# H. D. Kloosterman and His Work

*Tonny A. Springer*

April 9, 1900, marks the birth date of H. D. Kloosterman, one of the leading Dutch mathematicians of his generation. At the University of Leiden, where he spent most of his academic life, the centenary of his birth was the occasion for a small celebration, with a discussion by several experts of Kloosterman's work on number theory, modular forms, and representation theory; and of its place in contemporary mathematics. Although he was not a prolific writer, his work had a significant impact and is still of considerable interest.

The centenary is also a good occasion for a brief discussion of Kloosterman's work for a wider public. This is the raison d'être of the present article.

## Biography

Hendrik Douwe Kloosterman was born in Rottevalle, a small farming village in the northern part of the Netherlands. He moved to the city of The Hague, where he finished high school in 1918. In the same year he started his study in mathematics at the nearby University of Leiden. He was a very fast student, passing the examination for the final degree (comparable to a master's degree) in 1922.

The mathematics education in Leiden was solid, but it did not bring a brilliant student like Kloosterman in close contact with current research. The situation was different in physics. In the first decades of the twentieth century, high-level research in physics was in the air at the University of Leiden. There were the Nobel Prize winners H. Kamerlingh Onnes and H. A. Lorentz, and Einstein had a permanent visiting professorship. Lorentz's successor (after an early retirement from formal duties) was P. Ehrenfest, who was in very close contact with the developments in theoretical physics. He took great pains with the scientific education of his students, particularly his own Ph.D. students. But Kloosterman seems also to have come under Ehrenfest's wings, as a result of which Kloosterman was able to continue his studies abroad—first in Copenhagen with H. Bohr and then in Oxford with G. H. Hardy.

Kloosterman obtained his Ph.D. at the University of Leiden in 1924. Supported by a Rockefeller scholarship, he spent the years 1926–1928 at the Universities of Göttingen and Hamburg, after which, from 1928 to 1930, he had a position at the University of Münster.

In 1930 he returned to Leiden to take the position of "lector" (comparable to an associate professorship). This was mainly a teaching position. Kloosterman turned out to be an exceptional teacher. He was able to expose with great clarity and great economy the essentials of a piece of mathematics, be it elementary or advanced.

His teaching obligations were on the elementary level, but he also gave more advanced courses on topics outside the regular curriculum or on topics of then current interest. The notes for most of these courses, beautifully handwritten, have been preserved. In his later years Kloosterman intended to publish lectures on modular forms, but he did not complete the manuscript.

*Tonny A. Springer is professor emeritus of mathematics at the University of Utrecht. His e-mail address is* springer@math.uu.nl. *The author is grateful to D. R. Heath-Brown, P. Sarnak, and P. Kutzko for information and comments.*

In 1947 he was promoted to a full professorship, which he held until his death in 1968. Kloosterman visited the U.S. several times. For example, in the academic year 1955–56 he was a visiting professor at the University of Michigan in Ann Arbor.

## Quadratic Forms and the Hardy-Littlewood Method

Kloosterman's thesis [K1] contains an application of the Hardy-Littlewood method of analytic number theory to quadratic forms. Kloosterman's formal thesis adviser was the Leiden mathematician J. C. Kluyver, but the work of the thesis is in Hardy's sphere.

The thesis is in Dutch and is not easily accessible. It deals with the problem of finding the number of solutions in integers $x_i$ of the equation

$$(1) \qquad f(x_1, \ldots, x_s) = a_1 x_1^2 + \cdots + a_s x_s^2 = n,$$

where the $a_i$ and $n$ are integers $> 0$. Around 1920 Hardy and Littlewood devised an analytic method to attack similar problems (e.g., "Waring's problem").

One introduces the power series

$$F(z) = \sum_{x_i \in \mathbb{Z}} z^{f(x_1, \ldots, x_s)}.$$

It converges for $|z| < 1$, and the number of solutions of (1) equals the coefficient of $z^n$ in $F(z)$, which by a formula of Cauchy equals

$$(2) \qquad I(n) = (2\pi i)^{-1} \int_{|z|=r} F(z) z^{-n-1} dz,$$

where $r < 1$ is close to 1. To deal with (2), the circle of integration is subdivided into small arcs.

