

1. Нетрудно убедиться, что

$$h(z) = \frac{1}{2\pi i}\oint \frac{f(\zeta)}{\zeta - z}\,\frac{z^n}{\zeta^n}\,d\zeta$$

является аналитической функцией от z. Дифференцируя под знаком интеграла и пользуясь правилом Лейбница (т. I, стр. 229), находим, что

$$h^{(\mu)}(z) = \frac{1}{2\pi i}\sum_{\nu=0}^{\mu}\binom{\mu}{\nu}\nu!\,n(n-1)\ldots(n-\mu+\nu+1)\int\frac{f(\zeta)}{(\zeta-z)^{\nu+1}}\,\frac{z^{n-\mu+\nu}}{\zeta^n}\,d\zeta =$$

$$= \frac{\mu!}{2\pi i}\sum_{\nu=0}^{\mu}\binom{n}{\mu-\nu}\int\frac{f(\zeta)}{(\zeta-z)^{\nu+1}}\,\frac{z^{n-\mu+\nu}}{\zeta^n}\,d\zeta.$$

Отличны от нуля только те члены, для которых $\mu-\nu \leqslant n$, ибо в противном случае $\binom{n}{\mu-\nu}$ обращается в нуль. С другой стороны, член, для которого $\mu-\nu < n$, обращается в нуль при $z=0$; если $\mu < n$, то других членов не будет, так что $h^{(\mu)}(0) = 0$. Если же $\mu \geqslant n$, то остается лишь член, для которого $\mu - \nu = n$, так что

$$h^{(\mu)}(0) = \frac{\mu!}{2\pi i}\oint\frac{f(\zeta)}{(\zeta-z)^{\mu+1}}\,d\zeta = f^{(\mu)}(0).$$

2. $|a_k| = \left|\frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta^{k+1}}\,d\zeta\right| \leqslant \frac{1}{2\pi}\,\frac{M}{\rho^{k+1}}\,2\pi\rho$, где интеграл берется по окружности C: $|z| = \rho$.

3. $\oint_C \frac{f'(z)}{f(z)}\,dz$ равен сумме вычетов функции $\frac{f'}{f}$ относительно всех полюсов, лежащих внутри C. Если $f(z)$ имеет нуль n-го порядка в точке z_0, то

$$f(z) = (z - z_0)^n\,\varphi(z),$$

причем $\varphi(z_0) \neq 0$. Поэтому

$$\frac{f'(z)}{f(z)} = \frac{n\varphi(z) + (z - z_0)\,\varphi'(z)}{(z - z_0)\,\varphi(z)},$$

так что вычет функции $\frac{f'(z)}{f(z)}$ в ее полюсе z_0 равен $2\pi i n$.

4. а) Согласно упр. 3, число корней уравнения $P(z)+\theta Q(z)=0$ внутри C равно

$$\frac{1}{2\pi i}\oint_C \frac{P'(z)+\theta Q'(z)}{P(z)+\theta Q(z)}\,dz.$$

Знаменатель не обращается в нуль ни в какой точке кривой C при всяком значении θ, удовлетворяющем неравенству $0\leqslant\theta\leqslant 1$; поэтому наш интеграл является непрерывной функцией от θ, а так как он всегда равен целому числу, то он имеет постоянное значение, а стало быть, одно и то же значение при $\theta=0$ и при $\theta=1$.

б) Если

$$|a|<r^4-\frac{1}{r},$$

то

$$r>1;$$

поэтому уравнение

$$z^5+1=0$$

имеет пять корней внутри окружности $|z|=r$. Полагая

$$P(z)=z^5+1,\quad Q(z)=az,$$

имеем на окружности $|z|=r$

$$|Q(z)|=|a|\,r<r^5-1<|z^5+1|=|P(z)|,$$

т. е.

$$|Q(z)|<|P(z)|.$$

5. Ср. доказательство упр. 3.

§ 5, стр. 582.

