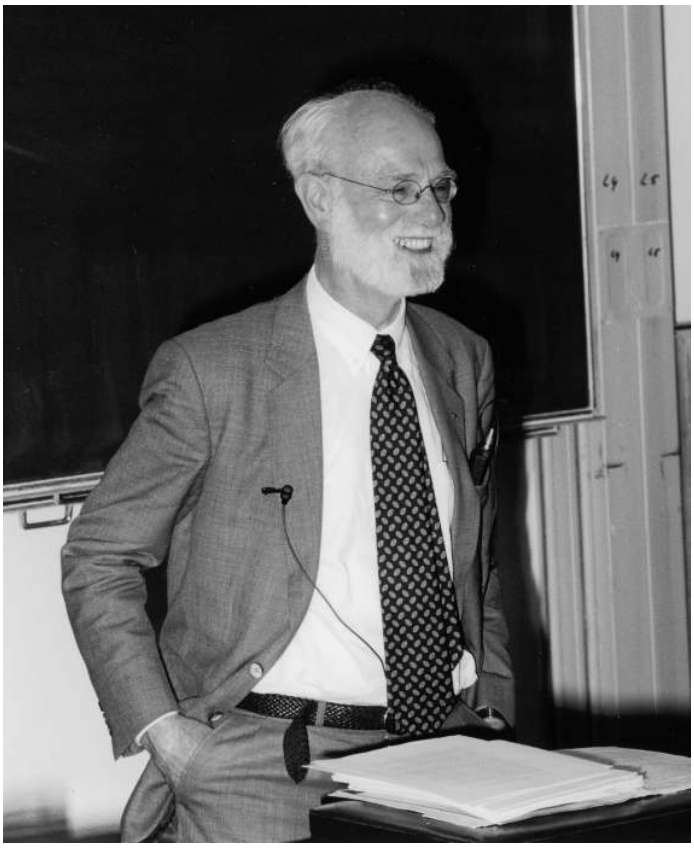
T.A. Springer
Mathematisch Instituut, Universiteit Utrecht
Budapestlaan 6, 3584 CD Utrecht
springer@math.uu.nl

Kloosterman's work on representations of finite modular groups

This is the written version of a lecture delivered at the Universiteit Leiden on April 7, 2000, at the celebration of the hundredth birthday of H.D. Kloosterman.



T.A. Springer during his lecture

At this celebration it is natural that those of us who have known Kloosterman personally — necessarily persons of a rather advanced age — consult their memories and try to evoke him in their minds. I am one of those and I do have some very clear recollections of my contacts with Kloosterman, but also quite a few blurred ones.

I am grateful to the organizers for the invitation to speak here about Kloosterman's work on representations of finite modular groups. I shall try to give a sketch of this work, in the context of the time of publication, but also in a present-day context. In the course of the talk I shall have occasion to insert some recollections and personal remarks. Although I never worked on the subject-matter of the work I am going to talk about, this work did influence my own mathematical interests. Today's celebration gives me an opportunity to acknowledge my indebtedness to Kloosterman.

The work to be discussed is contained in two long papers published in 1946 (see [5] and the résumé in [6]). Before going into the specifics of the papers I want to say a few words about Kloosterman's personal history.

Compared to other Dutch mathematicians of his generation, Kloosterman travelled widely. In the years between 1922 and 1930 he was abroad most of the time, in Kopenhagen, Oxford, Göttingen, Hamburg and Münster. He studied with prominent mathematicians of his time (H. Bohr, G.H. Hardy, E. Landau, E. Hecke) and he gathered a wide mathematical culture.

Kloosterman's thesis (1924) belongs in the sphere of Hardy. In the thesis Kloosterman applies the "Hardy-Littlewood method" — then a very new tool — to problems of the theory of quadratic forms. We shall hear more about this in other talks[1], and also about the "Kloosterman sums", which he used in a refinement of the Hardy-Littlewood method. Much of Kloosterman's later work is about modular forms. Special modular forms already appear in the thesis, namely classical theta functions.

Later, during his stay in Hamburg, Kloosterman came into contact with Hecke, whose work, I think, did impress Kloosterman. In Hamburg he got interested in general modular forms. This interest must have been stimulated by Hecke, as is shown by the paper [4] — about an application of Kloosterman's version of the Hardy-Littlewood method in the theory of modular forms — which had its origin in a suggestion of Hecke. Let me remind you in passing that Hecke's work is held in very high regard nowadays. For one thing, his work of the thirties on the connection between Dirichlet series satisfying a functional equation and modular forms has led to vast later developments, notably in the work of Langlands. (Kloosterman knew this work of Hecke well.)

