IDENTIFYING CONGRUENCE SUBGROUPS  
OF THE MODULAR GROUP

TIM HSU

(Communicated by Ronald M. Solomon)

Abstract. We exhibit a simple test (Theorem 2.4) for determining if a given (classical) modular subgroup is a congruence subgroup, and give a detailed description of its implementation (Theorem 3.1). In an appendix, we also describe a more “invariant” and arithmetic congruence test.

1. Notation

We describe (conjugacy classes of) subgroups $\Gamma \subset \text{PSL}_2(\mathbb{Z})$ in terms of permutation representations of $\text{PSL}_2(\mathbb{Z})$, following Millington [11, 12] and Atkin and Swinnerton-Dyer [1].

We recall that a conjugacy class of subgroups of $\text{PSL}_2(\mathbb{Z})$ is equivalent to a transitive permutation representation of $\text{PSL}_2(\mathbb{Z})$. Such a representation can be defined by transitive permutations $E$ and $V$ which satisfy the relations

$$1 = E^2 = V^3.$$  

The relations (1.1) are fulfilled by

$$E = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}. $$  

Alternately, such a representation can be defined by transitive permutations $L$ and $R$ which satisfy

$$1 = (LR^{-1}L)^2 = (R^{-1}L)^3,$$

with the relations being fulfilled by

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. $$  

One can also use permutations $E$ and $L$ such that

$$1 = E^2 = (L^{-1}E)^3,$$

with $E$ and $L$ corresponding to the indicated matrices in (1.2) and (1.4), respectively.
The various notations can be translated using the following conversion table:

\[(1.6) \quad E = LR^{-1}L, \quad V = R^{-1}L, \]
\[(1.7) \quad L = EV^{-1}, \quad R = EV^{-2}, \]
\[(1.8) \quad R = E^{-1}L^{-1}E. \]

**Example 1.1.** The permutations
\[(1.9) \quad E = (1 \ 2)(3 \ 4)(5 \ 6)(7 \ 8)(9 \ 10), \quad V = (1 \ 3 \ 5)(2 \ 7 \ 4)(6 \ 8 \ 9), \]
or, alternately,
\[(1.10) \quad L = (1 \ 4)(2 \ 5 \ 9 \ 10 \ 8)(3 \ 7 \ 6), \quad R = (1 \ 7 \ 9 \ 10 \ 6)(2 \ 3)(4 \ 5 \ 8), \]
describe a conjugacy class of subgroups of index 10 in \(\text{PSL}_2(\mathbb{Z})\).

**Remark 1.2.** Note that any concrete method of specifying a modular subgroup can easily be converted to permutation form. For instance, one way in which a modular subgroup \(\Gamma\) might be specified is by a list of generators. Such a list can be converted into permutations as follows: First, use the Euclidean algorithm to express each generator matrix as a product of \(L\)'s and \(R\)'s, where \(L\) and \(R\) are the elements in (1.4). Then enumerate the cosets of \(\Gamma\) in terms of these generators and presentation (1.3). This coset enumeration is easily converted into appropriate permutations \(L\) and \(R\). Similarly, any reasonable membership test for \(\Gamma\) can be used to enumerate the cosets of \(\Gamma\), with the same results as before.

## 2. Congruence Subgroups and the Level

We recall the following definitions.

**Definition 2.1.** \(\Gamma(N)\) is defined to be the group
\[(2.1) \quad \{ \gamma \in \text{PSL}_2(\mathbb{Z}) \mid \gamma \equiv \pm I \pmod{N} \}. \]
\(\Gamma(N)\) is the kernel of the natural projection from \(\text{PSL}_2(\mathbb{Z})\) to \(\text{SL}_2(\mathbb{Z}/N)/\{\pm I\}\).
We say that a modular subgroup \(\Gamma\) is a *congruence subgroup* if \(\Gamma\) contains \(\Gamma(N)\) for some integer \(N\). Otherwise, we say \(\Gamma\) is a *non-congruence subgroup*.

An important invariant of (conjugacy classes of) modular subgroups is the following.

**Definition 2.2.** The *level* of a modular subgroup \(\Gamma\), as specified by permutations \(L\) and \(R\), is defined to be the order of \(L\) (or the order of \(R\), since \(L\) is conjugate to \(R^{-1}\)).

We need the following result, sometimes known as Wohlfahrt’s Theorem (Wohlfahrt [13]).

**Theorem 2.3.** Let \(N\) be the level of a modular subgroup \(\Gamma\). \(\Gamma\) is a congruence subgroup if and only if it contains \(\Gamma(N)\).

**Proof.** This amounts to proving that, for congruence subgroups, our definition of the level is the same as the classical definition of the level. See Wohlfahrt [13]. □
Theorem 2.4. Let $\Gamma$ be a modular subgroup of level $N$, and let
\begin{equation}
(L, R | r_1, r_2, \ldots)
\end{equation}
be a presentation for $\text{SL}_2(\mathbb{Z}/N)/\{\pm I\}$ which is compatible with (1.4). Then $\Gamma$ is a congruence subgroup if and only if the representation of $\text{PSL}_2(\mathbb{Z})$ induced by $\Gamma$ respects the relations $\{r_i\}$.

