АКАДЕМИЯ НАУК СССР

МАТЕМАТИЧЕСКИЙ СБОРНИК

Т. 84 (126)

(ОТДЕЛЬНЫЙ ОТТИСК)

4

МОСКВА · 1971

УДК 517.432+517.862

Уравнения Винера — Хопфа в четверти плоскости, дискретные группы и автоморфные функции

В. А. Малышев (Москва)

§ 1. Введение

В настоящей работе рассматриваются уравнения вида

$$\sum_{k,\,l=0}^{\infty} a_{i-k,j-l}\xi_{kl} = \eta_{ij}, \quad i, j = 0, 1, 2, \ldots . \tag{1.1}$$

При этом предполагается, что $\sum\limits_{i,\,j=0}^{\infty} |\eta_{ij}| < \infty$, $\sum\limits_{p,q=-\infty}^{\infty} |a_{pq}| < \infty$, и ищется решение, также принадлежащее пространству l_1 последовательностей $\xi = \{\xi_{kl}\}_{k,\,l=0}^{\infty}$. Рассмотрим оператор A в банаховом пространстве l_1: $A\xi=\xi'$, где

$$\xi'_{ij} = \sum_{k,l=0}^{\infty}{}' a_{i-k,\,j-l}\xi_{kl}.$$

Следующее утверждение выделяет случаи, когда A является оператором Нётера ($\dim \operatorname{Ker} A < \infty$, $\dim \operatorname{Coker} A < \infty$).

Теорема 1.1 (Симоненко И. Б. [1], [2]). *Оператор A является оператором Нётера тогда и только тогда, когда выполняются следующие два условия:*

$$a(x, y) = \sum_{p,q=-\infty}^{\infty} a_{pq}x^p y^q \neq 0 \ \text{при} \ |x| = |y| = 1, \tag{1.2}$$

$$\operatorname*{ind}_{|x|=1} a(x, 1) = \operatorname*{ind}_{|y|=1} a(1, y) = 0. \tag{1.3}$$

Стандартным приемом доказывается следующее утверждение (см. [3]).

Лемма 1.1. *В условиях теоремы* 1.1 $\operatorname{ind} A=0$.

Методы доказательства этих двух утверждений не могут дать условий обратимости оператора, а тем более аналитических свойств решения и его явного представления.

Здесь будет дано исследование возможности получения явного представления решения подобного класса уравнений.

Теория в нётеровском случае для уравнений, инвариантных относительно сдвига, во всем пространстве и на полупрямой является классической (см. [4], [5]), а в полупространстве решение строится также с помощью метода Винера — Хопфа ([6]). Во всех этих случаях имеется компактная явная формула для решения. Подобная формула, основанная на идеях Винера — Хопфа, для уравнений в четверти плоскости может быть получена лишь в очень частных случаях (см. [15], [16]). Ситуация для общего случая резко усложняется. Возникают классы разрешимости уравнений тем или иным методом. В настоящей работе излагается новый подход к исследованию подобных уравнений и для одного класса строятся в некотором смысле полные основы их теории.

Мы будем говорить, что система уравнений (1.1) имеет тип $(n_1, n_2; m_1, m_2)$, где $-\infty \leqslant n_1, m_1 < \infty, -\infty < n_2, m_2 \leqslant \infty$, если a_{pq} могут быть отличны от нуля лишь при

$$n_1 \leqslant p \leqslant n_2, \quad m_1 \leqslant q \leqslant m_2. \tag{1.4}$$

Мы будем говорить, что тип $(n_1, n_2; m_1, m_2)$ точный, если область типа (1.4) невозможно сузить.

Замечание 1.1. Существуют уравнения, удовлетворяющие условиям (1.2) и (1.3), произвольного точного типа $(n_1, n_2; m_1, m_2)$. Например, можно положить $a(x, y) = a(x)\tilde{a}(y)$, где $x^{n_1} a(x)$ — многочлен от x степени n_1+n_2, имеющий n_1 нулей внутри единичного круга и n_2 вне его. Аналогично строится $y^{m_1}\tilde{a}(y)$, и для получения нетривиальных примеров достаточно пошевелить коэффициенты $a(x, y)$.

Если уравнение (1.1) имеет один из следующих типов: $(0, \infty; -\infty, \infty)$, $(-\infty, \infty; 0, \infty)$, $(-\infty, 0; -\infty, \infty)$, $(-\infty, \infty; -\infty, 0)$,— то оно разрешимо методом Винера — Хопфа, что легко следует из работы [15]. В данной работе рассматриваются уравнения типа $(-1, \infty; -1, \infty)$.

В случае уравнений (1.1) метод факторизации Винера — Хопфа, на применении которого по существу все кончается в одномерном случае, составляет необходимый предварительный этап исследования (§ 2). В § 3 доказывается обратимость соответствующего оператора в l_1. Первый шаг нового метода (перенос на риманову поверхность) рассматривается в §§ 4, 5. В случае уравнений типа $(-1; 1; -1, 1)$ род соответствующей римановой поверхности равен 0 или 1. Уравнения этого типа подробно исследуются в §§ 6—8. Основная конструкция для рода $g \geqslant 2$ проводится в § 9. Она связана с построением автоморфизмов Галуа на универсальной накрывающей и последующим преобразованием уравнений на универсальной накрывающей с их помощью. Здесь существенно используется техника дискретных групп движений плоскости Лобачевского. В § 10 выясняется аналитическое поведение решения (строится риманова область существования символа решения). В §§ 11, 12 решение уравнений сводится к задаче Карлемана на римановой поверхности, откуда и получается интегральное представление решения.

Некоторые важные черты метода сохраняются в случае произвольного рационального символа.

§ 2. Факторизация

Рассмотрим кольцо R рядов вида $r(x, y) = \sum_{i,j=-\infty}^{\infty} r_{ij}x^i y^j$, абсолютно сходящихся при $|x| = |y| = 1$. Мы будем рассматривать следующие подкольца этого кольца: $R_{++} = R_{\substack{++\\xy}}$ — множество функций $r(x, y)$ с $r_{ij} = 0$, если либо

$i<0$, либо $j<0$; R_{+-} — множество функций с $r_{ij}=0$, если либо $i<0$, либо $j\geqslant 0$; R_{-+}, $R_{\underset{x}{+}}$ и т. д. Введем операторы проектирования на эти подкольца: P_{++}, P_{+-}, $P_{\bar{y}}$, ... Например,

$$P_{\bar{y}}\left[\sum_{i,j=-\infty}^{\infty} r_{ij}x^iy^j\right]=\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{-1} r_{ij}x^iy^j.$$

Лемма 2.1. *Если* $r(x,y)\in R$, $r(x,y)\neq 0$ *при* $|x|=|y|=1$ *и*

$$\operatorname*{ind}_{|x|=1} r(x,1)=\operatorname*{ind}_{|y|=1} r(1,y)=0, \tag{2.1}$$

то $\ln r(x,y)\in R$.

Доказательство. Условия (2.1) гарантируют возможность выделения однозначной ветви $\ln r(x,y)$. Далее, если, например, $r(x,y)$ — кусочно гладкая функция, то доказательство проводится элементарно. В полной общности надо воспользоваться обобщенной теоремой Винера о локально аналитических функциях на пространстве максимальных идеалов кольца R (см. [7]).

Теперь можно положить

$$\ln a(x,y)=\sum_{i,j=-\infty}^{\infty} b_{ij}x^iy^j,$$

$$a_{\underset{x}{+}}(x,y)=\exp[P_{\underset{x}{+}}\ln a(x,y)],$$

$$a_{\bar{x}}(x,y)=\frac{a(x,y)}{a_{\underset{x}{+}}(x,y)},$$

. . . .

Пусть теперь уравнение (1.1) имеет точный тип $(-1,\infty;-1,\infty)$, причем выполняются условия теоремы 1.1. Положим $A(x,y)=xya(x,y)$. Тогда для любого y, $|y|=1$, имеем $\operatorname*{ind}_{|x|=1} A(x,y)=1$, и, аналогично, для любого x, $|x|=1$, $\operatorname*{ind}_{|y|=1} A(x,y)=1$. Поэтому для любого y, $|y|=1$, существует единственный нуль функции $A(x,y)$ внутри единичного круга, обозначаемый далее через $x_0(y)$. Он имеет первый порядок. Аналогично вводится $y_0(x)$.

Если $a(x,y)$ имеет тип $(-1,n;-1,m)$, где $n,m<\infty$, то $A(x,y)$ — многочлен. Пусть $A=A_1A_2\ldots A_k$ — его разложение на простые множители в кольце $\mathbf{C}[x,y]$.

Пусть, например, $A_1(x_0(y),y)=0$. Тогда из приведенных выше соображений следует, что $A_i(x_0(y),y)\neq 0$ при $i=1$ и $|y|=1$. При этом возможны два случая:

1) (далее называемый случаем неприводимости) для любого x $(|x|)=1)$

$$A_1(x,y_0(x))=0;$$

2) (случай приводимости) для любого x $(|x|=1)$ $A_1(x,y_0(x))\neq 0$. В этом случае мы будем предполагать, что $A_2(x,y_0(x))=0$.

Компоненты факторизации $a(x, y)$ имеют вид:

$$a_{\bar{x}}(x, y) = 1 - \frac{x_0(y)}{x}; \quad a_{\bar{y}}(x, y) = 1 - \frac{y_0(x)}{y}.$$

§ 3. Обратимость

Рассмотрим следующие функции-символы ($|x| = |y| = 1$):

$$a(x, y) = \sum_{p,q=-\infty}^{\infty} a_{pq} x^p y^q, \quad \eta(x, y) = \sum_{i,j=0}^{\infty} \eta_{ij} x^i y^j,$$

$$\xi(x, y) = \sum_{k,l=0}^{\infty} \xi_{kl} x^k y^l.$$

Умножая соотношения (1.1) на $x^i y^j$ и суммируя, получим

$$\eta(x, y) = \sum_{i,j=0}^{\infty} \sum_{k,l=0}^{\infty} a_{i-k,j-l} x^{i-k} y^{j-l} \xi_{kl} x^k y^l = \sum_{p,q} \sum a_{pq} x^p y^q \xi_{kl} x^k y^l,$$

где $p = i - k$, $q = j - l$, и вторая сумма берется по всем неотрицательным k и l при условиях $p + k \geqslant 0$, $q + l \geqslant 0$.

Полагая

$$b_{--}(x, y) = a_{-1,-1} \xi_{00} \frac{1}{xy},$$

$$b_{+-}(x, y) = \sum_{k=0}^{\infty} \xi_{k0} x^k \sum_{p \geqslant -k} a_{p,-1} x^p \frac{1}{y},$$

$$b_{-+}(x, y) = \sum_{l=0}^{\infty} \xi_{0l} y^l \sum_{q \geqslant -l} a_{-1,q} y^q \frac{1}{x},$$

получим

$$\eta(x, y) = a(x, y)\xi(x, y) - b_{+-}(x, y) - b_{-+}(x, y) - b_{--}(x, y). \tag{3.1}$$

Вводя функции

$$\pi(x) = \sum_{k=0}^{\infty} \xi_{k0} x^k \sum_{p \geqslant -k} a_{p,-1} x^{p+1} + a_{-1,-1} \xi_{00},$$

$$\tilde{\pi}(y) = \sum_{l=0}^{\infty} \xi_{0l} y^l \sum_{q \geqslant -l} a_{-1,q} y^{q+1},$$

перепишем уравнение (3.1) следующим образом:

$$A(x, y)\xi(x, y) - \pi(x) - \tilde{\pi}(y) = xy\eta(x, y). \tag{3.2}$$

Заметим, что все слагаемые в этом уравнении принадлежат R_{++}.

Замечание 3.1. Если известна функция $\pi(x) = \sum_{k=0}^{\infty} \pi_k x^k$, то ξ_{k0} получаются рекуррентным образом из системы уравнений

$$\pi_0 = a_{-1,-1}\xi_{00},$$
$$\pi_1' = a_{0,-1}\xi_{00} + a_{-1,-1}\xi_{10},$$
$$\dots\dots\dots\dots\dots$$

Аналогично получаются ξ_{0l}, если известна функция $\widetilde{\pi}(y)$.

Теорема 3.1. *Оператор A, соответствующий уравнению* (1.1) *точного типа* $(-1, n; -1, m)$, *где* $1 \leqslant n$, $m \leqslant \infty$, *в условиях теоремы* 1.1 *является обратимым в пространстве* l_1.

В силу леммы 1.1 достаточно доказать, что $\operatorname{Ker} A = 0$. В § 2 для любого y, $|y|=1$, определена функция $x_0(y)$ такая, что $|x_0(y)|<1$ и $a(x_0(y), y)=0$. Тогда из (3.2) следует, что

$$|\pi(x_0(y))| = |\widetilde{\pi}(y)|. \tag{3.3}$$

Аналогично

$$|\pi(x)| = |\widetilde{\pi}(y_0(x))|. \tag{3.4}$$

Пусть, например,

$$M = \max_{|x|=1} |\pi(x)| \geqslant \max_{|y|=1} |\widetilde{\pi}(y)| = \widetilde{M}.$$

Тогда из принципа максимума модуля и соотношения (3.4) следует, что, во-первых, $M = \widetilde{M}$, и, во-вторых, $\widetilde{\pi}(y) \equiv M$. Отсюда и $\pi(x) \equiv M$. Но из явного вида (3.1) для $\widetilde{\pi}(y)$ следует тогда, что $M=0$.

Следствие 3.1. *Если операторы A и A_n, $n = 1, 2, \dots$, имеют тип* $(-1, \infty; -1, \infty)$, *причем для их символов* $a(x, y)$, $a_n(x, y)$

$$\lim_{n\to\infty} \max_{|x|=|y|=1} |a(x, y) - a_n(x, y)| = 0,$$

то $\|A_n^{-1} - A^{-1}\| \to 0$. *При этом, если* $\|A - A_n\| \cdot \|A\| < 1$, *то*

$$\|A_n^{-1} - A^{-1}\| \leqslant \frac{\|A - A_n\|}{1 - \|A - A_n\| \cdot \|A\|}.$$

Замечание. Отсюда следует, что решения уравнения с нерациональным ядром можно получать, используя уравнения с рациональными ядрами.

§ 4. Операция проектирования на аналитическое множество

Далее до конца работы будет предполагаться, что уравнение (1.1) имеет произвольный точный тип $(-1, n; -1, m)$ с $1 \leqslant n$, $m < \infty$. Положим $A = \{(x, y) : |x|<1,\ |y|<1,\ A(x, y)=0\}$, D — внутренность единичного круга и Γ — его граница. Пусть $\overline{A}$ — замыкание главного аналитического множества A, лежащего в $D \times D$, т. е.

$$\overline{A} = \{(x, y) : |x| \leqslant 1, |y| \leqslant 1, A(x, y) = 0\}.$$

Тогда ясно, что решение уравнения (3.2) должно удовлетворять условию

$$xy\eta(x, y)+\pi(x)+\widetilde{\pi}(y)=0 \tag{4.1}$$

при $(x, y)\in\overline{A}$ и, в частности, на границе $\overline{A}\cap[(D\times\Gamma)\cup(\Gamma\times D)]$. Докажем обратное утверждение.

Лемма 4.1. *Если существуют непрерывные в единичном круге $\overline{D}$ и аналитические внутри него функции $\pi(x)$ и $\widetilde{\pi}(y)$, удовлетворяющие соотношению (4.1) на $\overline{A}\cap[(D\times\Gamma)\cup(\Gamma\times D)]$, то решение уравнения (3.2) определится по формуле*

$$\xi(x, y)=\frac{\pi(x)+\widetilde{\pi}(y)+xy\eta(x, y)}{xya(x, y)}. \tag{4.2}$$

Доказательство (другой метод см. в работе [14]). Вводя функции $\pi_1(x)=\frac{\pi(x)-a_{-1,-1}\xi_{00}}{x}$ и $\widetilde{\pi}_1(y)=\frac{\widetilde{\pi}(y)}{y}$, перепишем уравнение (3.2) так:

$$a(x, y)\xi(x, y)=\pi_1(x)\frac{1}{y}+\widetilde{\pi}_1(y)\frac{1}{x}+\frac{a_{-1,-1}\xi_{00}}{xy}+\eta(x, y)=0. \tag{4.3}$$

Отсюда

$$a_{\underset{x}{+}}(x, y)\xi(x, y)=\frac{1}{a_{\underset{x}{-}}(x, y)}\left[\frac{\pi_1(x)}{y}+\frac{\widetilde{\pi}_1(y)}{x}+\frac{a_{-1,-1}\xi_{00}}{xy}+\eta(x, y)\right] \tag{4.4}$$

и

$$0=P_{\overline{x}}\left[\frac{1}{a_{\underset{x}{-}}(x, y)}\left(\frac{\pi_1(x)}{y}+\frac{\widetilde{\pi}_1(y)}{x}+\frac{a_{-1,-1}\xi_{00}}{xy}+\eta(x, y)\right)\right]. \tag{4.5}$$

Аналогично получим

$$0=P_{\overline{y}}\left[\frac{1}{a_{\underset{y}{-}}(x, y)}\left(\frac{\pi_1(x)}{y}+\frac{\widetilde{\pi}_1(y)}{x}+\frac{a_{-1,-1}\xi_{00}}{xy}+\eta(x, y)\right)\right]. \tag{4.6}$$

Обратно, если выполнены условия (4.5) и (4.6), то функция в правой части формулы (4.4) принадлежит кольцу $R_{\underset{x}{+}}$. Обозначив ее $\psi(x, y)$ и положив $\xi(x, y)=\frac{\psi(x, y)}{a_{\underset{x}{+}}(x, y)}$, имеем формулу (4.2). Нам остается показать только, что операции проектирования в формулах (4.5) и (4.6) эквивалентны операции проектирования на аналитическое множество A.

Лемма 4.2. *Соотношение (4.1) выполняется тогда и только тогда, когда выполняются соотношения (4.5) и (4.6).*

Докажем сначала формулу

$$P_{\overline{y}}\left[\frac{1}{1-\frac{y_0(x)}{y}}\omega(y)\right]=\omega(y_0(x))\frac{\frac{y_0(x)}{y}}{1-\frac{y_0(x)}{y}}, \tag{4.7}$$

где $\omega(y)\in R_{+\atop y}$. Действительно, для любого x, $|x|=1$,

$$P_{\bar{y}}\left[\frac{1}{1-\frac{y_0(x)}{y}}y^i\right]=P_{\bar{y}}\left[\sum_{k=0}^{\infty}\left(\frac{y_0(x)}{y}\right)^k y^i\right]=y_0^i\sum_{k=1}^{\infty}\left(\frac{y_0(x)}{y}\right)^k.$$

Воспользовавшись теперь линейностью оператора $P_{\bar{y}}$, получим формулу (4.7). Рассматривая теперь, например, формулу (4.6), имеем

$$\frac{\pi_1(x)}{y}\cdot\frac{1}{1-\frac{y_0(x)}{y}}+\frac{\tilde{\pi}_1(y_0(x))}{x}\cdot\frac{\frac{y_0(x)}{y}}{1-\frac{y_0(x)}{y}}+\frac{a_{-1,-1}\xi_{00}}{xy\left(1-\frac{y_0(x)}{y}\right)}+$$
$$+\eta(x,y_0(x))\frac{y_0(x)}{y}\cdot\frac{1}{1-\frac{y_0(x)}{y}}=0,$$

что совпадает с (4.1) при $|x|=1$, $|y_0(x)|<1$.

Из леммы 4.1 следует, что все сводится к вычислению функций $\pi(x)$ и $\tilde{\pi}(y)$. Приведем здесь один из результатов такого рода: пусть R_1 и R_2 — римановы поверхности функций $\pi(x)$ и $\tilde{\pi}(y)$ соответственно (введенные ниже) с заданными накрытиями $h_1: R_1\to\mathbf{P}_x$ и $h_2: R_2\to\mathbf{P}_y$, где $\mathbf{P}_x$ и $\mathbf{P}_y$ — одномерные проективные пространства. Тогда $\xi(x,y)$ естественно определена на $R_1\times R_2$ и является мероморфной функцией на этом комплексном многообразии.

§ 5. Уравнения на римановой поверхности

Далее будет предполагаться, если не оговорено противное, что $a(x,y)$ рациональна и реализуется случай неприводимости (см. § 2). Пусть $A_1(x,y)$ — неприводимый многочлен, для которого $A_1(x_0(y),y)=$ $=A_1(x,y_0(x))=0$; n — степень $A_1(x,y)$ по x, m — степень $A_1(x,y)$ по y. Уравнение

$$A_1(x,y)=0 \tag{5.1}$$

определяет тогда алгебраические функции $y(x)$ и $x(y)$, римановы поверхности которых связны, компактны и конформно эквивалентны. Обозначим соответствующую абстрактную риманову поверхность через S. Как риманова поверхность алгебраической функции $y(x)$, S естественным образом реализуется как разветвленное m-листное накрытие $h_1: S\to\mathbf{P}$ комплексной сферы $\mathbf{P}$. Аналогично определяется накрытие $h_2: S\to\mathbf{P}$, соответствующее функции $x(y)$ и имеющее n листов.

Пусть $\mathbf{C}_A(x,y)$ — поле алгебраических функций, определенное уравнением (5.1). $\mathbf{C}_A(x,y)$ является конечным алгебраическим расширением поля $\mathbf{C}(x)$ рациональных функций от x, а также поля $\mathbf{C}(y)$ рациональных функций от y. $\mathbf{C}_A(x,y)$ естественно изоморфно $\mathbf{C}(S)$, полю мероморфных функций на римановой поверхности S. Функция $x(s)$, $s\in S$, соответствующая при этом изоморфизме функции $x(y)$, обладает следующим свойством: если $x(s_1)=x(s_2)$, s_1, $s_2\in S$, то $h_1s_1=h_1s_2$, и наоборот. Аналогичным свойством обладает функция $y(s)$, соответствующая функции $y(x)$.

Очевидно, что $h_1^{-1}(\Gamma)$ является границей открытого множества $h_1^{-1}(D)$, а $h_2^{-1}(\Gamma)$ — границей для $h_2^{-1}(D)$. Если над Γ нет точек ветвления h_i, $i=1, 2$, то $h_i^{-1}(\Gamma)$ состоит из простых аналитических замкнутых непересекающихся кривых. Точки самопересечения и недифференцируемости могут появляться в точках ветвления h_i на $h_i^{-1}(\Gamma)$. Ввиду наших предположений, при $|x|=1$ множество $h_1^{-1}(x) \cap h_2^{-1}(D)$ состоит из одной точки. Следовательно, пересечение $h_1^{-1}(\Gamma)$ с $h_2^{-1}(D)$ не содержит точек ветвления накрытия h_1.

Заметим, что в силу условия (1.2) $h_1^{-1}(\Gamma) \cap h_2^{-1}(\Gamma)=0$, и, таким образом, $h_1^{-1}(\Gamma) \cap h_2^{-1}(D)$ является аналитической простой замкнутой кривой. Обозначим ее Γ_0.

Аналогично определяется $\widetilde{\Gamma}_0 = h_2^{-1}(\Gamma) \cap h_1^{-1}(D)$. Положим $G = h_1^{-1}(D) \cap \cap h_2^{-1}(D)$. Ясно, что $\Gamma_0 \cup \widetilde{\Gamma}_0$ является границей G.