In the papers which are the subject of my talk modular forms also appear. In these papers Kloosterman is, eventually, aiming at determining the irreducible representations of finite groups $\Gamma_N = SL_2(\mathbf{Z}/N\mathbf{Z})$ of 2×2-matrices with entries in the finite ring $\mathbf{Z}/N\mathbf{Z}$ and determinant 1, N being any integer > 1. The case that N is a prime power is the crucial one.

When N is a prime p, i.e. when the ring is a finite field with p elements, the irreducible characters had been determined by Schur in 1907 (extending slightly anterior work of Frobenius), and in the 1930's the case $N = p^2$ had also been studied.

Kloosterman was the first to attack the general case. The method he uses is analytic. The basic idea goes back to Hecke and is as follows. Modular forms satisfying suitable natural conditions form a finite dimensional vector space, on which some group Γ_N acts, providing a representation of that group. Hecke had shown that irreducible representations of Γ_p occur in this manner in certain spaces of modular forms.

I should perhaps point out that the problem of constructing explicitly the irreducible representations or characters of a concrete finite group is a non-trivial one (as anybody knows who has had a look at the representation theory of symmetric groups).

The modular forms Kloosterman is dealing with in his papers are theta functions, of a very general kind. Let V be a real vector space of even dimension $2k$, and let Q be a positive definite quadratic form on V. For $x, y \in V$ we have

$$Q(x+y) = Q(x) + Q(y) + (x, y),$$

where $(\ ,\ )$ is a non-degenerate symmetric bilinear form on V. Assume that L is a lattice in V (a free abelian group generated by some basis of V) such that Q takes integral values on L. Then the matrix S of our bilinear form, relative to any basis of L, is positive definite, with integral entries and even diagonal elements. Let

$$L^\vee = \{x \in V \mid (x, L) \subset \mathbf{Z}\}.$$

This is a lattice containing L and the index of L in $L^\vee$ equals det(S). Let $\mathcal{H} = \{z \in \mathbf{C} \mid \mathrm{Im}(z) > 0\}$ be the complex upper half plane. For $z \in \mathcal{H}$ and $u \in L^\vee$ define

$$\theta_Q(z, u) = \sum_{x \in L} e^{2\pi i Q(u+x)z}.$$

Let $\Gamma = SL_2(\mathbf{Z})$. If f is a holomorphic function on $\mathcal{H}$ and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ put

$$(\gamma^{-1} \cdot_k f)(z) = (cz+d)^{-k} f((az+b)(cz+d)^{-1}).$$

This defines a representation of Γ in the space of holomorphic functions on $\mathcal{H}$, for any natural number k. If now Q and k are as above, the functions $\theta_Q(z, u)$ span a finite dimensional subspace Σ of the space of holomorphic functions which is stable under the Γ-action. This is a consequence of the "transformation formulas of theta functions". In the first paper [5] these formulas are established, in a more general situation. I have followed a simplified version, given in [14], which seems to go back to Eichler. (Kloosterman also deals with the case that V has odd dimension, and his theta functions involve an extra parameter in V. He also admits positive definite symmetric matrices S with arbitrary diagonal entries.)

Let $N = N(Q)$ be the smallest positive integer such that NS^{-1} is an integral matrix with even diagonal entries. Then one shows that if $\gamma \equiv 1 \pmod{N}$ we have $\gamma \cdot_k f = f$ for all $f \in \Sigma$. As a consequence we obtain a representation of the group Γ_N in Σ. There is a great deal of freedom in the construction, as Q is arbitrary, so far.

Assume that p is an odd prime. In the second paper [5] Kloosterman considers the case that Q is a two-dimensional quadratic form, so $k = 1$. Following [14], we assume that the matrix S has the form

