§5. The symmetric Hilbert modular group for primes p 1 mod 4

Either the surface is rational, or the three curves with $N = 2, 3$ can be blown down. Then S_0 can be blown down and S_1 and S_{-1} give two exceptional curves which intersect in two points. Thus the surface is rational.

Observe that in general the rationality of $Y(\mathfrak{o}_K, \mathfrak{B})$ implies the rationality of $\hat{Y}(\mathfrak{o}_K, \mathfrak{B})$ (Lüroth's theorem [64], Chap. III, § 2). We could show this directly by using our curves in $\hat{Y}(\mathfrak{o}_K, \mathfrak{B})$.

Exercise. Let $K = \mathbf{Q}(\sqrt{69})$. Calculate the arithmetic genera of $\overline{\mathfrak{H}^2/G}$ and $\overline{\mathfrak{H}^2/\hat{G}}$. Prove that the surface $\overline{\mathfrak{H}^2/\hat{G}}$ is rational !

In all cases where we know that the arithmetic genus equals 1 *we have proved rationality.*

§ 5. The symmetric Hilbert modular group for primes $p \equiv 1 \bmod 4$

5.1. Let S be a compact connected non-singular algebraic surface. The fixed point set D of a holomorphic involution T of S (different from the identity) consist of finitely many isolated fixed points $P_1, \ldots, P_r$ and a disjoint union of connected non-singular curves $D_1, \ldots, D_s$.

If there are no isolated fixed points P_j, then S/T is non-singular and the arithmetic genera of S and S/T are related by the formula

$$\chi(S/T) = \frac{1}{2}\left(\chi(S) + \frac{1}{4} c_1[D]\right) \tag{1}$$

where $D = \sum D_i$ and c_1 is the first Chern class of S (see [40], § 3).

Furthermore, if F is a curve on S (not necessarily irreducible) with $T(F) = F$ and F not contained in D and if $\tilde{F}$ is the image curve on S/T, then

$$\tilde{c}_1[\tilde{F}] = \tfrac{1}{2}(c_1[F] + F \cdot D), \quad \text{where } c_1 = \text{first Chern class of } S/T. \tag{2}$$

Proof. If $\pi : S \to S/T$ is the natural projection, then $c_1 = \pi^* \tilde{c}_1 - d$ where $d \in H^2(S, \mathbf{Z})$ is the Poincaré dual of the branching divisor D. Thus

$$(c_1 + d)[F] = \tilde{c}_1[2\tilde{F}].$$

5.2. Let p be a prime $\equiv 1 \bmod 4$. We consider the field $K = \mathbf{Q}(\sqrt{p})$ and its Hilbert modular group G. We make these restrictions throughout § 5 though some of our results are valid more generally.

The involution $(z_1, z_2) \mapsto (z_2, z_1)$ induces an involution T of $\mathfrak{H}^2/G$ and of $\overline{\mathfrak{H}^2/G}$. As mentioned before (4.5), it can be lifted to an involution T of our non-singular model $Y(p)$ because this was obtained by the canonical minimal resolution of all singularities in $\overline{\mathfrak{H}^2/G}$.

We shall study the algebraic surface $Y(p)/T$ (the isolated fixed points of T give rise to quotient singularities of type $(2; 1, 1)$ of this surface), calculate its arithmetic genus and determine for which p the surface is rational (see [39], [40]).

Equivalently we can consider the *symmetric Hilbert modular* group G_T which is an extension of index 2 of G by the involution $(z_1, z_2) \mapsto (z_2, z_1)$ and study $\mathfrak{H}^2/G_T$:

The surface $Y(p)/T$ (with the quotient singularities resolved) is a non-singular model of the compactification of $\mathfrak{H}^2/G_T$.

5.3. The field K has a unit of negative norm. Therefore, the groups G and $\hat{G}$ coincide (1.7) .The class number of K is odd. The ideal class groups C and C^+ are equal and the homomorphism Sq in 3.7 (42) is an isomorphism. Therefore for any ideal $\mathfrak{b} \subset \mathfrak{o}_K$ we can find a matrix $A \in \mathbf{GL}_2^+(K)$ (see 1.3) such that

$$A^{-1}\,\mathbf{SL}_2(\mathfrak{o}_K)\,A = \mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b}) \tag{3}$$

(see 3.7 (40) and 4.1 (7)). If A_1, A_2 are matrices satisfying (3), then, for $B = A_1 A_2^{-1} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have $B\mathbf{SL}_2(\mathfrak{o}_K)\,B^{-1} = \mathbf{SL}_2(\mathfrak{o}_K)$.

Proposition.

If $B \in \mathbf{GL}_2^+(K)$ *and* $B\,\mathbf{SL}_2(\mathfrak{o}_K)\,B^{-1} = \mathbf{SL}_2(\mathfrak{o}_K)$, *then*

$$\sqrt{\det B} \in K, \quad \frac{1}{\sqrt{\det B}} \cdot B \in \mathbf{SL}_2(\mathfrak{o}_K) \tag{4}$$

Proof (compare Maaß [54]). Put $h^2 = \det B$. We may assume that B is an integral matrix. Since

$$\mathbf{SL}_2(\mathfrak{o}_K) \ni B\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)B^{-1} = \begin{pmatrix} 1 - ac/h^2 & a^2/h^2 \\ -c^2/h^2 & 1 + ac/h^2 \end{pmatrix}$$

and a similar formula holds for $B\left(\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right)B^{-1}$, we see that $\frac{1}{h}B$ has coefficients which are algebraic integers. Thus the ideal (a, c) of $\mathfrak{o}_K$ consists exactly of those elements x in $\mathfrak{o}_K$ such that x/h is an algebraic integer. This implies that $(a, c)^2$ equals the principal ideal $(\det B)$. In our case, the ideal class group has odd order. Thus (a, c) is principal and $\det B$ multiplied with a totally positive unit is a square in $\mathfrak{o}_K$. But every totally positive unit is a square of a unit. Therefore $h \in \mathfrak{o}_K$. For the algebraic number theory needed, see [30], § 37.

An ideal is called *admissible* if it is not divisible by any natural number > 1. For any admissible ideal $\mathfrak{b} \subset \mathfrak{o}_K$ we have (4.1) a curve $C(\mathfrak{b})$ on $Y(\mathfrak{o}_K, \mathfrak{b}) = \hat{Y}(\mathfrak{o}_K, \mathfrak{b})$. In view of (3) we have a curve (which we also call $C(\mathfrak{b})$) on our Hilbert modular surface $Y(p)$. The curve is given in $\mathfrak{H}^2/G$ by

$$z_1 = A\zeta,\ z_2 = A'\zeta,\ \zeta \in \mathfrak{H}. \tag{5}$$

Because of (4) it does not depend on the choice of A. (Multiplication of A from the left by an element of $\mathbf{SL}_2(\mathfrak{o}_K)$ does not change the curve.)