2. Левая часть формулы есть сумма вычетов функции $\frac{z^m}{f(z)}$, деленная на $2\pi i$; поэтому она равна $\frac{1}{2\pi i}\oint\frac{z^m}{f(z)}\,dz$, где интеграл берется по окружности с центром в начале координат, содержащей внутри себя все корни α_k. Но этот интеграл стремится к нулю, когда радиус окружности стремится к бесконечности (причем центр ее остается неизменным).

Смешанные упражнения к гл. VIII, стр. 589.

1. $\frac{z_2-z_1}{z_3-z_1}$ должно быть действительным числом.

2. $\Delta=\frac{z_1-z_3}{z_2-z_3}:\frac{z_4-z_1}{z_2-z_4}$ должно быть действительным числом. В самом деле, окружность, проходящую через точки z_1, z_2 и z_3, можно отобразить на действительную ось с помощью линейного преобразования вида

$$\zeta=\frac{\alpha z+\beta}{\gamma z+\delta}$$

(ср. упр. 6, стр. 559); согласно упр. 5, стр. 559, при этом Δ не изменяется. Но тогда для того, чтобы точка z_4 лежала на той же окружности, что и точки z_1, z_2 и z_3, изображение точки z_4 должно лежать на действительной оси, а это эквивалентно условию, что Δ — действительное число.

3. Задача состоит в доказательстве тождества

$$|z_2-z_1|\cdot|z_4-z_3|+|z_3-z_2|\cdot|z_4-z_1|=|z_3-z_1|\cdot|z_4-z_2|$$

или

$$1+\frac{|z_2-z_1|\cdot|z_4-z_3|}{|z_3-z_2|\cdot|z_4-z_1|}=\frac{|z_3-z_1|\cdot|z_4-z_2|}{|z_3-z_2|\cdot|z_4-z_1|}.$$

Но обе дроби инвариантны относительно дробно-линейного преобразования (ср. упр. 5 и 6, стр. 559). Если при помощи подходящего дробно-линейного преобразования отобразить данную окружность на действительную ось, то придется доказать тождество $AB\cdot CD+BC\cdot AD=AC\cdot BD$ для четырех точек действительной оси, а это уже элементарно просто.

4. Функция $\zeta=e^{iz}$ принимает любое значение, кроме $\zeta=0$, что нетрудно вывести из соотношения $e^{iz}=e^{-y}(\cos x+i\sin x)$. Нашей задачей является выбрать ζ таким образом, чтобы было

$$\cos z=\frac{1}{2}\left(\zeta+\frac{1}{\zeta}\right)=c,$$

но это приводится к квадратному уравнению, которое всегда имеет решение

$$\zeta=c\pm\sqrt{c^2-1},$$

и это решение не равно нулю, так что требуемое значение z существует.

5. Ср. упр. 4. Если $\zeta=e^{iz}$, то

$$\operatorname{tg} z=\frac{1}{i}\,\frac{\zeta-\frac{1}{\zeta}}{\zeta+\frac{1}{\zeta}}=c, \qquad \text{откуда} \qquad \zeta=\sqrt{\frac{1+ic}{1-ic}}.$$

Стало быть, конечное $\zeta\neq 0$ получается лишь при $c\neq\pm i$. Следовательно, уравнение $\operatorname{tg} z=c$ имеет решение лишь в том случае, если c не равно ни $+i$ ни $-i$.

6. $\cos z=\cos(x+iy)=\cos x\cos iy-\sin x\sin iy=\cos x\operatorname{ch} y-i\sin x\operatorname{sh} y$, $\sin z=\sin(x+iy)=\sin x\cos iy+\cos x\sin iy=\sin x\operatorname{ch} y+i\cos x\operatorname{sh} y$.

Отсюда видно, что $\cos z$ имеет действительное значение, если $x=n\pi$ или $y=0$, а $\sin z$ имеет действительное значение, если $x=\frac{\pi}{2}+n\pi$ или $y=0$ (где n — любое целое число).

7. а) $r=1$ (при $|z|>1$ общий член стремится к ∞, при $|z|<1$ сравнить с геометрическим рядом);

б) $r=0$; в) $r=1$.

8. Ср. т. I, стр. 204—205.

9. а) Интегрировать $\frac{e^{iz}}{1+z^4}$ вдоль верхней полуокружности.