$$S = p^\lambda q \begin{pmatrix} 2 & 1 \\ 1 & \frac{1}{2}(1+r) \end{pmatrix},$$

where q and r are odd integers, with $r \equiv 3 \pmod 4$ and such that p, q, r are mutually coprime. We then have $N(Q) = p^\lambda qr$. So we get a representation of $\Gamma_{p^\lambda qr}$ in our vector space Σ. One shows that for any $r \in L^\vee$ the $\theta(z, r + p^{-\lambda} l)$ $(l \in L)$ generate a subspace Σ_r of Σ which is stable under Γ_{qr}. Hence we get a representation of $\Gamma_{p^\lambda} \simeq \Gamma_{qr}/\Gamma_{p^\lambda qr}$ in Σ_r. This is, essentially, the representation studied by Kloosterman (r being chosen suitably). These representations are not irreducible. To decompose them, a device used by Kloosterman is the introduction of an abelian automorphism group A of Σ_r centralizing the representation. Decomposing the space into isotypical subspaces for the characters of A, one obtains a decomposition into Γ_{p^λ}-stable subspaces, affording irreducible representations (in favorable cases). The group A is constructed using automorphisms modulo p^λ of the quadratic form Q. However, Kloosterman did not obtain all irreducible representations via this construction. That a slightly more general construction does produce all irreducibles was established later by Nobs and Wolfart in [14]. They also use the quadratic form associated with

1 The talks referred to were by D.R. Heath-Brown on *Arithmetic applications of Kloosterman sums* and by P. Sarnak on *Kloosterman, quadratic forms and modular forms.*

a matrix S of the form

$$S = p^{\lambda-1}q \begin{pmatrix} 2 & 1 \\ 1 & \frac{1}{2}(1+pr) \end{pmatrix},$$

q and r being as before, with $pr \equiv 3 \pmod 4$.

Kloosterman gives formulas for the characters of the irreducible representations which he does construct. It is a nice observation that these character values are sometimes Kloosterman sums.

Before continuing with the story of the representation theory of the groups Γ_{p^λ} I want to digress a bit.

From 1947 till 1951 I was Kloosterman's assistant. This was a job involving some teaching obligations (e.g. an exercise class for Kloosterman's course in elementary analysis), but I had a great deal of freedom. Kloosterman did not impose himself on his assistant. This does not mean that he left me to my own devices. I had many mathematical conversations with him, from which I learned a great deal. For example, he made me aware of the importance of algebraic geometry. Probably, the occasion was a conversation about Weil's proof of the Riemann hypothesis for function fields and the application to the proof of the good estimates for Kloosterman sums. (Understanding these difficult things was another matter...)

Kloosterman gave me reprints of his Annals papers, and I tried to study them. The first paper starts off with the transformation theory of theta functions, handled by Kloosterman with great technical skill. Being considerably less skillful, I found the papers quite hard. But I learned from them that the problem of finding explicitly the irreducible representations of a concrete finite group could be a very interesting one. I studied the literature about these matters, and found out that there was not too much. This led me to the question whether one could find the representations, or more modestly, the characters of other concrete groups, looking less formidable than the groups Γ_{p^λ}. Obvious candidates were linear groups over finite fields. After some experimentation I found the characters of $GL_3(k)$ and $GL_4(k)$, where k is a finite field. But it turned out that the same thing had been done in Toronto by a Ph.D. student of Richard Brauer, named Robert Steinberg (who later became a good mathematical friend).

A preliminary to the description of characters of a finite group is a study of its conjugacy classes and centralizers. Thus I was led to study conjugacy classes in linear groups, and this led to a thesis on the conjugacy classes in symplectic groups (in 1951) and to a subsequent interest in conjugacy classes in linear algebraic groups. This turned out to be a very fruitful subject indeed, with many ramifications (see [2]). As to the representation theory of finite linear groups, this has evolved into the representation theory of finite groups of Lie type, which is nowadays a vast subject (dominated by the work of George Lusztig, see [1]). What I wished to make clear here is that Kloosterman's Annals papers had a definite influence on my later interests.

In his talk [6] at the Cambridge Congress of 1950 Kloosterman states that all irreducible representations of Γ_{p^λ} (p odd) can be obtained via his constructions, using more general two-dimensional quadratic forms than those of his paper. He did not publish proofs of this statement, though. (As was mentioned above this was later established in [14].) His Ph.D. student J. van der Mark took up this general case in 1955, but did not completely solve the problem of determining all irreducible representations. Another of Kloosterman's Ph.D. students, A. Menalda, constructed representations of finite groups of the form $SL_2(\mathfrak{o}/\mathfrak{p}^\lambda)$, where $\mathfrak{o}$ is the ring of integers of a totally real number field, and $\mathfrak{p}$ is an odd prime ideal in $\mathfrak{o}$, using theta functions over such number fields. (Nowadays one would view such finite groups as groups of the same form, but with $\mathfrak{o}$ the ring of integers of a local field of characteristic zero, and $\mathfrak{p}$ its maximal ideal.)

Kloosterman's work was completed in the 1970s. This later work was along several lines. First, there was the work of [14], which was already mentioned. It gives a simplified version of Kloosterman's theta function approach.

