Roots of Modular Units

A. Beeson

May 2, 2014

Abstract

Let $p$ be a prime. We prove that if a modular unit has a $p^{th}$ root that is again a modular unit then the level of that root is at most $p$ times the level of the original unit.

1 Introduction

We prove the result in section 2, but because the literature contains inconsistent definitions, in section 3 we give a summary of the relevant definitions. The reader interested in learning more about the theory of modular functions should see [DS05].

Our main result is that the $p^{th}$ root of a modular unit, if it is again a modular unit, has the level that one would expect. The source of this question was the study of the Siegel functions and their square roots. See [Kub81]. The result could likely be extended using similar techniques to the general case of a cyclic Galois group.

2 On the level of a root of a modular function

Theorem 2.1. Let $\mathcal{F}_N$ be the field of modular functions of level $N$ with Fourier coefficients in $\mathbb{Q}(\zeta_N)$; that is, $\mathcal{F}_N$ is the fixed field of $\Gamma(N)$ inside $\mathcal{F}$, the full field of modular functions. If $p$ is a rational prime and $f(z) \in \mathcal{F}_N \setminus \mathcal{F}_N^p$ is a modular unit and $f(z)^{1/p}$ has level $M$ for some $M \in \mathbb{N}$ with $N | M$ then, in fact, $f(z)^{1/p}$ has level $pN$. 

Proof. We assume \( f(z)^{1/p} \) is known to be invariant under the subgroup \( \Gamma(M) \subseteq \Gamma(N) \). We will show that \( f(z)^{1/p} \) is invariant under \( \Gamma(pN) \).

Let \( \Gamma_1 \) be the subgroup of \( \Gamma(N) \) that fixes \( f(z)^{1/p} \); i.e., for all \( A \in \Gamma_1 \) and \( z \in \mathcal{H}^* \),

\[
f(A \circ z)^{1/p} = f(z)^{1/p}.
\]

Because \( \mathcal{F}_N(f(z)^{1/p}) \) is a degree \( p \) extension of \( \mathcal{F}_N \), the index \( [\Gamma(N) : \Gamma_1] = p \), and, thus, \( \Gamma_1 \) is a finite index subgroup of \( \Gamma \) as well. Furthermore, because \( f(z)^{1/p} \) is of level \( M \), \( \Gamma(M) \subseteq \Gamma_1 \). So we have the linear ordering of fields

\[
\mathbb{Q}(j(z)) \subseteq \mathcal{F}_N \subseteq \mathcal{F}_{\Gamma_1} \subseteq \mathcal{F}_M \subseteq \mathcal{F}
\]

where \( \mathcal{F}_{\Gamma_1} \) denotes the fixed field of \( \Gamma_1 \) inside \( \mathcal{F} \).

Let \( D \) be a fundamental domain for \( \Gamma \); so \( D \) is a simply connected subset of \( \mathcal{H}^* \) such that \( D \) contains precisely one point from each \( \Gamma \)-orbit. If \( D_1 \) is a fundamental domain for the subgroup of finite index \( \Gamma_1 \subseteq \Gamma \) then it is made up of translates of \( D \) by a full set of coset representatives for \( \Gamma_1 \) inside \( \Gamma \). Such a translate is called a modular triangle. Define the fan width of a fundamental domain at a cusp \( \alpha \) to be the order of the cyclic group that permutes the \( \Gamma_1 \)-inequivalent modular triangles meeting at \( \alpha \). Schoeneberg proves in [Sho74] that the conductor of a group \( \Gamma_1 \) is equal to the least common multiple of the fan widths at the rational cusps.

The index of \( \Gamma(N) \) in \( \Gamma \) is the size of \( SL_2(\mathbb{Z}/N\mathbb{Z})/\{\pm I\} \), which is

\[
l = [\Gamma : \Gamma(N)] = \frac{1}{2} N^3 \prod_{p|N} \left( 1 - \frac{1}{p^2} \right).
\]

As observed above, \( [\Gamma(N) : \Gamma_1] = p \), hence the group of automorphisms of \( \mathcal{F}_{\Gamma_1} \) fixing \( \mathcal{F}_N \) is cyclic. If we let \( \sigma \) be a generator then \( \Gamma(N) \) decomposes as the disjoint union

\[
\Gamma(N) = \Gamma_1 \cup \sigma \Gamma_1 \cup \sigma^2 \Gamma_1 \cup \cdots \cup \sigma^{p-1} \Gamma_1.
\]