We can also say that the surfaces $\mathfrak{H}^2/\mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b})$ are canonically identified and the curves $C(\mathfrak{b})$ are the diagonals in the different representations of $\mathfrak{H}^2/G$ as $\mathfrak{H}^2/\mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b})$. If we change A by multiplying from the right by a rational matrix with positive determinant, we get the same curve, because we make just a change of the parameter $\zeta \in \mathfrak{H}$. This implies that $C(\mathfrak{b}_1) = C(\mathfrak{b}_2)$ if there exists a matrix $A_0 \in \mathbf{GL}_2^+(\mathbf{Q})$ such that $A_0\,\mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b}_1)\,A_0^{-1} = \mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b}_2)$.

Lemma I. If $\mathfrak{b}_1, \mathfrak{b}_2$ are admissible ideals in $\mathfrak{o}_K$, then the curves $C(\mathfrak{b}_1), C(\mathfrak{b}_2)$ coincide if and only if $N(\mathfrak{b}_1) = N(\mathfrak{b}_2)$.

Proof. If $N(\mathfrak{b}_1) = N(\mathfrak{b}_2) = N$, then put $d = N/N((\mathfrak{b}_1, \mathfrak{b}_2))$. We have $(d, N/d) = 1$, because the ideals are admissible. Thus there exists a *rational* matrix of determinant d of the form

$$A_0 = \begin{pmatrix} \alpha_0 d & \beta_1 \\ \gamma_0 & \delta_0 d \end{pmatrix},\ \gamma_0 \equiv 0 \bmod N \tag{6}$$

where $\alpha_0, \beta_1, \gamma_0, \delta_0$ are integers. (Such a matrix occurred in a related context in 4.1). Then, for any A_0 with these properties,

$$A_0\,\mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b}_1)\,A_0^{-1} = \mathbf{SL}_2(\mathfrak{o}_K, \mathfrak{b}_2)$$

which shows that the curves coincide. If the curves coincide, then the norms are equal. (We leave the proof to the reader.)

A natural number $N \geqq 1$ is called admissible (with respect to p) if it is the norm of an admissible ideal. The prime ideal theory of quadratic fields which we always have used tacitly yields the following lemma.

Lemma II. The natural number $N \geqq 1$ is admissible with respect to p if and only if N is not divisible by p^2 and not by any prime $q \neq p$ with $\left(\frac{q}{p}\right) = -1$.

Definition.

In view of Lemma I we have a well-defined curve for any admissible natural number N. This curve on the surface $Y(p)$ will be called F_N.

Lemma III. For the involution T of $Y(p)$ and any admissible N we have $T(F_N) = F_N$.

Proof. If $N = N(\mathfrak{b})$, then $F_N = C(\mathfrak{b})$ is given in $\mathfrak{H}^2/G$ by (5) where A is as in (3). Therefore $T(F_N)$ is the curve $z_1 = A'\zeta$, $z_2 = A\zeta$. But his is $C(\mathfrak{b}')$ which equals $C(\mathfrak{b})$ by lemma I.

Remark. If $N \not\equiv 0 \bmod p$, then $N((\mathfrak{b}, \mathfrak{b}')) = 1$ and the involution T on F_N can be given by the matrix $A_0 = \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$ (see (6)) if we lift T to the non-singular model $\overline{\mathfrak{H}/\Gamma_0(N)}$ of F_N. Thus $\overline{\mathfrak{H}/\Gamma_*(N)}$ is the non-singular model of F_N/T. (see 4.1). In particular, T is not the identity on F_N if $N \not\equiv 0 \bmod p$ and $N > 1$.

5.4. The curves F_1 and F_p (*considered as curves in* $\mathfrak{H}^2/G$) are the only curves which are fixed pointwise under T, (see [14] Part II, [62]). The curve F_p belongs to the ideal $(\sqrt{p}\, e_0)$ where e_0 is a unit of negative norm and can be given by $z_1 = \sqrt{p}\, e_0\zeta$, $z_2 = -\sqrt{p}\, e_0'\zeta$ or by $z_1 = e_0^2 z_2$.

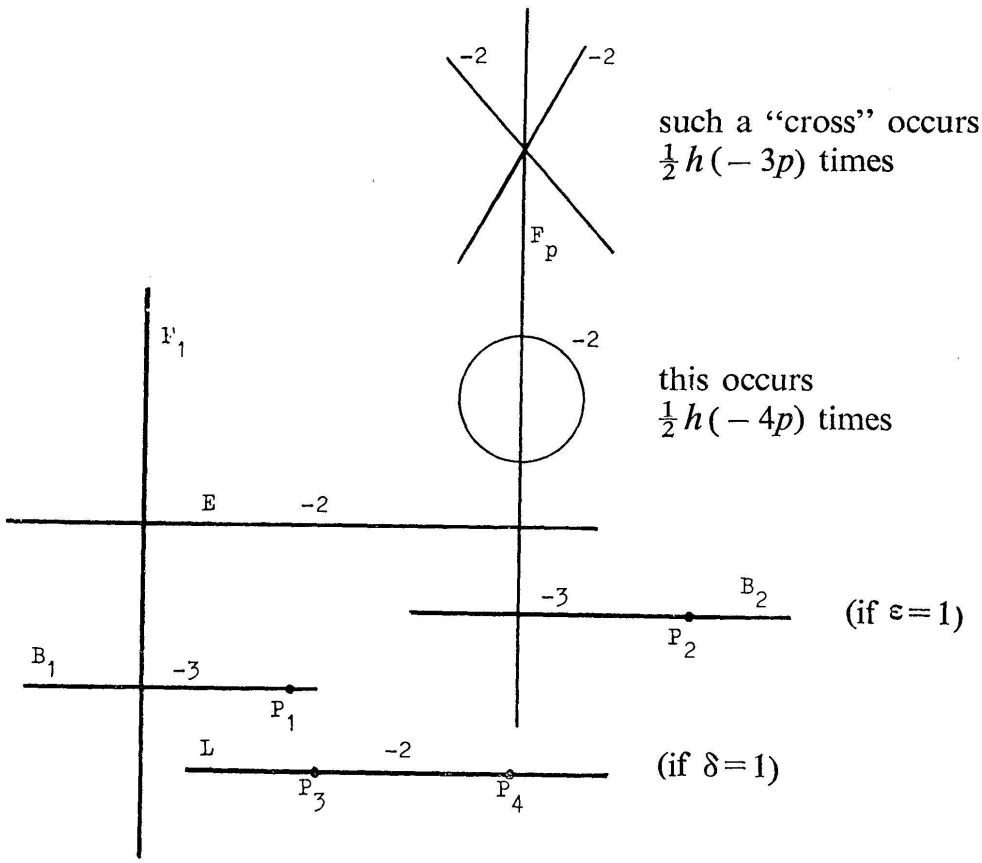
The involution T acts on the quotient singularities of $\mathfrak{H}^2/G$. The description of this action [62] depends on the residue class of p mod 24. Therefore we define

(7) $\qquad \varepsilon = 1$ for $p \equiv 1 \bmod 3$, $\varepsilon = 0$ for $p \equiv 2 \bmod 3$

$\qquad\quad\ \ \delta = 1$ for $p \equiv 1 \bmod 8$, $\delta = 0$ for $p \equiv 5 \bmod 8$

