A PRESENTATION OF THE GROUPS PSL(2, p)

H. BEHR AND J. MENNICKE

1. In the present paper, we shall prove the following result.

THEOREM A. The groups PSL(2, p) can be presented by the following system of generators and relations:

(1.1)
$$S^p = T^2 = (ST)^3 = (S^2 T S^{\frac{1}{2}(p+1)} T)^3 = 1 \quad (p > 2).$$

This theorem considerably improves earlier results of Bussey, Frasch, and Todd (cf. 2, pp. 93–96). The presentation of Frasch reads as follows:

(1.2)
$$S^p = T^2 = (ST)^3 = 1,$$
 $U^{-1}SU = S^{\alpha^2},$ $(UT)^2 = 1,$
 $U^{\frac{1}{2}(p-1)} = 1,$ $(TUS^{\alpha})^3 = 1,$ where α is a primitive root mod p .

By using simple properties of PSL(2, p), such as the existence of a Bruhat decomposition, it is not difficult to verify that (1.2) defines PSL(2, p). It would be desirable to have a similar direct proof of Theorem A.

Our proof proceeds indirectly. We adopt a general method for proving the finite presentation of generalized unit groups (cf. 1). After a suitable specialization, we obtain the following theorem.

THEOREM B. Let $\mathbf{Z}^{(2)} = (x/2^t, x, t \in \mathbf{Z})$. The group $SL(2, \mathbf{Z}^{(2)})$ can be presented as follows:

$$(1.3) \quad (AB)^3 = (UB)^2 = (UA^2B)^3 = B^2, \qquad B^4 = 1, \qquad U^{-1}AU = A^4.$$

The relations (1.3) are fulfilled by the elements

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

In (3) it was shown that any subgroup of finite index in $SL(2, \mathbb{Z}^{(2)})$ contains a full congruence subgroup. From this result and from Theorem B one can deduce Theorem A. This is carried out in the next section; §§ 3–5 contain the proof of Theorem B.

2. In this section, we shall deduce Theorem A from Theorem B. Let

$$(2.1) G = \operatorname{SL}(2, \mathbf{Z}^{(2)})$$

be the group defined in Theorem B. G is generated by the elements A, B, and U given in Theorem B, and it is presented by the system (1.3).

Received June 1, 1967.

1432

We shall define some subgroups of G. Let m be odd. By Q_m we denote the normal closure of the element

in G. Let N_m be the full congruence subgroup modulo m of G. In a previous paper (3), it was shown that

$$(2.3) N_m = Q_m$$

Let SL(2, m) be the special linear group of rank 1 over the ring Z/mZ. From Theorem B and from (2.3) we deduce that the following is an abstract presentation of the group SL(2, m):

$$(2.4) A^{m} = 1, \quad (AB)^{3} = (UB)^{2} = (UA^{2}B)^{3} = B^{2}, \quad B^{4} = 1, \quad U^{-1}AU = A^{4}.$$

We shall now simplify the relations (2.4). The exponents of A can be looked at as elements of the ring $\mathbb{Z}/m\mathbb{Z}$. Eliminate U in (2.4):

(2.5)
$$U = A^{\frac{1}{2}} B A^{\frac{1}{2}} B A^{\frac{1}{2}} B^{-1}.$$

Clearly, the following relations form a system which is equivalent to (2.4):

(2.6)
$$A^m = 1, \quad (AB)^3 = B^2, \quad B^4 = 1,$$

(2.7)
$$UBU = B, \qquad U^{-1}A^{\frac{1}{2}}U = A^{2},$$

(2.8)
$$U = A^{\frac{1}{2}} B A^{\frac{1}{2}} B A^{\frac{1}{2}} B^{-1}.$$

We shall modify the relations (2.7). After eliminating U, the first relation of (2.7) reads as follows:

(2.9)
$$(A^{\frac{1}{2}}BA^{2}BA^{\frac{1}{2}})^{2} = B^{2},$$

and the second relation of (2.7):

(2.10)
$$A^{\frac{1}{2}}BA^{\frac{1}{2}}BA^{\frac{1}{2}}B = BA^{\frac{1}{2}}BA^{\frac{1}{2}}BA^{\frac{1}{2}}.$$

