HECKE GROUPS AND CONTINUED FRACTIONS

DAVID ROSEN AND THOMAS A. SCHMIDT

The Hecke groups

$$G_q = \left\langle \begin{pmatrix} 1 & \lambda_q \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle; \lambda_q = 2 \cos \frac{\pi}{q}, q \geqslant 3 \text{ in } \mathbf{Z},$$

are Fuchsian groups of the first kind. In an interesting analogy to the use of ordinary continued fractions to study the geodesics of the modular surface, the λ–continued fractions (λF) introduced by the first author can be used to study those on the surfaces determined by the G_q. In this paper we focus on periodic continued fractions, corresponding to closed geodesics, and prove that the period of the λF for periodic $\sqrt{D}/C$ has nearly the form of the classical case. From this, we give: (1) a necessary and sufficient condition for $\sqrt{D}/C$ to be periodic; (2) examples of elements of $\mathbb{Q}(\lambda_q)$ which also have such periodic expansions; (3) a discussion of solutions to Pell's equation in quadratic extensions of the $\mathbb{Q}(\lambda_q)$; and (4) Legendre's constant of diophantine approximation for the G_q, that is, γ_q such that $\left| \alpha - \frac{P}{Q} \right| < \gamma_q/Q^2$ implies that P/Q of "reduced finite λF form" is a convergent of real $\alpha \notin G_q(\infty)$.

1. INTRODUCTION

A new class of continued fractions was introduced by the first-named author in [15], as a technique for studying the Hecke groups, G_q. Recall that

$$(1.1) \qquad G_q = \left\langle \begin{pmatrix} 1 & \lambda_q \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle; \lambda_q = 2 \cos \frac{\pi}{q}, q \geqslant 3 \text{ in } \mathbb{Z}$$

are Fuchsian groups of the first kind. The G_q act upon the Poincaré upper half-plane, $\mathfrak{H}$, via fractional linear transformations. When $q = 3$, G_3 is the full modular group and in what follows, we restrict to $q > 3$, referring to $q = 3$ as the classical case.

The link between continued fractions and the geometry of surfaces has arisen independently several times in this century. The connection between regular continued

fractions and the modular surface $S = G_3 \backslash \mathfrak{H}$ was first noted by L. Ford [7]. It was
E. Artin who exploited the connection to produce the first striking result: almost all
geodesics on S are quasi-ergodic; that is, pass near every point in every direction [3].
This led to a very fruitful line of inquiry as much stronger ergodicity pertains in much
more general settings [10]. Contemporaneous with Artin, J. Nielsen and M. Morse de-
veloped a symbolic dynamics of geodesics that, in principal, could be used to study the
geodesics. For S this symbolic dynamics amounts to the regular continued fractions
[19]. Ergodic theorists have found this connection to be an effective way of studying
the iteration of the continued fraction transformation as viewed as a map of the unit
interval to itself [1, 2]. As well, physicists have utilised this connection to construct
physical systems with chaotic trajectory behaviour [8]. However, S is the only non-
compact surface for which all of the above is explicitly understood. We expect that
the G_q in conjuction with the aforementioned class of continued fractions will provide
a whole new family of such examples.

The geometrical aspects of the continued fractions for the G_q were first utilised
by Lehner [12], to investigate Diophantine approximation by the Hecke groups. He
found the Hurwitz numbers of the groups with even q and gave a good bound on those
corresponding to odd q. Haas and Series [9] followed with further investigations of the
Markov spectra, using hyperbolic geometric techniques. Sheingorn [20] has emphasised
the simplicity of these continued fractions for hyperbolic geometric computations.

It is a general phenomenon, also remarked upon by Artin, that closed geodesics
of a surface correspond to periodic continued fractions. Furthermore, the set of closed
geodesics is basic to the dynamics of the surface, as evidenced in the Selberg trace
formula. Periodic continued fractions being quadratic in nature, we lay the foundation
for the study of the set of closed geodesics of the Hecke triangle surfaces by identifying
the set of the periodic $\sqrt{D}/C$.

Such periodic values are related to Pell equations. In the case of G_3, a single
solution to $P^2 - DQ^2 = 1$ is necessary and sufficient to ensure the periodicity of $\sqrt{D}$.
This can be shown as in [21], where one forms the matrix $\begin{pmatrix} P & DQ \\ Q & P \end{pmatrix}$ and proceeds
to conclude the periodicity. However, for $q > 3$, the existence of a solution, although
necessary, is not sufficient for periodicity. By [17], the elements of G_q are comprised of
numerators and denominators of consecutive convergents. For the traditional period-
icity argument to be successful, one needs a solution to the Pell equation with DQ/P
and P/Q being consecutive convergents. Significant complications are caused by the
units and class group of $\mathbb{Q}(\lambda_q)$. Indeed, this is the case throughout the study of these
continued fractions. We sketch the effects in this instance. If some P/Q is a cusp,
that is, an image of infinity under the action of the group, then the continued fraction
algorithm uniquely determines a convergent P_n/Q_n as defined in our equations (1.3) of

P/Q such that $P/Q = P_n/Q_n$. *A priori*, with relatively prime P and Q and unique factorisation in the ring of integers, one knows only that $P_n = uP$ and $Q_n = uQ$ for some unit u. (There are many of these: $\mathbb{Q}(\lambda_q)$ is a totally real field of degree $\phi(2q)/2$ over $\mathbb{Q}$, where ϕ is Euler's ϕ-function; it thus has a unit group of $\mathbb{Z}$ rank $\phi(2q)/2 - 1$.) Of course, one cannot say even this much if there is not unique factorisation. In all cases, a solution to the Pell equation may be transformed into convergents which no longer satisfy the equation.

A periodic $\sqrt{D}$ does give a solution to a Pell equation. Analogously to the important use of the classical continued fraction to find the fundamental unit of a real quadratic extension of $\mathbb{Q}$, one finds a non-trivial unit in $\mathbb{Q}\left(\lambda_q, \sqrt{D}\right)$. As mentioned in Section 4, there can be no more than a single infinite cyclic group of units which can arise from the convergents.