In $\mathfrak{H}^2/G$ the following holds [62]: Of the $h(-4p)$ quotient singularities of order 2, half of them lie on F_p and not on F_1, and one of them lies on F_1 and F_p and is the only intersection point of F_1 and F_p in $\mathfrak{H}^2/G$. There are in addition δ quotient singularities of order 2 which are fixed under T. "They" lie neither on F_1 nor on F_p. The remaining order 2 singularities are interchanged pairwise under T. Of the $h(-3p)$ quotient singularities of order 3, exactly half of them are of type (3; 1, 2). They lie on F_p. There is one singularity of type (3; 1, 1) which lies on F_1 whereas ε such singularities lie on F_p. The remaining singularities of type (3; 1, 1) are interchanged pairwise. For $p = 5$, the two singularities of order 5 are interchanged under T. The involution T acts freely outside F_1, F_p and the quotient singularities. If we pass to the non-singular model $Y(p)$ of $\overline{\mathfrak{H}^2/G}$, we get the following configuration of curves. We omit the curves coming from the quotient singularities which are pairwise interchanged and only show the intersection behaviour outside of the resolved cusp singularities.

(8)



The curves F_1, F_p are pointwise fixed under the involution T of $Y(p)$, therefore they are non-singular curves on $Y(p)$. All curves in the diagram are non-singular and (except F_p) rational. F_p is rational if and only if $p = 5, 13, 17, 29, 41$ (see 5.7). The points P_1, and P_2 if $\varepsilon = 1$, and P_3, P_4

if $\delta = 1$ are the only isolated fixed points of T on $Y(p)$ outside the resolved cusp singularities.

The following lemma is easy to prove and very useful for deducing from Prestel's results [62] that the configuration on $Y(p)$ is as indicated in (8).

Lemma. If S is a compact complex manifold of dimension 2 and T an involution on S which carries the non-singular rational curve C over into itself, then T is the identity on C or T has exactly two fixed points P and Q on C. In the latter case the following holds:

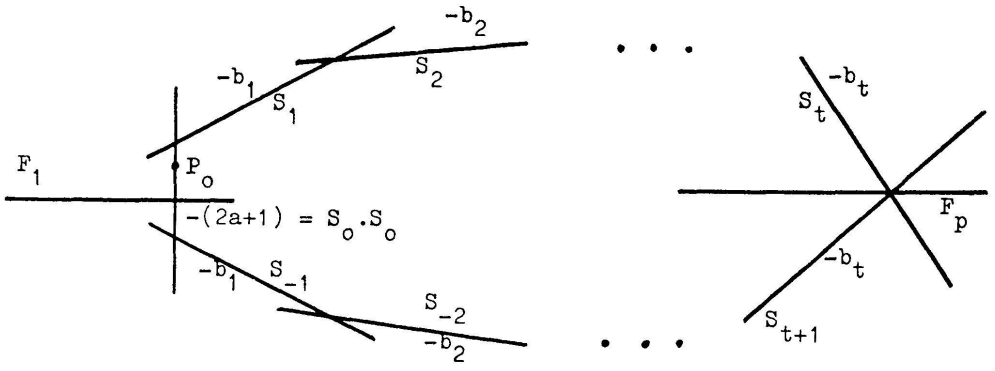
If $C \cdot C$ is odd, then one of the points P, Q is an isolated fixed point of T, the other one is a transversal intersection point of C with one of the (non-singular) curves which are pointwise fixed under T. If $C \cdot C$ is even, then P and Q both are isolated fixed points of T or both are such transversal intersection points with a curve pointwise fixed under T.

The class number h of $K = \mathbf{Q}(\sqrt{p})$ is odd. There are h cusp singularities corresponding to the h ideal classes (see 3.7). The involution T on $\overline{\mathfrak{H}^2/G}$ leaves one cusp fixed and interchanges the others pairwise. T maps the cusp of type (M, U^2) where M is a fractional ideal representing an ideal class to the cusp of type (M', U^2). If M is the $\mathbf{Z}$-module $\mathbf{Z} \cdot w + \mathbf{Z} \cdot 1$ (with $0 < w' < 1 < w$), then M' is strictly equivalent to $\mathbf{Z}\frac{1}{w'} + \mathbf{Z} \cdot 1$. The resolution of (M, U^2) is given by the primitive cycle of the purely periodic) continued fraction of w, the resolution of (M', U^2) by the primitive cycle of $\frac{1}{w'}$ which is the same cycle in opposite order. The involution on $Y(p)$ maps the cycle of curves in the resolution of (M, U^2) onto the cycle of curves in the resolution of (M', U^2). The fixed cusp is of type (M, U^2) where $M = \mathbf{Z}w_0 + \mathbf{Z} \cdot 1$ and where $w_0 = \frac{1}{2}(\{\sqrt{p}\} + \sqrt{p})$, see 4.5 (29). It is the cusp at ∞.

THEOREM. *The length r of the cycle of $w_0 = \frac{1}{2}(\{\sqrt{p}\} + \sqrt{p})$ is an odd number $r = 2t + 1$. The involution T on $Y(p)$ maps the curve S_k to the curve S_{-k}* (see 4.5). *The curve F_1 intersects S_0 transversally. It has the characteristic* $(0|\, 0, 1)$. *The curve F_p intersects S_{-t} and S_t, it has the characteristic* $(-t|\, 1, 1)$. *We put* $\{\sqrt{p}\} = 2a + 1$. *The intersection*

behaviour of the cycle of curves with F_1 and F_p is illustrated by the following diagram.

(9)



$$w_0 = [[2a+1, b_1, \ldots, b_t, b_t, \ldots, b_1]]$$

The point P_0 indicates an isolated fixed point of T. The points P_0, P_1, and P_2 (if $\varepsilon = 1$), and P_3, P_4 (if $\delta = 1$) are all the isolated fixed points of T. The curves F_1, F_p are the only one-dimensional components of the fixed point set.

Proof. As in 2.5 and 3.10 we denote ordinary continued fractions by $[a_0, a_1, a_2, \ldots]$. Then, since $a = \left[\frac{1+\sqrt{p}}{2}\right]$,

$$(10) \qquad w_0 = \frac{2a+1+\sqrt{p}}{2} = [2a, \overline{a_1, \ldots, a_m, a_m, \ldots, a_1, 2a-1}]$$

(See [60], § 30. Because there exists a unit of negative norm, the length of the primitive period in (10) is odd.)

If one applies the formula which transforms the continued fraction (10) into a continued fraction of our type (see 2.5 (19)) one has to go twice over the period in (10). We have

$$(11) \quad w_0 = [[\overline{2a+1, \underbrace{2, \ldots, 2}_{a_1-1}, \ldots, a_1+2, \underbrace{2, \ldots, 2}_{2a-2}, a_1+2, \ldots, \underbrace{2, \ldots, 2}_{a_1-1}}]]$$

Thus the length r of the primitive cycle of w_0 is odd ($r = 2t+1$). In fact, $t = a_1 + \ldots + a_m + a - 1$. Under the involution T only S_0 (self-intersection number $-(2a+1)$) is carried over into itself. *The only symmetric characteristics are* $(0 \mid 0, 1)$ *and* $(-t \mid 1, 1)$. The existence of the isolated fixed point P_0 follows from the preceding lemma. Q.E.D.