Лемма 5.1. *Пусть род римановой поверхности S не равен нулю и реализуется случай неприводимости. Тогда G является связной областью.*

Мы покажем, что если G несвязна, то Γ_0 и $\widetilde{\Gamma}_0$ гомотопически эквивалентны нулю на S. Окончание доказательства проводится так же, как доказательство п. 2 леммы 7.1. Действительно, мы уже заметили, что $h_1^{-1}(x) \cap \cap h_2^{-1}(D)$ при $|x|=1$ состоит из одной точки. Кроме того, если G несвязна, то она состоит из двух связных компонент, ограничиваемых соответственно кривыми Γ_0 и $\widetilde{\Gamma}_0$. Компонента G, ограничиваемая кривой Γ_0, накрывает единичный круг $\{x: |x| \leqslant 1\}$. Это накрытие, в силу сделанного выше замечания, однолистно на границе. Следовательно, оно однолистно везде. Отсюда и следует искомое утверждение.

Обозначим через Δ объединение связных компонент множества $h_1^{-1}(D) \cup \cup h_2^{-1}(D)$ таких, что они имеют с G непустое пересечение. Ясно, что в условиях леммы 5.1 Δ является связной областью. Пусть $\Gamma_1, \Gamma_2, \ldots, \Gamma_k \subset \subset h_1^{-1}(\Gamma)$ и $\widetilde{\Gamma}_1, \widetilde{\Gamma}_2, \ldots, \widetilde{\Gamma}_l \subset h_2^{-1}(\Gamma)$—связные компоненты границы области Δ. В случае, если на них нет точек ветвления накрытий h_1 и h_2 соответственно, они являются простыми замкнутыми непересекающимися аналитическими кривыми.

Пусть теперь дана функция $f(x)$, мероморфная в некоторой области D комплексной сферы $\mathbf{P}$, и риманова поверхность S вместе с разветвленным накрытием $h: S \to \mathbf{P}$. Тогда функция $f(x)$ может быть поднята на область $h^{-1}(D)$ следующим образом:

$$f_h(s) = f(hs), \quad s \in h^{-1}(D).$$

При этом $f_h(s)$ (обозначаемая далее просто $f(s)$) является мероморфной функцией в $h^{-1}(D)$, что для точек, где накрытие разветвлено, проверяется, например, по теореме об устранимых особенностях.

Перенося $\pi(x)$ на $h_1^{-1}(D) \cap \Delta$ посредством h_1^{-1}, а $\widetilde{\pi}(y)$ посредством h_2^{-1} на $h_2^{-1}(D) \cap \Delta$, получим функции $\pi(s)$ и $\widetilde{\pi}(s)$, определенные на $h_1^{-1}(D) \cap \Delta$ и $h_2^{-1}(D) \cap \Delta$ соответственно, аналитические в этих областях и непрерывные на границах. Далее для возможности проведения аналитического исследования уравнений мы всегда, если не оговорено противное, будем предполагать, что $\eta(x, y)$ в уравнении (3.2) есть многочлен. Тогда можно считать функции $x(s)$, $y(s)$, $\eta(s) = \eta(x(s), y(s))$ определенными на всей S. Функции $\pi(s)$ и $\tilde{\pi}(s)$ удовлетворяют на Γ_0 и $\tilde{\Gamma}_0$ одному и тому же уравнению

$$\pi(s) + \tilde{\pi}(s) = -x(s)\,y(s)\,\eta(s), \tag{5.2}$$

которое получается применением h_1^{-1} и h_2^{-1} к уравнению (3.2).

Уравнение (5.2) допускает, очевидно, аналитическое продолжение на область G, где определены функции $\pi(s)$ и $\tilde{\pi}(s)$ одновременно. Используя уравнение (5.2), функции $\pi(s)$ и $\tilde{\pi}(s)$ можно мероморфно продолжить на области $h_1^{-1}(D) \cap \Delta$ и $h_2^{-1}(D) \cap \Delta$ соответственно. Иначе говоря, $\pi(s)$ и $\tilde{\pi}(s)$ мероморфны в области Δ, причем в этой области выполняется уравнение (5.2). Обратное утверждение также верно.

Лемма 5.2. *Если в условиях леммы* 5.1 *существуют решения уравнения* (5.2) $\pi(s)$ *и* $\tilde{\pi}(s)$, *аналитические соответственно в областях* $h_1^{-1}(D) \cap \Delta$ *и* $h_2^{-1}(D) \cap \Delta$, *непрерывные на границе этих областей и такие, что* $\pi(s_1) = \pi(s_2)$ *при* $x(s_1) = x(s_2)$ *и* $\tilde{\pi}(s_1) = \tilde{\pi}(s_2)$ *при* $y(s_1) = y(s_2)$, *то решение уравнения* (3.2) *дается функциями* $\pi(x) = \pi(x(s))$ *и* $\tilde{\pi}(y) = \tilde{\pi}(y(s))$.

Действительно, определим $\xi(x, y)$ по формуле (4.2). Тогда, заметив, что выполняется уравнение (4.1), для доказательства достаточно использовать лемму 4.1.

§ 6. Точный тип (—1,1; —1,1). Род 0

В этом и в двух следующих параграфах мы полностью решим уравнение (1.1) точного типа (—1,1; —1,1). В этом случае, если многочлен $A(x, y)$ неприводим, то род римановой поверхности S равен нулю или единице. В данном параграфе мы разберем случай нулевого рода.

Представим $A(x, y)$ в виде $a(x)y^2 + b(x)y + c(x)$. Тогда для того, чтобы S имела род 0, необходимо и достаточно, чтобы дискриминант $D(x) = b^2(x) - 4a(x)c(x)$ имел либо вторую степень по x, либо имел кратный корень в $\mathbf{C}$.

Лемма 6.1. *На* $h_i^{-1}(\Gamma)$ *нет точек ветвления накрытия* h_i, $i = 1, 2$, *и множество* $h_i^{-1}(\Gamma)$ *несвязно, т. е. в* D *лежит четное число точек ветвления как накрытия* h_1, *так и накрытия* h_2. *Таким образом, так как* h_i *двулистно, то* $h_i^{-1}(\Gamma)$ *состоит из двух непересекающихся аналитических простых замкнутых кривых.*

Доказательство. Если бы множество $h_1^{-1}(\Gamma)$ было связным, то оно целиком принадлежало бы множеству $h_2^{-1}(D)$ (ибо $h_1^{-1}(\Gamma) \cap h_2^{-1}(\Gamma) = 0$), что в силу условия (1.3) невозможно.

Таким образом, возможны следующие четыре случая взаимного расположения $h_1^{-1}(D)$ и $h_2^{-1}(D)$.

А) В D лежит ровно две точки ветвления накрытия h_i, $i = 1, 2$, т. е. $h_1^{-1}(D)$ и $h_2^{-1}(D)$ связны (рис. 1).

B) В D нет точек ветвления накрытий h_1 и h_2, т. е. $h_1^{-1}(D)$ и $h_2^{-1}(D)$ несвязны (рис. 2).

C) В D лежит две точки ветвления h_1 и нет точек ветвления h_2, т. е. $h_1^{-1}(D)$ связна, а $h_2^{-1}(D)$ несвязна.

D) В D есть две точки ветвления h_2, но нет точек ветвления h_1, т. е. $h_2^{-1}(D)$ связна, а $h_1^{-1}(D)$ несвязна (рис. 3).

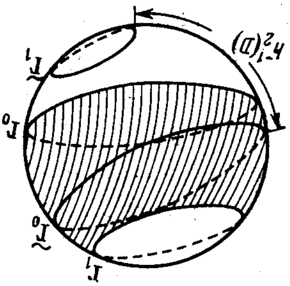

Рис. 1

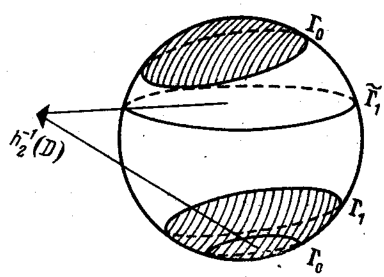

Рис. 2

Приведем примеры, когда эти случаи имеют место:

A) $a_{1,0} = a_{1,1} = a_{0,1} = 0$, B) $a_{-1,0} = a_{-1,-1} = a_{0,-1} = 0$,

C) $a_{10} = a_{1,-1} = a_{0,-1} = 0$, D) $a_{01} = a_{-1,1} = a_{-1,0} = 0$.

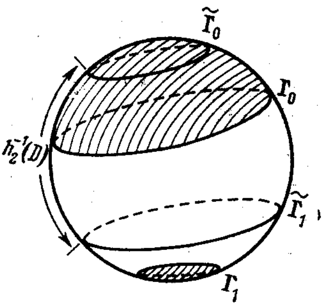

Рис. 3

При этом везде $a_{00} = -1$, а остальные a_{ij} положительны, причем их сумма меньше единицы (вероятностный аналог этих примеров см. в работе [14]).

Теорема 6.1. *В случаях* C) *и* D) *решение* $\xi(x, y)$ *является рациональной функцией.*

Доказательство. Достаточно доказать, что $\pi(x)$ и $\tilde{\pi}(y)$ рациональны. Для этого, в свою очередь, достаточно показать, что одна из функций $\pi(s)$ или $\tilde{\pi}(s)$ может быть мероморфно продолжена на всю риманову поверхность S. Тогда и вторая оказывается мероморфной на S в силу основного уравнения (5.2).

Основной прием, которым мы здесь и далее будем пользоваться, это продолжение мероморфных функций с помощью автоморфизмов Галуа. В случае типа $(-1, 1; -1, 1)$ поле $\mathbf{C}_A(x, y) \sim \mathbf{C}(S)$ является расширением Галуа как поля $\mathbf{C}(x)$ рациональных функций от x, так и поля $\mathbf{C}(y)$ рациональных функций от y. Группа Галуа поля $\mathbf{C}_A(x, y)$ над $\mathbf{C}(x)$ является циклической группой второго порядка. Обозначим ее нетривиальный элемент через $\tilde{\xi}$. Аналогично пусть $\tilde{\eta}$ — нетривиальный элемент группы Галуа поля $\mathbf{C}_A(x, y)$ над $\mathbf{C}(y)$. Автоморфизмам Галуа поля $\mathbf{C}(S)$ соответствуют конформные автоморфизмы ξ и η римановой поверхности S:

$$\tilde{\xi} f(s) = f(\xi s), \quad \tilde{\eta} f(s) = f(\eta s), \quad f \in \mathbf{C}(S).$$

Явный вид автоморфизмов Галуа ξ и η в случае точного типа $(-1, 1; -1, 1)$ дается формулами

$$\xi y = \frac{a_{1,-1}x^2 + a_{0,-1}x + a_{-1,-1}}{y(a_{11}x^2 + a_{01}x + a_{-1,1})}, \quad \eta x = \frac{a_{-1,1}y^2 + a_{-1,0}y + a_{-1,-1}}{x(a_{11}y^2 + a_{10}y + a_{1,-1})}.$$

Перейдем к доказательству теоремы 1 в случае C). Обозначим через $\widetilde{D}_0$ и $\widetilde{D}_1$ компоненты $h_2^{-1}(D)$, ограниченные соответственно $\widetilde{\Gamma}_0$ и $\widetilde{\Gamma}_1$. Область $\Delta = h_1^{-1}(D) \bigcup \widetilde{D}_0$ связна, и так же, как в § 5, получаем, что $\pi(s)$ и $\widetilde{\pi}(s)$ мероморфны в Δ. Заметим теперь, что $\xi\Gamma_0 = \Gamma_1$. Поэтому $\xi(\Delta \setminus h_1^{-1}(D)) = S \setminus \Delta$, и тогда $\Delta \bigcup \xi\Delta = S$. Из последнего соотношения вытекает, ввиду инвариантности $\pi(s)$ относительно ξ, что $\pi(s)$ продолжается на всю S. Случай D) рассматривается аналогично.

Явный вид решения в случае C) и D).

В случае C) обозначим через $s_1, \ldots, s_k$ $(k \leqslant 4)$ — полюса функции $-xy\eta(x, y)$ правой части формулы (5.2) в области $\Delta \setminus h_1^{-1}(D)$, а через $\varphi(s - s_i)$ — главные части этой функции в точках s_i (будем считать, что s — униформизирующая переменная на S). Тогда

$$\pi(s) = \sum_{i=1}^{k} \varphi_i(s - s_i) + \sum_{i=1}^{k} \xi\varphi_i(s - s_i), \tag{6.1}$$

$$\widetilde{\pi}(s) = -x(s)y(s)\eta(s) - \pi(s). \tag{6.2}$$

Действительно, $\pi(s)$ должна быть аналитична в $h_1^{-1}(D)$, но может иметь полюса в $\Delta \setminus h_1^{-1}(D)$, так как $\widetilde{\pi}(s)$ в этой области аналитична. Так как $h_1^{-1}(D)$ связна, то $\pi(s)$ инвариантна относительно ξ (что и использовалось в доказательстве теоремы 6.1), откуда и следует формула (6.1). Заметим, что $\widetilde{\pi}(s)$ не обязана быть инвариантной относительно η.

При этом в формуле (6.1), конечно, $\pi(s)$ определена с точностью до аддитивной константы, которая определяется из условия, чтобы $\widetilde{\pi}(y) = 0$ при $y = 0$.

Переходим к рассмотрению случая B).

Теорема 6.2. *Пусть D_1 — одна из связных компонент области $h_1^{-1}(D) \bigcup \bigcup h_2^{-1}(D)$, и пусть функции $\pi(s)$ и $\widetilde{\pi}(s)$, определенные на этой компоненте, удовлетворяют уравнению* (5.2). *Тогда $\pi(s)$ и $\widetilde{\pi}(s)$ мероморфно продолжимы на всю S с выколотой точкой s_1, являющейся неподвижной точкой автоморфизма $\xi\eta$ и не принадлежащей D_1. При этом (D_1 ограничена Γ_1):*

$$\pi(s) = -x(s)y(s)\eta(s) - \widetilde{\pi}(s), \tag{6.3}$$

$$\widetilde{\pi}(s) = \sum_{k=1}^{\infty} [A(\xi\delta^{-k}s) - A(\delta^{-k}s)] + \mathrm{const}, \tag{6.4}$$

$$A(s) = -x(s)y(s)\eta(s),$$

причем ряд в правой части формулы (6.4) *абсолютно сходится в любой ограниченной части плоскости* $S \setminus \{s_1\}$, *за исключением конечного числа точек в этой части, которые являются полюсами одного из членов ряда.*

Доказательство. Заметим сначала, что группа $\varkappa$ автоморфизмов S, порожденная ξ и η, является некоторой подгруппой группы дробнолинейных преобразований S. Обозначим через $D_0, D_1, \widetilde{D}_0, \widetilde{D}_1$ области, принадлежащие $h_1^{-1}(D) \cup h_2^{-1}(D)$ и ограниченные $\Gamma_0, \Gamma_1, \widetilde{\Gamma}_0, \widetilde{\Gamma}_1$ соответственно. Имеем (см. рис. 2): $D_0 \subset \widetilde{D}_1$, $\widetilde{D}_0 \subset D_1$, $\xi D_0 = D_1$, $\eta \widetilde{D}_0 = \widetilde{D}_1$. Построим индуктивно две последовательности областей:

$$D_0 \subset \widetilde{D}_1 \subset \eta D_1 \subset \eta\xi\widetilde{D}_1 \subset \ldots \subset \eta(\xi\eta)^k D_1 \subset (\eta\xi)^{k+1}\widetilde{D}_1 \subset \ldots,$$

$$\widetilde{D}_0 \subset D_1 \subset \xi\widetilde{D}_1 \subset \xi\eta D_1 \subset \ldots \subset \xi(\eta\xi)^k\widetilde{D}_1 \subset (\xi\eta)^{k+1} D_1 \subset \ldots.$$

Отношения принадлежности в этих цепочках легко проверяются по индукции.

Заметим теперь, что преобразование $\xi\eta$ (а следовательно, и $(\xi\eta)^n$) является либо гиперболическим, либо локсодромическим. Действительно, при параболическом или эллиптическом преобразовании образ ограниченного связного открытого множества (им можно считать D_1) не может строго содержать свой прообраз. При этом одна из неподвижных точек преобразования $\xi\eta$, точка s_1, принадлежит D_0, а вторая — s_2, принадлежит $\widetilde{D}_0$; $\xi s_1 = \eta s_1 = s_2$. Отсюда следует, что $\bigcup_{k=1}^{\infty} (\eta\xi)^k \widetilde{D}_1 = S \setminus \{s_2\}$ и $\bigcup_{k=1}^{\infty} (\xi\eta)^k D_1 = S \setminus \{s_1\}$. Обозначим поднятия функций $\pi(x)$ и $\widetilde{\pi}(y)$ на D_1 через $\pi(s)$ и $\widetilde{\pi}(s)$, а на $\widetilde{D}_1$ через $\Pi(s)$ и $\widetilde{\Pi}(s)$ соответственно. Тогда имеем $\pi(\xi s) = \Pi(s)$, $\widetilde{\pi}(\eta s) = \widetilde{\Pi}(s)$, $s \in \widetilde{D}_1$. Используя эти соотношения и уравнение (5.2), по индукции мероморфно продолжаем $\pi(s)$ и $\widetilde{\pi}(s)$ на области $(\xi\eta)^k D_1$, а Π и $\widetilde{\Pi}$ на $(\eta\xi)^k \widetilde{D}_1$, откуда и получается первое утверждение теоремы.

Для получения явного представления решения в случае B) рассмотрим уравнение

$$\pi(s) + \widetilde{\pi}(s) = A(s),$$

или

$$\Pi(\xi s) + \widetilde{\pi}(s) = \mathrm{A}(s). \tag{6.5}$$

Но

$$\Pi(\xi s) + \widetilde{\Pi}(\xi s) = A(\xi s),$$

или

$$\Pi(\xi s) + \widetilde{\pi}(\eta\xi s) = A(\xi s). \tag{6.6}$$

Сравнивая уравнения (6.5) и (6.6), получим

$$\widetilde{\pi}(\eta\xi s) - \widetilde{\pi}(s) = A(\xi s) - A(s). \tag{6.7}$$

Формальным решением последнего уравнения является ряд

$$\widetilde{\pi}(s) = \sum_{k=1}^{\infty} [A(\xi\delta^{-k}s) - A(\delta^{-k}s)] + \text{const},$$

где $\delta=\eta\xi$. Докажем, что этот ряд определяет мероморфную функцию на $S\setminus\{s_1\}$. Для этого достаточно показать, что этот ряд сходится в любой точке, отличной от полюса любого из членов этого ряда. Считая $\widetilde{\pi}(s_2)=0$, перепишем соотношение (6.7) следующим образом:

$$\widetilde{\pi}(s) - \widetilde{\pi}(\delta^{-1}s) = A(\xi\delta^{-1}s) - A(\delta^{-1}s),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\widetilde{\pi}(s) - \widetilde{\pi}(\delta^{-k}s) = \sum_{i=1}^{k} [A(\xi\delta^{-i}s) - A(\delta^{-i}s)].$$

Для любого $s \neq s_1$ имеем $\lim\limits_{k\to\infty} \widetilde{\pi}(\delta^{-k}s) = \widetilde{\pi}(s_2) = 0$. Поэтому ряд в правой части (6.4) сходится.

Заметим далее, что $A(\xi s) - A(s)$ не имеет полюсов в области $\widetilde{D}_0$. Для $A(s)$ это следует из того, что ни π, ни $\widetilde{\pi}$ не могут иметь полюсов в $\widetilde{D}_0$, а для $A(\xi s)$ — из того, что ни Π, ни $\widetilde{\Pi}$ не могут иметь полюсов в области $D_0 = \xi D_1$. Положим $s_2 = 0$, $s_1 = \infty$. Тогда, если $B(s) = A(\xi s) - A(s)$, то $B(0) = 0$, и так как $B(s)$ не имеет полюсов в точках 0 и ∞, то ее можно представить в виде линейной комбинации членов вида $\dfrac{cs^m}{(b-s)^n}$, $n > m \geqslant 1$.

Но тогда мероморфность ряда $\sum\limits_{k=0}^{\infty} B(r^{-k}s)$, $|r| > 1$, очевидна.

Симметричным образом получаем

$$\Pi(s) = \sum_{k=1}^{\infty} [A(\eta\widetilde{\delta}^{-k}s) - A(\widetilde{\delta}^{-k}s)] + \text{const}, \quad \widetilde{\delta} = \xi\eta.$$

Положим теперь $\pi(s) = \Pi(\xi s)$. После этого проверка соотношения (5.2) производится очевидным образом.

Случай А) рассматривается, в основном, аналогично предыдущему. Поэтому мы ограничимся формулировкой следующей теоремы.

Теорема 6.3. *Обозначим через s_1 и s_2 точки, обладающие свойством $\xi s_1 = \eta s_1 = s_2$ (или $A(x(s), y(s)) = A_x(x(s), y(s)) = A_y(x(s), y(s))$). Тогда $\pi(s)$ и $\widetilde{\pi}(s)$ мероморфно продолжаются на всю S с выколотыми точками s_1 и s_2.*

Для приведенного выше примера А) с $a_{10}=a_{11}=a_{01}=0$, отметим также следующий интересный факт. В этом случае $\pi(x)$ и $\widetilde{\pi}(y)$ рациональны (более того, являются многочленами от x и y соответственно). Действительно, уравнение

$$\pi(x) + \widetilde{\pi}(y) = x^n y^m \operatorname{mod}\ (a_{-1,-1} + a_{-1,0}y + a_{0,-1}x + a_{00}xy + a_{-1,1}y^2 + a_{1,-1}x^2)$$

разрешимо для любых $m, n \geqslant 0$. Будем проводить индукцию по $k=n+m$. Пусть это утверждение доказано для всех n, m таких, что $n+m<k$. Докажем, что тогда $x^n y^m$ при $n+m=k$ можно выразить линейно по $\operatorname{mod} A(x, y)$ через x^k, y^k и члены $x^n y^m$ с $n+m<k$ (обозначаемые далее через $o(k)$). Для этого запишем

$$a_{1,-1}x^{n+1}y^{k-n-1}+a_{00}x^n y^{k-n}+a_{-1,1}x^{n-1}y^{k-n+1}=o(k). \tag{6.8}$$

Для данного $n=n_0$, кроме соотношения (6.8) для $n=n_0$, из соотношений (6.8) для $n>n_0$ можно получить

$$x^{n_0}y^{k-n_0}+c_1 x^{n_0-1}y^{k-n_0+1}=o(k)+y^k,$$

а из соотношений (6.8) для $n<n_0$

$$x^{n_0}y^{k-n_0}+c_1 x^{n_0+1}y^{k-n_0-1}=o(k)+x^k.$$

Из соотношения (6.8) и двух последних соотношений непосредственно следует утверждение. Особый случай, возникающий при $n=2, 3$, также рассматривается без труда.

§ 7. Точный тип (—1,1; —1,1). Род 1

Если S имеет род 1, то на $\mathbf{P}$ имеется четыре точки ветвления как накрытия h_1, так и h_2 ($\mathbf{P}$ — комплексная сфера).