The main result of our paper is on the form of the expansion of a periodic $\sqrt{D}/C$. Knowledge of the shape of such a period leads us to: (1) a necessary and sufficient condition for $\sqrt{D}/C$ to be periodic; (2) the fact that elements of $\mathbb{Q}(\lambda_q)$ can have such periodic expansions; (3) a discussion of solutions to Pell's equation in quadratic extensions of the $\mathbb{Q}(\lambda_q)$; and (4) the Legendre constant of diophantine approximation for the G_q.

Numerical examples lead us to conjecture that many quadratics over $\mathbb{Q}(\lambda_q)$ do not have periodic expansions. We furthermore conjecture that this is also true for certain elements of $\mathbb{Q}(\lambda_q)$ which are not cusps. (We add: This is true and a discussion thereof, based upon our Theorem 3.2, will appear in a forthcoming joint paper of the second-named author with M. Sheingorn. Other recent results include the use by the second-named author [18] of the fractions to obtain results on the rational period functions, which are related to the Eichler cohomology, for the G_q. In this last reference there is also a treatment of the metrical properties of the interval map induced by the continued fractions for $q = 4$.) There was an intense study of the cusps of the G_q initiated by the first-named author [16] and furthered by the German school of Leutbecher *et al.*: [14, 5]. However there is a paucity of cases in which one knows the exact cusp set of a given G_q. It may very well be that a study based upon the periodic values, which are the hyperbolic fixed points of the group, as well as upon a judicious selection of elliptic fixed points as in [20], could give a technique for a satisfactory determination of the cusps, the fixed points of the parabolics.

1.2 NOTATION

For notational simplicity, we often write λ for λ_q. We call our continued fractions λ-fractions, denoted λF. They arise from a nearest-integer algorithm [15]; every $\alpha \in \mathbb{R}$

can be represented as

$$(1.2) \qquad \alpha = r_0\lambda + \cfrac{\varepsilon_1}{r_1\lambda + (\varepsilon_2/r_2\lambda + \dots)} = (r_0\lambda, \varepsilon_1/r_1\lambda, \varepsilon_2/r_2\lambda, \dots),$$

$$\text{where } r_0 \in \mathbb{Z} \text{ and } r_i \in \mathbb{Z}^+ \text{ for } i > 0 \text{ and } \varepsilon_i = \pm 1.$$

We recall that the convergents P_n/Q_n obtained by truncating the λF of α after n steps satisfy:

$$(1.3) \qquad P_n = r_n\lambda P_{n-1} + \varepsilon_n P_{n-2}, \quad Q_n = r_n\lambda Q_{n-1} + \varepsilon_n Q_{n-2}; \quad n > 1;$$

and the determinant relation

$$(1.4) \qquad P_n Q_{n-1} - Q_n P_{n-1} = (-1)^{n-1}\varepsilon_1\varepsilon_2\dots\varepsilon_n.$$

In all that follows, we shall assume that the λF considered are in their unique *reduced* form, as defined in [15, p.555]. We recall that the definition of reduced λF restricts the appearance of the sequences $B(h)$, where $h = \lfloor q - 3/2 \rfloor$ and for any natural number s, $B(s) = \left(\lambda, \underbrace{-1/\lambda, \dots, -1/\lambda}_{s-1 \text{ times}} \right).$

2. PERIODIC $\sqrt{D}/C$

We let C and D be elements of $\mathbb{Z}[\lambda_q]$ such that D is square-free up to units. Let the purely periodic λF

$$\beta = (r_{s+1}\lambda, \varepsilon_{s+2}/r_{s+2}, \dots, \varepsilon_n/r_n\lambda, \varepsilon/\beta).$$

We write

$$\left(r_0\lambda, \varepsilon_1/r_1\lambda, \dots, \varepsilon_s/r_s\lambda, \varepsilon_{s+1}\overline{/r_{s+1}\lambda, \dots, \varepsilon_n/r_n\lambda; \varepsilon} \right)$$

for the periodic λF

$$\alpha = (r_0\lambda, \varepsilon_1/r_1\lambda, \dots, \varepsilon_s/r_s\lambda, \varepsilon_{s+1}/\beta).$$

The acyclic part of α is then $(r_0\lambda, \varepsilon_1/r_1\lambda, \dots, \varepsilon_s/r_s\lambda, \varepsilon_{s+1})$.

The next few results follow work in the classical case, as in [6]; however there are interesting difficulties in the present setting.

LEMMA 2.1. *If $\alpha = \sqrt{D}/C$ is greater than $2/\lambda_q$, then $\lambda F(\alpha)$ is not purely periodic.*

PROOF: If $\alpha = \left(\overline{r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_n/r_n\lambda; \varepsilon}\right)$, let $p_n/q_n = (r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_n/r_n\lambda)$. Thus,

$$\alpha = (r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_n/r_n\lambda, \varepsilon/\alpha)$$

(2.1)
$$= \frac{\alpha p_n + \varepsilon p_{n-1}}{\alpha q_n + \varepsilon q_{n-1}}.$$

From this,

(2.2)
$$\alpha^2 q_n + (\varepsilon q_{n-1} - p_n)\alpha - \varepsilon p_{n-1} = 0.$$

Since $\alpha \notin \mathbb{Q}(\lambda_q)$, but $\alpha^2 \in \mathbb{Q}(\lambda_q)$, it follows that

(2.3)
$$\varepsilon q_{n-1} - p_n = 0, \qquad \alpha^2 = \varepsilon \frac{p_{n-1}}{q_n}.$$

Thus, $\varepsilon = 1$, $p_n = q_{n-1}$, and $\alpha^2 = p_{n-1}/q_n$. But, $q_n \geqslant q_{n-1}$, hence

(2.4)
$$\alpha^2 \leqslant \frac{p_{n-1}}{q_{n-1}} = \frac{p_{n-1}}{p_n}.$$

As no $r_i = 0$, equation (1.3) gives $p_n/p_{n-1} = (r_n\lambda, \varepsilon_n/r_{n-1}, \ldots, \varepsilon_1/r_0\lambda)$. From our use of the nearest-integer algorithm, $p_n/p_{n-1} \geqslant \lambda_q/2$. Thus, $\alpha^2 \leqslant 2/\lambda_q$. Since $2/\lambda_q > 1$, $\alpha < 2/\lambda_q$ and hence does not meet our hypotheses. $\qquad\square$

PROPOSITION 2.2. *If $\alpha = \sqrt{D}/C$ is greater than $2/\lambda_q$ and has a periodic λF, then this λF has no more than one acyclic term.*

PROOF: We suppose that

(2.5)
$$\alpha = \left(r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_s/r_s\lambda, \varepsilon_{s+1}/\overline{r_{s+1}\lambda, \ldots, \varepsilon_n/r_n\lambda; \varepsilon}}\right),$$

with the period of α not representable as beginning at r_j, $j \leqslant s$.