(2.9) and (2.10) imply that

(2.11)
$$(A^2BA^{\frac{1}{2}}B)^3 = 1.$$

Conversely, (2.9) and (2.11) imply (2.10); thus, we can drop (2.10). The relation (2.9) is a consequence of (2.6). Therefore, we have proved that the set of relations

(2.12)
$$A^m = 1$$
, $(AB)^3 = B^2$, $B^4 = 1$, $(A^2BA^{\frac{1}{2}}B)^3 = 1$

is equivalent to the set (2.4). (2.12) is a presentation of the group SL(2, m). By specializing m = p a prime, we obtain a presentation of SL(2, p). Finally, by adding the relation $B^2 = 1$, we obtain a presentation of the group PSL(2, p):

(2.13)
$$A^{p} = (AB)^{3} = B^{2} = (A^{2}BA^{\frac{1}{2}(p+1)}B)^{3} = 1.$$

Thus, assuming Theorem B, we have completed the proof of Theorem A.

3. Finite presentation of $SL(n, Z^{(P)})$.* It was shown in (1) that for some classical groups G, the group $G(Z_S)$, where Z_S denotes a Hasse domain in the field of rationals, can be finitely presented. Especially, we obtain a finite system of defining relations for $SL(2, Z^{(p)})$, but it seems to be very hard to describe it explicitly and to reduce it. By earlier specialization to G = SL and by some improvements of the method of (1), we obtain a better system, but then it seems necessary to sketch the whole proof and to refer to (1) only for some details.

3.1. Lattices. We denote by Q_p the field of *p*-adics, \mathbf{Z}_p the ring of *p*-adic integers, and, a \mathbf{Z}_p -module in Q_p^n , which contains *n* linear independent vectors, will be called a lattice. Let L_0 be the lattice spanned by the unit vectors and for some $g \in \mathrm{SL}(n, \mathbf{Z}^{(p)})$ let $L = gL_0$ be the lattice spanned by the columns of the matrix *g* and \mathfrak{L} the set of all lattices gL_0 for $g \in \mathrm{SL}(n, \mathbf{Z}^{(p)})$. We define a distance *d* in the set \mathfrak{L} by

$$d(L_1, L_2) = \min\{n \mid p^n L_1 \subseteq L_2 \subseteq p^{-n} L_1\}, \qquad (L_1, L_2 \in \mathfrak{L}).$$

d is invariant under the group $SL(n, \mathbb{Z}^{(p)})$. For $g \in SL(n, \mathbb{Z}^{(p)})$ we set

$$|g| = d(L_0, gL_0).$$

Then we have that

 $|g^{-1}| = |g|, \qquad |g_1g_2| \le |g_1| + |g_2|.$

LEMMA. Let L_1 , L_2 , $M \in \mathfrak{X}$ with $d(L_1, L_2) = d \neq 0$. There exists a lattice $L \in \mathfrak{X}$ with

(a) $d(L_1, L) = d - 1, d(L, L_2) = 1,$

(b) $d(L, M) \leq \max\{d(L_1, M), d(L_2, M)\}.$

The lemma is essentially a consequence of the elementary divisor theorem (cf. 1, pp. 131–132).

3.2. Generators of SL $(n, \mathbf{Z}^{(p)})$. It is well known that SL (n, \mathbf{Z}) can be finitely generated; choose a finite system E_0 of generators, which, with g, also contains g^{-1} . The set $\mathfrak{N} = \{L \in \mathfrak{X} | d(L, L_0) = 1\}$ (neighbours of L_0) is finite; for each $L \in \mathfrak{R}$, choose an element $g \in SL(n, \mathbf{Z}^{(p)})$ with $L = gL_0$ and call the set of these elements and its inverses E_p . It is easily seen that $E = E_0 \cup E_p$ is a set of generators of SL $(n, \mathbf{Z}^{(p)})$ (cf. 1, Satz 1).