For the number w_0 in (11) we wish to calculate w_{t+k} (where $k = 1, \ldots, a$), see 4.2. The continued fraction $[[\ldots]]$ of w_{t+k} begins with $a-k$ two's. Using again formula 2.5 (19) we obtain

$$0 - \frac{1}{w_{t+k}} = [-1, a-k+1, a_1, a_2, \ldots]$$

$$= -1 + \frac{1}{w_0 - a - k + 1}$$

which yields

$$w_{t+k} = \frac{\sqrt{p} - (2k-3)}{\sqrt{p} - (2k-1)} = \frac{M_{t+k} + \sqrt{p}}{2N_{t+k}}$$

where

$$(12) \qquad N_{t+k} = \frac{1}{4}(p - (2k-1)^2),\ M_{t+k} = 2N_{t+k} + (2k-1)$$

F_p has the characteristic $(-t \mid 1, 1) = (t+1 \mid 1, 1)$ which was obtained in the above proof by a symmetry argument.

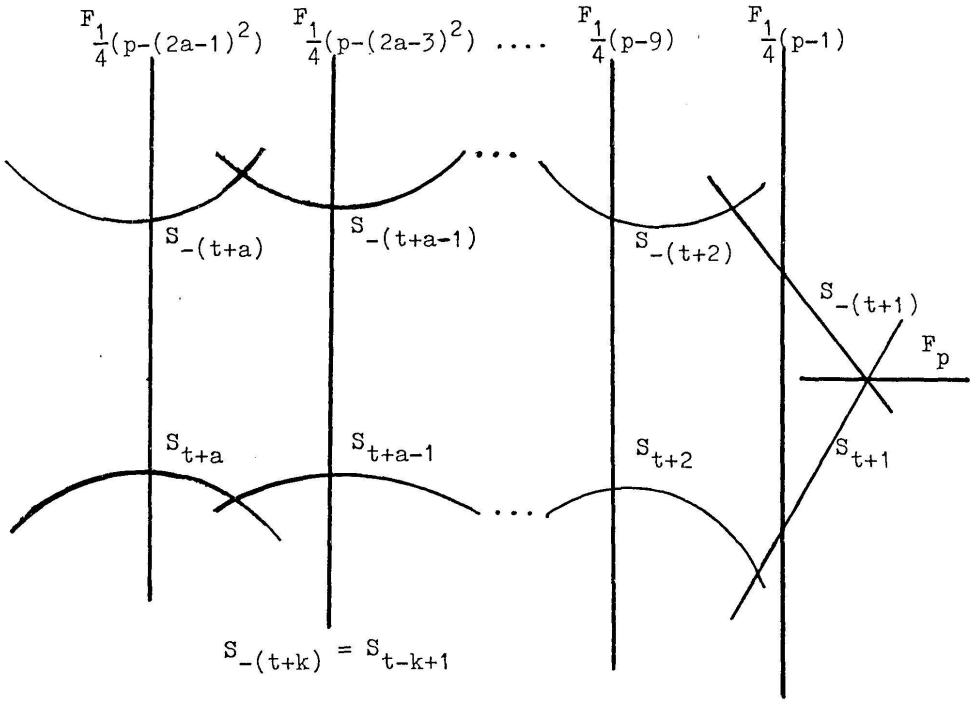
It follows also from the theorem in 4.1, because

$$N_{t+1} + N_t + M_t = N_{t+1} + N_{t+1} + M_{t+1} = 4N_{t+1} + 1 = p.$$

In view of (12) and the theorem in 4.1 we have the following proposition.

Proposition. On the Hilbert modular surface $Y(p)$ the cusp at ∞ gives the following configuration of curves $\left(a = \left[\frac{1+\sqrt{p}}{2}\right]\right)$

(13)



We have $S_{t+k} \cdot S_{t+k} = S_{-(t+k)} \cdot S_{-(t+k)} = -2$ *for* $1 \leqq k \leqq a-1$ *and* $S_{t+a} \cdot S_{-(t+a)} = -(a_1+2)$. *If* $p = (2a-1)^2 + 4$, *then* $S_{-(t+a)} = S_{t+a} = S_0$, *the curve* $F_{\frac{1}{4}(p-(2a-1)^2)}$ *equals* F_1 *and the diagram has to be changed accordingly. In this case we have*

$$w_0 = [2a, \overline{2a-1}] = [[\overline{2a+1, \underbrace{2, \ldots, 2}_{2a-2}}]]$$

and

$$S_{t+a} \cdot S_{t+a} = S_0 \cdot S_0 = -(2a+1).$$

We do not claim that the F_N are non-singular and do not indicate their mutual intersections nor their intersections with F_p. The intersections indicated are transversal.

5.5. The curve F_1 on $Y(p)$ is non-singular. It follows from (8) and 4.3 that it is exceptional. In general, we do not know whether F_N is non-singular. In view of 4.3 (24) the curves F_2, F_3, F_4 are candidates for exceptional curves. In fact, it follows from Corollaries I, II in 4.4 that they are exceptional if $Y(p)$ is not rational. $Y(p)$ is rational if and only if $p = 5$, 13, 17. Thus we have

Lemma. If p *is a prime* $\equiv 1 \bmod 4$ *and* > 17, *then the curves* F_N *on the Hilbert modular surface* $Y(p)$ *are exceptional for* $N = 1, 2, 3, 4$ *provided* N *is admissible (see* 5.3*):*

We always have the curve F_1. *The curves* F_2, F_4 *exist for* $p \equiv 1 \bmod 8$. *The curve* F_3 *exists for* $p \equiv 1 \bmod 3$.

For the following discussion we assume $p > 17$. The curves F_1, E, B_1 in diagram (8) can be blown down successively. In view of corollary III in 4.4, the curves F_2, F_3, F_4 are disjoint and do not intersect any of the curves F_1, E, B_1. According to the lemma in 5.4 the curves F_2, F_3, F_4 pass through exactly one of the isolated fixed points of the involution T.

For F_3 the value $c_1[F_3]$ equals 1, therefore by 4.3 it meets in $\mathfrak{H}^2/G$ exactly one quotient singularity of type (3; 1, 1), thus it must be the one which is fixed under T. It intersects B_2 (see (8)) only in P_2 and transversally because otherwise we would have $c_1[F_3] > 1$. The curve F_4 has the model $\overline{\mathfrak{H}/\Gamma_0(4)}$ which has three cusps. Therefore F_4 must intersect the curves of the resolved cusps of $\overline{\mathfrak{H}^2/G}$ in three points. One of them is fixed under T. Thus F_4 passes through P_0.

The curve F_2 passes through P_3 or P_4 in diagram (8), say P_3. It intersects L transversally in P_3 and does not intersect L in any other point, because otherwise L would give in the surface with F_2 blown down a curve $\hat{L}$ with $c_1[\hat{L}] \geqq 2$. The curves F_2, L can be blown down successively. Therefore L is disjoint to any exceptional curve different from F_2.