Let P_j/Q_j be the convergents of α and $\beta = \left(\overline{r_{s+1}\lambda, \ldots, \varepsilon_n/r_n\lambda; \varepsilon}}\right)$.
Then,

(2.6)
$$\alpha = \frac{\beta p_s + \varepsilon_{s+1} p_{s-1}}{\beta q_s + \varepsilon_{s+1} q_{s-1}}$$
$$= \frac{\beta p_n + \varepsilon p_{n-1}}{\beta q_n + \varepsilon q_{n-1}}.$$

Eliminating β, one obtains

$$(2.7) \qquad (-\varepsilon_{s+1}q_n q_{s-1} + \varepsilon\, q_s q_{n-1})\alpha^2$$
$$+ (\varepsilon_{s+1}q_n p_{s-1} + \varepsilon_{s+1}q_{s-1}p_n - \varepsilon\, p_{n-1}q_s - \varepsilon\, p_s q_{n-1})\alpha$$
$$+ (\varepsilon p_{n-1}p_s - \varepsilon_{s+1}p_{s-1}p_n) = 0$$

Again, $\alpha \notin \mathbb{Q}(\lambda_q)$. Hence,

$$(2.8) \qquad \begin{cases} \varepsilon_{s+1}(q_n p_{s-1} + q_{s-1}p_n) - \varepsilon\,(q_s p_{n-1} + q_{n-1}p_s) = 0, \\ \dfrac{\varepsilon\, p_{n-1}p_s - \varepsilon_{s+1}p_{s-1}p_n}{\varepsilon_{s+1}q_n q_{s-1} - \varepsilon\, q_s q_{n-1}} = \alpha^2. \end{cases}$$

From the first of these equations, $\varepsilon = \varepsilon_{s+1}$. The second can then be expressed as

$$(2.9) \qquad \frac{p_{n-1}p_{s-1}}{q_{n-1}q_{s-1}}\left(\frac{p_s/p_{s-1} - p_n/p_{n-1}}{q_n/q_{n-1} - q_s/q_{s-1}}\right) = \alpha^2.$$

Clearly, the second factor in the left hand side of equation (2.9) must be positive. For ease of notation, when $u < t$, we define

$$(2.10) \qquad f_{t,u} := (r_{t-1}\lambda, \varepsilon_{t-1}/r_{t-2}\lambda, \ldots, \varepsilon_u/r_u\lambda).$$

Since $\alpha > \lambda/2$ implies that all r_i are non-zero, $p_n/p_{n-1} = r_n\lambda + \varepsilon_n/f_{n,0}$ and similarly for p_s/p_{s-1}. Also, $q_n/q_{n-1} = r_n\lambda + \varepsilon_n/f_{n,1}$ and similarly for q_s/q_{s-1}. Thus the second factor of (2.9) gives

$$(2.11) \qquad \frac{-[(r_n - r_s)\lambda + (\varepsilon_n/f_{n,0} - \varepsilon_s/f_{s,0})]}{(r_n - r_s)\lambda + (\varepsilon_n/f_{n,1} - \varepsilon_s/f_{s,1})} > 0.$$

Since $\varepsilon = \varepsilon_{s+1}$, if r_s were equal to r_n, the period of α would begin at r_s, contrary to our assumption. Thus, $r_s \neq r_n$. Furthermore, $1/f_{t,u} \leqslant 2/\lambda$ whenever $t > u$. Thus, if $|r_s - r_n| \geqslant 2$, then

$$(2.12) \qquad |(r_n - r_s)\lambda| \geqslant 2\lambda \geqslant 4/\lambda \geqslant |1/f_{t,u} \pm 1/f_{t',u}|.$$

In this case, inequality (2.11) is not satisfied. Hence, we consider the case of $|r_s - r_n| = 1$. If $\varepsilon_s = \varepsilon_n$, then $|\varepsilon_n/f_{n,0} - \varepsilon_s/f_{s,0}| \leqslant 2/\lambda_q \leqslant \lambda_q$, thus inequality (2.11) would not hold. Furthermore,

$$(2.13) \qquad r_n - r_s = \varepsilon_s = -\varepsilon_n,$$

or again, inequality (2.11) is not satisfied.

In order to complete the proof of our Proposition, we show that inequality (2.11) is never satisfied. In fact, we shall show that $1/f_{t,u} + 1/f_{t',u'} < \lambda_q$ whenever $t \neq t'$, $\lambda_q F(\alpha)$ has more than $\max\{t, t'\}$ terms, $\varepsilon_t \varepsilon_{t'} = -1$, and $\alpha > 2/\lambda$.

When q is even, $\alpha > 2/\lambda$ implies that $f_{t,u} > B(h+1) = 2/\lambda_q$. Therefore,

$$(2.14) \qquad\qquad 1/f_{t,u} + 1/f_{t',u'} < \lambda_q \ (q \text{ even}).$$

For q odd, we note that for $\lambda F(\alpha)$ to be reduced and $\varepsilon_t \varepsilon_{t'} = -1$, not both $f_{t,u}$ and $f_{t',u}$ can begin $(B(h), -1/2\lambda, -1/B(h), \dots)$, let us say that $f_{t',u}$ does not.