And if \( \{A_1, \ldots, A_l\} \) is a complete set of representative for \( \Gamma/\Gamma(N) \) then

\[
\{A_1, \ldots, A_l, \sigma A_1, \ldots, \sigma A_l, \sigma^2 A_1, \ldots, \sigma^2 A_l, \ldots, \sigma^{p-1} A_1, \ldots, \sigma^{p-1} A_l\}
\]

is a complete set of representatives for \( \Gamma/\Gamma_1 \). Recalling our notation \( C(\Gamma(N)) \) for the cusps of \( \mathcal{H}/\Gamma(N) \), we see that

\[
C(\Gamma(N)) \subset C(\Gamma_1)
\]
and if \( \alpha \) is a cusp of \( \Gamma_1 \) then \( \sigma^i \alpha \in C(\Gamma(N)) \) for some \( i \) with \( 1 \leq i \leq p \).

Choose \( \{A_1, \ldots, A_l\} \) to be a complete set of representatives for \( \Gamma/\Gamma(N) \) such that
\[
\mathcal{D}_N = \cup_i A_i(\mathcal{D})
\]
is a fundamental domain for \( \mathcal{H}/\Gamma(N) \). Let \( \mathcal{D}_1 \) be a fundamental domain for \( \mathcal{H}/\Gamma_1 \).

Schoeneberg’s theorem implies that the least common multiple of the fan widths for \( \mathcal{D}_N \), \( \Gamma \) is \( N \). We will use this to show that, for any cusp \( \alpha \) of \( \Gamma_1 \), the fan width of \( \mathcal{D}_1 \) at \( \alpha \) divides \( pN \) so, by Schoeneberg’s theorem, the conductor of \( \Gamma_1 \) divides \( pN \). Because \( f(z)^{1/p} \notin \mathcal{F}_N \), \( f(z)^{1/p} \) must have level \( pN \).

Let \( \alpha \) be a cusp of \( \Gamma_1 \). As observed above, this implies \( \sigma^i \alpha \) is a cusp of \( \Gamma(N) \) for at least one \( i \) with \( 1 \leq i \leq p \). Letting \( \beta = \sigma^i \alpha \) be a translate of \( \alpha \) that is a cusp of \( \mathcal{D}_N \), we see that multiplication by \( \sigma^i \) is a homeomorphism between a neighborhood of \( \alpha \) and a neighborhood of \( \beta \). Thus, it suffices to prove the result for the cusps of \( \mathcal{D}_1 \) that are also cusps of \( \mathcal{D}_N \).

**Lemma 2.2.** If \( \alpha \) is a cusp of \( \Gamma_1 \) and of \( \Gamma(N) \) that is of fan width \( n \) for \( \mathcal{D}_N \) then its width for \( \mathcal{D}_1 \) is \( n \) or \( pn \).

**Proof.** Recall \( \Gamma_1 \subseteq \Gamma(N) \) so \( \Gamma \backslash \Gamma(N) \subseteq \Gamma \backslash \Gamma_1 \) and that the fan width \( n \) of \( \mathcal{D}_1 \) at the cusp \( \alpha \) is the order of the cyclic group that permutes the \( n \) \( \Gamma_1 \)-inequivalent triangles meeting at \( \alpha \).

If two triangles are \( \Gamma_N \)-inequivalent then they are \( \Gamma_1 \)-inequivalent so, assuming the width for \( \mathcal{D}_N \) is \( n \), the width for \( \Gamma_1 \) is at least \( n \). Then since \( [\Gamma(N) : \Gamma_1] = p \), we see that the width of a triangle for \( \Gamma_1 \) is no more than \( pn \).

This concludes the proof of the theorem so any \( p \text{th} \) root of a level \( N \) modular function that has a level, in fact, has level \( N \) or \( pN \). As \( f(z)^{1/p} \) is not level \( N \) by assumption, it must be level \( pN \).

The special case we are currently most interested in is when \( p = 2 \), in which case we have the following theorem and its corollary.

**Theorem 2.3.** If \( f(z) \in \mathcal{F}_N \setminus \mathcal{F}_N^2 \) is a modular unit and \( \sqrt{f(z)} \) has level \( M \) for some \( M \in \mathbb{N} \) with \( N|M \) then, in fact, \( \sqrt{f(z)} \) has level \( 2N \).

**Theorem 2.4.** If \( f(z) \in \mathcal{F}_N \setminus \mathcal{F}_N^2 \) is a modular unit with \( \sqrt{f(z)} \notin \mathcal{F}_N(\sqrt{j(z)} - 1728) \) then \( \sqrt{f(z)} \) is not level \( M \) for any \( M \in \mathbb{N} \).
Proof. By the index, there is a unique quadratic extension between $F_N$ and $F_{2N}$. We observe that $F_N(\sqrt{j(z) - 1728})$ is such an extension since $j(z) - 1728$ has a holomorphic $PSL_2(\mathbb{Z})$-invariant square root on $H$.