Let $N$ be a positive integer (to be chosen later), and consider the set $\mathcal{F}$ of all rational numbers in $[0, 1]$ with denominator $\leq N$. These rationals determine a decomposition of $[0, 1]$ into smaller segments. The ordered set of endpoints is the *Farey series* of order $N$. Using properties of Farey series, one deduces from this decomposition another decomposition such that each $f \in \mathcal{F}$ lies in the interior of exactly one segment of the new decomposition. The new decomposition defines a decomposition into arcs $\mathcal{D}$ of the integration circle.

The integral (2) is now written as the sum of integrals over the arcs of $\mathcal{D}$, each of which is analyzed by using properties of Jacobi theta functions (in the problems studied by Hardy and Littlewood the analysis is technically more difficult). The result of the analysis is a formula

$$I(n) = S(n) + J(n),$$

where $S(n)$ is computable, at least in principle. The outcome of the analysis is an infinite product of "local factors"



**H. D. Kloosterman**

$$S(n) = \prod S_p(n),$$

$p$ running through the primes, including $p = \infty$. The other term, $J(n)$, is a "junk" term. The analysis of $J(n)$ leads to an estimate for its order of magnitude as a function of $n$. If that order of magnitude is strictly smaller than the order of $S(n)$ (assuming $S(n)$ to be tractable), then one obtains an asymptotic formula for the number of solutions of (1) for large $n$. From this one infers that (1) then has a solution for all sufficiently large $n$.

Kloosterman shows that for $s \geq 5$ this program can be carried out, leading to the result that for $s \geq 5$, (1) is solvable if $n$ is sufficiently large and if the $a_i$ satisfy suitable congruence conditions.

But for $s = 4$ the original Hardy-Littlewood method does not lead to a good estimate of $J(n)$. Kloosterman notes the curious fact that the powerful analytic method does not seem to be able to prove Lagrange's old theorem that any positive integer is a sum of four squares!

## Kloosterman Sums

Kloosterman took up the challenge posed by the case $s = 4$ in [K2]. He refined the Hardy-Littlewood method so as to give a better estimate for $J(n)$, enabling him to obtain asymptotic formulas in that case.

In his analysis of the integrals over the arcs of $\mathcal{D}$ there appear what now are called "Kloosterman sums", to be discussed below.

Kloosterman's modification of the Hardy-Littlewood method has turned out to be quite significant. In 1928, during his stay in Hamburg, he already used his approach to obtain nontrivial estimates for Fourier coefficients of modular forms (that is, estimates towards the so-called Ramanujan conjecture). This work had its origin in a suggestion of E. Hecke, a mathematician who had an inspiring influence on Kloosterman's work.

Hecke's name is held in very high regard nowadays. For one thing, Hecke's work of the 1930s on the connection between modular forms and Dirichlet series satisfying a functional equation has led to vast developments, notably in the work of R. P. Langlands. Kloosterman knew this work of Hecke well.

Kloosterman sums are trigonometric sums of the form

$$(3) \qquad S(a,b;c) = \sum e^{2\pi i c^{-1}(ax+by)},$$

where $a, b, c$ are integers with $c > 0$, the summation being over the integers $x$ and $y$ with $1 \le x, y \le c$ and $xy \equiv 1 \pmod{c}$. It is easy to see that $S(a,b;c)$ is a real number. One can show that the determination of $S(a,b;c)$ can be reduced to the case that $c$ is a prime $p$. (In that case a Kloosterman sum can be viewed as a Bessel function for the field $\mathbb{Z}/p\mathbb{Z}$.) In the case $c = p$ Kloosterman proves in [K2] that there is a constant $C$ such that

$$(4) \qquad |S(a,b;c)| \le C p^{\frac{3}{4}}.$$

The fact that in the right-hand side an exponent $< 1$ appears enables Kloosterman to obtain a sufficiently good estimate for $J(n)$.

Inequality (4) is a consequence of the following inequality, which can be proved in an elementary way:

$$\sum_{a=1}^{p} S(a,1;p)^4 \le 3p^3.$$

Inequality (4) was improved later. The optimal estimate is

$$(5) \qquad |S(a,b;c)| \le 2p^{\frac{1}{2}}.$$

This follows from the Riemann hypothesis for function fields, proved by A. Weil in a 1941 paper.

## Applications of Kloosterman Sums

Since their first appearance in 1926, Kloosterman sums have shown up in number theory in many places. (At some time there was even talk of "Kloostermania".) Here is an example from elementary number theory.