Отв. $\frac{\pi\sqrt{2}}{4}e^{-\frac{\sqrt{2}}{2}}\left(\sin\frac{\sqrt{2}}{2}+\cos\frac{\sqrt{2}}{2}\right)$.

б) Интегрировать $\frac{z^2e^{iz}}{1+z^4}$ вдоль верхней полуокружности.

Отв. $\frac{\pi\sqrt{2}}{4}e^{-\frac{\sqrt{2}}{2}}\left(\cos\frac{\sqrt{2}}{2}-\sin\frac{\sqrt{2}}{2}\right)$.

в) Интегрировать $\frac{e^{iz}}{c^2+z^2}$ вдоль верхней полуокружности.

Отв. $\frac{\pi}{2c}e^{-c}$.

г) Интегрировать $\frac{z^{a-1}}{(z+1)(z+2)}$ вдоль контура области, ограниченной достаточно большой окружностью с центром в начале координат и разрезанной вдоль положительной вещественной полуоси.

Отв. $\frac{\pi(2^{\alpha-1}-1)}{\sin\pi\alpha}$.

10. а) Вычет $+2\pi i$ относительно полюсов $z=2n\pi$, вычет $-2\pi i$ относительно полюсов $z=(2n+1)\pi$.

б) Вычет $+2\pi i$ относительно полюсов $z=\frac{3\pi}{2}+2n\pi$, вычет $-2\pi i$ относительно полюсов $z=\frac{\pi}{2}+2n\pi$.

в) Воспользоваться функциональным уравнением

$$\Gamma(z)=\frac{1}{z(z+1)\dots(z+k)}\Gamma(z+k+1).$$

Отв. Вычет $\frac{(-1)^n}{n!}2\pi i$ относительно полюсов $z=-n$.

г) Вычет $2\pi i$ относительно полюсов $z=n\pi i$.

11. Подынтегральную функцию представим в виде

$$\frac{\operatorname{ctg}\pi t}{t-z}=\frac{\operatorname{ctg}\pi t}{t}+\frac{z\operatorname{ctg}\pi t}{t(t-z)};$$

$\operatorname{ctg}\pi t$ ограничен на сторонах квадратов C_n, а интегралы от $\frac{\operatorname{ctg}\pi t}{t}$ по противоположным сторонам квадрата почти компенсируют друг друга (уточнить!). Поэтому

$$\lim_{n\to\infty}\oint_{C_n}\frac{\operatorname{ctg}\pi t}{t-z}\,dt=\lim_{n\to\infty}\oint_{C_n}\frac{z\operatorname{ctg}\pi t}{t(t-z)}\,dt=0.$$

При составлении суммы всех вычетов мы сначала сложим попарно вычеты в полюсах, симметричных относительно начала; получится сходящийся бесконечный ряд, сумма которого равна нулю, а из этого равенства вытекает

$$\operatorname{ctg}\pi z=\frac{2z}{\pi}\left(\frac{1}{2z^2}+\frac{1}{z^2-1^2}+\frac{1}{z^2-2^2}+\dots\right)$$

(ср. т. I, стр. 518).

12.

$$\frac{1}{1+t}=1-t+t^2-\dots+(-1)^{n-1}t^{n-1}+(-1)^n\frac{t^n}{1+t}.$$

Интегрируем от 0 до z:

$$\ln(1+z)=z-\frac{z^2}{2}+\frac{z^3}{3}-\dots+(-1)^{n-1}\frac{z^n}{n}+R_n,$$

где

$$R_n=(-1)^n\int_0^z\frac{t^n}{1+t}\,dt.$$

На окружности $|z|=1$ положим $z=e^{i\theta}$ и за путь интегрирования примем отрезок прямой от начала до точки $e^{i\theta}$; тогда $t=se^{i\theta}$, $dt=e^{i\theta}ds$. При $e^{i\theta}\neq-1$ имеем

$$|R_n|=\left|\int_0^z \frac{t^n}{1+t}\,dt\right|\leqslant\int_0^1 \frac{s^n}{|1+se^{i\theta}|}\,ds\leqslant\frac{1}{m}\int_0^1 s^n\,ds=\frac{1}{m(n+1)},$$

где m обозначает наименьшее значение знаменателя $|1+se^{i\theta}|$ при $0\leqslant s\leqslant 1$. Отсюда видно, что если $z=e^{i\theta}\neq-1$, то $R_n\to 0$.