The theorem says that if the square root of a modular unit of level $N$ is a modular function on a congruence subgroup then it is level $2N$. Thus, because the Siegel units $\phi_{u,v}$ for $(u, v) \in \frac{1}{N}\mathbb{Z}$ are level $12N^2$, it suffices to show the square roots of Siegel functions are not level $24N^2$ in order to conclude that they do not, in fact, have a level at all. For definitions and further discussion of the Siegel units see [KL81].

3 Background

3.1 Modular functions

Let $\mathcal{H} = \{z \in \mathbb{C} | \text{Im}(z) > 0\}$ denote the complex upper half plane; let $\mathcal{H}^* = \mathcal{H} \cup \mathbb{P}^1(\mathbb{Q})$ be the extended upper half plane and $\hat{\mathbb{C}}$ the compactified complex plane. Let $\Gamma$ denote the (inhomogeneous) modular group, or the group of all fractional linear transformations mapping $\mathcal{H}$ to itself. Then $\Gamma$ is naturally identified with the matrix group

$$\Gamma = PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z})/\{\pm I\} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \big| a, b, c, d \in \mathbb{Z}, ad - bc = 1 \right\}/\{\pm I\},$$

which is generated by $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. The action of $\Gamma$ on an upper half-plane variable $z \in \mathcal{H}$ is given via fractional linear transformation:

$$A \circ z = \frac{az + b}{cz + d}.$$

A map $f : \mathcal{H}^* \to \hat{\mathbb{C}}$ is called a modular function (of level one) if

1. $f$ is meromorphic on $\mathcal{H}$,

2. $f(A \circ z) = f(z)$ for all $A \in \Gamma$ and $z \in \mathcal{H}^*$,
3. there is an $a > 0$ so that for $\text{Im}\{z\} > a$, $f(z)$ has an expansion in the local variable at $i\infty$, $q = e^{2\pi i z}$, of the form

$$f(z) = \sum_{n \geq n_0} a_n q^n, \ n \in \mathbb{Z}, \ a_{n_0} \neq 0.$$ 

so $n_0$ determines the behavior of $f$ as $z \to \infty$. If $n_0 < 0$ then $f(i\infty) = \infty$; if $n_0 = 0$ then $f(i\infty) = a_0$; and if $n_0 > 0$ then $f(i\infty) = 0$. In the last case, we call $f$ a cusp form.

Fix a natural number $N > 2$. Let $\Gamma(N) \leq \Gamma$ be the (inhomogeneous) principal congruence subgroup modulo $N$, or the kernel of the reduction mod $N$ map. In other words,

$$\begin{array}{c}
1 \longrightarrow \Gamma(N) \longrightarrow \Gamma \longrightarrow PSL_2(\mathbb{Z}/N\mathbb{Z}) \longrightarrow 1
\end{array}$$

is a short exact sequence.

By convention, we take $\Gamma(1) = \Gamma$. The upper half-plane modulo the action of $\Gamma$ (written, by abuse of notation, $\mathcal{H}/\Gamma$) is a singular surface whose one-point compactification by the image of the point $i\infty$ under the stereographic projection is homeomorphic to the Riemann sphere. The completed nonsingular curve is denoted $X(1)$. Similarly, $\mathcal{H}/\Gamma(N)$ can be compactified by adding finitely many points, the cusps of $\Gamma(N)$, or the translates of $i\infty$ under a full set of coset representatives for $PSL_2(\mathbb{Z}/N\mathbb{Z})$ in $\Gamma$. In this case, the curve is denoted $X(N)$.

If $H$ is a finite index subgroup in $\Gamma$ the set of cusps, or translates of $i\infty$ under a full set of coset representative for $H$ in $\Gamma$, will hereafter be denoted $C(H)$. A finite index subgroup of $\Gamma$ defined by congruence conditions is called a congruence subgroup. The conductor of a congruence subgroup $H$ is the largest $N$ for which $\Gamma(N) \subseteq H$.

### 3.2 Modular functions of level $N$

A modular function for a congruence subgroup $\Gamma(N)$ is a function, $f(z) : \mathcal{H}^* \to \hat{\mathbb{C}}$ such that

1. $f$ is meromorphic on $\mathcal{H}$,

2. $f(A \circ z) = f(z)$ for all $A \in \Gamma(N)$ and $z \in \mathcal{H}^*$,
3. \( f(z) \) has an expansion at each of the cusps in the local variable \( q = e^{2\pi iz} \) of the form

\[
f(z) = \sum_{n \geq n_0} a_n q^n, \quad n \in \mathbb{Z}, a_{n_0} \neq 0.
\]

If \( f \) is modular for \( \Gamma(N) \), we say \( f \) has level \( N \). A modular function of level \( N \) descends to a well-defined holomorphic function on \( X(N) \). As before, if \( n_0 > 0 \) for all \( \alpha \in C(\Gamma(N)) \) then \( f(z) \) is called a cusp form for \( \Gamma(N) \).