We now treat the cases of $h \geqslant 2$ and $h = 1$ separately. For the former,

$$(2.15) \qquad \begin{aligned} f_{t,u} &> f := \left(B(h), -1/\overline{2\lambda, -1/B(h), -1/2\lambda, -1/B(h-1); -1} \right) \\ f_{t',u} &> f' := \left(\overline{B(h), -1/2\lambda, -1/B(h-1), -1/2\lambda; -1} \right). \end{aligned}$$

Let $\beta = \left(\overline{2\lambda, -1/B(h), -1/2\lambda, -1/B(h-1); -1} \right)$. Thus, $1/f' = 2\lambda - \beta$ and $f = (B(h), -1/\beta)$. In [13], (3.16) defining β has a typographical error of a missing -1 at the end of its period, but (3.26) correctly shows that β satisfies

$$(2.16) \qquad\qquad -\beta^2 + (3\lambda - 2)\beta - 2\lambda^2 + 2\lambda + 1 = 0.$$

We evaluate f as follows. Since $1/(\lambda - 1) = B(h)$, which equals $\lambda - 1/B(h-1)$, is palindromic in its λF form,

$$(2.17) \qquad B(h) = \frac{X}{Y}, \quad B(h-1) = \frac{Y}{Z}, \quad Y = (\lambda - 1)X, \quad Z = (\lambda^2 - \lambda - 1)X.$$

Thus, for arbitrary $\phi = \pm 1$ and $\omega \in \mathbb{R}^+$ we can evaluate

$$(2.18) \qquad \begin{aligned} (B(h), \phi/\omega) &= \frac{\omega X + \phi Y}{\omega Y + \phi Z} \\ &= \frac{\omega + \phi(\lambda - 1)}{\omega(\lambda - 1) + \phi(\lambda^2 - \lambda - 1)}. \end{aligned}$$

We note that since $\lambda_5^2 - \lambda_5 - 1 = 0$, (2.18) also holds true when $h = 1$.

Applying the above to the case at hand, we find that

$$\frac{1}{f} = \frac{(\beta(\lambda - 1) - (\lambda^2 - \lambda - 1))}{(\beta - (\lambda - 1))},$$

and thus

$$(2.19) \qquad \frac{1}{f} + \frac{1}{f'} = \frac{-\beta^2 + (4\lambda - 2)\beta - 3\lambda^2 + 3\lambda + 1}{\beta - (\lambda - 1)},$$

which by equation (2.14) becomes

$$(2.20) \qquad \frac{1}{f} + \frac{1}{f'} = \lambda.$$

Hence, $1/f_{t,u} + 1/f_{t',u} < \lambda_q$, and we have proved our Proposition for the case of $h \geqslant 2$.

When $h = 1$ and $q = 5$, we must replace (2.15) by

$$(2.21) \qquad \begin{aligned} f_{t,u} &\geqslant f := \left(\lambda, -1/\overline{2\lambda, -1/\lambda, -1/2\lambda; -1} \right) \\ f_{t',u} &\geqslant f' := \left(\overline{\lambda, -1/2\lambda, -1/2\lambda; -1} \right). \end{aligned}$$

The analogue of β is $\gamma := \left(\overline{2\lambda, -1/\lambda, -1/2\lambda; -1} \right)$, so that $1/f = \gamma/(\gamma\lambda - 1)$ and $1/f' = 2\lambda - \gamma$. Thus,

$$(2.22) \qquad \frac{1}{f} + \frac{1}{f'} = \frac{-\lambda\gamma^2 + (2\lambda^2 + 2)\gamma - 2\lambda}{\lambda\gamma - 1}.$$

Let P_i/Q_i be the convergents of γ, so that $\gamma = (\gamma P_2 - P_1)/(\gamma Q_2 - Q_1)$. Using $\lambda^2 - \lambda - 1 = 0$, one now finds

$$(2.23) \qquad (2\lambda + 1)\gamma^2 - (5\lambda + 4)\gamma + 2\lambda + 1 = 0,$$

that is, $-\lambda^3\gamma^2 + (5\lambda + 4)\gamma - \lambda^3 = 0$, and we have

$$\frac{1}{f} + \frac{1}{f'} = \frac{\lambda^4\gamma - \lambda^3}{\lambda^2(\lambda\gamma - 1)} = \lambda.$$

We have proven our Proposition.   □

3. PERIODIC $\sqrt{D}/C$ CONTINUED

We now find the exact form of the period for periodic $\sqrt{D}/C$. As our examples will indicate, this includes possibilities which do not occur for the classical simple continued fractions. This discrepancy is due to the fact that a λF in reverse order may not be reduced.

THEOREM 3.1. *If $\alpha = \sqrt{D}/C$ is greater than $2/\lambda_q$ and has a periodic λF, then*

$$\sqrt{D}/C = \left(r_0\lambda, \varepsilon_1/\overline{r_1\lambda, \ldots, \varepsilon_n/r_n\lambda; \varepsilon_1} \right),$$

where $(r_n\lambda, \varepsilon_n/r_{n-1}, \ldots, \varepsilon_2/r_1\lambda) = (2r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_{n-1}/r_{n-1}\lambda)$.

PROOF: By Lemma 2.1 and Proposition 2.1,

(3.1)
$$\alpha = \left(r_0\lambda, \varepsilon_1/\overline{r_1\lambda, \ldots, \varepsilon_n/r_n\lambda; \varepsilon_1}\right)$$
$$= \left(r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_n/r_n\lambda, +1/\left(\frac{1}{\alpha - r_0\lambda}\right)\right).$$

This gives

(3.2)
$$\overline{\alpha = (1/\alpha - r_0\lambda)p_n + p_{n-1}}$$
$$(1/\alpha - r_0\lambda)q_n + q_{n-1}.$$

Hence, since α is not in $\mathbb{Q}(\lambda_q)$,

(3.3)
$$\begin{cases} q_n - r_0\lambda\, q_{n-1} - p_{n-1} = 0 \\ p_n - r_0\lambda\, p_{n-1} - q_{n-1}\alpha^2 = 0. \end{cases}$$

Therefore,

(3.4)
$$\begin{cases} \dfrac{q_n}{q_{n-1}} = r_0\lambda + \dfrac{p_{n-1}}{q_{n-1}}, \\ \alpha^2 = \dfrac{p_n - r_0\lambda\, p_{n-1}}{q_{n-1}}. \end{cases}$$

From the first of these equations,

(3.5) $\qquad (r_n\lambda, \varepsilon_n/r_{n-1}\lambda, \ldots, \varepsilon_2/r_1\lambda) = (2r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_{n-1}/r_{n-1}\lambda).$

Thus, our Theorem is proved.   □

Examples of the above include the classical form of the period ending with $r_n = 2r_0$ and the initial section of the period being palindromic. In the λF setting, the ε_i are also to be considered in the palindrome.