We have found an exceptional curve passing through P_0 only for $p \equiv 1 \bmod 8$. But there exists such a curve F for any $p > 17$.

For the cusp at ∞ we put as before $w_0 = \frac{1}{2}(\{\sqrt{p}\} + \sqrt{p}) = \frac{1}{2}(2a+1+\sqrt{p})$. The involution T is given in the coordinate system (u_0, v_0) by

$$(u_0, v_0) \mapsto (u_0^{-1}, u_0^{-(2a+1)} \cdot v_0), \tag{14}$$

as follows from 2.3 (9). The isolated fixed point P_0 of T has the coordinates $(-1, 0)$. Thus it lies on the curve $F \subset Y(p)$ given by $u_0 = -1$ which can be presented in $\mathfrak{H} \times \mathfrak{H}$ by

$$z_1 = \zeta + \frac{w_0}{2},\ z_2 = \zeta + \frac{w_0'}{2},\ (\zeta \in \mathfrak{H}). \tag{15}$$

Let Γ be the subgroup of those matrices $\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ of

$$\begin{pmatrix} 1 & -w_0/2 \\ 0 & 1 \end{pmatrix} \mathbf{SL}_2(\mathfrak{o}_K) \begin{pmatrix} 1 & w_0/2 \\ 0 & 1 \end{pmatrix}$$

which, when acting on $\mathfrak{H}^2$ carry the diagonal into itself. The curve $\overline{\mathfrak{H}/\Gamma}$ is a non-singular model of F. The group Γ is characterized by 4.1 (1), but the second condition is impossible. Thus Γ is the subgroup of $\mathbf{SL}_2(\mathbf{Q})$ of matrices $\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ for which

$$\begin{pmatrix} \alpha + \gamma w_0/2 & -\alpha w_0/2 + \beta - \gamma w_0^2/4 + \delta w_0/2 \\ \gamma & \delta - \gamma w_0/2 \end{pmatrix}$$

is integral. Since $w_0, 1$ is a $\mathbf{Z}$-base of $\mathfrak{o}_K$, we get that α, δ are integers and γ is an even integer. We have

$$\begin{aligned} &-\alpha w_0/2 + \beta - \gamma w_0^2/4 + \delta w_0/2 = \\ &(-\alpha/2 - \gamma(2a+1)/4 + \delta/2)\, w_0 + \beta + \gamma((2a+1)^2 - p)/16 \end{aligned} \tag{16}$$

If $p \equiv 1 \bmod 8$, then β is an integer and $\alpha\delta - \beta\gamma = 1$ implies $\alpha \equiv \delta \bmod 2$ and $\gamma \equiv 0 \bmod 4$, because the coefficient of w_0 in (16) must be integral. Thus $\Gamma = \Gamma_0(4)$ in this case.

If $p \equiv 5 \bmod 8$, then $\Gamma_0(4) \subset \Gamma$. We put $\gamma = 2\gamma^*$ and $\beta = \beta^*/2$. Then γ^*, β^* are integers which are congruent modulo 2. We have $\alpha + \delta \equiv \gamma^* \bmod 2$.

The matrix $\left(\begin{smallmatrix} 1 & -\frac{1}{2} \\ 2 & 0 \end{smallmatrix}\right)$, whose third power is $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, satisfies these conditions. Γ is a normal extension of index 3 of $\Gamma_0(4)$. The three cusp of $\mathfrak{H}/\Gamma_0(4)$ are identified. $\overline{\mathfrak{H}/\Gamma}$ is a rational curve. Put $\tilde{\Gamma} = \Gamma/\{1, -1\}$. We have $a_3(\tilde{\Gamma}) = 2$ ($a_r(\tilde{\Gamma}) = 0$ otherwise) and $\sigma(\tilde{\Gamma}) = 1$.

Therefore $c_1(\tilde{\Gamma}) = 1$ (see the definition in 4.3), and the curve F is exceptional. It passes through the isolated fixed point P_0 of T. For $p \equiv 1 \bmod 8$, the curve F equals F_4 because two different exceptional curves do not intersect. We have $T(F) = F$.

We can now state the following proposition.

Proposition. If we blow down the curves F_1, E, B_1, F, and F_2, L (for $\delta = 1$), and F_3 (for $\varepsilon = 1$) on the surface $Y(p)$ for $p > 17$, then we obtain a non-singular algebraic surface $Y^0(p)$. The involution T is also defined on $Y^0(p)$. It does not have any isolated fixed point. The curve F_p has a non-singular image F_p^0 in $Y^0(p)$ which is the complete fixed point set of T.

5.6. If c_1 is again the first Chern class of $Y(p)$, then

$$c_1[F_p] = -\frac{p+1}{6} + \frac{\varepsilon}{3} + 2 \tag{17}$$

This follows from 4.3 (19), because $[\mathbf{SL}_2(\mathbf{Z}) : \Gamma_0(p)] = p + 1$ and $[\Gamma^*(p) : \Gamma_0(p)] = 2$. We further use (8) and (9).

Let us now assume that $Y(p)$ is not rational which is the case for $p > 17$. In $Y(p)$ we have blown down $3 + 1 + 2\delta + \varepsilon$ curves and obtained the surface $Y^0(p)$ on which T has the fixed point set F_p^0. Let c_1^0 be the first Chern class of $Y^0(p)$. Then

$$c_1^0[F_p^0] = -\frac{p+1}{6} + \frac{\varepsilon}{3} + 2 + 2 + 1 + 2\delta + \varepsilon. \tag{18}$$

This follows from 4.4 (25a) using that F, F_2, F_3 intersect F_p transversally in exactly one point (see the lemma in 5.4). By 5.1 (1) the number $c_1^0[F_p^0]$ must be divisible by 4. We have

$$\frac{1}{4} c_1^0[F_p^0] = -\left[\frac{p-29}{24}\right], \tag{19}$$

since $\frac{1}{4}\left(\frac{\varepsilon}{3}+2\delta+\varepsilon\right)<1$. The surface $Y^0(p)/T$ is a non singular model for the compactification of $\mathfrak{H}^2/G_T$ (see 5.2). The arithmetic genus of $Y^0(p)/T$ will be denoted by $\chi_T(p)$. In 3.12 we have given a formula for the arithmetic genus of $Y(p)$ which we shall call here $\chi(p)$. Then

$$\chi(p)=\frac{1}{2}\zeta_K(-1)+\frac{h(-4p)}{8}+\frac{1}{6}h(-3p), \tag{20}$$

where $K=\mathbf{Q}(\sqrt{p})$. By 5.1 (1) and (19) the arithmetic genera $\chi(p)$ and $\chi_T(p)$ are related by the formula [40]

$$\chi_T(p)=\frac{1}{2}\left(\chi(p)-\left[\frac{p-29}{24}\right]\right), \tag{21}$$

(compare [14], Part II, Satz 2).