The second line started in A. Weil's paper [16]. In that paper he constructs by analytic means a representation — the Weil representation — of a double cover of a group $SL_2(A)$, where A is either a locally compact field or an adèle ring. Weil suggests [16](footnote, p. 2) the possibility of viewing the problems studied by Kloosterman in the context of his work, for the case that A is finite. (Weil's paper was published in 1964, a few years before Kloosterman's death in 1968. Kloosterman was aware of Weil's work. In a second paper, Weil applies his results to the proof of a general version of a formula of Siegel. I think that Kloosterman knew about this, too.)

Weil's suggestion was taken up by several authors. Already in 1966 by S. Tanaka [15], who constructed Weil representations for the groups Γ_{p^λ} with p odd. Subsequently. Nobs and Wolfart in [12] did the same for all p (the case $p = 2$ is really more complicated), and described the irreducible representations of all Γ_{p^λ}.

This work provides a complete solution of the problem of describing the representation theory of the groups Γ_{p^λ}. Another — purely algebraic — solution had already been given some years before by P. Kutzko in his thesis (see [7]). Subsequently, he also dealt with the groups $SL_2(\mathfrak{o}/\mathfrak{p}^\lambda)$ where $\mathfrak{o}$ is the ring of integers of a local field (of any characteristic), and $\mathfrak{p}$ its maximal ideal. But his work was not published in detail.

Although the problem attacked by Kloosterman in his Annals papers was solved in the seventies, it had not lost its interest. At around the same time, probably under the impetus of the work of Jacquet and Langlands [3], the problem came to be viewed in another perspective, namely from the point of view of the representation theory of p-adic groups. This is a topic of considerable present-day interest. I want to give a rough sketch of that perspective.

Denote $\mathbf{Q}_p$ the field of p-adic numbers and by $\mathbf{Z}_p$ the ring of p-adic integers. More generally, let k be a finite extension of $\mathbf{Q}_p$. Denote by $\mathfrak{o}$ its ring of integers and by $\mathfrak{p}$ its maximal ideal. The group $K = SL_2(\mathfrak{o})$ is a compact topological group, and $\Gamma_{\mathfrak{p}^\lambda} = SL_2(\mathfrak{o}/\mathfrak{p}^\lambda)$ can be viewed as the quotient of K by an open subgroup. In fact K is isomorphic to the projective limit of the finite groups $\Gamma_{\mathfrak{p}^\lambda}$, relative to the obvious homomorhisms $\Gamma_{\mathfrak{p}^\mu} \to \Gamma_{\mathfrak{p}^\lambda}$ $(\mu \geq \lambda)$. It is easy to see that every finite dimensional continous representation of K factors through one of the $\Gamma_{\mathfrak{p}^\lambda}$. Consequently, the representation theory of $\Gamma_{\mathfrak{p}^\lambda}$ is contained in the representation theory of K. This obvious observation is in itself of no great interest. But things become more interesting if one also introduces the group

$G = SL_2(k)$ and its representation theory. This is a noncompact topological group, and K is a maximal compact subgroup of G. The representations of G which one considers — the *admissible* representations — are not finite dimensional any more. They are defined to be the representations of G in a complex vector space V such that:
(a) each $v \in V$ is fixed by a compact open subgroup U of G,
(b) for each such U the subspace of V fixed by all elements of U has finite dimension.
These conditions are local finiteness conditions (the notion of admissible representation was introduced by Jacquet and Langlands [3]). For an admissible representation one can introduce matrix coefficients. An admissible representation is said to be *supercuspidal* if it is irreducible — in the usual algebraic sense — and if its matrix elements have compact support in G. (The definitions make sense for a much larger category of groups, for example for the groups $G = GL_n(k)$. In general the definition of supercuspidality is slightly different: one has to require compactness modulo the center of the support of matrix coefficients.)

Given a finite dimensional representation of K one can define an induced representation of G which is admissible, in the same way as in the case of finite groups.

In the seventies several people came to the insight (it is a bit hard to disentangle priorities) that a supercuspidal representation of G is induced by an irreducible representation of K, which is unique (up to isomorphism), see e.g. [8]. Not all irreducible representations of K arise in this fashion, but it is known how to describe the remaining ones (although the description does not seem to be in the literature).

It seems that the approach to the representation theory of finite modular groups via the infinite dimensional representation theory of the p-adic modular group is at present the best one.