EXAMPLE 3.1. Let $q = 5$ and $\alpha = \sqrt{6}$. One finds

$$\sqrt{6} = \left(2\lambda_5, -1/\overline{\lambda_5, -1/2\lambda_5, -1/2\lambda_5, -1/2\lambda_5, -1/\lambda_5, -1/4\lambda_5; -1}\right).$$

A different phenomenon arises when a truncated λF of α in reverse order is not reduced.

EXAMPLE 3.2. Consider $q = 6$ and $\alpha = \sqrt{7}$, so that

$$\sqrt{7} = \left(2\lambda_6, -1/\overline{\lambda_6, -1/\lambda_6, 1/3\lambda_6}; -1\right).$$

Here $r_0\lambda_6 + P_2/Q_2 = (4\lambda_6, -1/\lambda_6, -1/\lambda_6) = 7\lambda_6/2$. The other side of equation (3.5) is $Q_3/Q_2 = (3\lambda_6, 1/\lambda_6, -1/\lambda_6) = 7\lambda_6/2$.

Since *reduced* λF are unique, these "non-classical" periods are due to conditions defining reduced which are not symmetric with respect to reverse order. Indeed, the only manner in which these periods can arise is for $\lambda/2$ to be surrounded by oppposite signs in $\lambda F(\alpha)$.

In the following, we shall need the

DEFINITION: We say that $P/Q \in \mathbb{Q}(\lambda_q)$ is in *reduced finite λF form* if $\lambda F(P/Q)$ is of finite length and P and Q are exactly the numerator and denominator of the ultimate convergent of P/Q.

THEOREM 3.2. *If $\alpha = \sqrt{\overline{D}}/C$ is greater than $2/\lambda_q$, then $\lambda F(\alpha)$ is periodic if and only if there exist $x > y$ non-zero in $\mathbb{Z}[\lambda]$ such that $\begin{pmatrix} \nu x & \alpha^2 y \\ \nu y & x \end{pmatrix} \in G_q$, for $\nu = \pm 1$.*

PROOF: Suppose that α is periodic, so that $\lambda F(\alpha)$ has the form of Theorem 3.1.

Now define $\beta = (r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_n/(r_n - r_0)\lambda)$, as well as $P_n/Q_n = \beta$. We have $P_{n-1} = Q_n$ and

$$\begin{aligned}
\alpha &= (r_0\lambda, \varepsilon_1/r_1\lambda, \ldots, \varepsilon_n/(r_n - r_0)\lambda, +1/(1/\alpha)) \\
&= \frac{(1/\alpha)P_n + P_{n-1}}{(1/\alpha)Q_n + Q_{n-1}} \\
&= \frac{(1/\alpha)P_n + Q_n}{(1/\alpha)Q_n + Q_{n-1}}.
\end{aligned}$$

(3.6)

One solves to find $P_n = \alpha^2 Q_{n-1}$. By [17, Theorem 2], $\begin{pmatrix} \nu P_{n-1} & P_n \\ \nu Q_{n-1} & Q_n \end{pmatrix} \in G_q$, for $\nu = \pm 1$. Furthermore, $Q_n \geqslant Q_{n-1}$, therefore, $x > y$.

In the other direction, suppose that $\begin{pmatrix} \nu x & \alpha^2 y \\ \nu y & x \end{pmatrix} \in G_q$. From [17, Theorem 2], both $\alpha^2 y/x$ and x/y are in reduced finite λF form and are the consecutive convergents of some real θ. As $Q_n \geqslant Q_{n-1}$, $x > y$ implies that a choice of $\alpha^2 y/x = P_n/Q_n$ forces $x/y = P_{n-1}/Q_{n-1}$. We note that then $P_n/Q_{n-1} = \alpha^2$.

Since $P_{n-1} = Q_n$, one has $P_n/P_{n-1} = P_n/Q_n$, hence if

$$\frac{P_n}{Q_n} = (R_0\lambda, \varepsilon_1/R_1\lambda, \ldots, \varepsilon_n/R_n\lambda),$$

(3.7) $(R_n\lambda, \varepsilon_n/R_{n-1}, \ldots, \varepsilon_1/R_0\lambda) = (R_0\lambda, \varepsilon_1/R_1\lambda, \ldots, \varepsilon_n/R_n\lambda).$

Now letting

(3.8)
$$\xi = \left(R_0\lambda, \varepsilon_1/\overline{R_1\lambda, \ldots, \varepsilon_n/(R_n + R_0)\lambda; \varepsilon_1}\right)$$
$$= (R_0\lambda, \varepsilon_1/R_1\lambda, \ldots, \varepsilon_n/R_n, +1/(1/\xi)),$$

we find $\xi^2 = P_n/Q_{n-1} = \alpha^2$. Hence, α equals ξ, and is periodic.   □

EXAMPLE 3.3. We note that α might actually be in $\mathbb{Q}(\lambda_q)$ and have the above form. In particular, let $q = 6$ and take $\alpha = 7$. One finds

(3.9) $7 = \left(4\lambda_6, 1/\overline{8\lambda_6; +1}\right).$

   COROLLARY 3.1. *If $\alpha = \sqrt{D}/C$ is greater than $2/\lambda_q$, then $\lambda F(\alpha)$ is non-periodic if and only if $G(\alpha, x, y) = \left\langle \begin{pmatrix} \nu x & \alpha^2 y \\ \nu y & x \end{pmatrix}, G_q \right\rangle$ is non-Fuchsian for all x, $y \in \mathbb{Z}[\lambda_q]$, $\nu = \pm 1$.*

   PROOF: By Beardon [4, Theorem 11.3.3], a Hecke group is never a proper subgroup of a Fuchsian group.   □

   Let $U_1(D)$ be the group of all units of the field $\mathbb{Q}(\lambda_q)\left(\sqrt{D}\right)$ with relative norm to $\mathbb{Q}(\lambda_q)$ being 1. We shall denote elements of $U_1(D)$ by $x + \sqrt{D}y$.