This formula is also valid for $p=5, 13, 17$. In these cases the surface $Y(p)$ and therefore also $Y^0(p)/T$ are rational and (21) reduces to $1=\frac{1}{2}(1+1)$. It was shown in [40] that

$$\chi_T(p)>\frac{p^{3/2}}{1440}-\frac{p+1}{48} \qquad \text{(compare 3.12),}$$

and explicit calculations gave the result that $\chi_T(p)=1$ for exactly 24 primes, namely for all primes ($\equiv 1 \bmod 4$) smaller than the prime 193 and for $p=197, 229, 269, 293, 317$.

We wish to show in the next sections that the surfaces $Y^0(p)/T$ are rational for these primes. Since the rationality is already known for $p=5, 13, 17$ it remains to consider 21 primes. Since the first Betti number of $Y(p)$ vanishes (3.6), the same holds for $Y^0(p)/T$. Thus the rationality criteria of 4.4 (Corollaries I, II, III) can be applied.

5.7. The curve F_N in $Y(p)$ (for an admissible natural number $N>4$) projects down to a curve F_N^0 in $Y^0(p)$ and to a curve $F_N^*=F_N^0/T$ in $Y^0(p)/T$. If N is not divisible by p, then F_N^* has $\overline{\mathfrak{H}/\Gamma^*(N)}$ as non-singular model (see the remark in 5.3). We have a commutative diagram:

$$\begin{array}{ccccc}
\overline{H/\Gamma_0(N)} & \to & F_N^0 & \subset & Y_0(p) \\
\downarrow & & \downarrow & & \downarrow \\
\overline{H/\Gamma_*(N)} & \to & F_N^* & \subset & Y_0(p)/T
\end{array}$$

There is an involution τ on $\overline{\mathfrak{H}/\Gamma_0(N)}$ compatible with T and having $\overline{\mathfrak{H}/\Gamma_*(N)}$ as orbit space. Recall that F_p^0 is the fixed point set of T on $Y^0(p)$. Thus the intersection number $F_N^0 \cdot F_p^0$ is greater or equal to the number $\mathrm{fix}(\tau)$ of fixed points of τ on $\overline{\mathfrak{H}/\Gamma_0(N)}$:

$$F_N^0 \cdot F_p^0 \geqq \mathrm{fix}(\tau) = 2e\left(\overline{\mathfrak{H}/\Gamma_*(N)}\right) - e\left(\overline{\mathfrak{H}/\Gamma_0(N)}\right) \tag{22}$$

Let c_1^* be the first Chern class of $Y^0(p)/T$. By 5.1 (2) we get

$$c_1^*[F_N] = \tfrac{1}{2}\left(c_1^0[F_N^0] + F_N^0 \cdot F_p^0\right)$$

Since $c_1^0[F_N^0] \geqq c_1[F_N] \geqq c_1(N)$, see 4.3 and 4.4 (25a), the following estimate is obtained:

$$c_1^*[F_N^*] \geqq \tfrac{1}{2}c_1(N) + e\left(\overline{H/\Gamma_*(N)}\right) - \tfrac{1}{2}e\left(\overline{H/\Gamma_0(N)}\right) \tag{23}$$

The right side of (23) only depends on N. We shall denote it by $c_1^*(N)$ and have

$$c_1^*[F_N^*] \geqq c_1^*(N). \tag{24}$$

There are explicit formulas for the Euler numbers or equivalently the genera of the curves $\overline{\mathfrak{H}/\Gamma_*(N)}$, see [16], p. 357, and [13]. Helling [32] has shown that there are exactly 37 values $N \geqq 2$ for which $\overline{\mathfrak{H}/\Gamma_*(N)}$ is rational. (In [16], p. 367, Fricke omits the value $N = 59$). We shall give a list of the $c_1^*(N)$ for the 34 values $\geqq 5$.

By the definition of $c_1(N)$ we get:

If $\overline{\mathfrak{H}/\Gamma_*(N)}$ is rational, then (for $N \geqq 5$)

$$c_1^*(N) = 3 - g_0(N) - \tfrac{1}{2}\left(a_2(N) + a_3(N) + \sigma_0(N)\right). \tag{25}$$

Using [13] we obtain the following list:

$$e\left(\overline{\mathfrak{H}/\Gamma_*(N)}\right) = 2$$

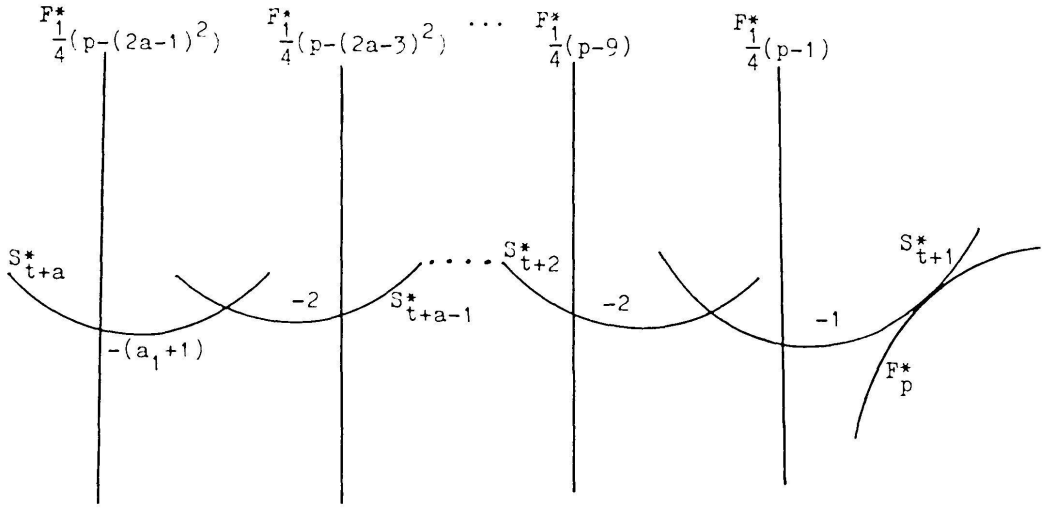
(26)

N	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
$c_1^*(N)$	1	1	1	1	1	0	1	0	0	0	0	0	0	−1	0	−1	−1

N	23	24	25	26	27	29	31	32	35	36	39	41	47	49	50	59	71
$c_1^*(N)$	0	−2	−1	−2	−1	−1	−1	−2	−2	−4	−3	−2	−2	−3	−6	−3	−4

5.8. The curves F_N will be used for rationality proofs. Consider the diagram (13) for $p > 17$. We have $\frac{1}{4}(p-(2a-3)^2) \geqq 5$. It follows from 4.2 (15) that the exceptional curves F_1, F, F_2, F_3 do not intersect S_{t+k} and $S_{-(t+k)}$ for $1 \leqq k \leqq a-1$. These exceptional curves also do not meet S_{t+a} and $S_{-(t+a)}$ if $\frac{1}{4}(p-(2a-1)^2) \geqq 5$. In this case, the configuration (13) is not changed by passing to $Y^0(p)$. If we apply the involution T we get the following configuration on $Y^0(p)/T$.

(27)



If $\frac{1}{4}(p-(2a-1)^2) < 5$, the diagram has to be changed. But the subdiagram of (27) obtained by not showing $F^*_{\frac{1}{4}(p-(2a-1)^2)}$ and S^*_{t+a} exists on the surface $Y^0(p)/T$ for any $p > 17$.