3.3. Defining relations of $SL(n, \mathbb{Z}^{(p)})$. We now consider relations between elements of E and relate with each of them a sequence of lattices in \mathfrak{X} . Let $r: a_1a_2 \ldots a_n = 1, a_i \in E$, be a relation; then we call *path of* r the sequence $P(r) = (L_0, L_1 \ldots, L_n)$, defined by $L_i = a_1a_2 \ldots a_i L_0$ $(L_n = a_1a_2 \ldots a_n L_0 = L_0)$. We call $D(r) = \max\{d(L_i, L_0), i = 0, 1, \ldots, n\}$ the distance of r and the number of pairs (L_i, L_{i+1}) with $L_i \neq L_{i+1}$ $(i = 0, 1, \ldots, n-1)$ the length of r.

1434

^{*}The definition of $Z^{(p)}$ is analogous to the definition of $Z^{(2)}$ given in the introduction.

We now construct, by induction with respect to the distance D, a finite set R_p of relations of length less than or equal to 6, so that we can reduce each relation by means of relations in R_p to a relation which contains only elements of E_0 and we know that $SL(n, \mathbb{Z})$ can be finitely presented (cf. 1, Satz 4).

If D(r) = 1, then for each point L_i of the path P(r) we have that $d(L_i, L_0) \leq 1$, which means that $|a_1 \dots a_i| \leq 1$ and therefore there exists an element $b_i \in E$ with $a_i^{-1} \dots a_1^{-1} L_0 = b_i L_0$. It follows that the products $b_i^{-1}a_{i+1}b_{i+1}$ keep L_0 fixed, we multiply each of them by a (fixed) product of elements of E_0 , and we obtain a finite set of relations of length less than or equal to 3.

Now we assume that D(r) > 1 and take a pair (L_i, L_{i+1}) of P(r) with $\max\{d(L_i, L_0), d(L_{i+1}, L_0)\} = D(r)$. By the lemma, there exist lattices M_i and M_{i+1} in \mathfrak{X} with $d(L_i, M_i) = d(L_{i+1}, M_{i+1}) = 1$ and $\max\{d(M_i, L_0), d(M_{i+1}, L_0)\} < D(r)$. We can choose b_i and $b_{i+1} \in E$ such that $M_i = a_1 \ldots a_i b_i L_0$ and $M_{i+1} = a_1 \ldots a_{i+1} b_{i+1} L_0$. We have that $d(M_i, M_{i+1}) = d(a_1 \ldots a_i b_i L_0, a_1 \ldots a_{i+1} b_{i+1} L_0) = d(L_0, b_i^{-1} a_{i+1} b_{i+1} L_0) \leq 3$. By repeated application of the lemma we obtain lattices N_1, \ldots, N_r $(r \leq 2)$ with $d(M_i, N_1) = d(N_j, N_{j+1}) = d(N_r, M_{i+1}) = 1$ and $d(N_j, L) \leq \max\{d(M_i, L), d(M_{i+1}, L)\}$ for $L \in \mathfrak{X}$, especially for $L = L_0$; thus, we have that $d(N_j, L_0) < D(r)$. Again we have elements c_1, \ldots, c_s $(s \leq 3)$ with

$$N_j = a_1 \dots a_i b_i c_1 \dots c_j L_0 \ (1 \leq j \leq r)$$

and

$$M_{i+1} = a_1 \ldots a_i b_i c_1 \ldots c_s L_0.$$

On the other hand, we have that $M_{i+1} = a_1 \ldots a_i a_{i+1} b_{i+1} L_0$; therefore, $b_{i+1}^{-1}a_{i+1}^{-1}b_i c_1 \ldots c_s$ keeps L_0 fixed. We multiply it by a (fixed) product of elements of E_0 so that we obtain a relation r_i of length less than or equal to 6. If we do this for each pair with maximal distance D(r) (using the same lattice M_i and the same b_i for the pairs (L_{i-1}, L_i) and (L_i, L_{i+1})) and go into rwith the relations $b_{i+1}r_i b_{i+1}^{-1}$ we finally obtain a relation r' with D(r') < D(r).