Лемма 7.1. *У накрытия h_i ($i=1, 2$) ровно две точки ветвления лежат внутри D и ровно две — вне D. При этом для любого i $h_i^{-1}(\Gamma)$ состоит из двух непересекающихся аналитических простых замкнутых кривых.*

Доказательство. То, что внутри D лежит четное число точек ветвления h_i, а также вторая часть леммы, доказывается так же, как лемма 6.1.

Если внутри D нет ни одной точки ветвления h_1, то обе компоненты $h_1^{-1}(\Gamma)$, обозначаемые через Γ_0 и Γ_1, ограничивают соответственно стягиваемые области D_0 и D_1 на торе. При этом у h_2 могут быть следующие возможности для числа точек ветвления внутри D.

1. Две точки ветвления h_2 лежат внутри D. Но тогда, как следует из построения накрывающей римановой поверхности $h_2: S\to\mathbf{P}$, кривые $\widetilde{\Gamma}_0$ и $\widetilde{\Gamma}_1$ не гомотопны нулю и ни одна из них не может целиком принадлежать D_0 или D_1.

2. Ни одной точки ветвления внутри D. Тогда

$$D_0\subset\widetilde{D}_1,\quad \widetilde{D}_0\subset D_1,\quad \xi D_0=D_1,\quad \eta\widetilde{D}_0=\widetilde{D}_1.$$

Введем на торе риманову метрику, индуцированную метрикой на универсальной накрывающей. Тогда автоморфизмы Галуа ξ и η должны сохранять площади, что противоречит соотношениям (7.1).

3. Четыре точки ветвления внутри D. Тогда $\widetilde{\Gamma}_0$ и $\widetilde{\Gamma}_1$ также гомотопны нулю и ограничивают стягиваемые области $\widetilde{D}_0'$ и $\widetilde{D}_1'$, причем $\widetilde{D}_0'\cup\widetilde{D}_1'=S\setminus h_2^{-1}(D)$. Аналогично должно быть $\widetilde{\Gamma}_0\subset D_0$, $D_1\subset\widetilde{D}_1'$ и, следовательно, $\widetilde{D}_0'\subset D_0$. Теперь мы приходим к противоречию так же, как и в случае 2.

Остается случай, когда h_1 и h_2 имеют внутри D четыре точки ветвления. В очевидных обозначениях тогда должно быть $D_1' \subset D_0$, $D_1 \subset D_0'$, и далее аналогично случаю 2.

Замечание 7.1. Это доказательство можно было бы провести на универсальной накрывающей. Аналогичным образом, учитывая построения § 9, завершается доказательство леммы 5.1.

Следствие 7.1. *Кривые* $\Gamma_0, \Gamma_1, \widetilde{\Gamma}_0, \widetilde{\Gamma}_1$ *гомотопны одному из элементов нормального базиса гомологий на торе.*

Таким образом, расположение областей $h_1^{-1}(D)$ и $h_2^{-1}(D)$ имеет вид, изображенный на рис. 4.

Доказательство. То, что Γ_0 и Γ_1 гомотопны одному из элементов нормального базиса гомологий на торе, очевидно из построения накрывающей римановой поверхности. Аналогичное утверждение верно для $\widetilde{\Gamma}_0$ и $\widetilde{\Gamma}_1$. Докажем, что Γ_0 гомотопно $\widetilde{\Gamma}_0$. Но это следует из того, что Γ_0 не пересекается с $\widetilde{\Gamma}_0$ и, следовательно, не может быть гомотопно *другому* элементу нормального базиса гомологий на торе.

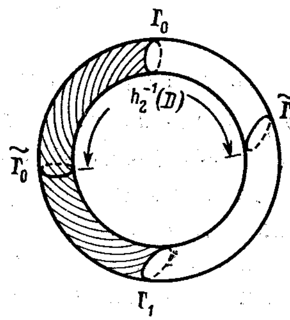

Рис. 4

Напомним, что так же, как и в случае рода 0, группы Галуа расширений $\mathbf{C}_A(x, y)$ над $\mathbf{C}(x)$ и $\mathbf{C}_A(x, y)$ над $\mathbf{C}(y)$ являются циклическими второго порядка, причем явный их вид задается формулами на стр. 509.

Универсальной накрывающей для S является комплексная плоскость $\mathbf{C}$. Пусть при этом фиксировано накрытие (неразветвленное) $\lambda: \mathbf{C} \to S$. Из теории униформизации известно, что S можно рассматривать как комплексную группу Ли, являющуюся фактор-группой аддитивной группы $\mathbf{C}$ по дискретной подгруппе $\{n\omega_1 + m\omega_2\}$, где периоды ω_1 и ω_2 линейно независимы над полем вещественных чисел $\mathbf{R}$, n и m — целые.

При накрытии λ любой отрезок длины $|\omega_i|$ и параллельный вектору ω_i проектируется в замкнутую кривую на S, класс гомологий которой является одним из элементов нормального базиса на S. Мы предположим без ущерба для общности, что $\lambda([0, \omega_1])$ гомологична Γ_0, а следовательно, по следствию 7.1 и всем $\Gamma_1, \widetilde{\Gamma}_0, \widetilde{\Gamma}_1$. Рассмотрим некоторую полосу

$$\Pi = \{\omega : \omega = \mu\omega_1 + \nu\omega_2;\ \mu, \nu \in \mathbf{R},\ 0 \leqslant \mu < 1\}.$$

Прообраз $\lambda^{-1}\Delta$ будет состоять из счетного числа связных криволинейных полос, сдвинутых друг относительно друга на вектора, кратные ω_2, а $\lambda^{-1}\Delta \cap \Pi$ состоит из некоторой связной области $(\lambda^{-1}\Delta)_0 \cap \Pi$ и ее сдвигов $[(\lambda^{-1}\Delta)_0 \cap \Pi] + n\omega_2$, где n — любое целое число. Область $(\lambda^{-1}\Delta)_0$ мы далее фиксируем и обозначаем через Δ_0. $\Delta_0 \cap \Pi$ является одной из связных компонент $\lambda^{-1}\Delta \cap \Pi$ в полосе Π (далее везде, где это не будет вести к недоразумениям, прообразы кривых и областей в Δ_0 будут обозначаться без индекса λ^{-1}).

Любая функция, заданная в области Δ, может быть поднята на область Δ_0 по формуле

$$f_\lambda(\omega)=f(\lambda\omega), \quad \lambda\omega\in\Delta.$$

Перенесем этим способом функции $x(s)$, $y(s)$, $\pi(s)$, $\tilde{\pi}(s)$ на Δ_0. Функции $x(\omega)$, $y(\omega)$ и $\eta(\omega)=\eta(x,y)$ оказываются при этом эллиптическими с периодами ω_1 и ω_2, мероморфными на всем $\mathbf{C}$ (индекс λ у перенесенных таким образом функций будем в дальнейшем опускать). При этом $\pi(\omega)$ и $\tilde{\pi}(\omega)$ определены в действительности на всем Δ_0, где они удовлетворяют уравнениям

$$\pi(\omega+\omega_1)=\pi(\omega), \tag{7.1}$$

$$\tilde{\pi}(\omega+\omega_1)=\tilde{\pi}(\omega), \tag{7.2}$$

$$\pi(\omega)+\tilde{\pi}(\omega)=-x(\omega)y(\omega)\eta(\omega). \tag{7.3}$$

Пусть x_1 и x_2 — точки ветвления алгебраической функции $y(x)$, лежащие внутри единичного круга. Прообразы точек $h_1^{-1}(x_1)$ и $h_2^{-1}(x_2)$ в области Δ_0 обозначим a_1 и a_2. Произвольный конформный автоморфизм $\tilde{\zeta}$ римановой поверхности S может быть продолжен до конформного автоморфизма $\zeta=\lambda^{-1}\zeta\lambda$ универсальной накрывающей (см. [10]). Это продолжение, конечно, неоднозначно, но оно становится однозначным, если фиксировать образ некоторой точки $\omega\in\mathbf{C}$ при автоморфизме ζ (этот образ должен принадлежать множеству $\{\lambda^{-1}\tilde{\zeta}\lambda\omega\}$). Поэтому для автоморфизма Галуа ξ мы потребуем, чтобы точка a_1 была неподвижной точкой этого автоморфизма: $\xi a_1=a_1$. Тогда $\xi\omega=-\omega$ в системе координат с началом в точке a_1. Действительно, известно, что самый общий конформный автоморфизм ζ комплексной плоскости $\mathbf{C}$ представляется в виде: $\zeta\omega=a\omega+b$. Так как начало координат является неподвижной точкой, то $b=0$. Но $\xi^2=1$, откуда $a^2=1$ и $a=-1$.

Аналогично определим $b_i=\Delta_0\cap\Pi\cap\{\lambda^{-1}h_2^{-i}(y_i)\}$, $i=1,2$, где y_1, y_2 — точки ветвления алгебраической функции $x(y)$, лежащие внутри единичного круга, и поднятие автоморфизма Галуа $\tilde{\eta}$ на универсальную накрывающую в системе координат с началом в точке b_1: $\eta\omega=-\omega$.

Лемма 7.2.

$$\xi\eta\omega=\omega+\omega_3,\ \text{где}\ \omega_3=2(a_1-b_1),$$

т. е. произведение автоморфизмов Галуа есть сдвиг на вектор, равный удвоенному расстоянию между их неподвижными точками.

Замечание 7.2. Аналогично можно доказать, что

$$a_1-a_2=\pm\frac{\omega_1}{2}, \qquad b_1-b_2=\pm\frac{\omega_1}{2}.$$

Теорема 7.1. *Функции $\pi(\omega)$ и $\tilde{\pi}(\omega)$ мероморфно продолжаются на всю универсальную накрывающую.*

Доказательство. Ввиду соотношений (7.1) и (7.2) достаточно ограничиться полосой Π или рассматривать некомпактную риманову поверхность S', являющуюся фактором $\mathbf{C}$ по $\{n\omega_1\}$ и топологически эквивалентную бесконечному цилиндру. Так как $\eta\Gamma_0 \subset h_2^{-1}(D)$, то $\eta\xi\Gamma_1 \subset \Delta_0$, и, следовательно, $\Delta_0 \cap (\eta\xi)\Delta_0 \neq 0$. Ввиду того, что $\eta\xi$ есть сдвиг на $-\omega_3$, объединение областей $(\eta\xi)^n\Delta_0$, $n = \ldots, -1, 0, 1, \ldots$, покрывает всю плоскость $\mathbf{C}$. Отсюда можно показать, что $\pi(\omega)$ и $\tilde{\pi}(\omega)$ продолжаются вдоль любого пути на $\mathbf{C}$, и с помощью теоремы о монодромии завершить доказательство теоремы.

Рассмотрим риманову поверхность $\tilde{S}$, являющуюся фактором $\mathbf{C}$ по $\{n\omega_1 + m\omega_3\}$, т. е. область $\tilde{\Delta}_0$ между Γ_1 и $\eta\Gamma_0$, принадлежащую $\Delta_0 \cap \Pi$, с отождествленными границами. Будем далее отождествлять Γ_1 и $\eta\Gamma_0$. Определим кусочно аналитическую функцию на $\tilde{S}$:

$$\Pi(\omega) = \begin{cases} \pi(\omega), & \omega \in h_1^{-1}(D) \setminus \overline{G}, \\ -\tilde{\pi}(\omega), & \omega \in \tilde{\Delta}_0 \cap h_2^{-1}(D). \end{cases}$$

Таким образом, получаем краевую задачу Римана на римановой поверхности $\tilde{S}$ (см., например, [8]): найти кусочно аналитическую функцию $\Pi(\omega)$ со скачками

$$\begin{aligned} \Pi_+(\omega) - \Pi_-(\omega) &= x(\xi\omega)\, y(\xi\omega)\, \eta(\xi\omega), \quad \omega \in \Gamma_1; \\ \Pi_+(\omega) - \Pi_-(\omega) &= -x(\omega)\, y(\omega)\, \eta(\omega), \quad \omega \in \tilde{\Gamma}_0. \end{aligned} \tag{7.4}$$

(Мы выбираем направление обхода $\Gamma_1 = \eta\Gamma_0$ и $\tilde{\Gamma}_0$ на $\tilde{S}$ по направлению ω_1 и через Π_+ обозначаем значения слева, а через Π_- — справа от соответствующей кривой.)

Рассмотрим теперь ζ-функцию Вейерштрасса

$$\zeta(u) = \frac{1}{u} + \sum_{n^2+m^2 \neq 0} \left[\frac{1}{(u - n\omega_1 - m\omega_3)} + \frac{1}{n\omega_1 + m\omega_3} + \frac{u}{(n\omega_1 + m\omega_3)^2} \right].$$

Выберем при этом начало координат в точке a_1; тогда $\zeta(u)$ нечетна относительно a_1, a_2, b_1, b_2.

Теорема 7.2.

$$\Pi(\omega) = \frac{1}{2\pi i} \int_{\Gamma_1 \cup \tilde{\Gamma}_0} \zeta(\omega - \tau) A(\tau)\, d\tau. \tag{7.5}$$

Кроме того, если $\omega \in h_1^{-1}(D) \setminus \overline{G}$, *то*

$$\pi(\omega) = \frac{1}{2\pi i} \int_{\Gamma_\delta} [-\zeta(\omega - \tau) + \zeta(\omega + \tau)]\, x(\tau)\, y(\tau)\, \eta(\tau)\, d\tau, \tag{7.6}$$

а функция $A(\tau)$ *определяется по формуле*

$$A(\tau) = \begin{cases} x(\xi\tau)\, y(\xi\tau)\, \eta(\xi\tau), & \tau \in \Gamma_1, \\ -x(\tau)\, y(\tau)\, \eta(\tau), & \tau \in \tilde{\Gamma}_0. \end{cases}$$

Доказательство. Функция $\Pi(\omega)$, определяемая выражением (7.5), является кусочно аналитической, имеет разрывы на кривых

$\Gamma_1+n\omega_3$, $\tilde{\Gamma}_0+n\omega_3$, скачки на которых согласуются с формулами (7.4), что легко проверяем, используя формулы Сохоцкого. Достаточно доказать, таким образом (для возможности отождествления Γ_1 и $\eta\Gamma_0$), что $\Pi(\omega)$ имеет периоды ω_1 и ω_3. Так как $\zeta(\omega+\omega_1)=\zeta(\omega)+\eta_1$, $\zeta(\omega+\omega_3)=\zeta(\omega)+\eta_3$, где η_1 и η_3 — константы, то для этого достаточно доказать, что из первого интегрального представления (формула (7.5)) следует второе (формула (7.6)), ибо ядро второго интегрального представления есть эллиптическая функция. Заметив, что $r(\tau)=x(\tau)y(\tau)\eta(\tau)$ не имеет полюсов на G, имеем

$$\int\limits_{\tilde{\Gamma}_0}\zeta(\omega-\tau)A(\tau)\,d\tau=-\int\limits_{\tilde{\Gamma}_0}\zeta(\omega-\tau)r(\tau)\,d\tau=-\int\limits_{\Gamma_0}\zeta(\omega-\tau)r(\tau)\,d\tau,$$

$$\int\limits_{\Gamma_1}\zeta(\omega-\tau)r(\xi\tau)\,d\tau=-\int\limits_{\Gamma_1}\zeta(-\omega-\xi\tau)r(\xi\tau)\,d\tau=$$

$$=-\int\limits_{\Gamma_0}\zeta(-\omega-\tau)r(\tau)\,d\tau=\int\limits_{\Gamma_0}\zeta(\omega+\tau)r(\tau)\,d\tau.$$

Решение задачи (7.4) единственно с точностью до константы, ибо разность двух подобных решений аналитична на всей $\tilde{S}$. Эта константа определяется из условия $\tilde{\pi}(y(s))=0$ при $y(s)=0$.

Из формулы (7.6) легко можно получить выражение вида

$$\pi(x)=\int\limits_{|x|=1}A(x,x')\,xy_0(x)\,\eta(x,y_0(x))\,dx,$$

где ядро $A(x,x')$ можно задать в явном виде, делая замену переменных в (7.6).

§ 8. Точный тип $(-1,1;\ -1,1)$. Случай приводимости

Пусть $A(x,y)$ приводим. Тогда с точностью до перестановки x и y возможны следующие случаи.

1) $A(x,y)=A(x)\tilde{A}(y)$, где $A(x)$ и $\tilde{A}(y)$ зависят лишь от x и от y соответственно. Этот случай легко решается методом факторизации (см. [15])

2) $A(x,y)=B(x,y)C(y)$, где $C(y)$ зависит лишь от y и имеет по y первую степень. В этом случае также применим метод факторизации. Действительно, если $C(y)=y+c$, где $|c|>1$, то $B(x,y)=b_1(x)y+b_2(x)$, и можно выбрать факторизацию так, чтобы $a_{-+}(x,y)=1$. Отсюда и следует утверждение (см. [15]). Случай $|c|<1$ разбирается аналогично.

3) $A(x,y)=A_1(x,y)A_2(x,y)$, где A_1 и A_2 — многочлены первой степени как по x, так и по y.

Пусть сначала

3А) $A_1(x_0(y),y)=A_1(x,y_0(x))=0$. Если $A_1(x,y)=(cx+d)y+ax+b$, то это возможно, например, когда $|c|$ велико по сравнению с модулями остальных коэффициентов.

Л е м м а 8.1. *В случае* 3А) $\pi(x)$ *и* $\tilde{\pi}(y)$ *рациональны.*

Доказательство. Рассмотрим риманову поверхность S рода 0, определенную уравнением $A_1(x, y)=0$, с однолистными неразветвленными накрытиями $h_1: S \to \mathbf{P}$ и $h_2: S \to \mathbf{P}$. Области $h_i^{-1}(D)$ однолистно накрывают D, и так как $h_1^{-1}(\Gamma) \subset h_2^{-1}(D)$, $h_2^{-1}(\Gamma) \subset h_1^{-1}(D)$, то $h_1^{-1}(D) \cup h_2^{-1}(D) = S$. Отсюда и следует рациональность π и $\tilde{\pi}$.

Явный вид решения в случае 3А).

Функция $-xy\eta(x, y)$ должна иметь ровно два полюса на S, причем один $y(s)=y_0$ — в области $h_1^{-1}(D)$, а второй $x(s)=x_0$ — в области $h_2^{-1}(D)$. Обозначая главные части в этих полюсах через $\tilde{\pi}(y)$ и $\pi(x)$ соответственно, будем иметь решение с точностью до констант, определяемых, например, из условия $\tilde{\pi}(0)=0$.

3Б) $A_1(x_0(y), y) = A_2(x, y_0(x)) = 0$. В этом случае рассмотрим две римановы поверхности: S_1, определенную уравнением $A_1(x, y)=0$, с накрытиями $h_{11}: S_1 \to \mathbf{P}$ и $h_{12}: S_2 \to \mathbf{P}$; S_2, определяемую уравнением $A_2(x, y)=0$, с накрытиями $h_{21}: S_2 \to \mathbf{P}$ и $h_{22}: S_2 \to \mathbf{P}$. Имеем, очевидно, $h_{12}^{-1}(D) \subset h_{11}^{-1}(D)$, $h_{21}^{-1}(D) \subset h_{22}^{-1}(D)$. Поднятия $\pi(x)$ и $\tilde{\pi}(y)$ на S_1 обозначим через $\pi(s)$ и $\tilde{\pi}(s)$, а на S_2 — $\Pi(t)$ и $\tilde{\Pi}(t)$ соответственно. $h_{11}^{-1}(D)$ и $h_{21}^{-1}(D)$ естественно конформно эквивалентны (как прообразы D). Обозначим этот изоморфизм $\xi: h_{11}^{-1}(D) \to h_{21}^{-1}(D)$. Аналогично вводится $\eta: h_{22}^{-1}(D) \to h_{12}^{-1}(D)$. Будем считать ξ и η продолженными до конформных изоморфизмов $\xi: S_1 \to S_2$ и $\eta: S_2 \to S_1$.

Так же, как и в случае В) § 6, $\xi\eta$ оказывается гиперболическим или локсодромическим, и строится последовательность областей

$$D_1 = h_{11}^{-1}(D) \subset D_2 \subset \ldots \subset D_k \subset \ldots,$$

где $D_k = (\eta\xi)^k D_1$.

При этом *функции $\pi(s)$ и $\tilde{\pi}(s)$ мероморфно продолжаются на $S_1 \setminus \{s_1\}$, где s_1 — неподвижная точка автоморфизма $\eta\xi$, лежащая вне D_1.* Явный вид функций получается аналогично случаю В) § 6, и мы не будем его выписывать.

§ 9. Построение автоморфизмов Галуа на универсальной накрывающей

Далее всегда будет предполагаться, что реализуется случай неприводимости (см. § 2) и род римановой поверхности S' $g \geqslant 2$. Далее используются терминология и результаты монографии [9] (см. также [10] и [11]).

Пусть $\mathbf{D}$ — неевклидова плоскость Лобачевского, реализованная как внутренность единичного круга. $\mathbf{D}$ является универсальной накрывающей для S с фиксированным накрытием $\lambda: \mathbf{D} \to S$ и группой преобразований наложения F. F является дискретной фуксовой группой первого рода движений плоскости Лобачевского, состоящей лишь из гиперболических элементов.

Рассмотрим прообраз $\lambda^{-1}\Delta$ связной (см. лемму 5.1) области Δ и одну из его связных компонент. Именно, фиксируем прообраз ω_0 некоторой точки $s_0 \in \Delta$ и рассмотрим компоненту $\Delta_0 \subset \lambda^{-1}\Delta$, содержащую ω_0.

Пусть точка ω_1 такова, что $\lambda\omega_1 = \lambda\omega_0 = s_0$, а C_{01} — произвольная кривая на $\mathbf{D}$, соединяющая ω_0 и ω_1. Тогда обозначим через $[\lambda C_{01}]$ элемент фундаментальной группы $\pi_1(S, s_0)$, соответствующий кривой λC_{01}. Очевидно, что этот элемент не зависит от C_{01} при данном ω_1, но различен для разных ω_1. Таким образом, получаем изоморфизм $\varphi: F \to \pi_1(S, s_0)$, причем если $\omega_1 = f\omega_0$, $f \in F$, то $\varphi(f) = [\lambda C_{01}]$.

Рассмотрим теперь всевозможные точки $\omega_i \in \Delta_0$ такие, что $\lambda\omega_i = s_0$. Множество $f_i \in F$ таких, что $f_i\omega_0 = \omega_i$, образует подгруппу $F_0 \subset F$, причем имеется канонический гомоморфизм $\psi: F_0 \to \pi_1(\Delta, s_0)$, и коммутативна следующая диаграмма:

$$\begin{array}{ccc} F_0 & \xrightarrow{\psi} & \pi_1(\Delta, s_0) \\ \cap\downarrow & & \downarrow \\ F & \xrightarrow{\varphi} & \pi_1(S, s_0). \end{array}$$

(Очевидно, что F_0 не зависит от выбора точки s_0.) Граница области Δ_0 состоит из кусочно аналитических кривых, являющихся прообразами $\Gamma_1, \Gamma_2, \ldots, \Gamma_k, \tilde{\Gamma}_1, \ldots, \tilde{\Gamma}_l$ (см. § 5). Эти прообразы мы будем обозначать через $\lambda_0^{-1}\Gamma_1, \ldots, \lambda_0^{-1}\Gamma_k, \ldots, \lambda_0^{-1}\tilde{\Gamma}_l$ соответственно. Если на $\Gamma_1, \ldots, \Gamma_k, \tilde{\Gamma}_1, \ldots, \tilde{\Gamma}_l$ нет точек ветвления накрытий h_1 или h_2 соответственно, то $\lambda_0^{-1}\Gamma_1, \ldots, \lambda_0^{-1}\Gamma_k, \ldots, \lambda_0^{-1}\tilde{\Gamma}_l$ аналитичны.