We do not know whether the curves $F^*_{\frac{1}{4}(p-(2k-1)^2)}$ are non-singular and do not claim anything about their mutual intersection behaviour. The S^*_{t+k} are the image of S_{t+k} and $S_{-(t+k)}$. They are non-singular. The equation $S^*_{t+1} \cdot S^*_{t+1} = -1$ or equivalently $c^*_1[S^*_{t+1}] = 1$ follows from 5.1 (2). The curves S^*_{t+k} $(1 \leqq k \leqq a-1)$ can be blown down successively. Then $F^*_{(p-(2k-1)^2)/4}$ gives in the resulting surface a curve for which the value of the first Chern class of the new surface on this curve is greater or equal to $c^*_1((p-(2k-1)^2)/4) + a - k$.

Proposition. Let p be a prime $\equiv 1 \bmod 4$ (and $p > 17$). The non-singular model $Y^0(p)/T$ for the symmetric Hilbert modular group is rational if there exists a natural number k with $1 \leqq k \leqq a-1 = \left[\frac{\sqrt{p}-1}{2}\right]$ such that

$$c^*_1((p-(2k-1)^2)/4) + a - k \geqq 2$$

This is a consequence of corollary I in 4.4. For the above proposition one does not need any assumption about the genus of F_N where $N = \frac{1}{4}(p-(2k-1)^2)$. However, we shall try to get through using the N listed in 5.7 for which the curves F_N are rational.

The tables in 5.7 give immediately

$$c_1^*\left(\frac{p-1}{4}\right) + a - 1 \geqq 2$$

for $p = 29, 37, 41, 53, 61, 73, 97, 101, 109, 197$.

We find

$$c_1^*\left(\frac{p-9}{4}\right) + a - 2 \geqq 2$$

for $p = 89, 137, 293$.

For $p = 173$ we have

$$c_1^*\left(\frac{p-81}{4}\right) + a - 5 = c_1^*(23) + 7 - 5 = 2$$

For the remaining 7 primes 113, 149, 157, 181, 229, 269, 317 we shall try to use the following lemma.

Lemma. We keep the notations of the preceding proposition. Suppose there exist two natural numbers k_1, k_2 with $1 \leqq k_1 < k_2 \leqq a-1$ such that

$$c_1^*\left(\frac{p-(2k_i-1)^2}{4}\right) + a - k_i = 1 \qquad \text{for } i = 1,2$$

Then $Y^0(p)/T$ is rational.

Proof. Blowing down $S^*_{t+1}, \ldots, S^*_{t+a-1}$ in $Y^0(p)/T$ gives a surface in which the images of $F^*_{N_i}$ $(N_i = (p-(2k_i-1)^2)/4,\ i = 1, 2)$ are exceptional curves or the surface is rational (4.4, Corollary II). If we have the two exceptional curves, then they intersect and the surface is rational by Corollary III in 4.4.

The assumptions of the lemma are true for $p = 113$ and $k_1 = 2$, $k_2 = 4$, for $p = 149$ and $k_1 = 4$, $k_2 = 5$, for $p = 157$ and $k_1 = 4$ and $k_2 = 5$, for $p = 181$ and $k_1 = 5$, $k_2 = 6$, for $p = 229$ and $k_1 = 6$, $k_2 = 7$, for $p = 317$ and $k_1 = 5$, $k_2 = 8$.

For $p = 269$ we have $a = 8$. The curve S^*_{t+8} has self-intersection number -3. It intersects F^*_{11}, since $11 = \frac{269 - 15^2}{4}$. Either the surface is rational or F^*_{11} is exceptional. If F^*_{11} is exceptional, then we blow down $F^*_{11}, S^*_{t+1}, \ldots, S^*_{t+8}$. The curve F^*_{47} $(k = 5)$ gives in the resulting surface $\tilde{Y}$ a curve $\tilde{D}$ with $\tilde{c}_1 [D] \geqq 2$ where $\tilde{c}_1$ is the first Chern class of $\tilde{Y}$.

We have proved the desired result.

THEOREM. *Let p be a prime $\equiv$ a mod 4. Let G_T be the symmetric Hilbert modular group for $K = \mathbf{Q}(\sqrt{p})$. Then the surface $\overline{\mathfrak{H}^2/G_T}$ is rational, (or equivalently the field of meromorphic automorphic functions with respect to G_T is a purely transcendental extension of $\mathbf{C}$), if and only if $p < 193$ or $p = 197, 229, 269, 293, 317$.*

5.9. Example. If the prime $p \equiv 1 \bmod 4$ is of the form

$$p = (2a-1)^2 + 4,$$

then $$w_0 = \frac{2a + 1 + \sqrt{p}}{2} = [[\overline{2a+1, \underbrace{2, \ldots, 2}_{2a-2}}]]$$

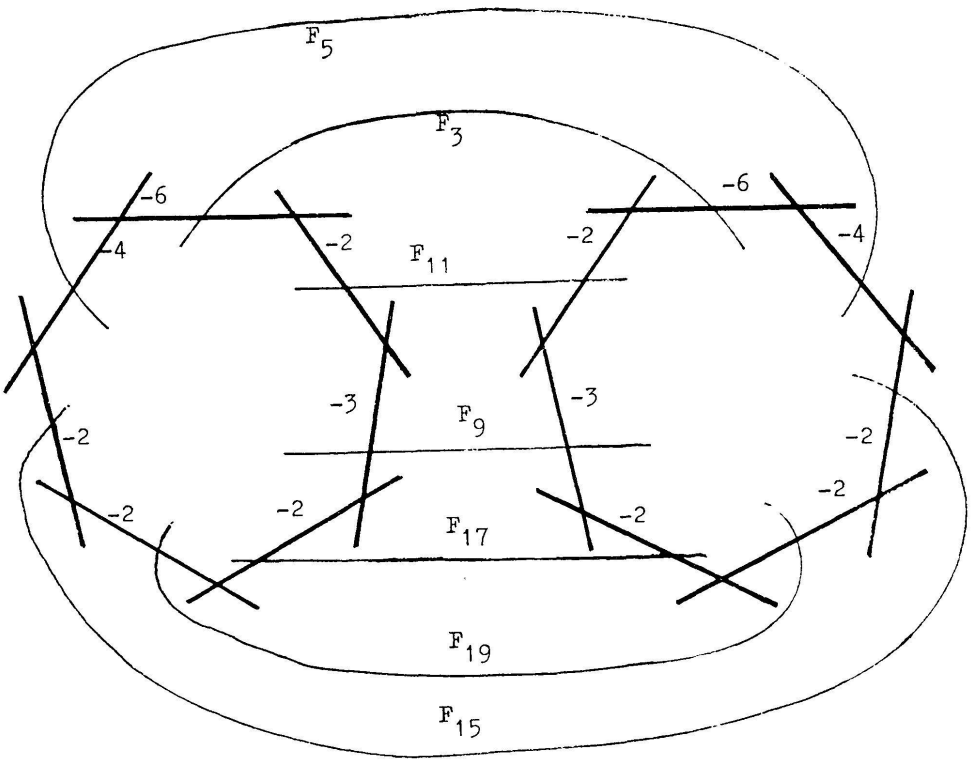
and we have in diagram (13) that $S_{t+a} = S_{-(t+a)} = S_0$. Since $(p - (2a-3)^2)/4 = 2a - 1$, the smallest admissible $N > 1$ which can be written in the form $x^2 N_k + xy M_k + y^2 N_{k-1}$ (with integers $x, y \geqq 0$) equals $2a - 1$ (see 4.2 and 5.4 (12)). Any divisor d of $2a - 1$ is admissible. If d is a prime dividing $2a - 1$ and $1 < d < 2a - 1$, then the curve F_d has two cusps and does not pass through the cusp at ∞ of $\overline{\mathfrak{H}^2/G}$. Thus there must be other cusps of $\overline{\mathfrak{H}^2/G}$. We have proved