Let $p$ be a prime, and let $a$ be an integer prime to $p$. Solve the congruence

$$mn \equiv a \pmod{p}$$

in natural numbers $m, n$, with $m \le n$, and let $M(a)$ be the smallest possible value of $n$. How large can $M(a)$ be? The analysis of this question leads to Kloosterman sums, and via an application of (5) one proves that

$$M(a) < 2(\log p)p^{\frac{3}{4}}.$$

It is not known to what extent this estimate can be improved.

Consider the Kloosterman sum (3), with $c = p$, a prime number. A general method to study such trigonometric sums is provided by "$l$-adic cohomology", in characteristic $p$. The sums can be expressed in terms of the cohomology of certain locally constant $l$-adic sheaves on an algebraic variety $X$ and the action of a "Frobenius map" on the cohomology. The case of Kloosterman sums was taken up by Deligne (1978). In that case $X$ is the line minus the point $\{0\}$. A deep study of the "Kloosterman sheaves" on $X$, in particular of their monodromy, was made by N. Katz in [Ka]. As a consequence of his study he obtained a result about the distribution of the values of Kloosterman sums.

It follows from (5) that for $a$ prime to $p$ there is a unique $\theta_{p,a} \in [0, \pi]$ such that

$$S(a,1;p) = 2\cos\theta_{p,a}.$$

The distribution function $f_p$ of the numbers $\theta_{p,a}$ for fixed $p$ is the function on $[0, \pi]$ whose value at $x$ is the fraction of the number of $\theta_{p,a}$ lying in $[0, x]$. It is shown in [Ka, p. 241] that $f_p$ tends to the distribution function defined by the measure $2\pi^{-1} \sin^2 \theta \, d\theta$, as $p$ tends to $\infty$.

However, about the distribution of the $\theta_{p,a}$ for fixed $a$ (say $a = 1$) for $p \le x$ as $x \to \infty$, nothing seems to be known. It is conjectured that the limit distribution is again the one defined by $2\pi^{-1} \sin^2 \theta \, d\theta$.

Another continuation of the study of Kloosterman sums was along analytic lines, involving the study of nonholomorphic modular forms, whose definition will not concern us. This started with ideas of Linnik (1962) about "averages" of Kloosterman sums. These led Selberg (1965) to study Dirichlet series of the form

$$Z(a,b;s) = \sum_{c=1}^{\infty} S(a,b;c)c^{-\frac{1}{2}-s}.$$

This formula defines a holomorphic function in the half plane $\text{Re}(s) > 1$. Selberg proved that it has a meromorphic continuation to the complex plane.

An important result about the average

$$\Sigma(a,b;x) = \sum_{c=1}^{x} S(a,b;c)c^{-\frac{1}{2}}$$

is due to Kuznetsov [Kuz]. He derived a tracelike formula giving an explication of Selberg's results and allowing him to prove that there exist positive constants $C > 0$ and $\theta < 1$ such that

(6) $$|\Sigma(a, b : x)| \leq Cx^\theta.$$

In fact, $\theta$ may be taken to be any number $> \frac{2}{3}$. It is conjectured that $\frac{2}{3}$ can be replaced by $\frac{1}{2}$.

The connection between Kloosterman sums and nonholomorphic modular forms has been well developed (by Iwaniec and others) and has been used in many proofs of striking results in number theory. The following is a typical example.

Let $p$ be a prime. An integer $q$ is a *primitive root* for $p$ if $q + p\mathbb{Z}$ generates the multiplicative group of $\mathbb{Z}/p\mathbb{Z}$. For this to be true with $p > 3$, it is necessary that $q \neq 0, \pm1$ and that $q$ is not a square. There is a conjecture of Artin that any such $q$ is a primitive root for infinitely many primes. This conjecture remains unsolved but has been shown to be true if $q$ is a prime, with at most two exceptions.

## Representations of $SL_2(\mathbb{Z}/N\mathbb{Z})$

During the German occupation of the Netherlands in World War II, the University of Leiden was closed from 1941 to 1945. In that period Kloosterman turned to research, the results of which were published in the two long papers [K3]. In these papers Kloosterman is, eventually, aiming to determine the irreducible representations of concrete finite groups.