Мы будем предполагать далее, что *поле* $\mathbf{C}(S)$ *мероморфных функций на* S *является нормальным расширением как поля* $\mathbf{C}(x)$ *так и поля* $\mathbf{C}(y)$. Аналогично можно было бы рассмотреть случай, когда существует конечное алгебраическое расширение поля $\mathbf{C}(S)$, нормальное над $\mathbf{C}(x)$ и над $\mathbf{C}(y)$. Для этого надо было бы поднять предварительно уравнения на риманову поверхность этого расширения.

Лемма 9.1. *Для любых* Γ_i *и* $\tilde{\Gamma}_j$ *существует неевклидово движение (гиперболическое или эллиптическое)* g_i *такое, что* $g_i(\lambda_0^{-1}\Gamma_0) = \lambda_0^{-1}\Gamma_i$, *а также неевклидово движение* $\tilde{g}_j$ *такое, что* $\tilde{g}_j(\lambda_0^{-1}\tilde{\Gamma}_0) = \lambda_0^{-1}\tilde{\Gamma}_j$.

Докажем, что $\lambda_0^{-1}\Gamma_1, \ldots, \lambda_0^{-1}\tilde{\Gamma}_1, \ldots, \lambda_0^{-1}\tilde{\Gamma}_l$ аналитичны. Действительно, в §5 было показано, что Γ_0 является аналитической простой замкнутой кривой. Обозначим через γ_i конформный автоморфизм римановой поверхности S такой, что $\gamma_i\Gamma_0 = \Gamma_i$. Легко показать, что он существует и индуцируется соответствующим автоморфизмом Галуа поля $\mathbf{C}(S)$ над $\mathbf{C}(x)$. Аналогично введем автоморфизм $\tilde{\gamma}_j$ такой, что $\tilde{\gamma}_j\tilde{\Gamma}_0 = \tilde{\Gamma}_j$.

Для доказательства леммы 9.1 нам достаточно, таким образом, доказать следующую лемму и положить в ней $\gamma_i = \tilde{g}$, $g_i = g$.

Лемма 9.2. *Пусть* $\tilde{g}$ — *конформный автоморфизм* S, *причем для некоторых точек* $s_0, s \in S$ *имеет место* $\tilde{g}(s_0) = s$. *Тогда, если* $\lambda\omega = s$, $\lambda\omega_0 = s_0$, *то существует конформный автоморфизм g универсальной накрывающей* $\mathbf{D}$ *такой, что* $\lambda g = \tilde{g}\lambda$ *и* $\tilde{g}\omega_0 = \omega$. *При этом g является эллиптическим или гиперболическим.*

Действительно, если s_0 и s — две точки на S, не являющиеся точками ветвления накрытия $h_1: S \to \mathbf{P}$ и такие, что $x(s_0) = x(s)$, то существует и единственен конформный автоморфизм $\tilde{g}$ римановой поверхности S такой, что $\tilde{g}s_0 = s$. Далее мы не будем делать различия между $\tilde{g}$ и соответствующим автоморфизмом Галуа.

Произвольный конформный автоморфизм $\tilde{g}$ поверхности S_x может быть поднят на универсальную накрывающую $\mathbf{D}$: $g = \lambda^{-1}\tilde{g}\lambda$. Это перенесение неоднозначно, но становится однозначным, если фиксировать образ $g\omega$ одной точки $\omega \in \mathbf{D}$. При этом должно выполняться только условие $g\omega \in \{\lambda^{-1}g\lambda\omega\}$. Автоморфизм g является неевклидовым движением плоскости Лобачевского (см. [10]).

Докажем, что g является либо гиперболическим, либо эллиптическим. Если $\tilde{g}$ имеет неподвижную точку на S, то можно поднять его на $\mathbf{D}$ так, чтобы g имел неподвижную точку внутри $\mathbf{D}$. Тогда g является эллиптическим. Пусть теперь g не имеет неподвижных точек в $\mathbf{D}$. Это будет иметь место, в частности, когда $\tilde{g}$ не имеет неподвижных точек на S. Докажем, что тогда

$$d = \inf_{\omega \in \mathbf{D}} \rho(\omega, g\omega) > 0, \tag{9.1}$$

где $\rho(\omega, \omega')$ — расстояние между двумя точками на плоскости Лобачевского.

Пусть сначала $\tilde{g}$ не имеет неподвижных точек на S. Введем на S риманову метрику, порожденную неевклидовой метрикой на $\mathbf{D}$. Из компактности S следует, что

$$\inf_{s \in S} \rho(s, \tilde{g}s) > 0. \tag{9.2}$$

Отсюда легко следует и формула (9.1). Пусть теперь $\tilde{g}$ имеет неподвижную точку на S, а g не имеет неподвижных точек на $\mathbf{D}$. Тогда проведем для каждой неподвижной точки на S достаточно малую неевклидову окружность с центром в этой точке. Ограничиваемую ею открытую окрестность этой точки обозначим через O_i. Тогда $S \setminus [\bigcup_i O_i]$ инвариантна относительно $\tilde{g}$, и так же, как и в предыдущем случае, можно доказать, что

$$\inf_{\omega \in \lambda^{-1}[S \setminus \{\bigcup_i O_i\}]} \rho(\omega, g\omega) > 0. \tag{9.3}$$

Точки же из любой компоненты O множества $\lambda^{-1}O_i$ под действием g могут перейти лишь в точки конгруэнтного множества fO для некоторого $f \in F$, $f \neq 1$, откуда

$$\inf_{\omega \in \lambda^{-1}O_i} \rho(\omega, g\omega) > 0. \tag{9.4}$$

Ввиду конечности множества индексов i из (9.3) и (9.4) следует формула (9.1).

Нам осталось доказать таким образом, что если выполнено неравенство (9.1), то g — гиперболический элемент. Легко доказывается теперь, что существует z такой, что $\rho(z, gz) = d$. Если мы докажем теперь, что точки z, gz, g^2z лежат на некоторой неевклидовой прямой l, то эта прямая инвариантна относительно g, и, таким образом, g — неевклидов перенос (гиперболический элемент). Для доказательства обозначим через ζ середину неевклидова сегмента $[z, gz]$. Тогда $g\zeta$ является серединой сегмента $[gz, g^2z]$ и $\rho(\zeta, g\zeta) \leqslant \rho(\zeta, gz) + \rho(gz, g\zeta) = d$. Но в то же время $\rho(\zeta, g\zeta) \geqslant d$, откуда и следует утверждение.

Для окончания доказательства леммы 9.1 остается заметить, что для любого i и произвольной точки $\omega \in \lambda_0^{-1}\Gamma_i$ найдется такая точка $\omega' \in \lambda_0^{-1}\Gamma_0$, что $x(\omega) = x(\omega')$, и воспользоваться приведенными выше построениями. При этом g_i оказывается поднятием на $\mathbf{D}$ соответствующего автоморфизма Галуа.

Определение. Группой уравнения (1.1), риманова поверхность S для которого имеет род $g \geqslant 2$, будем называть группу $\varkappa$ неевклидовых движений плоскости Лобачевского, порожденную группой F_0 и всеми автоморфизмами g_i и $\tilde{g}_j$.

§ 10. Теорема об аналитическом поведении решения

Функции $x(s)$, $y(s)$, $\pi(s)$ и $\tilde{\pi}(s)$, определенные на $\Delta \subset S$, поднимем на Δ_0 по формуле $\pi_\lambda(\omega) = \pi(\lambda\omega)$, $\lambda\omega \in \Delta$, $\omega \in \Delta_0$ (в дальнейшем индекс λ у π и $\tilde{\pi}$ опускаем). Очевидно, что $\pi(\omega)$ и $\tilde{\pi}(\omega)$ удовлетворяют следующим соотношениям при $\omega \in \Delta_0$ и $h \in F_0$:

$$\begin{gathered} \pi(h\omega) = \pi(\omega), \\ \tilde{\pi}(h\omega) = \tilde{\pi}(\omega), \\ \pi(\omega) + \tilde{\pi}(\omega) = -x(\omega)y(\omega)\eta(x(\omega), y(\omega)). \end{gathered} \tag{10.1}$$

Теорема 10.1. *Функции $\pi(\omega)$ и $\tilde{\pi}(\omega)$ допускают мероморфное продолжение на всю универсальную накрывающую* $\mathbf{D}$.

Докажем сначала следующую лемму.

Лемма 10.1. *Объединение областей $h\Delta_0$, где $h \in \varkappa$, покрывает всю неевклидову плоскость* $\mathbf{D}$.

Доказательство. Заметим, что $\lambda_0^{-1}\Gamma_0 \cup \lambda_0^{-1}\tilde{\Gamma}_0$ лежит на положительном неевклидовом расстоянии ε от границы $\dot{\Delta}_0$ в плоскости Лобачевского. Выберем произвольным образом точку $\omega_0 \in \Delta_0$. Область Δ_0 обладает следующими свойствами:

1) граница области Δ_0 принадлежит объединению конечного числа аналитических кривых;

2) пусть l — одна из таких граничных аналитических кривых (мы не требуем, чтобы $l \cap \dot{\Delta}_0$ было связным). Тогда существует $h \in \varkappa$ такой, что любая точка $l \cap \dot{\Delta}_0$ принадлежит $\Delta_0 \cup h\Delta_0$ вместе со своей неевклидовой

ε-окрестностью. По лемме 9.1, в качестве h можно выбрать одну из образующих группы $\varkappa$: g_i или $\widetilde{g}_j$.

Построим теперь по индукции последовательность областей

$$\Delta_0, \Delta_1, \ldots, \Delta_n, \ldots$$

и последовательность аналитических кривых

$$l_0, l_1, \ldots, l_n, \ldots$$

следующим образом. Положим $l_0 = \Gamma_0$, $l_1 = \widetilde{\Gamma}_0$; далее в произвольном порядке расположим аналитические кривые, имеющие общую дугу с границей $\dot{\Delta}_0$, затем аналитические кривые, имеющие общую дугу с границей $\dot{\Delta}_1$, не совпадающие с ранее построенными, и т. д.

Выберем $h_0 = g_i$ или $\widetilde{g}_j$ так, чтобы $l_2 = g_i l_0$ или $\widetilde{g}_j l_1$. Положим $\Delta_1 = \Delta_0 \bigcup h_0 \Delta_0$. Пусть мы уже построили Δ_n. Тогда существует $h_n = g_i$ или $\widetilde{g}_j$ такой, что $l_{n+2} = h_n l_{i_n}$ для некоторого $i_n \leqslant n+1$. Полагаем $\Delta_{n+1} = \Delta_n \bigcup h_n \Delta_n$. При этом l_{n+2} принадлежит Δ_{n+1} вместе со своей неевклидовой ε-окрестностью. Выполнение свойств 1) и 2) для Δ_{n+1} очевидно. Кроме того, по построению для любого n существует $N > n$ такое, что $\rho(\omega_0, \dot{\Delta}_N) > \rho(\omega_0, \dot{\Delta}_n) + \varepsilon$. Отсюда и следует, что объединение возрастающей последовательности связных областей Δ_n совпадает со всей $\mathbf{D}$.

Вернемся теперь к доказательству теоремы 1. Имеем, в силу сделанных в § 9 построений,

$$\pi(g_i\omega) = \pi(\omega), \qquad \widetilde{\pi}(\widetilde{g}_j\omega) = \widetilde{\pi}(\omega), \tag{10.2}$$

если $g_i\omega, \omega \in \Delta_0$ и $\widetilde{g}_j\omega, \omega \in \Delta_0$, соответственно. По индукции функции $\pi(\omega)$ и $\widetilde{\pi}(\omega)$ продолжаются на области Δ_n. Пусть они определены в области Δ_n. Тогда, если $h_n = g_i$, то продолжаем $\pi(\omega)$ на область Δ_{n+1} с помощью соотношения $\pi(h_n\omega) = \pi(\omega)$. Функция $\widetilde{\pi}(\omega)$ определяется тогда из уравнения (10.1); аналогично поступаем, если $h_n = \widetilde{g}_j$. Теперь можно легко доказать, что $\pi(\omega)$ и $\widetilde{\pi}(\omega)$ могут быть продолжены вдоль любой кривой на $\mathbf{D}$, и, следовательно, они оказываются однозначными на $\mathbf{D}$ (по теореме о монодромии).

Таким образом, из теоремы 1 и замечания 3.1 получается полное описание возможностей аналитического продолжения решения.

§ 11. Дискретность группы уравнения

Здесь мы изучим группу $\varkappa$ уравнения в случае рода $g \geqslant 2$ и случая неприводимости.

Лемма 11.1. *Если род римановой поверхности S $g \geqslant 2$ и реализуется случай неприводимости, то группа $\varkappa$ уравнения* (1.1) *дискретна.*

Доказательство. По теореме Пуанкаре (см. [9], стр. 99), для групп движений неевклидовой плоскости дискретность и разрывность эквивалентны. Таким образом, если $\varkappa$ недискретна, то она не может быть разрывной. Следовательно, каждая точка $\omega \in \mathbf{D}$ является предельной точкой $\varkappa$. Но по теореме 4А главы 3, [9], стр. 103, множество $\varkappa z$ плотно тогда в любой (предельной) точке $\omega \in \mathbf{D}$ для любого $z \neq \omega$, кроме, может быть, одного значения $z = z_0$, для которого $\{\varkappa z_0\} = \{z_0\}$, т. е. в нашем случае для любого $z \in \mathbf{D}$.

Известно (см. [10]), что группа конформных автоморфизмов E произвольной римановой поверхности S рода $g \geqslant 2$ конечна. Отсюда легко следует, что множество точек $\{\lambda^{-1} e_i \lambda \omega\}$ для любой точки $\omega \in \mathbf{D}$, где $E = \{e_0, \ldots, e_n\}$, дискретно. Тем более является дискретным и множество точек $\varkappa\omega$ для любого ω. Лемма доказана.

Лемма 11.2. *$\varkappa$ может состоять лишь из эллиптических и гиперболических элементов.*

Доказательство. Фиксируем некоторую точку $s_0 \in \Delta \subset S$ и один из ее прообразов $\omega_0 \in \{\lambda^{-1} s_0\} \cap \Delta_0 \subset \mathbf{D}$. Для любой точки $s \in \Delta$ обозначим

$$\rho_\Delta(s_0, s) = \inf_l \rho(l),$$

где $\rho(l)$ — длина спрямляемой кривой $l \subset \Delta$, соединяющей точки s_0 и s_1, в римановой метрике, индуцированной метрикой плоскости Лобачевского, а inf берется по всем таким спрямляемым кривым. Аналогично определяется $\rho_{\Delta_0}(\omega_0, \omega)$ для любой точки $\omega \in \Delta_0$. Легко видеть, что существует такая константа C_0, что для любой точки $\omega \in \Delta_0$

$$\inf_{1 \neq h \in F_0} \rho_{\Delta_0}(\omega, h\omega) < C_0. \tag{11.1}$$

Рассмотрим нормальный фундаментальный многоугольник N (см. [10]) относительно точки ω_0 в предположении, что ω_0 не является неподвижной точкой для $\varkappa$. Докажем, что N компактен. Действительно, в противном случае существовала бы точка $\omega \in \Delta_0$, удаленная от ω_0 на сколь угодно большое расстояние, что противоречит условию (11.1).

Из компактности нормальной фундаментальной области следует, что $\varkappa$ не содержит параболических элементов (см. [9], теорема 7Е, стр. 149).

При этом группа $\varkappa$ имеет образующие $h_1, h_2, \ldots, h_l$, $a_1, b_1, \ldots, a_m, b_m$ и определяющие соотношения:

$$h_1 \ldots h_l a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_m b_m a_m^{-1} b_m^{-1} = 1, \; h_i^{k_i} = 1, i = 1, \ldots, l$$

(см. [9], стр. 241).

Лемма 11.3. *Для некоторого n существует нормальный делитель $\varkappa_0$ группы $\varkappa$ с образующими $A_1, B_1, \ldots, A_n, B_n$ и единственным определяющим соотношением:*

$$A_1 B_1 A_1^{-1} B_1^{-1} \ldots A_n B_n A_n^{-1} B_n^{-1} = 1.$$

Доказательство. Сельберг доказал (см. [13]), что в любой матричной группе над полем комплексных чисел с конечным числом образующих существует нормальный делитель конечного индекса, не содержащий элементов конечного порядка. Этот результат, очевидно, можно применить к нашей группе $\varkappa$. Тогда такой нормальный делитель оказывается состоящим лишь из гиперболических элементов, а в силу конечности индекса его фундаментальная группа компактна. Откуда (см. [9]) и следует, что он имеет единственное указанное выше определяющее соотношение.

§ 12. Задача Карлемана для фундаментального многоугольника и явный вид решения

Несмотря на получаемое ниже явное представление для решения, очень важную роль играет метод аналитического продолжения из § 10. В частности, с его помощью мы получаем все полюса функций $\pi(\omega)$ и $\tilde{\pi}(\omega)$ на $\mathbf{D}$ и их главные части в этих полюсах. Процедура выделения главных частей очевидна из доказательства теоремы 10.1.

Замечание 12.1. В частности, отсюда вытекает, что асимптотическое поведение коэффициентов функций $\pi(x)$ и $\tilde{\pi}(y)$ определяется либо ближайшими к единичному кругу их алгебраическими точками ветвления, либо ближайшими к единичному кругу полюсами и может быть вычислено в конкретных случаях (то, что других особых точек у $\pi(x)$ и $\tilde{\pi}(y)$ нет, следует из теоремы 10.1).

Этот факт играет основную роль в вероятностных приложениях, но здесь мы на нем не останавливаемся.

Рассмотрим нормальный делитель $\varkappa_0$ из леммы 11.3. Используя соотношения (10.1), (10.2) и следующие из них соотношения

$$\pi(\tilde{g}_j\omega) - \tilde{\pi}(\omega) = x(\omega)\,y(\omega)\,\eta(\omega) - x(\tilde{g}_j\omega),\ y(\tilde{g}_j\omega)\,\eta(\tilde{g}_j\omega), \tag{12.1}$$

можно получить уравнения для функций π, $\tilde{\pi}$ и образующих $A_1, B_1, \ldots, A_n, B_n$ этого нормального делителя

$$\pi(A_i\omega) - \pi(\omega) = \alpha_i(\omega), \quad \pi(B_i\omega) - \pi(\omega) = \beta_i(\omega), \quad i = 1, \ldots, n, \tag{12.2}$$

где $\alpha_i(\omega)$ и $\beta_i(\omega)$ — линейные комбинации автоморфных функций относительно группы F, сдвинутых на элементы группы $\varkappa$.

Соотношения (12.2) представляют собой задачу Карлемана для фундаментального многоугольника $N: A_1B_1A_1^{-1}B_1^{-1} \ldots A_nB_nA_n^{-1}B_n^{-1} = 1$, который мы в дальнейшем фиксируем и который можно выбрать так, чтобы $\alpha_i, \beta_i, \pi, \tilde{\pi}$ не имели полюсов на его границе (относительно задачи Карлемана см. библиографию в [8]).

Рассмотрим ядро Беенке — Штейна $\mathfrak{A}(\omega, \omega')$ ([8]) для римановой поверхности $\mathbf{D}/\varkappa_0$, явную конструкцию которого приводил по существу еще Вейерштрасс ([12]). На $\mathbf{D}$ $\mathfrak{A}(\omega, \omega')$ представляет собой автоморфную функцию по ω и автоморфную форму по ω'.

Одно из автоморфных в $\mathbf{D}$ решений задачи (12.2) представляется в виде

$$\Pi(\omega) = \frac{1}{2\pi i}\sum_i \left[\int_{A_i} \mathfrak{A}(\omega, \omega')\,\alpha_i(\omega')\,d\omega' + \int_{B_i} \mathfrak{A}(\omega, \omega')\,\beta_i(\omega')\,d\omega'\right]. \tag{12.3}$$

Обозначим через $\Pi'(\omega)$ — сумму главных частей в полюсах $\Pi(\omega)$ в N, а через $\Pi''(\omega)$ — такую же сумму главных частей для $\pi(\omega)$ (см. замечания в начале параграфа).

Разность $\pi(\omega)$ и частного решения $\Pi(\omega)$ неоднородной задачи (12.2) должна совпадать в N с автоморфной функцией $\tilde{\Pi}(\omega)$ относительно группы $\varkappa_0$. При этом $\tilde{\Pi}(\omega)=\pi(\omega)-\Pi(\omega)$ не имеет полюсов на границе N. $\tilde{\Pi}(\omega)$ определяется тем фактом, что сумма главных частей $\tilde{\Pi}(\omega)$ для полюсов, лежащих в N, должна совпадать с $\Pi''(\omega)-\Pi'(\omega)$. Сформулируем полученный результат.

Теорема 12.1. *В фундаментальном многоугольнике N функция $\pi(\omega)$ имеет следующее явное представление:*

$$\pi(\omega)=\frac{1}{2\pi i}\sum_i\left(\int_{A_i}\mathfrak{A}(\omega,\omega')\,\alpha_i(\omega')\,d\omega'+\int_{B_i}\mathfrak{A}(\omega,\omega')\,\beta_i(\omega')\,d\omega'\right)+\tilde{\Pi}(\omega)+\text{const},$$

где $\tilde{\Pi}(\omega)$ — автоморфная относительно группы $\varkappa_0$ функция, определенная выше. Аналогичное представление имеет функция $\tilde{\pi}(\omega)$, а константы в их правых частях определяются из условия $\tilde{\pi}(y)=0$ при $y=0$.

Замечание 12.2. Можно было бы построить аналогичное интегральное представление непосредственно для группы $\varkappa$. При этом выражение стало бы более явным, но ввиду наличия эллиптических элементов формулы были бы более громоздкими.

В заключение хочу поблагодарить И. И. Пятецкого-Шапиро за ценную консультацию по дискретным группам.

(Поступила в редакцию 24/III 1970 г. и 7/VII 1970 г.)

Литература

1. И. Б. Симоненко, Операторы типа свертки в конусах, Матем. сб., **74** **(116)** (1967), 298—313.
2. И. Б. Симоненко, О многомерных дискретных свертках, Матем. исследования, Кишинев, 3, № 1 (1968), 108—122.
3. В. А. Малышев, О решении дискретных уравнений Винера — Хопфа в четверти плоскости, ДАН СССР, **187**, **№ 6** (1969), 1243—1246.
4. N. Wiener, E. Hopf, Ueber Eine Klasse Singularen Integralgleichungen, Sitzungsb. Preus. Akad. Wiss., (1931), 696—706.
5. М. Г. Крейн, Интегральные уравнения на полупрямой с ядром, зависящим от разности аргументов, Успехи матем. наук, **XIII**, вып. **5** (83) (1958), 3—120.
6. И. Ц. Гохберг, Л. С. Гольденштейн, О многомерном интегральном уравнении на полупространстве с ядром, зависящим от разности аргументов, и его дискретном аналоге, ДАН СССР, **131**, **№ 1** (1960), 9—12.
7. И. М. Гельфанд, Д. А. Райков, Г. Е. Шилов, Коммутативные нормированные кольца, Москва, Физматгиз, 1960.
8. Э. И. Зверович, Г. С. Литвинчук, Краевые задачи со сдвигом для аналитических функций и сингулярные функциональные уравнения, Успехи матем. наук, **XXIII**, вып. **3** **(141)** (1968), 67—121.