   COROLLARY 3.2. *For D greater than $2/\lambda_q$ and square-free up to units in $\mathbb{Z}[\lambda_q]$, the maximal subgroup of $U_1(D)$ having every Dy/x of reduced finite λF form with x/y its penultimate convergent is at most infinite cyclic.*

   PROOF: Since $x^2 = 1 + Dy^2$,

(3.10) $trace\begin{pmatrix} x & Dy \\ y & x \end{pmatrix} = 2x > 2.$

   Thus, each such matrix is hyperbolic. Furthermore, each such fixes both $\sqrt{D}$ and $-\sqrt{D}$ under the usual Möbius action. However, Lehner [11, p.15] shows that a Fuchsian group has at most one cyclic subgroup of hyperbolic elements fixing the same point.

   The map ϕ from $U_1(D)$ to $SL(2, \mathbb{R})$,

(3.11) $\phi\left(x + \sqrt{D}y\right) := \begin{pmatrix} x & Dy \\ y & x \end{pmatrix}$

is an injective homomorphism. Therefore, our Corollary is proved.   □

   Of course, $U_1(D)$ exactly corresponds to the solution set of the Pell's equation

(3.12) $x^2 - Dy^2 = 1, \quad \text{with} \quad x, y \in \mathbb{Z}[\lambda_q].$

Thus, when $\lambda F\left(\sqrt{D}\right)$ is periodic, one finds one "$\mathbb{Z}$'s" worth of solutions to Pell's equation as convergents.

4. THE LEGENDRE CONSTANT OF DIOPHANTINE APPROXIMATION

From the above, we see that when $\sqrt{D}$ is periodic, there is a proper subgroup of the units of $\mathbb{Q}(\lambda)\left(\sqrt{D}\right)$ found as convergents to $\lambda F\left(\sqrt{D}\right)$. On the other hand, every $x + \sqrt{D}y$ in $U_1(D)$ allows one to satisfy:

$$(4.1) \qquad \left| \sqrt{D} - \frac{x}{y} \right| < 1/2y^2.$$

Thus [12, Theorem 3] seems to be false. Indeed, in [13], there is a correction to [12] which amounts to considering only x/y in reduced finite λF form. A correct version of the theorem follows.

THEOREM 4.1. *Let $\alpha \in \mathbb{R} \setminus G_q(\infty)$ and P/Q be in reduced finite λF form. Let*

$$|\alpha - P/Q| < 1/\gamma_q Q^2, \qquad where \qquad \gamma_q = \begin{cases} \dfrac{6\lambda^2 + 8}{\lambda^3 + 4\lambda}, & \text{if } q \text{ is even;} \\[3mm] \dfrac{\lambda^2 + 4}{\lambda^2 - \lambda + 2}, & \text{if } q \text{ is odd.} \end{cases}$$

Then P/Q is a convergent of the reduced $\lambda F(\alpha)$.

PROOF: We follow Lehner's adaptation of Hardy and Wright's classical proof, making the corrections necessary to avoid certain errors. Write

$$(4.2) \qquad \alpha - P/Q = \frac{\varepsilon \theta}{Q^2}, \qquad \text{where } 0 < \theta < 1/\gamma_q, \ \varepsilon = \pm 1.$$

We have:

$$(4.3) \qquad P/Q = (r_0\lambda, \varepsilon_1/r_1\lambda, \dots, \varepsilon_{n-1}/r_{n-1}\lambda).$$

Call the convergents P_i/Q_i; thus $P/Q = P_{n-1}/Q_{n-1}$. Define ε_n by $\varepsilon = \varepsilon_1\varepsilon_2 \dots \varepsilon_n(-1)^{n-1}$ and ω by $\alpha = (P_{n-1}\omega + \varepsilon_n P_{n-2})/(Q_{n-1}\omega + \varepsilon_n Q_{n-2})$.

As in Lehner, $(\varepsilon\theta/Q^2) = \alpha - P/Q$ leads to $\theta = Q_{n-1}/(Q_{n-1}\omega + \varepsilon_n Q_{n-2})$ and $\omega = (Q_{n-1} - \theta\varepsilon_n Q_{n-2})/(\theta Q_{n-1}) > 0$.

We expand ω in a reduced λF:

$$(4.4) \qquad \omega = (r_n\lambda, \varepsilon_{n+1}/r_{n+1}\lambda, \dots)$$

and consider

$$(4.5) \qquad \begin{aligned} \beta &= (r_0\lambda, \varepsilon_1/r_1\lambda, \dots, \varepsilon_{n-1}/r_{n-1}\lambda, \varepsilon_n/\omega) \\ &= (r_0\lambda, \varepsilon_1/r_1\lambda, \dots, \varepsilon_{n-1}/r_{n-1}\lambda, \varepsilon_n/r_n\lambda, \varepsilon_{n+1}/r_{n+1}\lambda, \dots). \end{aligned}$$

We obtain $\beta = (P_{n-1}\omega + \varepsilon_n P_{n-2})/(Q_{n-1}\omega + \varepsilon_n Q_{n-2})$. Therefore, $\beta = \alpha$. If β is reduced or if β reduces to a form which continues to have P/Q as a convergent, then we are done.

We consider the expansion of β *vis-à-vis* the defining Properties 1–5 of [**15**, p.555]. The pieces for this representation coming from P/Q and from ω are each separately reduced. Thus, we determine when their concatenation introduces an inadmissible sequence.