3.4. Reduction of the system of defining relations.

Length 6. We obtain relations of length 6 in our defining system only if $|b_i^{-1}a_{i+1}b_{i+1}| = 3$ (cf. § 3.3). Then we have that $d(N_2, L_i) \leq \max\{d(M_i, L_i), d(M_{i+1}, L_i)\} \leq 2$ (since $d(M_{i+1}, L_i) \leq d(M_{i+1}, L_{i+1}) + d(L_{i+1}, L_i) \leq 1 + 1$), which means that $d(a_1 \ldots a_i b_i c_1 c_2 L_0, a_1 \ldots a_i L_0) = d(b_i c_1 c_2 L_0, L_0) \leq 2$. There exist elements $d_1, d_2 \in E$ and a product e of elements of E_0 such that $b_i c_1 c_2 = ed_1 d_2$; this is a relation of length less than or equal to 5. If we substitute $b_i c_1 c_2$ by $ed_1 d_2$ in r_i we also have a relation of length less than or equal to 5.

Length 5. Modulo relations of length less than or equal to 3, we can assume that we have r: $a_1a_2a_3a_4a_5e = 1$, $|a_i| = 1$, e a product of elements of E_0 . If we do not have the case $|a_1a_2| = |a_2a_3| = |a_3a_4| = |a_4a_5| = 2$, we can immediately insert a relation of length less than or equal to 3 to obtain a relation of length less than or equal to 4.

Length 4. Again we can assume that $r: a_1a_2a_3a_4e = 1$, $|a_i| = 1$, e a product of elements of E_0 . If $|a_1a_2| = |a_2a_3| = 2$, we set $L_1 = a_1L_0$, $L_2 = a_1a_2L_0$,

 $L_3 = a_1 a_2 a_3 L_0$, and the lemma provides us with a lattice $M \in \mathfrak{X}$ with $d(M, L_0) = d(M, L_1) = d(M, L_2) = d(M, L_3) = 1$, and that means that we can reduce r by relations of length less than or equal to 3. If $|a_1 a_2| = 1$ or $|a_2 a_3| = 1$ we can at once reduce to relations of length less than or equal to 3.

3.5. *Remark.* All previous results are valid for each group for which the lemma holds.

4. The group $SL(2, \mathbb{Z}^{(P)})$. We shall now give the generators and describe the defining relations for n = 2. Some of them will be given explicitly; some series of relations will only be computed for p = 2 in the last section.

4.1. Generators of SL(2, $\mathbf{Z}^{(p)}$). We set

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} p & 0 \\ 0 & p^{-1} \end{pmatrix}.$$

A and B generate SL(2, **Z**); with the notations of § 3.2 we have, therefore, that $E_0 = \{A, B, A^{-1}, B^{-1}\}$. Furthermore, we have to consider the elements which transport the unit-lattice L_0 into a neighbour of it; these are matrices which have at least one coefficient $\epsilon \cdot p^{-1}$, $\epsilon a p$ -adic unit. For each neighbour, one has to choose such an element, we can take the following matrices:

$$\begin{pmatrix} p^{-1} & 0\\ xp^{-1} & p \end{pmatrix} \text{ for } x = 0, 1, \dots, p^2 - 1, \begin{pmatrix} p & 0\\ 0 & p^{-1} \end{pmatrix}, \begin{pmatrix} y & -p\\ p^{-1} & 0 \end{pmatrix} \text{ for } y = 1, 2, \dots, p - 1.$$

Thus, we have that

$$E_{p} = \{ (A^{x}U^{-1})^{\pm 1}, (B^{-1}A^{-yp}U^{-1})^{\pm 1}, x = 0, 1, \dots, p^{2} - 1, y = 1, 2, \dots, p - 1 \}$$

4.2. Relations of length 5. At first, we exclude the exceptional case of relations of length 5. For this purpose we give a list of products a_1a_2 , $a_i \in E_p$ with $|a_1a_2| = 2$.