9. J. Lehner, Discontinuous groups and automorphic functions (1964), Providence (American Math. Soc. Math. Survey No. 8).
10. Дж. Спрингер, Введение в теорию римановых поверхностей, Москва, ИЛ, 1960.
11. Л. Р. Форд, Автоморфные функции, Москва — Ленинград, ОНТИ, 1936.
12. K. Weierstrass, Vorlesungen über die theorie der Abelschen transzendenten, Math. Werke, Bd. 4 (1902), Berlin.
13. A. Selberg, On discontinuous groups in higher dimensional symmetric spaces, Tata Inst, of Fundam. Research (1960), 147—164.
14. В. А. Малышев, Аналитический метод в теории случайных блужданий в четверти плоскости: простое блуждание с косым отражением, Труды Советско-Японского симпозиума по теории вероятностей, Хабаровск, 1969, 176—184.
15. В. С. Рабинович, Многомерное уравнение Винера — Хопфа для конусов, Теория функций, функц. анализ и их приложения, вып. 5 (1957), 59—67.
16. В. А. Какичев, Краевые задачи линейного сопряжения для функций, голоморфных в бицилиндрических областях, Теория функций, функц. анализ и их приложения, вып. 5 (1967), 37—58.

Mat. Sbornik
Tom 84 (126) (1971), No. 4

Math. USSR Sbornik
Vol. 13 (1971), No. 4

WIENER-HOPF EQUATIONS IN A QUADRANT OF THE PLANE, DISCRETE GROUPS, AND AUTOMORPHIC FUNCTIONS

V. A. MALYŠEV

UDC 517.432 + 517.862

Abstract. Operators $A\,(l_1(Z_2^{++}) \to l_1(Z_2^{++}))$ of the form $(A\xi)(x) = \sum_{K\in Z_2^{++}} a(x-k)\,\xi(k)$, where $a \in l_1(Z_2)$ and Z_2 (Z_2^{++}) is the set of planar points with integral (nonnegative) coordinates, are considered. Basic results of the paper: invertibility of the operator A is proved, and an analysis is made of analytic properties of the symbol $F\xi$ of the solution of the equation $A\xi = \eta$.

Illustrations: 4. Bibliography: 16 titles.

§1. Introduction

Equations of the form

$$\sum_{k,l=0}^{\infty} a_{i-k,j-l}\xi_{kl} = \eta_{ij}, \quad i,j = 0, 1, 2, \ldots. \tag{1.1}$$

are examined in this paper. In this connection it is assumed that $\sum_{i,j=0}^{\infty} |\eta_{ij}| < \infty$, $\sum_{p,q=-\infty}^{\infty} |a_{pq}| < \infty$, and a solution is sought which also belongs to the space l_1 of sequences $\xi = \{\xi_{kl}\}_{k,l=0}^{\infty}$. Let us consider an operator A on the Banach space l_1 such that $A\xi = \xi'$, where

$$\xi'_{ij} = \sum_{k,l=0}^{\infty} a_{i-k,j-l}\xi_{kl}.$$

The following assertion distinguishes the case when A is a Noetherian operator ($\dim \operatorname{Ker} A < \infty$, $\dim \operatorname{Coker} A < \infty$).

Theorem 1.1. (I. B. Simonenko, [1], [2]). *The operator A is Noetherian if and only if the following two conditions are fulfilled*:

$$a(x,y) = \sum_{p,q=-\infty}^{\infty} a_{pq}x^p y^q \neq 0 \quad for \quad |x| = |y| = 1, \tag{1.2}$$

AMS 1970 *subject classifications*. Primary 47C05, 30A58; Secondary 20H10, 30A14.

$$\operatorname*{ind}_{|x|=1} a(x, 1) = \operatorname*{ind}_{|y|=1} a(1, y) = 0. \tag{1.3}$$

The following assertion is proved in the standard way (cf. [3]).

Lemma 1.1. *Under the conditions of Theorem* 1.1, ind $A = 0$.

The methods of proving these two assertions cannot yield conditions for invertibility of the operator nor, even less, analytic properties of the solution and its explicit representation.

A study will be made here of the possibility of obtaining an explicit representation of the solution of such a class of equations.

In the Noetherian case the theory for equations invariant under translation is classical when they are in the entire space and on a halfline (cf. [4], [5]), while in a half-space a solution is constructed with the aid, in addition, of the Wiener-Hopf ([6]) method. In all these cases there is a compact, explicit formula for the solution. A similar formula, based on Wiener-Hopf ideas, for equations in a quadrant of the plane can be obtained only in very special cases (cf. [15], [16]). The situation for the general case is extremely complicated. Solvability classes of equations arise in one way or another. In this paper a new approach to the study of such equations is set forth, and the foundations, in a sense complete, are laid for the theory of one class.

We shall say that the equation system (1.1) *is of type* $(n_1, n_2; m_1, m_2)$, where $-\infty \leq n_1, m_1 < \infty$, $-\infty < n_2, m_2 \leq \infty$, if a_{pq} can be nonzero only when

$$n_1 \leqslant p \leqslant n_2, \quad m_1 \leqslant q \leqslant m_2. \tag{1.4}$$

We shall say that type $(n_1, n_2; m_1, m_2)$ *is exact* if the region of type (1.4) cannot be contracted.

Remark 1.1. Equations of an arbitrary exact type $(n_1, n_2; m_1, m_2)$ exist which satisfy conditions (1.2) and (1.3). For example, it is possible to set $a(x, y) = a(x)\,\tilde{a}(y)$, where $x^{n_1} a(x)$ is a polynomial in x of degree $n_1 + n_2$ which has n_1 zeros inside the unit disk and n_2 outside it, $y^{m_1}\,\tilde{a}(y)$ is constructed analogously, and to obtain nontrivial examples it is sufficient to shuffle the coefficients of $a(x, y)$ a little.

If equation (1.1) is of one of the following types: $(0, \infty; -\infty, \infty)$, $(-\infty, \infty; 0, \infty)$, $(-\infty, 0; -\infty, \infty)$, $(-\infty, \infty; -\infty, 0)$, then it is solvable by the Wiener-Hopf method, as follows easily from [15]. Equations of type $(-1, \infty; -1, \infty)$ are considered in this paper.

In the case of equations (1.1) the Wiener-Hopf factorization method, upon application of which everything essentially reduces to the one-dimensional case, constitutes a necessary preliminary step of the analysis (§2). In §3 the invertibility of the corresponding operator in l_1 is proved. The first step of the new method (conversion to a Riemann surface) is considered in §§ 4 and 5. In the case of equations of type $(-1, 1; -1, 1)$ the genus of the corresponding Riemann surface equals 0 or 1.

Equations of this type are studied in detail in §§6–8. The basic construction for genus $g \geq 2$ is carried out in §9. It is associated with the construction of Galois automorphisms on a universal covering and the subsequent transfer of the equations by means of them to the universal covering. Essentially the technique of discrete groups of transformations of the Lobačevskiĭ plane is used here. In §10 the analytic behavior of a solution is explained. (The Riemann existence domain of the solution symbol is constructed.) In §§11 and 12 the solution of equations is reduced to the Carleman problem on a Riemann surface, from which an integral solution representation is obtained.

Certain important features of the method are preserved in the case of an arbitrary rational symbol.

§2. Factorization

Let us consider the ring R of series of the form $r(x, y) = \sum_{i,j=-\infty}^{\infty} r_{ij}x^i y^j$ which are absolutely convergent for $|x| = |y| = 1$. We shall consider the following subrings of the ring: $R_{++} = R_{+x+y}$, the set of functions $r(x, y)$ with $r_{ij} = 0$ if either $i < 0$ or $j < 0$; R_{+-}, the set of functions with $r_{ij} = 0$ if either $i < 0$ or $j \geq 0$; R_{-+}, R_{+x}, etc. Let us introduce projection operators on these subrings: P_{++}, P_{+-}, P_{-y}, $\cdots$. For example,

$$P_{\bar{y}}\left[\sum_{i,j=-\infty}^{\infty} r_{ij}x^i y^j\right] = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{-1} r_{ij}x^i y^j.$$

Lemma 2.1. *If* $r(x, y) \in R$, $r(x, y) \neq 0$ *for* $|x| = |y| = 1$ *and*

$$\operatorname*{ind}_{|x|=1} r(x, 1) = \operatorname*{ind}_{|y|=1} r(1, y) = 0, \tag{2.1}$$

then $\ln r(x, y) \in R$.

Proof. Conditions (2.1) guarantee the possibility of isolating a single-valued branch of $\ln r(x, y)$. Then, for example, if $r(x, y)$ is a piecewise smooth function, proof is carried out in an elementary way. For complete generality it is necessary to use Wiener's generalized theorem on locally analytic functions on the space of maximal ideals of the ring R (cf. [7]).

Now it is possible to set

$$\ln a(x, y) = \sum_{i,j=-\infty}^{\infty} b_{ij}x^i y^j,$$

$$a_{\underset{x}{+}}(x, y) = \exp[P_{\underset{x}{+}} \ln a(x, y)],$$

$$a_{\bar{x}}(x, y) = \frac{a(x, y)}{a_{\underset{x}{+}}(x, y)},$$

. . . .

Now let equation (1.1) have exact type $(-1, \infty; -1, \infty)$ with the conditions of Theorem 1.1 fulfilled. Let us set $A(x, y) = xya(x, y)$. Then for any y, $|y| = 1$, we have $\operatorname{ind}_{|x|=1} A(x, y) = 1$, and analogously for any x, $|x| = 1$, $\operatorname{ind}_{|y|=1} A(x, y) = 1$. Thus for $|y| = 1$ a single zero, hereafter denoted by $x_0(y)$, of the function $A(x, y)$ exists inside the unit disk. It is of first order. $y_0(x)$ is introduced analogously.

If $a(x, y)$ has type $(-1, n; -1, m)$, where $n, m < \infty$, then $A(x, y)$ is a polynomial. Let $A = A_1 A_2 \cdots A_k$ be its factorization into prime factors in the ring $\mathbf{C}[x, y]$.

Let, for example, $A_1(x_0(y), y) = 0$. Then the considerations mentioned above imply that $A_i(x_0(y), y) \neq 0$ for $i = 1$ and $|y| = 1$. Two cases are possible in this connection:

1) (the irreducibility case) for any x $(|x| = 1)$

$$A_1(x, y_0(x)) = 0;$$

2) (the reducibility case) for any x $(|x| = 1)$, $A_1(x, y_0(x)) \neq 0$. We shall assume $A_2(x, y_0(x)) = 0$ in this case.

The components of the factorization of $a(x, y)$ have the form

$$a_{\bar{x}}(x, y) = 1 - \frac{x_0(y)}{x}; \quad a_{\bar{y}}(x, y) = 1 - \frac{y_0(x)}{y}.$$

§3. Invertibility

Let us consider the following function symbols $(|x| = |y| = 1)$:

$$a(x, y) = \sum_{p,q=-\infty}^{\infty} a_{pq} x^p y^q, \quad \eta(x, y) = \sum_{i,j=0}^{\infty} \eta_{ij} x^i y^j,$$

$$\xi(x, y) = \sum_{k,l=0}^{\infty} \xi_{kl} x^k y^l.$$

Multiplying relations (1.1) by $x^i y^j$ and summing, we obtain

$$\eta(x, y) = \sum_{i,j=0}^{\infty} \sum_{k,l=0}^{\infty} a_{i-k,j-l} x^{i-k} y^{j-l} \xi_{kl} x^k y^l = \sum_{p,q} \sum a_{pq} x^p y^q \xi_{kl} x^k y^l;$$

where $p = i - k$, $q = j - l$, and the second sum is taken with respect to all nonnegative k and l under the conditions $p + k \geq 0$, $q + l \geq 0$.

Setting

$$b_{--}(x, y) = a_{-1,-1} \xi_{00} \frac{1}{xy},$$

$$b_{+-}(x, y) = \sum_{k=0}^{\infty} \xi_{k0} x^k \sum_{p \geq -k} a_{p,-1} x^p \frac{1}{y},$$

$$b_{-+}(x, y) = \sum_{l=0}^{\infty} \xi_{0l} y^l \sum_{q \geq -l} a_{-1,q} y^q \frac{1}{x},$$

we obtain

$$\eta(x, y) = a(x, y)\,\xi(x, y) - b_{+-}(x, y) - b_{-+}(x, y) - b_{--}(x, y). \tag{3.1}$$

Introducing the functions

$$\pi(x) = \sum_{k=0}^{\infty} \xi_{k0} x^k \sum_{p \geqslant -k} a_{p,-1} x^{p+1} + a_{-1,-1}\xi_{00},$$

$$\tilde{\pi}(y) = \sum_{l=0}^{\infty} \xi_{0l} y^l \sum_{q \geqslant -l} a_{-1,q} y^{q+1},$$

let us rewrite equation (3.1) in the following way:

$$A(x, y)\,\xi(x, y) - \pi(x) - \tilde{\pi}(y) = xy\eta(x, y). \tag{3.2}$$

We note that all addends in this equation belong to R_{++}.

Remark 3.1. If the function $\pi(x) = \Sigma_{k=0}^{\infty} \pi_k xk$ is known, then the ξ_{k0} are obtained recursively from the systemsof equations

$$\pi_0 = a_{-1,-1}\xi_{00},$$
$$\pi_1 = a_{0,-1}\xi_{00} + a_{-1,-1}\xi_{10},$$
$$\cdots\cdots\cdots\cdots\cdots$$

We obtain ξ_{0l} analogously if the function $\tilde{\pi}(y)$ is known.

Theorem 3.1. *Under the conditions of Theorem* 1.1 *the operator A corresponding to the equation* (1.1) *of exact type* $(-1, n; -1, m)$, *where* $1 \le n, m \le \infty$, *is invertible in the space* l_1.

By virtue of Lemma 1.1 it is sufficient to prove that $\operatorname{Ker} A = 0$. In §2 the function $x_0(y)$ is defined for any y, $|y| = 1$, such that $|x_0(y)| < 1$ and $a(x_0(y), y) = 0$. Then (3.2) implies

$$|\pi(x_0(y))| = |\tilde{\pi}(y)|. \tag{3.3}$$

Analogously

$$|\pi(x)| = |\tilde{\pi}(y_0(x))|. \tag{3.4}$$

For example, let

$$M = \max_{|x|=1} |\pi(x)| \geqslant \max_{|y|=1} |\tilde{\pi}(y)| = \tilde{M}.$$

Then the maximum modulus principle and relation (3.4) imply, firstly, $M = \tilde{M}$ and, secondly, $\tilde{\pi}(y) \equiv M$. Hence also $\pi(x) \equiv M$. But the explicit formula (3.1) for $\tilde{\pi}(y)$ then implies $M = 0$.

Corollary 3.1. *If the operators A and A_n, $n = 1, 2, \cdots$, are of type $(-1, \infty; -1, \infty)$, where for their symbols $a(x, y)$ and $a_n(x, y)$ we have*

$$\lim_{n\to\infty} \max_{|x|=|y|=1} |a(x, y) - a_n(x, y)| = 0,$$

then $\|A_n^{-1} - A^{-1}\| \to 0$. Moreover, if $\|A - A_n\| \cdot \|A\| < 1$, then

$$\|A_n^{-1} - A^{-1}\| \leqslant \frac{\|A - A_n\|}{1 - \|A - A_n\| \cdot \|A\|}.$$

Remark. Hence it follows that the solutions of an equation with an irrational kernel can be obtained by utilizing equations with rational kernels.

§4. Operation of projecting onto an analytic set

From here to the end of the paper it will be assumed that equation (1.1) is of arbitrary exact type $(-1, n; -1, m)$ with $1 \leq n, m < \infty$. Let us suppose that $A = \{(x, y): |x| < 1, |y| < 1, A(x, y) = 0\}$, and that D is the interior of the unit disk, and Γ its boundary. Let $\bar{A}$ be the closure of the principal analytic set A in $D \times D$, i.e.

$$\bar{A} = \{(x, y): |x| \leqslant 1, |y| \leqslant 1, A(x, y) = 0\}.$$

Then it is clear that a solution of equation (3.2) must satisfy the condition

$$xy\eta(x, y) + \pi(x) + \tilde{\pi}(y) = 0 \tag{4.1}$$

for $(x, y) \in \bar{A}$, particularly on the boundary $\bar{A} \cap [(D \times \Gamma) \cup (\Gamma \times D)]$. Let us prove the converse assertion.

Lemma 4.1. *If functions $\pi(x)$ and $\tilde{\pi}(y)$ exist which are continuous on the unit disk $\bar{D}$ and analytic inside it, and satisfy relation (4.1) on $\bar{A} \cap [(D \times \Gamma) \cup (\Gamma \times D)]$, then a solution of equation (3.2) is defined by the formula*

$$\xi(x, y) = \frac{\pi(x) + \tilde{\pi}(y) + xy\eta(x, y)}{xya(x, y)}. \tag{4.2}$$

Proof. (cf. [14] for another method). Introducing the functions

$$\pi_1(x) = \frac{\pi(x) - a_{-1,-1}\xi_{00}}{x} \text{ and } \tilde{\pi}_1(y) = \frac{\tilde{\pi}(y)}{y},$$

we rewrite equation (3.2) as

$$a(x, y)\xi(x, y) = \pi_1(x)\frac{1}{y} + \tilde{\pi}_1(y)\frac{1}{x} + \frac{a_{-1,-1}\xi_{00}}{xy} + \eta(x, y) = 0. \tag{4.3}$$

Hence

$$a_{\underset{x}{+}}(x, y)\xi(x, y) = \frac{1}{a_{\underset{x}{-}}(x, y)}\left[\frac{\pi_1(x)}{y} + \frac{\tilde{\pi}_1(y)}{x} + \frac{a_{-1,-1}\xi_{00}}{xy} + \eta(x, y)\right] \tag{4.4}$$

and

$$0 = P_{\underset{x}{-}}\left[\frac{1}{a_{\underset{x}{-}}(x, y)}\left(\frac{\pi_1(x)}{y} + \frac{\tilde{\pi}_1(y)}{x} + \frac{a_{-1,-1}\xi_{00}}{xy} + \eta(x, y)\right)\right]. \tag{4.5}$$

We obtain

$$0 = P_{\underset{y}{-}}\left[\frac{1}{a_{\underset{y}{-}}(x, y)}\left(\frac{\pi_1(x)}{y} + \frac{\tilde{\pi}_1(y)}{x} + \frac{a_{-1,-1}\xi_{00}}{xy} + \eta(x, y)\right)\right] \tag{4.6}$$

analogously. Conversely, if conditions (4.5) and (4.6) are fulfilled, then the function in the right side of (4.4) belongs to the ring R_{+_x}. Denoting this function by $\psi(x, y)$

and setting $\xi(x, y) = \psi(x, y)/a_{+_x}(x, y)$, we have formula (4.2). It remains for us only to show that the projection operations in formulas (4.5) and (4.6) are equivalent to the operation of projecting onto the analytic set A.

Lemma 4.2. *Relation* (4.1) *is fulfilled if and only if relations* (4.5) *and* (4.6) *are fulfilled.*

First let us prove the formula

$$P_{\bar{y}}\left[\frac{1}{1-\frac{y_0(x)}{y}}\,\omega(y)\right] = \omega(y_0(x))\frac{\frac{y_0(x)}{y}}{1-\frac{y_0(x)}{y}}, \tag{4.7}$$

where $\omega(y) \in R_{+_y}$. In fact, for any x, $|x| = 1$,

$$P_{\bar{y}}\left[\frac{1}{1-\frac{y_0(x)}{y}}\,y^i\right] = P_{\bar{y}}\left[\sum_{k=0}^{\infty}\left(\frac{y_0(x)}{y}\right)^k y^i\right] = y_0^i\sum_{k=1}^{\infty}\left(\frac{y_0(x)}{y}\right)^k.$$

Using the linearity of the operator P_{-_y} we obtain (4.7). Considering, for example, formula (4.6) now, we have

$$\frac{\pi_1(x)}{y}\cdot\frac{1}{1-\frac{y_0(x)}{y}} + \frac{\tilde{\pi}_1(y_0(x))}{x}\cdot\frac{\frac{y_0(x)}{y}}{1-\frac{y_0(x)}{y}} + \frac{a_{-1,-1}\xi_{00}}{xy\left(1-\frac{y_0(x)}{y}\right)}$$
$$+\,\eta(x, y_0(x))\frac{y_0(x)}{y}\cdot\frac{1}{1-\frac{y_0(x)}{y}} = 0,$$

which coincides with (4.1) for $|x| = 1$, $|y_0(x)| < 1$.

Lemma 4.1 implies that everything comes down to evaluating the functions $\pi(x)$ and $\tilde{\pi}(y)$. Let us cite a result of that kind here: let R_1 and R_2 be the Riemann surfaces (introduced below) of the functions $\pi(x)$ and $\tilde{\pi}(y)$, respectively, with given coverings $h_1\colon R_1 \to \mathbf{P}_x$ and $h_2\colon R_2 \to \mathbf{P}_y$, where $\mathbf{P}_x$ and $\mathbf{P}_y$ are one-dimensional projective spaces. Then $\xi(x, y)$ is defined naturally on $R_1 \times R_2$ and is a meromorphic function on this complex manifold.

§5. Equations on a Riemann surface

Hereafter it will be assumed, if not stipulated to the contrary, that $a(x, y)$ is rational and the irreducibility case (cf. §2) is realized. Let $A_1(x, y)$ be an irreducible polynomial for which $A_1(x_0(y), y) = A_1(x, y_0(x)) = 0$; let n be the degree of $A_1(x, y)$ in x; and m the degree of $A_1(x, y)$ in y.

Then the equation

$$A_1(x, y) = 0 \tag{5.1}$$

defines the algebraic functions $y(x)$ and $x(y)$, whose Riemann surfaces are connected, compact and conformally equivalent. Let us denote by S the corresponding abstract Riemann surface. As the Riemann surface of the algebraic function $y(x)$, S is realized in a natural way as a branched m-sheeted covering $h_1\colon S \to \mathbf{P}$ of the complex sphere $\mathbf{P}$.

The covering $h_2\colon S \dashrightarrow \mathbf{P}$ corresponding to the function $x(y)$ and having n sheets is defined analogously.

Let $\mathbf{C}_A(x, y)$ be the field of algebraic functions which is defined by equation (5.1). $\mathbf{C}_A(x, y)$ is a finite algebraic extension of the field $\mathbf{C}(x)$ of rational functions of x, as well as of the field $\mathbf{C}(y)$ of rational functions of y. $\mathbf{C}_A(x, y)$ is naturally isomorphic to the field $\mathbf{C}(S)$ of meromorphic functions on the Riemann surface S. The function $x(s)$, $s \in S$, which corresponds to the function $x(y)$ under this isomorphism possesses the following property: if $x(s_1) = x(s_2)$, $s_1, s_2 \in S$, then $h_2 s_1 = h_2 s_2$, and conversely. The function $y(s)$ corresponding to the function $y(x)$ possesses an analogous property.