We first show that our hypotheses imply that if $\varepsilon_n = -1$, then the expansion for ω cannot begin with both $r_n = 1$ and $\varepsilon_{n+1} = -1$. As in the above example, we find that

$$(4.6) \qquad \left| \alpha - \frac{P}{Q} \right| Q^2 = \frac{1}{|\omega + \varepsilon_n Q_{n-2}/Q_{n-1}|}.$$

If ω had its expansion beginning as above, then $\lambda > \omega > \lambda/2$. But each $\lambda < 2$. Thus $\gamma_q > 2 > \lambda > \omega - Q_{n-2}/Q_{n-1}$. If $Q_{n-2}/Q_{n-1} > \omega$, then $Q_{n-2}/Q_{n-1} - \omega \leqslant 1 - \omega \leqslant 1 - \lambda/2 < 1$. Thus we have our claim.

An argument almost identical to that above shows that if $\lambda F(P/Q)$ ends with $B(j)$, for j maximal, ε_n is -1 and ω begins with $r_n = 1$, then $|\alpha - P/Q| Q^2 > 1/2$. Since Properties 1 and 2 ensure that P/Q ends with at most a $B(h)$ if q is odd or a $B(h+1)$ if q is even, the above says that no $B(h+1)$ for q odd nor $B(h+2)$ for q even is formed. This then guarantees that β satisfies Property 1 when q is odd.

When q is even, the only remaining possibility for β not to satisfy Property 1 is if P/Q ends with $B(h+1)$ and $\varepsilon_n = -1$. Since only Properties 1 and 5 refer to even q, and Property 5 only to finite λF, we finish the proof for these q by considering this last case.

Property 5 implies that $P/Q = (\dots, \varepsilon/r\lambda, +1/B(h+1))$. Recall from [**15**, p.554] that

$$(4.7) \qquad (\dots, \varepsilon/r\lambda, +1/B(h+1), -1/s\lambda, \dots)$$
$$= (\dots, \varepsilon/(r+1)\lambda, -1/B(h+1), +1/(s-1)\lambda, \dots).$$

Note that truncating both sides of this equation after the indicated $B(h+1)$ gives the same value. From this, if β is as above, with $s \neq 2$, then β reduces to a β' which still has P/Q as a convergent. If however, $\omega = (2\lambda, -1/B(h), -1/t\lambda, \dots)$, then $\beta' = (\dots, \varepsilon/(r+1)\lambda, -1/B(h+1), +1/\lambda, -1/B(h), -1/t\lambda, \dots)$ is not reduced. One has $\beta' = \beta''$, where

$$(4.8) \qquad \beta'' := (\dots, \varepsilon/(r+1)\lambda, -1/B(h), -1/2\lambda, -1/B(h+1), -1/(t-1)\lambda, \dots),$$

which no longer has P/Q as a convergent. We note that this process can continue to β''', $\beta^{(4)}$, *et cetera* if $t = 2$ and one has $(-1/B(h), -1/\dots)$ in the succeeding position. However, $\beta^{(n)}$ will differ from β'' only after the $(\dots, \varepsilon/(r+1)\lambda, -1/B(h), -1/2\lambda)$ block. Similar comments apply when $t = 1$.

We now bound $|\alpha - P/Q|\, Q^2$ under the hypothesis of q even, $P/Q = (\dots, \varepsilon/r\lambda, +1/B(h+1))$, and $\omega = (2\lambda, -1/B(h), -1/\xi)$, with $\xi > 0$, $\lambda F(\xi)$ of infinite length:

$$(4.9) \qquad \omega < (2\lambda, -1/B(h)) = \lambda + B(h+1) = \frac{\lambda^2 + 2}{\lambda};$$

$$(4.10) \qquad \frac{Q_{n-1}}{Q_{n-2}} = (B(h+1), +1/r\lambda, \varepsilon) < (B(h+1), +1/(2/\lambda)).$$

Since $B(h+1) = 2/\lambda$, reasoning similar to that surrounding equation (2.18) gives

$$(4.11) \qquad (B(h+1), +1/(2/\lambda)) = \frac{\lambda^2 + 4}{\lambda^3}.$$

Hence,

$$(4.12) \qquad \omega - \frac{Q_{n-2}}{Q_{n-1}} < \gamma_q , \quad q \text{ even} .$$

We remark that our bounds in (4.9) and (4.10) are sharp.

We now consider q to be odd. Our claim just before equation (4.6) shows that if $\lambda F(P/Q)$ does not end with $B(h)$, then β has no conflict with Properties 1 or 2 of the definition of reduced. The same is true with respect to Property 3. Properties 4 and 5 refer to finite expansions and hence are trivially satisfied by β. All five properties being satisfied, we conclude that β is indeed in reduced form.

From the above statements, we know that the only manner in which β can be non-reduced is when P/Q ends with $B(h)$ and β does not satisfy Property 3. If P/Q ends with $(\dots, B(h), -1/2\lambda, -1/B(h))$ and β is non-reduced with ω not starting $(2\lambda, -1/B(h), \dots)$, then by [15, p.554], β can be reduced in such a way that P/Q remains a convergent.

We are left with the case of: $\beta = (\dots, \varepsilon/B(h), -1/\omega)$, where $\omega = (2\lambda, -1/B(h), -1/\xi)$, for $\xi > 0$. A sharp upper bound for ω is $2\lambda - 1/B(h)$. This bound is $\lambda + 1$, by [15, p.556]. Now, $Q_{n-1}/Q_{n-2} = (B(h), \varepsilon/\dots, \varepsilon_2/r_1\lambda)$. If $\varepsilon = -1$, then $B(h) > Q_{n-1}/Q_{n-2}$, from which our bound follows easily. If $\varepsilon = +1$, then a sharp upper bound is $(B(h), 1/(2/\lambda)) \geqslant Q_{n-1}/Q_{n-2}$. From equation (2.18),

$$(4.13) \qquad \begin{aligned} (B(h), 1/(2/\lambda)) &= \frac{2X/\lambda - Y}{2Y/\lambda - Z} \\ &= \frac{\lambda^2 - \lambda + 2}{\lambda^3 - \lambda^2 + \lambda - 2}. \end{aligned}$$

Hence,

$$(4.14) \qquad \lambda + 1 - \frac{\lambda^3 - \lambda^2 + \lambda - 2}{\lambda^2 - \lambda + 2} > \omega - \frac{Q_{n-2}}{Q_{n-1}}.$$

This now simplifies to give our γ_q.   ▢

  We give examples which truly require γ_q instead of 2 as the constant of Diophantine approximation.