(a)
$$A^{x}U^{-1} \cdot A^{x'}U^{-1} = \begin{pmatrix} p^{-2} & 0\\ xp^{-2} + x' & p^{2} \end{pmatrix},$$

(b)
$$B^{-1}A^{-\nu p}U^{-1} \cdot A^{x}U^{-1} = \begin{pmatrix} yp^{-1} - x & -p^{2} \\ p^{-2} & 0 \end{pmatrix}$$
,

(c)
$$UA^{-x} \cdot A^{x'}U^{-1} = \begin{pmatrix} 1 & 0 \\ (x' - x)p^{-2} & 1 \end{pmatrix}, \quad x' \neq x \mod p,$$

(d)
$$UA^{-x} \cdot B^{-1}A^{-yp}U^{-1} = \begin{pmatrix} yp & -p^2 \\ -xyp^{-1} + p^{-2} & x \end{pmatrix},$$

(e)
$$UA^{-x} \cdot UA^{-x'} = \begin{pmatrix} p^2 & 0 \\ -x - x'p^{-2} & p^{-2} \end{pmatrix},$$

(f)
$$UA^{-x} \cdot UA^{\nu p}B = \begin{pmatrix} 0 & p^2 \\ -p^{-2} & -x + yp^{-1} \end{pmatrix},$$

(g)
$$UA^{\nu p}B \cdot A^{x}U^{-1} = \begin{pmatrix} x & p^{2} \\ -p^{-2} + xyp^{-1} & yp \end{pmatrix}.$$

4.3. List of relations of length less than or equal to 3. We shall write on the left-hand side a product $a_1a_2a_3$, $a_i \in E$, with $|a_1a_2a_3| = 0$, on the right-hand side, a product of elements of E_0 (the choice of which is not unique) and in some cases, only an element of SL(2, **Z**). The list will not be complete, but if we add the inverse relations of the relations in the list we shall have all relations of length less than or equal to 3 which cannot be trivially reduced to relations of length 0, i.e., relations in SL(2, **Z**).

(A)
$$A^{x}U^{-1} \cdot A \cdot UA^{-x'} = A^{x-x'+p^{2}},$$

(B)
$$A^{x}U^{-1} \cdot A^{\pm 1} \cdot UA^{yp}B = A^{x+yp\pm p^{2}}B,$$

(C)
$$A^{x}U^{-1} \cdot B^{\pm 1} \cdot U^{-1} = A^{x}B^{\pm 1},$$

(D)
$$A^{x}U^{-1} \cdot B^{-1}A^{-yp}U^{-1} \cdot B^{-1}A^{-y'p}U^{-1} = \begin{pmatrix} (yy'-1)p^{-1} & -y \\ x(yy'-1)p^{-1} + y' & -xy - p \end{pmatrix}$$

for
$$yy' - 1 \equiv 0 \mod p$$
,

(E)
$$B^{-1}A^{-yp}U^{-1}\cdot A\cdot UA^{y'p}B = B^{-1}A^{p(y'-y)+p^2}B,$$

(F)
$$B^{-1}A^{-\nu p}U^{-1} \cdot B^{\pm 1} \cdot U^{-1} = B^{-1}A^{-\nu p}B^{\pm 1},$$

(G)
$$B^{-1}A^{-yp}U^{-1} \cdot B^{-1}A^{-y'p}U^{-1} \cdot B^{-1}A^{-y''p}U^{-1} = \begin{pmatrix} (yy'-1)y''-y & (1-yy')p \\ (y'y''-1)p^{-1} & -y' \end{pmatrix}$$

for $y'y''-1 \equiv 0 \mod p$,

(H)
$$UA^{-x} \cdot A \cdot A^{x'} U^{-1} = A^{(x'-x+1)p^{-2}}$$
 for $x' - x + 1 \equiv 0 \mod p^2$,

(I)
$$UA^{-x} \cdot B \cdot A^{x'} U^{-1} = \begin{pmatrix} x' & p^2 \\ -(1+xx')p^{-2} & -x \end{pmatrix}$$
 for $1 + xx' \equiv 0 \mod p^2$,