13. а) $f(z)=\sum\limits_{k=1}^{\infty}\left(\frac{1}{(2k-1)^z}-\frac{1}{(2k)^z}\right)$; модуль общего члена этого ряда

$$\left|\frac{1}{(2k-1)^z}-\frac{1}{(2k)^z}\right|=\left|z\int_{2k-1}^{2k}\frac{1}{t^{z+1}}\,dt\right|\leqslant\frac{|z|}{|(2k-1)^{z+1}|}=\frac{|z|}{(2k-1)^{1+x}},$$

а ряд $\sum\limits_{k=1}^{\infty}\frac{1}{(2k-1)^{1+x}}$ сходится абсолютно при $x>0$.

б) $$(1-2^{1-z})\,\zeta(z)=1+\frac{1}{2^z}+\frac{1}{3^z}+\frac{1}{4^z}+\ldots-\frac{2}{2^z}-\frac{2}{4^z}-\frac{2}{6^z}-\ldots=$$
$$=1-\frac{1}{2^z}+\frac{1}{3^z}-\frac{1}{4^z}+\ldots=f(z).$$

в) $\lim\limits_{z\to 1}(z-1)\,\zeta(z)=f(1)\cdot\lim\limits_{z\to 1}\frac{z-1}{1-2^{1-z}}=\frac{f(1)}{g'(1)}=1,$ где $g(z)=1-2^{1-z}$.

14. а) Значение интеграла на четверти окружности радиуса ε стремится к нулю при $\varepsilon\to 0$. На части контура, идущей вдоль оси x, имеем $z=x$, $dz=dx$; на единичной окружности $|z|=1$ положим $z=e^{i\theta}$, $dz=ie^{i\theta}\,d\theta$; на оси y имеем $z=iy$, $dz=i\,dy$. Тогда теорема Коши дает

$$0=\int_0^1\left(x+\frac{1}{x}\right)^m x^{n-1}dx+i\int_0^{\frac{\pi}{2}}(e^{i\theta}+e^{-i\theta})^m e^{i\theta n}\,d\theta-$$
$$-i\int_0^1\left(iy+\frac{1}{iy}\right)^m (iy)^{n-1}\,dy=$$
$$=\int_0^1\left(x+\frac{1}{x}\right)^m x^{n-1}\,dx+i\cdot 2^m\int_0^{\frac{\pi}{2}}\cos^m\theta e^{in\theta}\,d\theta-$$
$$-e^{\frac{i\pi(n-m)}{2}}\int_0^1\left(-y+\frac{1}{y}\right)^m y^{n-1}\,dy.$$

Приравнивая мнимые части, получим

$$2^m \int_0^{\frac{\pi}{2}} \cos^m\theta \cos n\theta d\theta = \sin \frac{\pi(n-m)}{2} \int_0^1 \left(-y+\frac{1}{y}\right)^m y^{n-1} dy =$$

$$= \frac{1}{2} \sin \frac{\pi(n-m)}{2} \int_0^1 (1-\eta)^m \eta^{\frac{n-m-2}{2}} d\eta =$$

$$= \frac{1}{2} \sin \frac{\pi}{2}(n-m) \, \mathrm{B}\left(m+1, \frac{n-m}{2}\right) \quad (\text{ср. стр. } 358),$$

где $y^2 = \eta$, $dy = \dfrac{d\eta}{2\sqrt{\eta}}$.