Proof. From Theorem 2.3, we only need to check if $\Gamma$ contains $\Gamma(N)$. Now, since $\Gamma(N)$ is normal in $\text{PSL}_2(\mathbb{Z})$, $\Gamma$ contains $\Gamma(N)$ if and only if the normal kernel of $\Gamma$ contains $\Gamma(N)$. However, the normal kernel of $\Gamma$ is exactly the kernel of the representation induced by $\Gamma$, and since the relations $\{r_i\}$ generate $\Gamma(N)$ as their normal closure, the theorem follows.

Compare Magnus [9, Ch. III], Britto [4], Wohlfahrt [13], and Larcher [8]. Lang, Lim, and Tan [7] have also developed a congruence test; see the related paper Chan, Lang, Lim, and Tan [5].

Example 2.5. Suppose $\Gamma$ is the conjugacy class of subgroups specified by (1.10). Since $L$ has order 30, we need to use a presentation for $\text{SL}_2(\mathbb{Z}/30)/\{\pm I\}$. We find that $\text{SL}_2(\mathbb{Z}/30)/\{\pm I\}$ has a presentation with defining relations
\begin{equation}
1 = L^{30},
\end{equation}
\begin{equation}
1 = [L^2, R^{15}] = [L^3, R^{10}] = [L^5, R^6]
\end{equation}
in addition to the relations in (1.3). (The commutator $[x, y]$ is defined to be $x^{-1}y^{-1}xy$, so $1 = [x, y]$ means “$x$ commutes with $y$”. ) Only the commutator relations (2.4) need to be checked. However,
\begin{equation}
L^2 = (2 9 8 5 10)(3 6 7),
\end{equation}
which does not commute with
\begin{equation}
R^{15} = (2 3),
\end{equation}
so $\Gamma$ is a non-congruence subgroup. (It is worth mentioning that Larcher’s results also imply that $\Gamma$ is non-congruence, since $L$ does not contain a 30-cycle.)

Remark 2.6. The results in this section extend essentially verbatim to the Bianchi groups $\text{SL}_2(\mathcal{O}_d)$, where $\mathcal{O}_d$ is the ring of algebraic integers of an imaginary quadratic field $\mathbb{Q}[\sqrt{-d}]$ with class number 1. (See Fine [6] for more on the Bianchi groups.) However, for practical use, one needs a uniform presentation of $\text{SL}_2(\mathcal{O}_d)/\mathfrak{A}$ for any ideal of $\mathcal{O}_d$.

3. Implementation

To assure the reader that the procedure described by Theorem 2.4 is practical, we provide the following detailed algorithm. Suppose we are given a subgroup $\Gamma$ of finite index in $\text{PSL}_2(\mathbb{Z})$.

1. Describe $\Gamma$ in terms of permutations $L$ and $R$. If necessary, use conversion (1.7), conversion (1.8), or another similar conversion. (See also Remark 1.2.)
2. Let $N$ be the order of $L$, and let $N = em$, where $e$ is a power of 2 and $m$ is odd.
3. We have three cases:
   (a) $N$ is odd: $\Gamma$ is a congruence subgroup if and only if the relation
   \begin{equation}
   1 = (R^2 L^{-\frac{1}{2}})^3
   \end{equation}
is satisfied, where $\frac{1}{2}$ is the multiplicative inverse of $2 \mod N$.

(b) $N$ is a power of $2$: Let $S = L^{20}R^2L^{-4}R^{-1}$, where $\frac{1}{5}$ is the multiplicative inverse of $5 \mod N$. $\Gamma$ is a congruence subgroup if and only if the relations

\[(LR^{-1}L)^{-1}S(LR^{-1}L) = S^{-1},\]

\[S^{-1}RS = R^{25},\]

\[1 = (SR^2LR^{-1}L)^3\]

are satisfied.

(c) Both $e$ and $m$ are greater than $1$:

(i) Let $\frac{1}{2}$ be the multiplicative inverse of $2 \mod m$, and let $\frac{1}{5}$ be the multiplicative inverse of $5 \mod e$.

(ii) Let $c$ be the unique integer mod $N$ such that $c \equiv 0 \pmod{e}$ and $c \equiv 1 \pmod{m}$, and let $d$ be the unique integer mod $N$ such that $d \equiv 0 \pmod{m}$ and $d \equiv 1 \pmod{e}$.

(iii) Let $a = L^c$, $b = R^c$, $l = L^d$, $r = R^d$, and let $s = l^{20}r^2l^{-4}r^{-1}$.