It is obvious that $h_1^{-1}(\Gamma)$ is the boundary of the open set $h_1^{-1}(D)$, while $h_2^{-1}(\Gamma)$ is the boundary of $h_2^{-1}(D)$. If there are no branch points of h_i, $i = 1, 2$, on Γ, then $h_i^{-1}(\Gamma)$ consists of simple, analytic, closed, nonintersecting curves. Points of self-intersection and nondifferentiability can appear at branch points of h_i on $h_i^{-1}(\Gamma)$. In view of our assumptions, for $|x| = 1$ the set $h_1^{-1}(x) \cap h_2^{-1}(D)$ consists of one point. Consequently the intersection of $h_1^{-1}(\Gamma)$ and $h_2^{-1}(D)$ does not contain a branch point of the covering h_1.

Let us note that, by virtue of condition (1.2), $h_1^{-1}(\Gamma) \cap h_2^{-1}(\Gamma) = 0$; thus $h_2^{-1} \cap h_2^{-1}(D)$ is an analytic simple closed curve. Let us denote it by Γ_0.

We define $\Gamma_0 = h_2^{-1}(\Gamma) \cap h_1^{-1}(D)$ analogously. Let us set $G = h_1^{-1}(D) \cap h_2^{-1}(D)$. It is clear that $\Gamma_0 \cup \widetilde{\Gamma}_0$ is the boundary of G.

Lemma 5.1. *Let the genus of the Riemann surface S not be equal to zero, and let the irreducibility case be realized. Then G is a connected region.*

We shall show that if G is disconnected, then Γ_0 and $\widetilde{\Gamma}_0$ are homotopically equivalent to zero on S. The end of the proof is developed exactly like the proof of Case 2 of Lemma 7.1. Actually we have already observed that $h_1^{-1}(x) \cap h_2^{-1}(D)$, $|x| = 1$, consists of a single point. Furthermore, if G is disconnected, then it consists of two connected components bounded respectively by the curves Γ_0 and $\widetilde{\Gamma}_0$. The component of G which is bounded by the curve Γ_0 covers the unit disk $\{x\colon |x| \le 1\}$. By virtue of a remark made above, this covering is one-sheeted on the boundary. Consequently it is one-sheeted everywhere. Hence the assertion follows.

Let us denote by Δ the union of those connected components of the set $h_1^{-1}(D) \cup h_2^{-1}(D)$ which have a nonempty intersection with G. It is clear that under the conditions of Lemma 5.1 Δ is a connected region. Let $\Gamma_1, \Gamma_2, \cdots, \Gamma_k \subset h_1^{-1}(\Gamma)$ and $\widetilde{\Gamma}_1, \widetilde{\Gamma}_2, \cdots, \widetilde{\Gamma}_i \subset h_2^{-1}(\Gamma)$ be connected components of the boundary of region Δ. They are simple, closed, nonintersecting, analytic curves if there are no branch points of the coverings h_1 and h_2, respectively, on them.

Now let a function $f(x)$ which is meromorphic in some region D of the complex sphere $\mathbf{P}$ be given, as well as the Riemann surface S together with the branched covering $h\colon S \dashrightarrow \mathbf{P}$. Then the function $f(x)$ can be lifted to the region $h^{-1}(D)$ in the following way:

$$f_h(s) = f(hs), \ s \in h^{-1}(D).$$

In this connection $f_h(s)$ (denoted henceforth simply as $f(s)$) is a meromorphic function in $h^{-1}(D)$, which is verified, for example, by the theorem on removable singularities for points where the covering is branched.

Carrying $\pi(x)$ to $h_1^{-1}(D) \cap \Delta$ by means of h_1^{-1} and $\widetilde{\pi}(y)$ to $h_2^{-1}(D) \cap \Delta$ by means of h_2^{-1}, we obtain the functions $\pi(s)$ and $\widetilde{\pi}(s)$, which are defined respectively on $h_1^{-1}(D) \cap \Delta$ and $h_2^{-1}(D) \cap \Delta$, are analytic in these regions, and are continuous on the boundaries. Henceforth, as a means of conducting the analytical investigation of the equations, we shall everywhere, unless stipulated to the contrary, assume that $\eta(x, y)$ in equation (3.2) is a polynomial. Then the functions $x(s)$, $y(s)$ and $\eta(s) = \eta(x(s), y(s))$ can be assumed defined on all of S. On Γ_0 and $\widetilde{\Gamma}_0$ the functions $\pi(s)$ and $\widetilde{\pi}(s)$ satisfy one and the same condition:

$$\pi(s) + \widetilde{\pi}(s) = -x(s)y(s)\eta(s), \tag{5.2}$$

which is obtained by applying h_1^{-1} and h_2^{-1} to equation (3.2).

Equation (5.2) obviously admits analytic extension to the region G where the functions $\pi(s)$ and $\widetilde{\pi}(s)$ are defined simultaneously. By utilizing equation (5.2), the functions $\pi(s)$ and $\widetilde{\pi}(s)$ can be meromorphically extended to the regions $h_1^{-1}(D) \cap \Delta$ and $h_2^{-1}(D) \cap \Delta$, respectively. In otherwords, if equation (5.2) is satisfied in Δ, $\pi(s)$ and $\widetilde{\pi}(s)$ are meromorphic in that region. The converse assertion also is true.

Lemma 5.2. *If under the conditions of Lemma* 5.1 *the solutions $\pi(s)$ and $\widetilde{\pi}(s)$ of equation* (5.2) *exist, and are analytic in the regions $h_1^{-1}(D) \cap \Delta$ and $h_2^{-1}(D) \cap \Delta$, respectively, continuous on the boundary of these regions, and such that $\pi(s_1) = \pi(s_2)$ for $x(s_1) = x(s_2)$, and $\widetilde{\pi}(s_1) = \widetilde{\pi}(s_2)$ for $y(s_1) = y(s_2)$, then the functions $\pi(x) = \pi(x(s))$ and $\widetilde{\pi}(y) = \widetilde{\pi}(y(s))$ afford a solution of equation* (3.2).

In fact, define $\xi(x, y)$ by formula (4.2). Then, observing that equation (4.1) is fulfilled, it is sufficient to utilize Lemma 4.1 to complete the proof.

§6. Exact type (– 1, 1; – 1, 1). Genus 0.

In this and the following two sections we shall completely solve equation (1.1) of exact type (1 –, 1; – 1, 1). If the polynomial $A(x, y)$ is irreducible, the genus of the Riemann surface S equals zero or one. In the present section we shall examine the case of zero genus.

Let us represent $A(x, y)$ as $a(x)y^2 + b(x)y + c(x)$. Then in order that S have genus 0 it is necessary and sufficient that the discriminant $D(x) = b^2(x) - 4a(x)c(x)$ be of second degree in x or have a multiple root in $\mathbf{C}$.

Lemma 6.1. *There are no branch points on $h_i^{-1}(\Gamma)$ of the covering h_i, $i = 1, 2$, and the set $h_i^{-1}(\Gamma)$ is disconnected, i.e. an even number of branch points of both the covering h_1 and the covering h_2 are in D. Thus, since h_i is two-sheeted, $h_i^{-1}(\Gamma)$ consists of two nonintersecting, analytic, simple closed curves.*

Proof. If the set $h_1^{-1}(\Gamma)$ were connected, it would belong wholly to the set $h_2^{-1}(D)$ (since $h_1^{-1}(\Gamma) \cap h_2^{-1}(\Gamma) = 0$), which is impossible by virtue of condition (1.3).

Thus the following four cases for the relative position of $h_1^{-1}(D)$ and $h_2^{-1}(D)$ are possible.

A) Exactly two branch points of the covering h_i, $i = 1, 2$, are in D, i.e. $h_1^{-1}(D)$ and $h_2^{-1}(D)$ are connected (Figure 1).

B) There are no branch points of the coverings h_1 and h_2 in D, i.e. $h_1^{-1}(D)$ and $h_2^{-1}(D)$ are disconnected (Figure 2).

C) In D there are two branch points of h_1 and no branch points of h_2, i.e. $h_1^{-1}(D)$ is connected, and $h_2^{-1}(D)$ is disconnected.

D) In D there are two branch points of h_2 but no branch points of h_1, i.e. $h_2^{-1}(D)$ is connected and $h_1^{-1}(D)$ is disconnected (Figure 3).

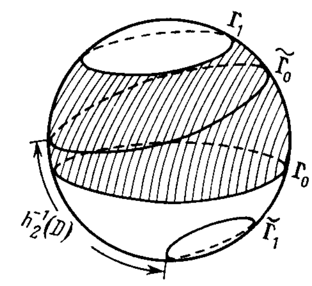

Figure 1

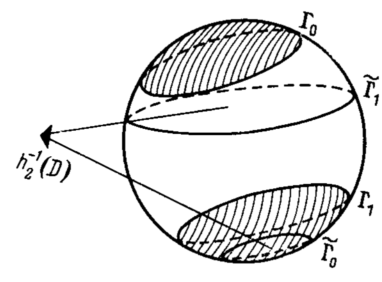

Figure 2

Let us cite examples of when these cases hold:

A) $a_{1,0} = a_{1,1} = a_{0,1} = 0$, B) $a_{-1,0} = a_{-1,-1} = a_{0,-1} = 0$,

C) $a_{10} = a_{1,-1} = a_{0,-1} = 0$, D) $a_{01} = a_{-1,1} = a_{-1,0} = 0$.

In this connection $a_{00} = -1$ everywhere, while the remaining a_{ij} are positive, their sum being less than unity. (For a probability analog of these examples cf. [14].)

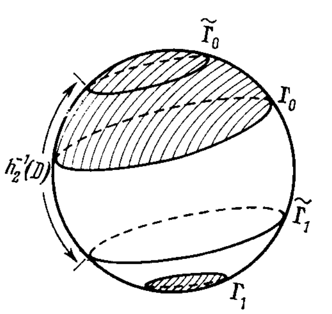

Figure 3

Theorem 6.1. *In cases* C) *and* D) *the solution* $\xi(x, y)$ *is a rational function.*

Proof. It is sufficient to prove that $\pi(x)$ and $\tilde{\pi}(y)$ are rational. For this, in turn, it is sufficient to show that one of the functions $\pi(s)$ or $\tilde{\pi}(s)$ can be meromorphically extended to the entire Riemann surface S. Then the other also proves to be meromorphic on S by virtue of the basic equation (5.2).

The basic method which we use here and later is the extension of meromorphic functions by means of Galois automorphisms. In the case of type $(-1, 1; -1, 1)$ the field $C_A(x, y) \sim C(S)$ is a Galois extension of both the field $C(x)$ of rational functions of x and the field $\mathbf{C}(y)$ of rational functions of y. The Galois group of the field $\mathbf{C}_A(x, y)$ over $\mathbf{C}(x)$ is a cyclic group of second order. Let us denote its nontrivial element by $\bar{\xi}$. Analogously, let $\bar{\eta}$ be the nontrivial element of the Galois group of the field $\mathbf{C}_A(x, y)$ over $\mathbf{C}(y)$. The conformal automorphisms ξ and η of the Riemann surface S correspond to the Galois automorphisms of the field $\mathbf{C}(S)$:

$$\tilde{\xi}f(s) = f(\xi s),\ \tilde{\eta}f(s) = f(\eta s),\ f \in \mathbf{C}(S).$$

The explicit form of the Galois automorphisms ξ and η in the case of exact type $(-1, 1; -1, 1)$ is given by the formulas

$$\xi y = \frac{a_{1,-1}x^2 + a_{0,-1}x + a_{-1,-1}}{y\,(a_{11}x^2 + a_{01}x + a_{-1,1})},\quad \eta x = \frac{a_{-1,1}y^2 + a_{-1,0}y + a_{-1,-1}}{x\,(a_{11}y^2 + a_{10}y + a_{1,-1})}. \tag{6.0}$$

Let us pass to the proof of Theorem 6.1 in case C). We denote by $\tilde{D}_0$ and $\tilde{D}_1$ the components of $h_2^{-1}(D)$ which are bounded respectively by $\tilde{\Gamma}_0$ and $\tilde{\Gamma}_1$. The region $\Delta = h_1^{-1}(D) \bigcup \tilde{D}_0$ is connected, and, just as in §5, we obtain the result that $\pi(s)$ and $\tilde{\pi}(s)$ are meromorphic in Δ. Now let us observe that $\xi\Gamma_0 = \Gamma_1$. Therefore $\xi(\Delta\backslash h_1^{-1}(D)) = S\backslash\Delta$, and then $\Delta \bigcup \xi\Delta = S$. In view of the invariance of $\pi(s)$ relative to ξ, the latter relation implies that $\pi(s)$ is extendable to all of S. Case D) is treated analogously.

Explicit solution form in cases C) *and* D).

In case C) denote by $s_1, \cdots, s_k$ $(k \leq 4)$ the poles in the region $\Delta\backslash h_1^{-1}(D)$ of the function $-xy\eta(x, y)$ in the right side of formula (5.2), and by $\phi(s - s_i)$ the principal parts of this function at the points s_i. (We assume that s is a uniformizing variable on S.) Then

$$\pi(s) = \sum_{i=1}^{k} \varphi_i(s - s_i) + \sum_{i=1}^{k} \xi\varphi_i(s - s_i), \tag{6.1}$$

$$\tilde{\pi}(s) = -x(s)\,y(s)\,\eta(s) - \pi(s). \tag{6.2}$$

Actually, $\pi(s)$ must be analytic in $h_1^{-1}(D)$, but it can have poles in $\Delta\backslash h_1^{-1}(D)$ since $\tilde{\pi}(s)$ is analytic in this region. Since $h_1^{-1}(D)$ is connected, $\pi(s)$ is invariant relative to ξ (which is exploited in the proof of Theorem 6.1), implying formula (6.1). Let us note that $\tilde{\pi}(s)$ is not constrained to be invariant relative to η.

Moreover, $\pi(s)$ is evaluated in formula (6.1), of course, with accuracy up to an additive constant, which is determined from the condition that $\tilde{\pi}(y) = 0$ for $y = 0$.

Let us now consider case B).

Theorem 6.2. *Let D_1 be one of the connected components of the region $h_1^{-1}(D) \bigcup h_2^{-1}(D)$, and let the functions $\pi(s)$ and $\tilde{\pi}(s)$, which are defined on this component, satisfy equation* (5.2). *Then $\pi(s)$ and $\tilde{\pi}(s)$ are meromorphically extendable to all of S with the deleted point s_1, which is a fixed point of the automorphism $\xi\eta$ and does not belong to D_1. Moreover* (D_1 *being bounded by* Γ_1):

$$\pi(s) = -x(s)y(s)\eta(s) - \tilde{\pi}(s), \tag{6.3}$$

$$\widetilde{\pi}(s) = \sum_{k=1}^{\infty} [A(\xi\delta^{-k}s) - A(\delta^{-k}s)] + \text{const},$$

$$A(s) = -x(s)y(s)\eta(s), \tag{6.4}$$

the series in the right side of formula (6.4) being absolutely convergent in any bounded part of the plane $S\backslash\{s_1\}$ except for a finite number of points in that part which are poles of one of the terms of the series.

Proof. First let us observe that the group κ of automorphisms of S which is generated by ξ and η is a subgroup of the group of bilinear transformations on S. Let us denote by D_0, D_1, $\widetilde{D}_0$, $\widetilde{D}_1$ regions belonging to $h_1^{-1}(D) \cup h_2^{-1}(D)$ and bounded by Γ_0, Γ_1, $\widetilde{\Gamma}_0$, $\widetilde{\Gamma}_1$, respectively. We have $D_0 \subset \widetilde{D}_1$, $\widetilde{D}_0 \subset D_1$, $\xi D_0 = D_1$ and $\eta\widetilde{D}_0 = \widetilde{D}_1$ (cf. Figure 2). Let us construct the two sequences of regions

$$D_0 \subset \widetilde{D}_1 \subset \eta D_1 \subset \eta\xi\widetilde{D}_1 \subset \ldots \subset \eta(\xi\eta)^k D_1 \subset (\eta\xi)^{k+1}\widetilde{D}_1 \subset \ldots,$$

$$\widetilde{D}_0 \subset D_1 \subset \xi\widetilde{D}_1 \subset \xi\eta D_1 \subset \ldots \subset \xi(\eta\xi)^k\widetilde{D}_1 \subset (\xi\eta)^{k+1}D_1 \subset \ldots$$

inductively. The membership relations in these chains are easily verified by induction.

Now let us observe that the transformation $\xi\eta$ (and consequently also $(\xi\eta)^n$) is either hyperbolic or loxodromic. In fact, under a parabolic or an elliptic transformation the image of a bounded, connected, open set (say D) cannot strictly contain its preimage. Moreover, the point s_1, one of the fixed points of the transformation $\xi\eta$, belongs to D_0, while a second, s_2, belongs to $\widetilde{D}_0$; $\xi s_1 = \eta s_1 = s_2$. This implies that

$$\bigcup_{k=1}^{\infty} (\eta\xi)^k \widetilde{D}_1 = S\backslash\{s_2\} \text{ and } \bigcup_{k=1}^{\infty} (\xi\eta)^k D_1 = S\backslash\{s_1\}.$$

Let us denote the lifting of the functions $\pi(x)$ and $\widetilde{\pi}(y)$ to D_1 by $\pi(s)$ and $\widetilde{\pi}(s)$, and to $\widetilde{D}_1$ by $\Pi(s)$ and $\widetilde{\Pi}(s)$, respectively. Then we have $\pi(\xi s) = \Pi(s)$ and $\widetilde{\pi}(\eta s) = \Pi(s)$, $s \in \widetilde{D}_1$. Exploiting these relationships and equation (5.2), we extend $\pi(s)$ and $\widetilde{\pi}(s)$ meromorphically by induction to the region $(\xi\eta)^k D_1$ and Π and $\widetilde{\Pi}$ to $(\eta\xi)^k\widetilde{D}_1$, whence the first assertion of the theorem is obtained.

To obtain an explicit solution representation in case B), let us consider the equation

$$\pi(s) + \widetilde{\pi}(s) = A(s),$$

or

$$\Pi(\xi s) + \widetilde{\pi}(s) = A(s). \tag{6.5}$$

But

$$\Pi(\xi s) + \widetilde{\Pi}(\xi s) = A(\xi s).$$

or

$$\Pi(\xi s) + \widetilde{\pi}(\eta\xi s) = A(\xi s). \tag{6.6}$$

Comparing equations (6.5) and (6.6), we obtain

$$\widetilde{\pi}(\eta\xi s) - \widetilde{\pi}(s) = A(\xi s) - A(s). \tag{6.7}$$

The series

$$\tilde{\pi}(s)=\sum_{k=1}^{\infty}[A(\xi\delta^{-k}s)-A(\delta^{-k}s)]+\text{const},$$

where $\delta = \eta\xi$, is the fomal solution of the latter. Let us prove that this series defines a meromorphic function on $S\backslash\{s_1\}$. It is sufficient to show that this series converges at any point different from a pole of any of the terms of this series. Taking $\tilde{\pi}(s_2) = 0$, let us rewrite relation (6.7) as follows:

$$\tilde{\pi}(s)-\tilde{\pi}(\delta^{-1}s)=A(\xi\delta^{-1}s)-A(\delta^{-1}s),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\tilde{\pi}(s)-\tilde{\pi}(\delta^{-k}s)=\sum_{i=1}^{k}[A(\xi\delta^{-i}s)-A(\delta^{-i}s)].$$

For any $s \neq s_1$ we have $\lim_{k\to\infty}\tilde{\pi}(\delta^{-k}s) = \tilde{\pi}(s_2) = 0$. Therefore the series in the right side of (6.4) converges.

Let us observe further that $A(\xi s) - A(s)$ does not have poles in the region $\tilde{D}_0$. For $A(s)$ this follows from the fact that neither π nor $\tilde{\pi}$ can have poles in $\tilde{D}_0$, and for $A(\xi s)$, from the fact that neither Π nor $\tilde{\Pi}$ can have poles in the region $D_0 = \xi D_1$. Let us set $s_2 = 0$, $s_1 = \infty$. Now, if $B(s) = A(\xi s) - A(s)$, then $B(0) = 0$, and since $B(s)$ does not have poles at the points 0 and ∞, it can be represented as a linear combination of terms of the form $cs^m/(b-s)^n$, $n > m \geq 1$. But then the meromorphy of the series $\Sigma_{k=0}^{\infty}B(r^{-k}s)$, $|r| > 1$, is obvious.

We obtain

$$\Pi(s)=\sum_{k=1}^{\infty}[A(\eta\tilde{\delta}^{-k}s)-A(\tilde{\delta}^{-k}s)]+\text{const},\quad \tilde{\delta}=\xi\eta$$

symmetrically. Now let us set $\pi(s) = \Pi(\xi s)$. Next, verification of relationship (5.2) is effected in the obvious way.

Case A) is on the whole treated analogously to the preceding one. Therefore we shall confine ourselves to formulating the following theorem.

Theorem 6.3. *Let us denote by* s_1 *and* s_2 *points possessing the property* $\xi s_1 = \eta s_1 = s_2$ (*or* $A(x(s), y(s)) = A_x(x(s), y(s)) = A_y(x(s), y(s))$). *Then* $\pi(s)$ *and* $\tilde{\pi}(s)$ *are meromorphically extendable to all of* S *with deleted points* s_1 *and* s_2.

Let us also note the following interesting fact with respect to example A) cited above with $a_{10} = a_{11} = a_{01} = 0$. In this case $\pi(x)$ and $\tilde{\pi}(y)$ are rational (and what is more, they are polynomials in x and y, respectively). In fact, the equation

$$\pi(x)+\tilde{\pi}(y)=x^ny^m \bmod (a_{-1,-1}+a_{-1,0}y+a_{0,-1}x+a_{00}xy+a_{-1,1}y^2+a_{1,-1}x^2)$$

is solvable for any $m, n \geq 0$. Let us carry out induction on $k = n + m$. Let this assertion be proved for all n, m such that $n + m < k$. Let us prove that then x^ny^m, $n+m=k$, can be expressed linearly mod $A(x, y)$ by x^k, y^k and terms x^ny^m with $n + m < k$ (hereafter denoted by $o(k)$). To this end let us write down

$$a_{1,-1}x^{n+1}y^{k-n-1}+a_{00}x^ny^{k-n}+a_{-1,1}x^{n-1}y^{k-n+1}=o(k). \tag{6.8}$$

For given $n = n_0$, besides relation (6.8) for $n = n_0$, from the relations (6.8) for $n > n_0$ it is possible to obtain

$$x^{n_0}y^{k-n_0} + c_1x^{n_0-1}y^{k-n_0+1} = o(k) + y^k$$

and from the relations (6.8) for $n < n_0$,

$$x^{n_0}y^{k-n_0} + c_1x^{n_0+1}y^{k-n_0-1} = o(k) + x^k.$$

The assertion follows immediately from relation (6.8) and the last two relations. The special case arising for $n = 2, 3$ is also treated without difficulty.

§7. Exact type $(-1, 1; -1, 1)$. Genus 1.