There is a further line of development in the story of the representation theory of $\Gamma_{\mathfrak{p}^\lambda}$. The algebraic group SL_2 is a semi-simple linear algebraic group of minimal possible dimension. What about other types of semi-simple (or reductive) groups? The first examples which present themselves are the groups GL_n. Here there is recent work by several people, for example by Bushnell and Kutzko (see [9]). They proved that a supercuspidal representation of $GL_n(k)$ is induced by a continuous irreducible representation of an open subgroup which is compact modulo the center, belonging to an explicit family of such subgroups including $GL_n(\mathfrak{o})$. It is conjectured that results of this kind will hold in general, i.e. for all reductive groups.

Thus, some irreducible representations of $GL_n(\mathfrak{o})$ (but not all of them) are connected with supercuspidal representations of $GL_n(k)$.

These recent developments contain information about the following problem, which generalizes Kloosterman's original problem: describe the representation theory of $GL_n(\mathfrak{o}/\mathfrak{p}^\lambda)$. A full answer does not seem to be known, even for the first interesting case $n = 3$ (for $n = 2$ one is very close to the original problem). More generally, one has the analogous problem for the finite groups $G(\mathfrak{o}/\mathfrak{p}^\lambda)$, where G is a smooth, reductive, affine group scheme over $\mathbf{Z}$. In this generality, very little seems to be known if $\lambda > 1$. (If $\lambda = 1$, the finite group in question is a finite group of Lie type.)

The classification of supercuspidal representations of $GL_n(k)$ has much to do with the "local Langlands conjecture" (recently proved by Harris-Taylor and Henniart), i.e. with non-abelian local class field theory (see [13]).

I stop with these recent developments, with which I have no technical familiarity. But I wanted to mention them briefly, in order to show how the problem of Kloosterman's Annals papers is related to questions of actual interest.

Acknowledgments
The author is indebted to Ph. Kutzko for information and comments.

References

1 R.W. Carter, Finite groups of Lie type: Conjugacy classes and representations, Wiley, 1985.

2 J.E. Humphreys, Conjugacy Classes in Semisimple Algebraic Groups, Math. Surveys and Monographs vol. 43, Amer. Math. Soc., 1995.

3 H. Jacquet and R.P. Langlands, Automorphic forms on GL_2, Lecture Notes in Math., vol. 114, Springer-Verlag, 1970.

4 H.D. Kloosterman, Asymptotische Formeln für die Fourierkoeffizienten ganzer Modulformen, Abh. Math. Sem. d. Hamb. Univ., 5 (1927), 337–352.

5 H.D. Kloosterman, The behaviour of general theta functions under the modular group and the characters of binary modular congruence groups, Ann. of Math. 47 (1946), 317–447.

6 H.D. Kloosterman, The characters of binary modulary congruence groups, Proc. Int. Congr. of Math. 1950, vol. I, p. 275–280.

7 P. Kutzko, The characters of the binary modular congruence group, Bull. Amer. Math. Soc. 79 (1973), 702–704.

8 P. Kutzko, On the supercuspidal representations of Gl_2, I, II, Amer. J. Math. 100 (1978), 43–60 and 705–716.

9 P. Kutzko, Smooth representations of reductive p-adic groups: an introduction to the theory of types. In: "Geometry and representation theory of real and p-adic groups" p. 175–196, Progr. in Math. vol. 158, Birkhäuser, 1998.

10 J. van der Mark, On the characters of binary modulary congruence groups, thesis, University of Leiden (1955).

11 A. Menalda, Theta-series and representations of modulary congruence groups for totally real algebraic number fields, thesis, University of Leiden (1964).

12 A. Nobs, Die irreduziblen Darstellungen der Gruppen $SL_2(\mathbf{Z}_p)$, insbesondere $SL_2(\mathbf{Z}_2)$ I, II (with J. Wolfart), Comm. Math. Helv, 51 (1976), 465–489 and 491–526.

13 J. Rogawski, The nonabelian reciprocity law for local fields, Notices Amer. Math. Soc. 47 (2000), 35–41.

14 A. Nobs and J. Wolfart, Darstellungen von $SL(2, \mathbf{Z}/p^\lambda \mathbf{Z})$ und Thetafunktionen, Math. Z. 138 (1974), 239–254.

15 S. Tanaka, On irreducible unitary representations of some special linear groups of the second order II, Osaka J. Math. 3 (1966), 229–242.

16 A. Weil, Sur certains groupes d'opérateurs unitaires, in: Oeuvres Scienfifiques vol. III, p.1–69, Springer-Verlag, 1980.