Let $\Gamma = SL_2(\mathbb{Z})$ be the group of $2 \times 2$-matrices

(7) $$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$. If $N$ is an integer $> 1$, denote by $SL_2(\mathbb{Z}, N)$ the normal subgroup of $\Gamma$ consisting of the $\gamma$ as in (7) with $a, d \equiv 1 \pmod{N}$ and $c, d \equiv 0 \pmod{N}$. The finite groups of [K3] are the quotients $\Gamma_N = SL_2(\mathbb{Z})/SL_2(\mathbb{Z}, N)$. Such a quotient can also be viewed as the group $SL_2(\mathbb{Z}/N\mathbb{Z})$ of $2 \times 2$-matrices with entries in the finite ring $\mathbb{Z}/N\mathbb{Z}$ and determinant 1. The group $\Gamma_N$ is the direct product of the $\Gamma_{p^\lambda}$, where $p^\lambda$ runs through the prime powers dividing $N$. We assume from now on that $N = p^\lambda$.

A *representation $R$* of a finite group $G$ is a homomorphism of $G$ into the group $GL(V)$ of invertible linear maps of a finite-dimensional complex vector space $V$. $R$ is said to be *irreducible* if the only subspaces of $V$ stable under all $R(g)$ for $g \in G$ are $\{0\}$ and $V$. It is always true that $V$ is a direct sum of $G$-stable, irreducible subspaces. The *character* of $R$ is the function $\chi$ on $G$ with $\chi(g) = \text{trace}(R(g))$. It determines $R$ uniquely, up to isomorphism. The theory of representations goes back to Frobenius around 1900.

The problem of describing the irreducible representations or, more modestly, the irreducible characters of a concrete group $G$ is a nontrivial one, as is already evident from Frobenius's study of the case that $G$ is a symmetric group.

As to the groups $\Gamma_{p^\lambda}$, the case $\lambda = 1$ was settled by Schur in 1907, extending slightly work of Frobenius; and in the 1930s the case $\lambda = 2$ had also been settled.

Kloosterman was the first to attack the case of an arbitrary prime power. The method he used is analytic. The basic idea is to use "holomorphic modular forms" and goes back to Hecke.

A *holomorphic modular form* of weight $k$ (an integer $\geq 0$) is a holomorphic function $f$ in the complex upper half-plane

$$\mathcal{H} = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$$

such that

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

for all $z \in \mathcal{H}$ and all

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

in some $SL_2(\mathbb{Z}, N)$. One also imposes a growth condition on $f$ as $z$ tends to $\infty$, but we shall not need a detailed formulation. For a fixed $N$ such modular forms span a finite-dimensional vector space of functions, on which $SL_2(\mathbb{Z})$ acts linearly by

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

where $\gamma$ (as before) lies in $SL_2(\mathbb{Z})$ and $f$ is in the vector space. Hecke had used these facts to construct explicit realizations (not easy to obtain using algebra) of certain irreducible representations of groups $\Gamma_p$ in spaces of modular forms.

Kloosterman in [K3] uses particular modular forms, namely, theta functions associated to quadratic forms. These (or rather specialized versions) can be described as follows.

Let $V$ be a real vector space of even dimension $2k$, and let $b(\cdot, \cdot)$ be a nondegenerate symmetric bilinear form on $V$, which we assume to be positive definite. Assume that $L$ is a lattice in $V$ (a free abelian group generated by some basis of $V$) such that $b(x, y)$ is an integer for $x, y \in L$ and is even if $x = y$. The matrix $S = (b(x_i, x_j))$ of our bilinear form, relative to any basis $(x_i)$ of $L$, is symmetric and positive definite, with integral entries and even diagonal elements. Let

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

This is a lattice containing $L$, and the index of $L$ in $L^\vee$ equals $\det(S)$.

For $z \in \mathcal{H}$ and $u \in L^\vee$, let

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

These series define holomorphic functions in $\mathcal{H}$ ($u$ running through $L^\vee$), which span a

finite-dimensional complex vector space $M$ of holomorphic functions.

They are, in fact, modular forms of weight $k$. The integer $N$ can be taken to be the smallest positive integer such that $NS^{-1}$ is an integral positive definite matrix with even diagonal elements. Moreover, $M$ is stable under the action of $SL_2(\mathbb{Z})$ on modular forms of weight $k$. So we obtain a representation of $\Gamma_N$ in $M$.