EXAMPLE 4.1. Let $q = 5$ and consider

$$(4.15) \qquad \frac{P}{Q} = (2\lambda, -1/\lambda, -1/2\lambda, 1/\lambda, 1/\lambda) = \frac{19\lambda + 14}{8\lambda + 5} = \frac{P_4}{Q_4}.$$

  Let $\omega = (2\lambda, -1/\lambda, -1/\sqrt{11})$ and

$$(4.16) \qquad \alpha = \frac{\omega P_4 - P_3}{\omega Q_4 - Q_3} = (2\lambda, -1/\lambda, -1/2\lambda, 1/2\lambda, -1/\lambda, \dots).$$

Since the unique reduced $\lambda F(P/Q)$ does not appear in this reduced λF expansion of α, P/Q is not a convergent of α. Now,

$$(4.17) \qquad \left| \alpha - \frac{P}{Q} \right| Q^2 = \left| 2.49187\cdots - \frac{19\lambda + 14}{8\lambda + 5} \right| (8\lambda + 5)^2 = 0.49\dots.$$

Of course, $1/\gamma_5 = 0.45\dots$ is less than the above.   ▢

EXAMPLE 4.2. Let $q = 6$ and consider

$$(4.18) \qquad \frac{P}{Q} = (2\lambda, 1/\lambda, -1/\lambda, 1/\lambda, -1/\lambda) = \frac{17\sqrt{3}}{7} = \frac{P_4}{Q_4}.$$

  Let $\omega = (2\lambda, -1/\lambda, -1/\sqrt{10})$ and

$$(4.19) \qquad \alpha = \frac{\omega P_4 - P_3}{\omega Q_4 - Q_3} = (2\lambda, 1/\lambda, -1/2\lambda, -1/\lambda, -1/2\lambda, \dots).$$

Then

$$(4.20) \qquad \left| \alpha - \frac{P}{Q} \right| Q^2 = \left| 4.19628\cdots - \frac{17\sqrt{3}}{7} \right| 49 = 0.49\dots.$$

On the other hand, $1/\gamma_6 = 0.466\dots$, which is less than the above.   ▢

REFERENCES

[1] R. Adler and L. Flatto, 'Cross section maps for geodesic flows.I (the modular surface)', in *Ergodic theory and dynamical systems* **2**, Progress in Math., pp. 103–161 (Birkhäuser, Boston, Basil and Stuttgart, 1980).

[2] R. Adler and L. Flatto, 'Geodesic flows, interval maps, and symbolic dynamics', *Bull. AMS* **25** (1991), 229-334.

[3] E. Artin, 'Ein mechanisches System mit quasiergodischen Bahnen', *Abh. Math. Sem. Univ. Hamburg* **3** (1924), 170–175.

[4] A. F. Beardon, *The Geometry of discrete groups* (Springer-Verlag, Berlin, Heidelberg, New York, 1983).

[5] W. Bohro and G. Rosenberger, 'Eine Bemerkung zur Hecke-Gruppe $G(\lambda)$', *Abh. Math. Sem. Univ. Hamburg* **39** (1973), 83–87.

[6] G. Crystal, *Algebra, part II* (Black Ltd., London, 1922).

[7] L. Ford, 'A geometrical proof of a theorem of Hurwitz', *Proc. Edinburgh Math. Soc.* **35** (1917), 59–65.

[8] M.C Gutzwiller, *Chaos in classical and quantum mechanics* (Springer-Verlag, Berlin, Heidelberg, New York, 1990).

[9] A. Haas and C. Series, 'The Hurwitz constant and Diophantine approximation of Hecke groups', *J. London Math. Soc.* **34** (1986), 219–234.

[10] G. Hedlund, 'On the metrical transitivity of the geodesics on closed surfaces of constant negative curvature', *Ann. of Math (2)* **35** (1934), 787–808.

[11] J. Lehner, *A short course in automorphic forms* (Holt, New York, 1966).

[12] J. Lehner, 'Diophantine approximation of Hecke groups', *Glasgow Math. J.* **27** (1985), 117–127.

[13] J. Lehner, 'The local Hurwitz constant and Diophantine approximation on Hecke groups', *Math Comp* **55** (1990), 765–781.

[14] A. Leutbecher, 'Über die Heckeschen Gruppen $G(\lambda)$ II', *Math Ann.* **211** (1974), 63–84.

[15] D. Rosen, 'A class of continued fractions associated to certain properly discontinuous groups', *Duke Math. J.* **21** (1954), 549–562.

[16] D. Rosen, 'An arithmetic characterisation of the parabolic points of $G(2\cos\pi/5)$', *Proc. Glasgow Math. Assoc.* **6** (1963), 88–96.

[17] D. Rosen, 'The substitutions of the Hecke group $\Gamma(2\cos\pi/5)$', *Arch. Math.* **46** (1986), 533–538.

[18] T. Schmidt, 'Remarks on the Rosen λ-continued fractions', (submitted).

[19] C. Series, 'The modular surface and continued fractions', *J. London Math. Soc.* **31** (1985), 69–80.

[20] M. Sheingorn, 'Low height Hecke triangle group geodesics', in *Grosswald memorial volume*, Editors M.I. Knopp and M. Sheingorn (to appear).

[21] A.J. van der Poorten, 'An introduction to continued fractions', in *Diophantine analysis: Lecture Notes in Math* 109, Editors J.H. Loxton and A.J. van der Poorten, pp. 99–138 (Cambridge Univ. Press, Cambridge, 1986).

Department of Mathematics
Swarthmore College
Swarthmore PA 19081
United States of America

Department of Mathematics
Widener University
Chester PA 19013
United States of America