If S has genus 1, there are four branch points of each of the coverings h_1 and h_2 on $\mathbf{P}$ (where $\mathbf{P}$ is the complex sphere).

Lemma 7.1. *The covering h_i $(i = 1, 2)$ has exactly two branch points inside D and exactly two outside D. Moreover, for any i, $h_i^{-1}(\Gamma)$ consists of two nonintersecting, analytic, simple closed curves.*

Proof. That inside D there are an even number of branch points of h_i, as well as the second part of the lemma, is proved exactly like Lemma 6.1.

If there is not a single branch point of h_1 inside D, both components of $h_1^{-1}(\Gamma)$, denoted by Γ_0 and Γ_1, bound contractible regions D_0 and D_1, respectively, on a torus. In this connection there can be the following possibilities for the number of branch points of h_2 inside D.

1. Two branch points of h_2 are inside D. But then, as the construction of the Riemann covering surface $h_2: S \to \mathbf{P}$ implies, the curves $\widetilde{\Gamma}_0$ and $\widetilde{\Gamma}_1$ are not homotopic to zero, and neither one of them can belong as a whole to D_0 or D_1.

2. Not a single branch point is inside D. Then

$$D_0 \subset \widetilde{D}_1, \quad \widetilde{D}_0 \subset D_1, \quad \xi D_0 = D_1, \quad \eta \widetilde{D}_0 = \widetilde{D}_1. \tag{7.0}$$

Let us introduce on the torus the Riemann metric induced by the metric on a universal covering. Then the Galois automorphisms ξ and η must be area preserving, which contradicts relations (7.0).

3. Four branch points are inside D. Then $\widetilde{\Gamma}_0$ and $\widetilde{\Gamma}_1$ are also homotopic to zero and bound the contractible regions $\widetilde{D}_0'$ and $\widetilde{D}_1'$, where $\widetilde{D}_0' \cup \widetilde{D}_1' = S \backslash h_2^{-1}(D)$. Analogously it must be the case that $\widetilde{\Gamma}_0 \subset D_0$, $D_1 \subset \widetilde{D}_1'$, and consequently $\widetilde{D}_0' \subset D_0$. Now we pass to the contradiction just as in case 2.

The case remains when h_1 and h_2 have four branch points inside D. Then, in notation which is obvious, it must be the case that $D_1' \subset D_0$, $D_1 \subset D_0'$, and so on, analogously to case 2.

Remark 7.1. The proof could be carried out on a universal covering. The proof of Lemma 5.1 can be effected in an analogous way by taking into account the constructions of §9.

Corollary 7.1. *The curves Γ_0, Γ_1, $\widetilde{\Gamma}_0$ and $\widetilde{\Gamma}_1$ are homotopic to the same element of the normal homology basis on the torus.*

Thus the arrangement of the regions $h_1^{-1}(D)$ and $h_2^{-1}(D)$ has the form represented in Figure 4.

Proof. That Γ_0 and Γ_1 are homotopic to the same element of the normal homology basis on the torus is obvious from the construction of the Riemann surface coverings. An analogous assertion is true for $\tilde{\Gamma}_0$ and $\tilde{\Gamma}_1$. Let us prove that Γ_0 is homotopic to $\tilde{\Gamma}_0$. But this follows from the fact that Γ_0 does not intersect $\tilde{\Gamma}_0$, and consequently cannot be homotopic to a *different* element of the normal homology basis on the torus.

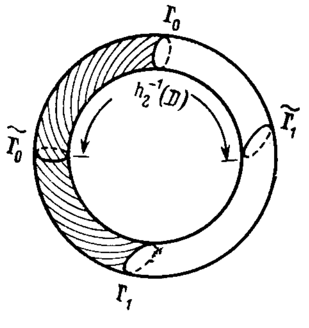

Figure 4

Let us recall that just as in the case of genus 0, the Galois groups of the extensions $C_A(x, y)$ over $C(x)$ and $C_A(x, y)$ over $C(y)$ are cyclic of second order, their explicit form being given by the formulas (6.0).

The complex plane C is a universal covering for S. In this connection let the covering (unbranched) λ: $C \to S$ be fixed. From uniformization theory it is known that S can be considered as a complex Lie group which is the factor group of the additive group C modulo the discrete subgroup $\{n\omega_1 + m\omega_2\}$, where the periods ω_1 and ω_2 are linearly independent over the field of real numbers R, n and m being integers.

Under the covering λ any segment of length $|\omega_i|$ parallel to the vector ω_i is projected onto S as a closed curve whose homology class is one of the elements of the normal basis on S. We assume without loss of generality that $\lambda([0, \omega_1])$ is homologous to Γ_0, and consequently, by Corollary 7.1, also to all of Γ_1, $\tilde{\Gamma}_0$ and $\tilde{\Gamma}_1$. Let us consider the strip

$$\Pi = \{\omega: \omega = \mu\omega_1 + \nu\omega_2;\ \mu, \nu \in \mathbf{R}, 0 \leqslant \mu < 1\}.$$

The preimage $\lambda^{-1}\Delta$ will consist of an even number of connected curvilinear strips translated relatively to each other by vector multiples of ω_2, while $\lambda^{-1}\Delta \cap \Pi$ consists of a connected region $(\lambda^{-1}\Delta)_0 \cap \Pi$ and its translations $[(\lambda^{-1}\Delta)_0 \cap \Pi] + n\omega_2$, where n is any integer. Hereafter we shall hold the region $(\lambda^{-1}\Delta)_0$ fixed and denote it by Δ_0. $\Delta_0 \cap \Pi$ is one of the connected components of $\lambda^{-1}\Delta \cap \Pi$ in the strip Π. (Hereafter, wherever it will not lead to ambiguity, preimages of curves and regions in Δ_0 will be denoted without the symbol λ^{-1}.)

Any function defined on the region Δ can be lifted to the region Δ_0 by the formula

$$f_\lambda(\omega) = f(\lambda\omega), \quad \lambda\omega \in \Delta.$$

Let us transfer the functions $x(s)$, $y(s)$, $\pi(s)$ and $\tilde{\pi}(s)$ to Δ_0 in this way. In this connection the functions $x(\omega)$, $y(\omega)$ and $\eta(\omega) = \eta(x, y)$ prove to be elliptic with periods ω_1 and ω_2 and meromorphic on all C. (The index λ will in the future be omitted for functions transferred in this manner.) Moreover, $\pi(\omega)$ and $\tilde{\pi}(\omega)$ are in fact defined on all of Δ_0, where they satisfy the equations

$$\pi(\omega+\omega_1)=\pi(\omega), \tag{7.1}$$

$$\tilde{\pi}(\omega+\omega_1)=\tilde{\pi}(\omega), \tag{7.2}$$

$$\pi(\omega)+\tilde{\pi}(\omega)=-x(\omega)y(\omega)\eta(\omega). \tag{7.3}$$

Let x_1 and x_2 be the branch points of the algebraic function $y(x)$ which are inside the unit disk. Let us denote the preimages of the points $h_1^{-1}(x_1)$ and $h_1^{-1}(x_2)$ in the region Δ_0 by a_1 and a_2. An arbitrary conformal automorphism $\tilde{\zeta}$ of the Riemann surface S can be extended to the conformal automorphism $\zeta=\lambda^{-1}\tilde{\zeta}\lambda$ of the universal covering (cf. [10]). This extension, of course, is not single-valued, but it becomes single-valued if the image of some point $\omega\in\mathbf{C}$ is fixed under the automorphism ζ. (This image must belong to the set $\{\lambda^{-1}\tilde{\zeta}\lambda\omega\}$.) As for the Galois automorphism ξ, we therefore require that the point a_1 be a fixed point of this automorphism: $\xi a_1=a_1$. Then $\xi\omega=-\omega$ in the coordinate system with origin at the point a_1. Indeed, it is well known that the very general conformal automorphism ζ of the complex plane $\mathbf{C}$ is representable as $\zeta\omega=a\omega+b$. Since the coordinate origin is a fixed point, $b=0$. But $\xi^2=1$, whence $a^2=1$ and $a=-1$.

Let us define analogously both $b_i=\Delta_0\cap\Pi\cap\{\lambda^{-1}h_2^{-i}(y_i)\}$, $i=1, 2$, where y_1, y_2 are the branch points of the algebraic function $x(y)$ which are inside the unit disk, and the lifting of the Galois automorphism η to the universal covering in the coordinate system with origin at the point b_1: $\eta\omega=-\omega$.

Lemma 7.2.

$$\xi\eta\omega=\omega+\omega_3,\ \omega_3=2(a_1-b_1),$$

i.e. the composition of the Galois automorphisms is a translation by a vector equal to twice the distance between their fixed points.

Remark 7.2. Analogously,

$$a_1-a_2=\pm\frac{\omega_1}{2},\qquad b_1-b_2=\pm\frac{\omega_1}{2}.$$

Theorem 7.1. *The functions $\pi(\omega)$ and $\tilde{\pi}(\omega)$ are meromorphically extendable to the entire universal covering.*

Proof. In view of relations (7.1) and (7.2), it is sufficient to confine ourselves to the strip Π or to consider the noncompact Riemann surface S' which is the factor of $\mathbf{C}$ modulo $\{n\omega_1\}$ and topologically equivalent to an infinite cylinder. Since $\eta\Gamma_0\subset h_2^{-1}(D)$, we have $\eta\xi\Gamma_1\subset\Delta_0$, and consequently $\Delta_0\cap(\eta\xi)\Delta_0\neq 0$. In view of the fact that $\eta\xi$ is a translation by $-\omega_3$, the union of the regions $(\eta\xi)^n\Delta_0$, $n=\cdots,-1,0,1,\cdots$, covers the entire plane $\mathbf{C}$. Hence it is possible to show that $\pi(\omega)$ and $\tilde{\pi}(\omega)$ are extendable along any path on $\mathbf{C}$ and by means of the monodromy theorem to complete the proof of the theorem.

Let us consider the Riemann surface $\tilde{S}$, which is the factor of $\mathbf{C}$ modulo $\{n\omega_1+m\omega_3\}$, i.e. the region $\tilde{\Delta}_0$ between Γ_1 and $\eta\Gamma_0$ which belongs to $\Delta_0\cap\Pi$ and has identified boundaries. Hereafter we shall identify Γ_1 and $\eta\Gamma_0$. Let us define the piecewise analytic function

$$\Pi(\omega) = \begin{cases} \pi(\omega), & \omega \in h_1^{-1}(D)\setminus\overline{G}, \\ -\widetilde{\pi}(\omega), & \omega \in \widetilde{\Delta}_0 \cap h_2^{-1}(D) \end{cases}$$

on $\widetilde{S}$. We thus obtain Riemann's boundary value problem on the Riemann surface S (cf., for example, [8]): find the piecewise analytic function $\Pi(\omega)$ with the jumps

$$\begin{aligned} &\Pi_+(\omega) - \Pi_-(\omega) = x(\xi\omega)\, y(\xi\omega)\, \eta(\xi\omega), \quad \omega \in \Gamma_1; \\ &\Pi_+(\omega) - \Pi_-(\omega) = -x(\omega)\, y(\omega)\, \eta(\omega), \quad \omega \in \widetilde{\Gamma}_0. \end{aligned} \tag{7.4}$$

(We choose the direction of traversal of $\Gamma_1 = \eta\Gamma_0$ and $\widetilde{\Gamma}_0$ on $\widetilde{S}$ in accordance with the direction of ω_1 and denote by Π_+ values from the left, and by Π_- values from the right, on the respective curves.)

Now let us consider the Weierstrass ζ-function

$$\zeta(u) = \frac{1}{u} + \sum_{n^2+m^2\neq 0} \left[\frac{1}{(u - n\omega_1 - m\omega_3)} + \frac{1}{n\omega_1 + m\omega_3} + \frac{u}{(n\omega_1 + m\omega_3)^2}\right].$$

In this connection let us choose the coordinate origin at the point a_1; then $\zeta(u)$ is odd relatively to a_1, a_2, b_1 and b_2.

Theorem 7.2.

$$\Pi(\omega) = \frac{1}{2\pi i} \int_{\Gamma_1 \cup \widetilde{\Gamma}_0} \zeta(\omega - \tau)\, A(\tau)\, d\tau. \tag{7.5}$$

In addition, if $\omega \in h_1^{-1}(D)\setminus\overline{G}$, *then*

$$\pi(\omega) = \frac{1}{2\pi i} \int_{\Gamma_0} [-\zeta(\omega - \tau) + \zeta(\omega + \tau)]\, x(\tau)\, y(\tau)\, \eta(\tau)\, d\tau, \tag{7.6}$$

and the function $A(\tau)$ *is defined by the formula*

$$A(\tau) = \begin{cases} x(\xi\tau)\, y(\xi\tau)\, \eta(\xi\tau), & \tau \in \Gamma_1, \\ -x(\tau)\, y(\tau)\, \eta(\tau), & \tau \in \widetilde{\Gamma}_0. \end{cases}$$

Proof. The function $\Pi(\omega)$ defined by expression (7.5) is piecewise analytic and has discontinuities on the curves $\Gamma_1 + n\omega_3$ and $\widetilde{\Gamma}_0 + n\omega_3$, at which the jumps conform to formulas (7.4), which we easily verify by utilizing the Sohockiĭ formula. Thus it is sufficient to prove (with respect to the possibility of identifying Γ_1 and $\eta\Gamma_0$) that $\Pi(\omega)$ has periods ω_1 and ω_3. For this, in turn, since $\zeta(\omega + \omega_1) = \zeta(\omega) + \eta_1$ and $\zeta(\omega + \omega_3) = \zeta(\omega) + \eta_3$, where η_1 and η_3 are constants, it is sufficient to prove that the first integral representation (formula (7.5)) implies the second (formula (7.6)), because the kernel of the second integral representation is an elliptic function. Having noted that $r(\tau) = x(\tau)y(\tau)\eta(\tau)$ does not have poles on G, we have

$$\int_{\widetilde{\Gamma}_0} \zeta(\omega - \tau)\, A(\tau)\, d\tau = -\int_{\widetilde{\Gamma}_0} \zeta(\omega - \tau)\, r(\tau)\, d\tau = -\int_{\Gamma_0} \zeta(\omega - \tau)\, r(\tau)\, d\tau,$$

$$\int_{\Gamma_1} \zeta(\omega - \tau)\, r(\xi\tau)\, d\tau = -\int_{\Gamma_1} \zeta(-\omega - \xi\tau)\, r(\xi\tau)\, d\tau =$$

$$= -\int_{\Gamma_0} \zeta(-\omega-\tau) r(\tau)\, d\tau = \int_{\Gamma_0} \zeta(\omega+\tau) r(\tau)\, d\tau.$$

The solution of problem (7.4) is unique with accuracy up to a constant because the difference of two similar solutions is analytic on all of $\widetilde{S}$. This constant is determined from the condition $\widetilde{\pi}(y(s)) = 0$, $y(s) = 0$.

An expression of the form

$$\pi(x) = \int_{|x|=1} A(x, x')\, x y_0(x)\, \eta(x, y_0(x))\, dx$$

is easily obtainable from formula (7.6), where the kernel $A(x, x')$ can be given in explicit form by making a change of variables in (7.6).

§8. Exact type $(-1, 1; -1, 1)$. Reducible case

Let $A(x, y)$ be reducible. Then with accuracy up to a permutation of x and y the following cases are possible.

1) $A(x, y) = A(x)\widetilde{A}(y)$, where $A(x)$ and $\widetilde{A}(y)$ depend only on x and y, respectively. This case is easily solved by the factorization method (cf. [15]).

2) $A(x, y) = B(x, y)C(y)$, where $C(y)$ depends only on y and is of first degree in y. In this case also we apply the factorization method. In fact, if $C(y) = y + c$, where $|c| > 1$, then $B(x, y) = b_1(x)y + b_2(x)$, and a factorization can be chosen so that $a_{-+}(x, y) = 1$. Hence the assertion follows (cf. [15]). The case of $|c| < 1$ is treated analogously.

3) $A(x, y) = A_1(x, y)A_2(x, y)$, where A_1 and A_2 are first degree polynomials in both x and y.

First let

3A) $A_1(x_0(y), y) = A_1(x, y_0(x)) = 0$. If $A_1(x, y) = (cx + d)y + ax + b$, then this is possible, for example, when $|c|$ is large compared to the moduli of the remaining coefficients.

Lemma 8.1. *In case* 3A) $\pi(x)$ *and* $\widetilde{\pi}(y)$ *are rational.*

Proof. Let us consider the Riemann surface S of genus 0 which is defined by the equation $A_1(x, y) = 0$, with one-sheeted, unbranched coverings $h_1\colon S \to \mathbf{P}$ and $h_2\colon S \to \mathbf{P}$. The regions $h_i^{-1}(D)$ are one-sheeted coverings of D, and since $h_1^{-1}(\Gamma) \subset h_2^{-1}(D)$ and $h_2^{-1}(\Gamma) \subset h_1^{-1}(D)$, we have $h_1^{-1}(D) \cup h_2^{-1}(D) = S$. Hence the rationality of π and $\widetilde{\pi}$ follows.

Explicit form of the solution in case 3A).

The function $-xy\eta(x, y)$ must have exactly two poles on S, one, $y(s) = y_0$, being in the region $h_1^{-1}(D)$ and the second, $x(s) = x_0$, being in the region $h_2^{-1}(D)$. Denoting the principal parts at these poles by $\widetilde{\pi}(y)$ and $\pi(x)$, respectively, we shall have a solution with accuracy up to constants determined, for example, by the condition $\widetilde{\pi}(0) = 0$.

3B) $A_1(x_0(y), y) = A_2(x, y_0(x)) = 0$. In this case let us consider two Riemann surfaces: S_1, defined by the equation $A_1(x, y) = 0$, with coverings $h_{11}: S_1 \to \mathbf{P}$ and $h_{12}: S_1 \to \mathbf{P}$; and S_2, defined by the equation $A_2(x, y) = 0$, with coverings $h_{21}: S_2 \to \mathbf{P}$ and $h_{22}: S_2 \to \mathbf{P}$. Obviously we have $h_{12}^{-1}(D) \subset h_{11}^{-1}(D)$ and $h_{21}^{-1}(D) \subset h_{22}^{-1}(D)$. Let us denote the lifting of $\pi(x)$ and $\tilde{\pi}(y)$ to S_1 by $\pi(s)$ and $\tilde{\pi}(s)$, and to S_2 by $\Pi(t)$ and $\tilde{\Pi}(t)$, respectively. Naturally, $h_{11}^{-1}(D)$ and $h_{21}^{-1}(D)$ are conformally equivalent (as preimages of D). Let us denote this isomorphism by $\xi: h_{11}^{-1}(D) \to h_{21}^{-1}(D)$. The isomorphism $\eta: h_{22}^{-1}(D) \to h_{12}^{-1}(D)$ is introduced analogously. We shall assume ξ and η extended to the conformal isomorphisms $\xi: S_1 \to S_2$ and $\eta: S_2 \to S_1$.

Just as in case B) of §6, $\xi\eta$ proves to be hyperbolic or loxodromic, and a sequence of regions

$$D_1 = h_{11}^{-1}(D) \subset D_2 \subset \ldots \subset D_k \subset \ldots$$

is constructed, where $D_k = (\eta\xi)^k D_1$.

Moreover, *the functions $\pi(s)$ and $\tilde{\pi}(s)$ are meromorphically extendable to $S_1 \setminus \{s_1\}$, where s_1 is a fixed point of the automorphism $\eta\xi$ which is outside D_1*. The explicit form of the functions is obtained analogously to case B) of §6, and we shall not write it out.

§9. Construction of Galois automorphisms on a universal covering

It will always be assumed hereafter that the irreducible case is realized (cf. §2) and the genus of the Riemann surface S is $g \geq 2$. The terminology and results of the monograph [9] (cf. also [10] and [11]) are used below.

Let $\mathbf{D}$ be the non-euclidean Lobačevskiĭ plane realized as the interior of the unit disk. $\mathbf{D}$ is a universal covering for S with fixed covering $\lambda: \mathbf{D} \to S$ and a group F of covering transformations. F is a discrete Fuchsian group (of the first kind) of transformations of the Lobačevskiĭ plane and consists of hyperbolic elements only.

Let us consider the preimage $\lambda^{-1}\Delta$ of the connected (cf. Lemma 5.1) region Δ and one of its connected components. That is, let us fix the preimage ω_0 of some point $s_0 \in \Delta$ and consider the component $\Delta_0 \subset \lambda^{-1}\Delta$ which contains ω_0.

Let the point ω_1 be such that $\lambda\omega_1 = \lambda\omega_0 = s_0$, and let C_{01} be an arbitrary curve on $\mathbf{D}$ which joins ω_0 and ω_1. Then let us denote by $[\lambda C_{01}]$ the element of the fundamental group $\pi_1(S, s_0)$ which corresponds to the curve λC_{01}. It is obvious that this element does not depend on C_{01}, given an ω_1, but is different for different ω_1. Thus we obtain an isomorphism $\phi: F \to \pi_1(S, s_0)$, with $\phi(f) = [\lambda C_{01}]$, $f \in F$, if $\omega_1 = f\omega_0$.

Now let us consider all possible points $\omega_i \in \Delta_0$ such that $\lambda\omega_i = s_0$. The set of $f_i \in F$ such that $f_i\omega_0 = \omega_i$ forms a subgroup $F_0 \subset F$ with the canonical homomorphism $\psi: F_0 \to \pi_1(\Delta, s_0)$ and the following commutative diagram:

$$\begin{array}{ccc} F_0 & \xrightarrow{\psi} & \pi_1(\Delta, s_0) \\ \cap\downarrow & & \downarrow \\ F & \xrightarrow{\varphi} & \pi_1(S, s_0). \end{array}$$

(It is obvious that F_0 does not depend on the choice of the point s_0.) The boundary of the region Δ_0 consists of piecewise analytic curves which are preimages of $\Gamma_1, \cdots, \Gamma_k, \tilde{\Gamma}_1, \cdots, \tilde{\Gamma}_l$ (cf. §5). We shall denote these preimages by $\lambda_0^{-1}\Gamma_1, \cdots, \lambda_0^{-1}\Gamma_k, \cdots, \lambda_0^{-1}\tilde{\Gamma}_l$, respectively. If there are no branch points of the coverings h_1 or h_2, respectively, on $\Gamma_1, \cdots, \Gamma_k, \tilde{\Gamma}_1, \cdots, \tilde{\Gamma}_l$, then $\lambda_0^{-1}\Gamma_1, \cdots, \lambda_0^{-1}\tilde{\Gamma}_l$ are analytic.

We shall assume hereafter that *the field* $\mathbf{C}(S)$ *of meromorphic functions on* S *is a normal extension of both the field* $\mathbf{C}(x)$ *and the field* $\mathbf{C}(y)$. The case of a finite algebraic extension of the field $\mathbf{C}(S)$ existing which is normal over $\mathbf{C}(x)$ and over $\mathbf{C}(y)$ could be treated analogously. In this case it would be necessary first to lift the equations to the Riemann surface of this extension.