These facts follow from the "transformation formulas of theta functions", dealt with at length (and in greater generality) in the first part of [K3].

There is a great deal of freedom in the construction, as $b$ is arbitrary so far. Assume that $p$ is an odd prime. Kloosterman studies in detail the case that $k = 1$ and that $S$ is of the form $p^\lambda q S'$, where $q$ and $\det(S')$ are prime to $p$. For a suitable choice of $S'$, he constructs a subspace $V$ of $M$ and a representation $R$ of $\Gamma_{p^\lambda}$ in $V$. To decompose $V$, he introduces an abelian automorphism group $A$ of $V$ that centralizes $R$; $A$ comes from automorphisms modulo $p^\lambda$ of the bilinear form defined by $S'$.

For each multiplicative character $\alpha$ of $A$ let $V_\alpha$ be the maximal subspace of $V$ on which any element $a \in A$ acts as scalar multiplication by $\alpha(a)$. Then $V$ is the direct sum of the nonzero $V_\alpha$. Each such space is $\Gamma_{p^\lambda}$-stable, and it turns out to be irreducible. Thus one obtains the decomposition of $V$ into $\Gamma_{p^\lambda}$-stable irreducible subspaces.

Kloosterman computes the characters of the irreducible representations so obtained. Somewhat surprisingly, Kloosterman sums turn up as character values.

It was shown later by Nobs and Wolfart [NW] that Kloosterman's method indeed produces all irreducible representations, provided one admits also the case that $p$ divides $\det(S')$ (confirming a prediction by Kloosterman).

## Further Developments

The story of the representation theory of $\Gamma_{p^\lambda}$ does not end here. Its ramifications are still of considerable interest. The story developed along several lines. First, there was the completion of Kloosterman's analytic approach, which was just mentioned.

A second line had its origin in the paper [W] of A. Weil in 1964, which can be viewed as a study of the transformation formula of theta functions in a more general context. Weil constructs a representation—now often called the Weil representation—of a double cover of groups $SL_2(A)$, where $A$ is either a locally compact field or an adele ring. Weil suggests in a footnote the possibility of viewing the problems studied by Kloosterman in the context of his work, for the case that $A$ is finite (see [W, p. 144]).

Weil's suggestion was taken up by several authors—first by Tanaka (1966) and subsequently, in more detail, by Nobs and Wolfart (1976) in [N]. Nobs and Wolfart also dealt with the more complicated case $p = 2$. This work provides a complete solution to the problem of describing the representation theory of the groups $\Gamma_{p^\lambda}$.

Another solution of the problem—purely algebraic—had been given for $p \neq 2$ a few years earlier in the thesis of P. Kutzko (1973). The problem has a natural extension to the case of finite groups $\Gamma_{p^\lambda} = SL_2(\mathfrak{o}/\mathfrak{p}^\lambda)$, where $\mathfrak{o}$ is the ring of integers of a local field $k$ and $\mathfrak{p}$ is the maximal ideal of $\mathfrak{o}$. Subsequently, Kutzko also described the irreducible representations of such groups. But his work has not been published in detail.

In the beginning of the 1970s, probably under the impetus of the work of Jacquet and Langlands, the problem came to be viewed in another perspective, namely, from the point of view of the representation theory of $p$-adic groups. Here is a rough sketch of that perspective.

Let $k$, $\mathfrak{o}$, and $\mathfrak{p}$ be as before. The group $K = SL_2(\mathfrak{o})$ is a compact topological group, and $\Gamma_{p^\lambda}$ can be viewed as a quotient of $K$ by an open subgroup. It is easy to see that every finite-dimensional continuous representation of $K$ factors through $\Gamma_{p^\lambda}$ for some $\lambda$. Consequently, the representation theory of $\Gamma_{p^\lambda}$ is contained in the representation theory of $K$. This in itself is an observation of no great interest. But things become more interesting if one also introduces the group $G = SL_2(k)$ and its representation theory.

This $G$ is a noncompact topological group, and $K$ is a maximal compact subgroup of $G$. The representations of $G$ that one considers—the *admissible representations*—are not finite dimensional any more. They are defined to be 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 pointwise by all elements of $U$ is finite dimensional.

For admissible representations one can define matrix coefficients, as in the case of a finite-dimensional representation. An admissible representation is *supercuspidal* if it is irreducible in the usual sense and if its matrix coefficients have compact support in $G$.