Proposition. If $p = (2a-1)^2 + 4$ (p prime) and if $2a - 1$ is not a prime, then $h(p) > 1$. (See [29], [51]).

The first example is $p = 229 = 15^2 + 4$. We have $h(p) = 3$. The number 229 is the only one of our 24 primes in the preceding theorem with class number greater than one. (If $(2a-1) \equiv \pm 2 \bmod 7$, then 7 is admissible for p. Thus, also in this case $h(p) > 1$ provided $2a - 1 > 7$. Example: $p = 1373 = 37^2 + 4$, $h(p) = 3$.)

The cycles for the 2 cusps not at ∞ of $Y(229)$ look as follows

(28)



We also have drawn some curves. The curve F_{15} has $\overline{\mathfrak{H}/\Gamma_0(15)}$ as non-singular model. This has 4 cusps corresponding to the fact that F_{15} also passes through the cusp at ∞ of $Y(p)$, namely through the curves S_1 and S_{-1} of this cusp. One can show that F_9 passes through S_0 of the cusp at ∞ in two points ($\overline{\mathfrak{H}/\Gamma_0(9)}$ has 4 cusps).

If N is admissible and is a product of k different primes ($\neq p$), then $\overline{\mathfrak{H}/\Gamma_0(N)}$ has 2^k cusps. The 2^k intersections of F_N with the resolved cusps in $Y(p)$ correspond to 2^k admissible ideals $\mathfrak{b}$ with $N(\mathfrak{b}) = N$ (see 5.3).

In general, it is possible to give a complete description of the intersection of F_N with the resolved cusps of $Y(p)$. The corresponding theory can be developed for any Hilbert modular surface.

ADDED IN PROOF:

A. Selberg has informed me that he has proved the following result. If Γ is a discrete irreducible subgroup of $(\mathbf{PL}_2^+(\mathbf{R}))^n$ such that $\mathfrak{H}^n/\Gamma$ has finite volume, but is not compact, then Γ is conjugate in $(\mathbf{PL}_2^+(\mathbf{R}))^n$ to a group commensurable with the Hilbert modular group of some totally real field K with $[K: \mathbf{Q}] = n$.

Thus Selberg's conjecture mentioned in the remark at the end of 1.5 is true. Actually, Selberg's results are more general. The proof has not been published yet. There is a sketch (still involving additional assumptions which could be eliminated later) in the Proceedings of the 15th Scandinavian

congress, Oslo 1968, Lecture Notes in Mathematics, Springer Verlag, vol. 118, in particular pp. 106-113.

REFERENCES

[1] ATIYAH, M. F. *Elliptic operators and compact groups.* Lectures, The Institute for Advanced Study (1971).

[2] —— and R. BOTT. A Lefschetz fixed point formula for elliptic complexes: I, II. *Ann. of Math. 86*, pp. 374-407 (1967) and *88*, pp. 451-491 (1968).

[3] —— and I. M. SINGER. The index of elliptic operators III. *Ann. of Math. 87*, 546-604 (1968).

[4] BAILY, W. L. and A. BOREL. Compactification of arithmetic quotients of bounded symmetric domains. *Ann. of Math. 84*, 442-528 (1966).

[5] BLUMENTHAL, O. Über Modulfunktionen von mehreren Veränderlichen. *Math. Ann. 56*, 509-548 (1903) und *58*, 497-527 (1904).

[6] BOREWICZ, S. I. und I. R. ŠAFAREVIČ. *Zahlentheorie.* Birkhäuser Verlag, Basel (1966).

[7] BOURBAKI, N. *Variétés différentielles et analytiques, Fascicule de résultats*, Hermann, Paris (1967).

[8] CARTAN, H. *Sur les groupes de transformations analytiques.* Act. Sc. Ind. *198*, Hermann, Paris (1935).

[9] —— Quotient d'un espace analytique par un groupe d'automorphismes. *Algebraic Geometry and Topology, Symposium S. Lefschetz*, pp. 90-102, Princeton University Press (1957).

[10] CHERN, S. S. On the curvatura integra in a Riemannian manifold. *Ann. of Math. 46*, 674-684 (1945).

[11] CHRISTIAN, U. Zur Theorie der Hilbert-Siegelschen Modulfunktionen. *Math. Ann. 152*, 275-341 (1963).

[12] COHN, H. Support polygons and the resolution of modular functional singularities *(submitted to Acta Arithmetica).*

[13] FELL, H., M. NEWMAN and E. ORDMAN. Tables of genera of groups of linear fractional transformations. *J. of Research of the Nat. Bureau of Standards. 67B*, 61-68 (1963).

[14] FREITAG, E. Über die Struktur der Funktionenkörper zu hyperabelschen Gruppen I, II. *Journal f. d. r. u. a. Math. (Crelle) 247*, 97-117 (1971) und *254*, 1-16 (1972).

[15] —— Lokale und globale Invarianten der Hilbertschen Modulgruppe. *Invent. Math. 17*, 106-134 (1972).

[16] FRICKE, R. *Die elliptischen Funktionen und ihre Anwendungen. Zweiter Teil.* B. G. Teubner, Leipzig und Berlin (1922).

[17] GRAUERT, H. Über Modifikationen und exzeptionelle analytische Mengen. *Math. Ann. 146*, 331-368 (1962).

[18] —— und R. Remmert. Komplexe Räume. *Math. Ann. 136*, 245-318 (1958).

[19] GUNDLACH, K.-B. Über die Darstellung der ganzen Spitzenformen zu den Idealstufen der Hilbertschen Modulgruppe und die Abschätzung ihrer Fourierkoeffizienten. *Acta math. 92*, 309-345 (1954).

[20] —— Quotientenraum und meromorphe Funktionen zur Hilbertschen Modulgruppe. *Nachr. Akad. Wiss. Göttingen, Nr. 3*, pp. 77-85 (1960).

[21] —— Some new results in the theory of Hilbert's modular group, *Contribution to function theory, Tata Institute*, pp. 165-180, Bombay (1960).

[22] —— Die Bestimmung der Funktionen zu einigen Hilbertschen Modulgruppen. *Journal f. d. r. u. a. Math. (Crelle) 220*, 109-153 (1965).