### 3.3 The full tower of modular functions \( \mathcal{F} \)

The set of modular functions invariant under the full modular group \( \Gamma \) is, in fact, a function field of genus one and is generated over \( \mathbb{C} \) by the classical \( j \)-function,

\[
j(z) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + 864299970q^3 + O(q^4).
\]

We write \( \mathcal{F}_1 = \mathbb{Q}(j(z)) \) and note that \( \mathcal{F}_1 \) is the full field of rational functions on \( X(1) \) whose Fourier coefficients are rational. The \( j \)-function is normalized so that its \( q \)-expansion at \( i\infty \) (which is the only cusp of \( \mathcal{H}/\Gamma \)) has integral coefficients. Thus, it is reasonable to define the ring of integers in this field to be \( \mathbb{Z}[j] \).

Furthermore, the set of level \( N \) functions together with the \( N^{th} \) roots of unity generate a field extension of \( \mathcal{F}_1 \), denoted \( \mathcal{F}_N \), which is a finite Galois extension of \( \mathcal{F}_1 \) with Galois group \( \text{PGL}_2(\mathbb{Z}/\mathbb{N}\mathbb{Z}) \cong \Gamma/\Gamma(N) \times (\mathbb{Z}/\mathbb{N}\mathbb{Z})^\times \). The Galois action is given by writing \( \text{PGL}_2(\mathbb{Z}/\mathbb{N}\mathbb{Z}) \cong \text{PSL}_2(\mathbb{Z}/\mathbb{N}\mathbb{Z}) \times (\mathbb{Z}/\mathbb{N}\mathbb{Z})^\times \) and letting \( \text{PSL}_2(\mathbb{Z}/\mathbb{N}\mathbb{Z}) \) act as usual as fractional linear transformations on \( z \in \mathcal{H} \). A matrix in \( \text{PGL}_2(\mathbb{Z}/\mathbb{N}\mathbb{Z}) \) with determinant \( d \in (\mathbb{Z}/\mathbb{N}\mathbb{Z})^\times \) acts on a (not necessarily primitive) \( N^{th} \) root of unity \( \zeta \) via \( \sigma_d : \zeta \mapsto \zeta^d \). In other words, if \( f \) has Fourier expansion

\[
f(z) = \sum_{n \geq n_0} a_n q^n
\]

then elements of the form \( A = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \in \text{PGL}_2(\mathbb{Z}/\mathbb{N}\mathbb{Z}) \) act as follows:

\[
f(A \circ z) = \sum_{n \geq n_0} \sigma_d(a_n) q^n
\]
and more general elements $A$ of determinant $d$ act as

$$f(A \circ z) = \sum_{n \geq n_0} \sigma_d(a_n)(A' \circ q)^n$$

where $A' = \frac{1}{\sqrt{d}} A \in PSL_2(\mathbb{Z}/N\mathbb{Z})$.

Taking the integral closure of $\mathbb{Z}[j]$ in $\mathcal{F}_N$, we get a ring $R_N$, whose units, $U_N$, are the modular units of level $N$. It is not uncommon, however, to extend scalars to $\mathbb{C}$, that is, to study $U_N \otimes \mathbb{C} \subseteq R_N \otimes \mathbb{C}$. In this setting the set of functions with multiplicative inverses coincide precisely with the set of function whose divisor of zeros and poles is supported at the cusps of $X(N)$.

Finally, the compositum of the $\mathcal{F}_N$ over all $N$ is called the full tower of modular functions $\mathcal{F}$. The set of units $U$ in the full tower of modular functions is the direct limit of the $U_N$ with respect to the natural inclusion maps.

References

[DS05] F. Diamond and J. Shurman. A First Course in Modular Forms, volume 228 of Graduate Texts in Mathematics. Springer, 2005.

[KL81] D. Kubert and S. Lang. Modular Units, volume 224 of Die Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, 1981.

[Kub81] D. Kubert. The square root of the siegel group. Proc. London Math. Soc., 43(2):193–226, 1981.

[Sho74] B. Shoeneberg. Elliptic Modular Functions: an Introduction, volume 203 of Die Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, 1974. Translated from the German by J. R. Smart and E. A. Schwandt.