If $G$ is a finite group and $K$ is a subgroup, one has an induction procedure (going back to Frobenius), associating to a representation of $K$ a representation of $G$. The procedure extends to the situation of our $G$ and $K$. Induction of a finite-dimensional continuous representation of $K$ leads to an admissible representation of $G$.

(These definitions and notions make sense for a wider class of groups, for example, for the groups $GL_n(k)$. In general the definition of supercuspidality is slightly different: one has to require compactness modulo the center of the support of matrix coefficients.)

In the 1970s several people came to the insight (it is a bit hard to disentangle priorities) that a supercuspidal representation of $G = SL_2(k)$ is induced by an irreducible continuous representation of $K$ that is unique up to isomorphism. Not all irreducible representations of $K$ arise in this fashion, but it is known how to describe the remaining ones. Thus we have an approach to the representation theory of the groups $\Gamma_{\mathfrak{p}^\lambda}$ via the infinite-dimensional representation theory of $SL_2(k)$. This is probably the most satisfactory approach.

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 semisimple linear algebraic group of minimal possible dimension. What about other types of semisimple (or reductive) groups? The first examples that present themselves are the groups $GL_n$. Here there is recent work by several people, for example, by Bushnell and Kutzko (see [Ku]). They proved that a supercuspidal representation of $GL_n(k)$ is induced by a continuous irreducible representation of an open subgroup that is compact modulo the center, belonging to an explicit finite set of such subgroups that includes $GL_n(\mathfrak{o})$. 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 $GL_n(k)$. At this point it should be observed that the classification of supercuspidal representations of $GL_n(k)$ has much to do with the recently proved (see [R]) "local Langlands conjecture" for $GL_n$, i.e., with nonabelian local class field theory. One gets into deep waters!

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 $\mathbb{Z}$. 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". Thanks to the efforts of several people, notably G. Lusztig, the representation theory and character theory of the latter groups are well understood (see [L]). One can wonder whether the powerful methods of algebraic topology used to deal with the finite groups of Lie type can also be used in the case of the groups $G(\mathfrak{o}/\mathfrak{p}^\lambda)\ldots$.

## References

[Ka] N. M. KATZ, *Gauss Sums, Kloosterman Sums, and Monodromy Groups*, Ann. of Math. Studies, vol. 116, Princeton Univ. Press, 1988.

[K1] H. D. KLOOSTERMAN, *Over het splitsen van geheele positieve getallen in een som van kwadraten*, thesis, Univ. of Leiden, 1924.

[K2] ____, On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$, *Acta Math.* **49** (1926), 407–464.

[K3] ____, The behaviour of general theta functions under the modular group and the characters of binary modular congruence groups I, II, *Ann. of Math.* **47** (1946), 317–447.

[Ku] P. KUTZKO, Smooth representations of reductive $p$-adic groups: An introduction to the theory of types, *Geometry and Representation Theory of Real and $p$-adic Groups*, Progress in Math., vol. 158, Birkhäuser, 1998, pp. 175–196.

[Kuz] N. V. KUZNETSOV, The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums (Russian), *Math. Sbornik* **111(153)** (1980), 334–383, 479.

[L] G. LUSZTIG, *Characters of Reductive Groups over a Finite Field*, Ann. of Math. Studies, vol. 107, Princeton Univ. Press, 1984.

[N] A. NOBS, Die irreduziblen Darstellungen der Gruppen $SL_2(\mathbf{Z}_p)$, insbesondere $SL_2(\mathbf{Z}_2)$ I, II (with J. Wolfart), *Comment. Math. Helv.* **51** (1976), 465–489, 491–526.

[NW] A. NOBS and J. WOLFART, Darstellungen von $SL(2, \mathbf{Z}/p^\lambda\mathbf{Z})$ und Thetafunktionen I, *Math. Z.* **138** (1974), 239–254.

[R] J. ROGAWSKI, The nonabelian reciprocity law for local fields, *Notices Amer. Math. Soc.* **47** (2000), 35–41.

[W] A. WEIL, Sur certains groupes d'opérateurs unitaires, *Acta Math.* **111** (1964), 143–211; *Oeuvres Scientifiques*, vol. III, Springer-Verlag, 1980, pp. 1–69.