б) Воспользоваться соотношением

$$\sin \frac{(n-m)\pi}{2} \Gamma\left(\frac{n-m}{2}\right) = \frac{\pi}{\Gamma\left(1-\frac{n-m}{2}\right)} \quad (\text{ср. стр. } 357).$$

15. Если $x \neq 0$ и если C' есть контур в той области, в которой $f(z)$ регулярна, который содержит y и не содержит начала координат, то, согласно упражнению на стр. 572,

$$\frac{d^n}{dy^n} \frac{yf(y)}{(y+a)^{n+1}} = \frac{n!}{2\pi i} \oint_{C'} \frac{tf(t)}{(t+a)^{n+1}(t-y)^{n+1}} dt.$$

Если подставить в интеграл $a = y = \sqrt{x}$, то получится

$$\frac{n!}{2\pi i} \oint_{C'} \frac{tf(t)}{(t^2-x)^{n+1}} dt,$$

а после замены переменной $t^2 = \tau$, $2t\,dt = d\tau$ интеграл преобразуется к виду

$$\frac{n!}{4\pi i} \oint_{C} \frac{f(\sqrt{\tau})}{(\tau-x)^{n+1}} d\tau,$$

где C есть контур, содержащий x, но не 0. Этот интеграл равен

$$\frac{1}{2} \frac{d^n}{dx^n} f(\sqrt{x}).$$

16. а) $|\operatorname{sh}(x+iy)|^2 = \left(\dfrac{e^{x+iy}-e^{-x-iy}}{2}\right)\left(\dfrac{e^{x-iy}-e^{-x+iy}}{2}\right) =$
$= \dfrac{1}{2}(\operatorname{ch} 2x - \cos 2y) \geq \dfrac{1}{2}(\operatorname{ch} 2x - 1) = A(x).$

б) Интегрировать эту функцию вдоль периметра квадрата со сторонами $x = \pm\pi\left(n+\frac{1}{2}\right)$ и $y = \pm\pi\left(n+\frac{1}{2}\right)$, где n — целое число. При $n \to \infty$ интеграл стремится к нулю; следовательно, сумма вычетов стремится к нулю.

17. Пусть притягивающие точки будут $(x_1, y_1), \ldots, (x_n, y_n)$. Тогда равнодействующая сила притяжения в точке (x, y) поля имеет проекции на координатные оси:

$$X = -\sum_{k=1}^{n} \frac{x - x_k}{(x - x_k)^2 + (y - y_k)^2}, \quad Y = -\sum_{k=1}^{n} \frac{y - y_k}{(x - x_k)^2 + (y - y_k)^2}.$$

(Коэффициент пропорциональности берем для простоты равным единице.) Введем комплексные величины $z_1 = x_1 + iy_1, \ldots, z_n = x_n + iy_n$, $z = x + iy$, $Z = X + iY$. Тогда

$$Z = -\sum_{k=1}^{n} \frac{1}{\bar{z} - \bar{z}_k} = -\overline{\frac{f'(z)}{f(z)}},$$

где $f(z)$ обозначает многочлен $(z - z_1)(z - z_2) \ldots (z - z_n)$, а черта над буквой или выражением обозначает комплексно-сопряженную величину. Положения равновесия определяются условием $Z = 0$, т. е. нулями многочлена $f'(z)$, а таких нулей может быть не более чем $n - 1$.

В предложенном частном случае четырех точек три положения равновесия: $(0, 0)$, $(\sqrt{a^2 - b^2}, 0)$, $(-\sqrt{a^2 - b^2}, 0)$.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

Р. Курант

Курс дифференциального
и интегрального исчисления

М., 1970 г., 672 стр. с илл.

Редактор *А. Ф. Лапко*

Техн. редактор *И. Ш. Аксельрод*

Корректоры *Н. Б. Румянцева* и
Г. С. Смоликова

Сдано в набор 16/VI 1969 г. Подписано
к печати 31/X 1969 г. Бумага $60 \times 90^1/_{16}$. Физ.
печ. л. 42. Условн. печ. л. 42. Уч.-изд. л. 44,8.
Тираж 50 000 экз.
Цена книги 1 р. 67 к. Заказ № 607.

Издательство «Наука»
Главная редакция
физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Ордена Трудового Красного Знамени
Ленинградская типография № 1 «Печатный
Двор» имени А. М. Горького
Главполиграфпрома Комитета по печати
при Совете Министров СССР,
Ленинград, Гатчинская, 26.