Lemma 9.1. *For any* Γ_i *and* $\tilde{\Gamma}_j$ *a non-euclidean transformation (hyperbolic or elliptic)* g_i *exists such that* $g_i(\lambda_0^{-1}\Gamma_0) = \lambda_0^{-1}\Gamma_i$, *as well as a non-euclidean transformation* $\tilde{g}_j$ *such that* $\tilde{g}_j(\lambda_0^{-1}\tilde{\Gamma}_0) = \lambda_0^{-1}\tilde{\Gamma}_j$.

Let us prove that $\lambda_0^{-1}\Gamma_1, \cdots, \lambda_0^{-1}\tilde{\Gamma}_l$ are analytic. As a matter of fact, that Γ_0 is an analytic simple closed curve was shown in §5. Let us denote by γ_i the conformal automorphism of the Riemann surface S such that $\gamma_i\Gamma_0 = \Gamma_i$. It is easy to show that it exists and is induced by the corresponding automorphism of the Galois field $\mathbf{C}(S)$ over $\mathbf{C}(x)$. Let us introduce analogously the automorphism $\tilde{\gamma}_j$ such that $\tilde{\gamma}_j\tilde{\Gamma}_0 = \tilde{\Gamma}_j$.

Thus to prove Lemma 9.1 it is sufficient for us to prove the following lemma and to set $\gamma_i = \tilde{g}$ and $g_i = g$ in it.

Lemma 9.2. *Let* $\tilde{g}$ *be a conformal automorphism of* S, $\tilde{g}(s_0) = s$ *holding for some points* $s_0, s \in S$. *Now, if* $\lambda\omega = s$ *and* $\lambda\omega_0 = s_0$, *a conformal automorphism* g *of the universal covering* $\mathbf{D}$ *exists such that* $\lambda g = \tilde{g}\lambda$ *and* $g\omega_0 = \omega$. *Moreover,* g *is elliptic or hyperbolic.*

In fact, if s_0 and s are two points on S which are not branch points of the covering h_1: $S \to \mathbf{P}$ and are such that $x(s_0) = x(s)$, then a unique conformal automorphism $\tilde{g}$ of the Riemann surface S exists such that $\tilde{g}s_0 = s$. Hereafter we shall not distinguish between $\tilde{g}$ and the corresponding Galois automorphism.

An arbitrary conformal automorphism $\tilde{g}$ of the surface S_x can be lifted to the universal covering $\mathbf{D}$ by $g = \lambda^{-1}\tilde{g}\lambda$. This transfer is not single-valued but becomes single-valued if the image $g\omega$ of a point $\omega \in \mathbf{D}$ is fixed. In this connection only the condition $g\omega \in \{\lambda^{-1}g\lambda\omega\}$ must be fulfilled. The automorphism g is a non-euclidean transformation of the Lobačevskiĭ plane (cf. [10]).

Let us prove that g is either hyperbolic or elliptic. If $\tilde{g}$ has a fixed point on S, then it can be lifted to $\mathbf{D}$ so that g has a fixed point inside $\mathbf{D}$. Then g is elliptic. Now let g not have fixed points in $\mathbf{D}$. This will hold in particular when g does not have fixed points on S. Let us prove that then

$$d = \inf_{\omega \in \mathbf{D}} \rho(\omega, g\omega) > 0, \tag{9.1}$$

where $\rho(\omega, \omega')$ is the distance between two points on the Lobačevskiĭ plane.

First let $\tilde{g}$ have no fixed points on S. Let us introduce on S the Riemann metric generated by the non-euclidean metric on $\mathbf{D}$. The compactness of S implies

$$\inf_{s\in S} \rho(s, \tilde{g}s) > 0. \tag{9.2}$$

Formula (9.1) follows easily from this. Now let $\tilde{g}$ have fixed points on S, while g has no fixed point on $\mathbf{D}$. Now let us draw a sufficiently small non-euclidean circle for each fixed point on S with center at that point. Let us denote by O_i an open neighborhood of the point bounded by a circle. Then $S\setminus[\bigcup_i O_i]$ is invariant relative to $\tilde{g}$, and just as in the previous case it is possible to prove that

$$\inf_{\omega\in\lambda^{-1}[S\setminus\{\bigcup_i O_i\}]} \rho \left|\overline{(\omega, g\omega) > 0.}\right| \tag{9.3}$$

However, the points of any component O of the set $\lambda^{-1}O_i$ can be mapped under the operation of g only to points of a congruent set fO for some $f \in F$, $f \neq 1$, whence

$$\inf_{\omega\in\lambda^{-1}O_i} \rho(\omega, g\omega) > 0. \tag{9.4}$$

In view of the finiteness of the set of indices i, (9.3) and (9.4) imply (9.1).

Thus it remains for us to prove that if inequality (9.1) is fulfilled, then g is a hyperbolic element. Now it is easily proved that there exists a z such that $\rho(z, gz) = d$. If we prove now that the points z, gz and g^2z are on some non-euclidean line l, this line will be invariant relatively to g, and thus g is a non-euclidean translation (a hyperbolic element). For purposes of the proof let us denote by ζ the middle of the non-euclidean segment $[z, gz]$. Then $g\zeta$ is the middle of the segment $[gz, g^2z]$, and $\rho(\zeta, g\zeta) \leq \rho(\zeta, gz) + \rho(gz, g\zeta) = d$. But at the same time $\rho(\zeta, g\zeta) \geq d$, which implies the assertion.

In order to complete the proof of Lemma 9.1 it remains to observe that for any i and arbitrary point $\omega \in \lambda_0^{-1}\Gamma_i$, there is a point $\omega' \in \lambda_0^{-1}\Gamma_0$ such that $x(\omega) = x(\omega')$, and to make use of the constructions cited above. In this connection g_i proves to be the lifting to $\mathbf{D}$ of the corresponding Galois automorphism.

Definition. The group κ of non-euclidean transformations of the Lobačevskiĭ plane which is generated by the group F_0 and all automorphisms g_i and $\tilde{g}_j$ will be called the group of equation (1.1) whose Riemann surface S has genus $g \geq 2$.

§10. Theorem on the analytic behavior of a solution

Let us lift the functions $x(s)$, $y(s)$, $\pi(s)$ and $\tilde{\pi}(s)$, which are defined on $\Delta \subset S$, to Δ_0 by the formula $\pi_\lambda(\omega) = \pi(\lambda\omega)$, $\lambda\omega \in \Delta$, $\omega \in \Delta_0$. (In what follows we omit the index λ on π and $\tilde{\pi}$.) It is obvious that $\pi(\omega)$ and $\tilde{\pi}(\omega)$ satisfy the following relations for $\omega \in \Delta_0$ and $h \in F_0$:

$$\begin{gathered}\pi(h\omega) = \pi(\omega),\\ \tilde{\pi}(h\omega) = \tilde{\pi}(\omega),\\ \pi(\omega) + \tilde{\pi}(\omega) = -x(\omega)y(\omega)\eta(x(\omega), y(\omega)).\end{gathered} \tag{10.1}$$

Theorem 10.1. *The functions $\pi(\omega)$ and $\widetilde{\pi}(\omega)$ admit meromorphic extension to the entire universal covering* $\mathbf{D}$.

First let us prove the following lemma.

Lemma 10.1. *The union of the regions $h\Delta_0$, where $h \in \kappa$, covers the entire non-euclidean plane* $\mathbf{D}$.

Proof. Let us observe that $\lambda_0^{-1}\Gamma_0 \cup \lambda_0^{-1}\widetilde{\Gamma}_0$ is at a positive non-euclidean distance ϵ from the boundary $\dot{\Delta}_0$ in the Lobačevskiĭ plane. Let us arbitrarily choose the point $\omega_0 \in \Delta_0$. The region Δ_0 possesses the following properties:

1) The boundary of the region Δ_0 belongs to the union of a finite number of analytic curves.

2) Let l be one of these analytic boundary curves. (We do not require that $l \cap \dot{\Delta}_0$ be connected.) Then there exists an $h \in \kappa$ such that any point of $l \cap \dot{\Delta}_0$ belongs to $\Delta_0 \cup h\Delta_0$ together with its non-euclidean ϵ-neighborhood. By Lemma 9.1 one of the generators g_i or $\widetilde{g}_j$ of the group κ can be chosen as h.

Now let us construct inductively the sequence of regions

$$\Delta_0, \Delta_1, \dots, \Delta_n, \dots$$

and the sequence of analytic curves

$$l_0, l_1, \dots, l_n, \dots$$

in the following manner. Let us set $l_0 = \Gamma_0$ and $l_1 = \widetilde{\Gamma}_0$. Next let us arrange in arbitrary order the analytic curves having an arc in common with the boundary $\dot{\Delta}_0$, then the analytic curves having an arc in common with the boundary $\dot{\Delta}_1$ which do not coincide with those processed earlier, etc.

Let us choose $h_0 = g_i$ or $\widetilde{g}_j$ so that $l_2 = g_i l_0$ or $\widetilde{g}_j l_1$. Let us set $\Delta_1 = \Delta_0 \cup h_0\Delta_0$. Suppose we have already constructed Δ_n. Then there exists an $h_n = g_i$ or $\widetilde{g}_j$ such that $l_{n+2} = h_n l_{i_n}$ for some $i_n \leq n+1$. Let us set $\Delta_{n+1} = \Delta_n \cup h_n\Delta_n$. In this connection l_{n+2} belongs to Δ_{n+1}, together with its non-euclidean ϵ-neighborhood. Fulfillment of properties 1) and 2) by $\dot{\Delta}_{n+1}$ is obvious. In addition, by construction, for any n there is an $N > n$ such that $\rho(\omega_0, \dot{\Delta}_N) > \rho(\omega_0, \dot{\Delta}_n) + \epsilon$. This implies that the union of the ascending sequence of connected regions Δ_n coincides with all of $\mathbf{D}$.

Now let us turn to the proof of Theorem 10.1. By virtue of the constructions made in §9,

$$\pi(g_i\omega) = \pi(\omega), \quad \widetilde{\pi}(\widetilde{g}_j\omega) = \widetilde{\pi}(\omega) \tag{10.2}$$

for $g_i\omega$, $\omega \in \Delta_0$, and $\widetilde{g}_j\omega$, $\omega \in \Delta_0$, respectively. The functions $\pi(\omega)$ and $\widetilde{\pi}(\omega)$ are extendable by induction to the regions Δ_n. Let them be defined in the region Δ_n. Then if $h_n = g_i$, let us extend $\pi(\omega)$ to the region Δ_{n+1} by means of the relation $\pi(h_n\omega) = \pi(\omega)$. The function $\widetilde{\pi}(\omega)$ is then determined from equation (10.1); we act analogously if $h_n = \widetilde{g}_j$. Now it is easily provable that $\pi(\omega)$ and $\widetilde{\pi}(\omega)$ can be extended along any curve on $\mathbf{D}$; consequently they prove to be single-valued on $\mathbf{D}$ (by the monodromy theorem).

Thus a complete description of the potentialities for analytic extension of a solution is obtained from Theorem 10.1 and Remark 3.1.

§11. Discreteness of the equation's group

Here we study the group κ of the equation in the distance of genus $g \geq 2$ and the irreducibility case.

Lemma 11.1. *If the genus of the Riemann surface S is $g \geq 2$ and the irreducibility case is realized, then the group κ of equation* (1.1) *is discrete.*

Proof. By a theorem of Poincaré (cf. [9], p. 99), discreteness and disconnectedness are equivalent with respect to the group of transformations of the non-euclidean plane. Thus if κ is indiscrete, it cannot be disconnected. Consequently every point $\omega \in \mathbf{D}$ is a limit point of κ. But by Theorem 4A of [9], p. 103, the set κz is then dense at any (limit) point $\omega \in \mathbf{D}$ for any $z \neq \omega$ except for, perhaps, one value $z = z_0$ for which $\{\kappa z_0\} = \{z_0\}$, i.e. for any $z \in \mathbf{D}$ in our case.

It is well known (cf. [10]) that the group of conformal automorphisms E of an arbitrary Riemann surface S of genus $g \geq 2$ is finite. Hence it follows easily that the set of points $\{\lambda^{-1} e_i \lambda\omega\}$ for any point $\omega \in \mathbf{D}$, where $E = \{e_0, \cdots, e_n\}$, is discrete. The set of points $\kappa\omega$ for any ω is all the more discrete. The lemma has been proved.

Lemma 11.2. *κ can consist of only elliptic and hyperbolic elements.*

Proof. Let us fix some point $s_0 \in \Delta \subset S$ and one of its preimages $\omega_0 \in \{\lambda^{-1} s_0\} \cap \Delta_0 \subset \mathbf{D}$. Let us denote

$$\rho_\Delta(s_0, s) = \inf_l \rho(l)$$

for any point $s \in \Delta$, where $\rho(l)$ is the length in the Riemann metric induced by the metric on the Lobačevskiĭ plane of the rectifiable curve $l \subset \Delta$ joining the points s_0 and s_1, and the infimum is taken with respect to all such rectifiable curves. $\rho_{\Delta_0}(\omega_0, \omega)$ is defined analogously for any point $\omega \in \Delta_0$. It is easy to see that a constant C_0 exists such that for any point $\omega \in \Delta_0$

$$\inf_{1 \neq h \in F_0} \rho_{\Delta_0}(\omega, h\omega) < C_0. \tag{11.1}$$

Let us consider the normal fundamental polygon N (cf. [10]) with respect to the point ω_0 on the asusmption that ω_0 is not a fixed point for κ. Let us prove that N is compact. Indeed, otherwise a point $\omega \in \Delta_0$ would exist which is removed from ω_0 by an arbitrarily great distance, which contradicts condition (11.1).

The compactness of the normal fundamental region implies that κ does not contain parabolic elements (cf. [9], Theorem 7E, p. 149).

In this connection the group κ has generators $h_1, \cdots, h_l, a_1, b_1, \cdots, a_m, b_m$ and the defining relations

$$h_1 \ldots h_l a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_m b_m a_m^{-1} b_m^{-1} = 1, \; h_i^{k_i} = 1, i = 1, \ldots, l$$

(cf. [9], p. 241).

Lemma 11.3. *For some n a normal divisor κ_0 of the group κ exists with generators $A_1, B_1, \cdots, A_n, B_n$ and the unique defining relation*

$$A_1B_1A_1^{-1}B_1^{-1}\ldots A_nB_nA_n^{-1}B_n^{-1}=1.$$

Proof. Selberg [13] shows that in any matrix group over the field of complex numbers which has a finite number of generators, a normal divisor of finite index exists which does not contain elements of finite order. This result can obviously be applied to our group κ. Then such a normal divisor turns out to consist of only hyperbolic elements, and by virtue of the finiteness of the index its fundamental group is compact. This implies (cf. [9]) that it has the aforementioned unique defining relation.

§12. Carleman's problem with respect to the fundamental polygon and an explicit form of the solution

Notwithstanding the explicit representation of the solution which is obtained below, the method of analytic extension in §10 plays a very important role. In particular, by means of it we obtain all poles of the functions $\pi(\omega)$ and $\widetilde{\pi}(\omega)$ on $\mathbf{D}$ and their principal parts at these poles. The procedure for isolating principal parts is obvious from the proof of Theorem 10.1.

Remark 12.1. This implies in particular that the asymptotic behavior of the coefficients of the functions $\pi(x)$ and $\widetilde{\pi}(y)$ is determined either by their algebraic branch points nearest the unit disk or by their poles nearest the unit disk, and can be computed in specific cases. (That $\pi(x)$ and $\widetilde{\pi}(y)$ do not have other singular points follows from Theorem 10.1.)

This fact plays a basic role in probability applications, but we do not go into it here.

Let us examine the normal divisor of κ_0 whose existence was asserted in Lemma 11.3. By utilizing relations (10.1) and (10.2) and the relations

$$\pi(\widetilde{g}_j\omega)-\pi(\omega)=x(\omega)\,y(\omega)\,\eta(\omega)-x(\widetilde{g}_j\omega)\,y(\widetilde{g}_j\omega)\,\eta(\widetilde{g}_j\omega), \tag{12.1}$$

which they imply, the equations

$$\pi(A_i\omega)-\pi(\omega)=\alpha_i(\omega),\quad \pi(B_i\omega)-\pi(\omega)=\beta_i(\omega),\quad i=1,\ldots,n, \tag{12.2}$$

can be obtained for the functions π, $\widetilde{\pi}$ and the generators $A_1, B_1, \cdots, A_n, B_n$ of this normal divisor, where $\alpha_i(\omega)$ and $\beta_i(\omega)$ are linear combinations of functions automorphic with respect to the group F which are converted to elements of the group κ.

Relations (12.2) represent Carleman's problem with respect to the fundamental polygon N: $A_1B_1A_1^{-1}B_1^{-1}\cdots A_nB_nA_n^{-1}B_n^{-1}=1$, which we shall fix in what follows and which can be chosen so that α_i, β_i, π and $\widetilde{\pi}$ do not have poles on its boundary. (Regarding Carleman's problem, cf. the bibliography in [8].)

Let us consider the Behnke-Stein kernel $\mathfrak{A}(\omega,\omega')$ [8] for the Riemann surface $\mathbf{D}/\kappa_0$, whose explicit construction Weierstrass has already in essence developed in [12]. On $\mathbf{D}$ $\mathfrak{A}(\omega,\omega')$ represents an automorphic function of ω and an automorphic form in ω'.

One of the solutions of problem (12.1) which is automorphic on D is representable as

$$\Pi(\omega) = \frac{1}{2\pi i}\sum_i \left[\int_{A_i} \mathfrak{A}(\omega, \omega')\alpha_i(\omega')\,d\omega' + \int_{B_i} \mathfrak{A}(\omega, \omega')\beta_i(\omega')\,d\omega'\right]. \tag{12.3}$$

Let us denote by $\Pi'(\omega)$ the sum of the principal parts as the poles of $\Pi(\omega)$ in N, and by $\Pi''(\omega)$ the same sum of principal parts for $\pi(\omega)$ (cf. the remarks at the beginning of this section).

The difference between $\pi(\omega)$ and the particular solution $\Pi(\omega)$ of the nonhomogeneous problem (12.2) must coincide in N with a function $\widetilde{\Pi}(\omega)$ which is automorphic with respect to the group κ_0. Moreover, $\widetilde{\Pi}(\omega) = \pi(\omega) - \Pi(\omega)$ does not have poles on the boundary of N. $\widetilde{\Pi}(\omega)$ is defined by the fact that the sum of the principal parts of $\widetilde{\Pi}(\omega)$ at poles in N must coincide with $\Pi''(\omega) - \Pi'(\omega)$. Let us formulate the result obtained.

Theorem 12.1. *The function $\pi(\omega)$ has the explicit representation*

$$\pi(\omega) = \frac{1}{2\pi i}\sum_i \left(\int_{A_i} \mathfrak{A}(\omega, \omega')\alpha_i(\omega')\,d\omega' + \int_{B_i} \mathfrak{A}(\omega, \omega')\beta_i(\omega')\,d\omega'\right) + \widetilde{\Pi}(\omega) + \text{const},$$

in the fundamental polygon N, where $\widetilde{\Pi}(\omega)$ is defined above, which is automorphic with respect to the group κ_0. The function $\tilde{\pi}(\omega)$ has an analogous representation. The constants in their right sides are determined by the condition $\tilde{\pi}(y) = 0$ for $y = 0$.

Remark 12.2. It might be possible to construct an analogous integral representation directly for the group κ. Moreover, the expression would become more explicit, but in view of the presence of elliptic elements the formulas would be more cumbersome.

In conclusion I wish to thank I. I. Pjateckiĭ-Šapiro for valuable consultation concerning discrete groups.

Received 24 MAR 70 and
7 JULY 70

BIBLIOGRAPHY

[1] I. B. Simonenko, *Operators of convolution type in cones*, Mat. Sb. 74 (116) (1967), 298–313 = Math. USSR Sb. 3 (1967), 279–294. MR 36 #5773.

[2] ———, *Multidimensional discrete convolutions*, Mat. Issled. 3 (1968), vyp. 1 (7), 108–122. (Russian) MR 41 #2412.

[3] V. A. Malyšev, *On the solution of discrete Wiener-Hopf equations in a quarter-plane*, Dokl. Akad. Nauk SSSR 187 (1969), 1243–1246 = Soviet Math. Dokl. 10 (1969), 1032–1035. MR 41 #7463.

[4] N. Wiener and E. Hopf, *Über eine Klasse singulärer Integralgleichungen*, S.-B. Preuss. Akad. Wiss. (= S.-B. Deutsch. Akad. Wiss. Berlin) 1931, 696–706.

[5] M. G. Kreĭn, *Integral equations on a half-line with kernel depending upon the difference of the arguments*, Uspehi Mat. Nauk 12 (1958), no. 5 (83), 3–120; English transl., Amer. Math. Soc. Transl. (2) 22 (1962), 163–288. MR 21 #1507.

[6] L. S. Gol'denšteĭn and I. C. Gohberg, *On a multidimensional integral equation on a half-space whose kernel is a function of the difference of the arguments, and on a discrete analogue of this equation*, Dokl. Akad. Nauk SSSR 131 (1960), 9–12 = Soviet Math. Dokl. 1 (1960), 173–176. MR 22 #8298.

[7] I. M. Gel'fand, D. A. Raĭkov and G. E. Šilov, *Commutative normed rings*, Fizmatgiz, Moscow, 1960; English transl., Chelsea, New York, 1964. MR 23 #A1242; 34 #4940.

[8] E. I. Zverovič and G. S. Litvinčuk, *Boundary value problems with shift for analytic functions, and singular functional equations*, Uspehi Mat. Nauk 23 (1968), no. 3 (141), 67–121 = Russian Math. Surveys 23 (1968), no. 3, 67–124. MR 37 #5405.

[9] J. Lehner, *Discontinuous groups and automorphic functions*, Math. Surveys, no. 8, Amer. Math. Soc., Providence, R. I., 1964. MR 20 #1332.

[10] G. Springer, *Introduction to Riemann surfaces*, Addison-Wesley, Reading, Mass., 1957; Russian transl., IL, Moscow, 1960. MR 19, 1169; 23 #A318.

[11] L. R. Ford, *Automorphic functions*, McGraw-Hill, New York, 1929; Russian transl., ONTI, Moscow, 1936.

[12] K. Weierstrass, *Vorlesungen über die Theorie der Abelschen Transzendenten. Mathematische Werke*. Band 4, Mayer and Müller, Berlin, 1902.

[13] A. Selberg, *On discontinuous groups in higher-dimensional symmetric spaces. Contributions to function theory*, Internat. Colloq. Function Theory (Bombay, 1960), Tata Institute of Fundamental Research, Bombay, 1960, pp. 147–164. MR 24 #A188.

[14] V. A. Malyšev, *Analytic methods in the theory of random walks in a quadrant of the plane: a simple walk with skewed image*, Proc. Soviet-Japanese Sympos. on Probability Theory, Khabarovsk, 1969, pp. 176–184. (Russian)

[15] V. S. Rabinovič, *The multidimensional Wiener-Hopf equation for cones*, Teor. Funkciĭ Funkcional. Anal. i Priložen. Vyp. 5 (1967), 59–67. (Russian) MR 37 #777.

[16] V. A. Kakičev, *Boundary value problem of linear conjugacy functions holomorphic in bicylindrical regions*, Teor. Funkciĭ Funkcional. Anal. i Priložen. Vyp. 5 (1967), 37–58. (Russian) MR 37 #1645.

Translated by:
F. M. Goldware