(J)
$$UA^{-p} \cdot B \cdot B^{-1}A^{-(p-1)p}U^{-1} = A^{-1},$$

(K)
$$UA^{-p} \cdot B^{-1} \cdot B^{-1}A^{-(p-1)p}U^{-1} = A^{-1}B^2,$$

(L)
$$UA^{-x} \cdot A^{x'}U^{-1} \cdot B^{-1}A^{-yp}U^{-1} = \begin{pmatrix} y & -p \\ ((x'-x)y+1)p^{-1} & -(x'-x)p^{-1} \end{pmatrix}$$

for $y(x'-x) \equiv -p \mod p^2$

(M)
$$UA^{yp}B \cdot A \cdot B^{-1}A^{-y'p}U^{-1} = \begin{pmatrix} 1 + y'p & -p^2 \\ (y - y')p^{-1} + yy' & 1 - yp \end{pmatrix}$$

for
$$y \equiv y' \mod p$$
,

(N)
$$UA^{yp}B \cdot B^{-1}A^{-y'p}U^{-1} \cdot B^{-1}A^{-y''p}U^{-1} = \begin{pmatrix} y'' & -p \\ (y - y')p^{-1}y'' + p^{-1} & y' - y \end{pmatrix}$$

for $y''(y - y') + 1 \equiv 0 \mod p$.

4.4. A system of defining relations of $SL(2, \mathbb{Z}^{(p)})$. $SL(2, \mathbb{Z})$ can be defined by the relations

(1)
$$B^2 = (AB)^3, \quad B^4 = 1.$$

The relations (A), (B), (E), (H), and (J) in §4.3 are equivalent to

(2) $U^{-1}A U = A^{p^2}.$

The relations (C) and (F) in §4.3 are equivalent to the relation

$$(3) (UB)^2 = B^2.$$

With the help of (1), (2), and (3) we can eliminate (K). The results of §§ 3 and 4 show that (1)–(3), (D), (G), (I), and (L)–(M) is a defining system of $SL(2, \mathbb{Z}^{(p)})$.

5. The group SL(2, $Z^{(2)}$). We shall now give all relations for p = 2 explicitly. There are two cases of (I) and (L), only one case of (D), (G), and (M), and no case of (N):

(D')
$$y = y' = 1: A^{x}U^{-1} \cdot B^{-1}A^{-2}U^{-1} \cdot B^{-1}A^{-2}U^{-1} = A^{x+2}B^{-1},$$

(G')
$$y' = y'' = 1$$
: $B^{-1}A^{-2y}U^{-1} \cdot B^{-1}A^{-2}U^{-1} \cdot B^{-1}A^{-2}U^{-1} = B^{-1}A^{2-2y}B^{-1}$,

(I'a)
$$x = 1, x' = 3: UA^{-1} \cdot B \cdot A^{3}U^{-1} = B^{-1}A^{4}B^{-1}A,$$

(I'b)
$$x = 3, x' = 1: UA^{-3} \cdot B \cdot A U^{-1} = A^{-1}B^{-1}A^{-4}B,$$

(L'a)
$$y = 1, x' - x = 2$$
: $UA^2U^{-1} \cdot B^{-1}A^{-2}U^{-1} = AB^{-1}A^2B$,

(L'b)
$$y = 1, x' - x = -2; UA^{-2}U^{-1} \cdot B^{-1}A^{-2}U^{-1} = B^{-1}A^{2}B,$$

(M')
$$y = y' = 1$$
: $UA^{2}B \cdot A \cdot B^{-1}A^{-2}U^{-1} = B^{-1}A^{-3}BAB^{-1}$.

The relations (D') and (G') are equivalent to the relation

(4)
$$(UA^2B)^3 = B^2.$$

With the help of (2) and (3), we can also reduce (L'a) and (L'b) to (4) and, finally, eliminate (I'a), (I'b), and (M') with the help of (1), (2), and (3).

We have thus proved that (1), (2) (for p = 2), (3), and (4) is a system of defining relations for the group SL(2, $\mathbf{Z}^{(2)}$), that is, Theorem B.

Added in proof. The preliminary notes (3) are superseded by (4).

References

- H. Behr, Über die endliche Definierbarkeit verallgemeinerter Einheitengruppen, J. Reine Angew. Math. 211 (1962), 123-135.
- 2. H. S. M. Coxeter and W. O. J. Moser, Generators and relations for discrete groups (Springer-Verlag, Berlin, 1957).
- 3. J. Mennicke, On Ihara's modular groups. I (mimeographed notes, University of Göttingen, April, 1967).

4. — On Ihara's modular group, Invent. Math. 4 (1967), 202-228.

Mathematisches Institut der Universität, Göttingen

1438