(iv) $\Gamma$ is a congruence subgroup if and only if the relations

\[1 = [a, r],\]

\[1 = (ab^{-1}a)^4,\]

\[(ab^{-1}a)^2 = (b^{-1}a)^3,\]

\[(ab^{-1}a)^2 = (b^2a^{-1})^3,\]

\[(lr^{-1}l)^{-1}s(lr^{-1}l) = s^{-1},\]

\[s^{-1}rs = r^{25},\]

\[(lr^{-1}l)^2 = (sr^5lr^{-1}l)^3\]

are satisfied.

**Theorem 3.1.** The above procedure determines if $\Gamma$ is a congruence subgroup.

Before proving Theorem 3.1, we need an algebraic trick (Lemma 3.2) and some known results (Lemma 3.3, due to Behr and Mennicke [2]; and Lemma 3.4, due to Mennicke [10]).

**Lemma 3.2** (Braid trick). Let $x$ and $y$ be elements which generate a group $G$ and satisfy the relation

\[(xyx)^2 = (yx)^3.\]

Then the element $(xyx)^2 = (yx)^3$ is central in $G$. Furthermore,

\[xyx = yxy\]

and

\[(xyx)^{-1}x(xy) = y.\]

We call this the “braid trick” because (3.2) is the defining relation for the 3-string braid group.

**Proof.** The elements $X = xyx$ and $Y = yx$ also generate $G$, and the element $Z = (xyx)^2 = (yx)^3 = X^2 = Y^3$ commutes with both $X$ and $Y$, so $Z$ is central. (3.2) and (3.3) follow from cancellation in $xyxxyx = yxyxyx$. \(\square\)
Lemma 3.3. Let $m$ be an odd integer, and let $\frac{1}{2}$ be the multiplicative inverse of $2 \mod m$. $\text{SL}_2(\mathbb{Z}/m)$ is isomorphic to
\[ G = \left\langle a, b \mid \begin{array}{c}
1 = a^m, \\
1 = (ab^{-1}a)^4, \\
(ab^{-1}a)^2 = (b^{-1}a)^3, \\
(ab^{-1}a)^2 = (b^2a^{-\frac{1}{2}})^3
\end{array} \right\rangle. \] (3.4) (3.5) (3.6) (3.7)

Relations (3.4)–(3.7) are fulfilled by $a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ in $\text{SL}_2(\mathbb{Z}/m)$.

Proof. $G$ is equivalent to Behr and Mennicke’s presentation [2, (2.12)] by the following Tietze transformations. Add generators $A = b$ and $B = ab^{-1}a$. Applying the braid trick to (3.6), we get that $B^2$ is central, and from (3.2), we also get that (3.8)
\[ BA = b^{-1}a. \]

(3.8) implies that $a = ABA$, which means that we can eliminate $a$ and $b$.

Using (3.3), (3.8), and the centrality of $B^2$, we see that (3.4)–(3.6) become (3.9) (3.10)
\[ 1 = A^m = B^4, \]
\[ B^2 = (AB)^3, \]
so it remains to convert (3.7) to Behr and Mennicke’s form. However, applying (3.3), we have (3.11)
\[ B^2 = (b^2a^{-\frac{1}{2}})^3 = (A^2B^{-1}A\frac{1}{2}B)^3, \]
so, using $1 = B^8$ and the centrality of $B^2$,
\[ 1 = (A^2B^{-1}A\frac{1}{2}B)^3B^6 = (A^2BA\frac{1}{2}B)^3. \] \[ \square \]

Lemma 3.4. Let $e = 2^n$, let $\frac{1}{5}$ be the multiplicative inverse of $5 \mod e$, and let $s = t^{20}r^{-1}t^{-4}r^{-1}$. $\text{SL}_2(\mathbb{Z}/e)$ is isomorphic to
\[ G = \left\langle l, r \mid \begin{array}{c}
1 = l^n, \\
1 = (lr^{-1}l)^4, \\
(lr^{-1}l)^2 = (r^{-1}l)^3, \\
(lr^{-1}l)^{-1}s(lr^{-1}l) = s^{-1}, \\
s^{-1}rs = r^{25}, \\
(lr^{-1}l)^2 = (sr^{5}lr^{-1}l)^3
\end{array} \right\rangle. \] (3.13) (3.14) (3.15) (3.16) (3.17) (3.18)

Relations (3.13)–(3.18) are fulfilled by $l = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $r = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, and $s = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$ in $\text{SL}_2(\mathbb{Z}/e)$.
Proof. As the reader may verify, the relations (3.13)–(3.18) and \( s = l^{20}r^{4}l^{-4}r^{-1} \) are satisfied in \( \text{SL}_2(\mathbb{Z}/e) \), so it suffices to show that \( G \) is a homomorphic image of Mennicke’s presentation [10, p. 210]. Add generators \( A = r, B = lr^{-1}l, \) and \( T = s \). Applying the braid trick to (3.15), we get that \( B^2 \) is central, \( BA = r^{-1}l \), and \( l \) is conjugate to \( A^{-1} \). As in the proof of the previous lemma, we can then eliminate generators \( l \) and \( r \). Then (3.13), (3.14), (3.15), (3.16), (3.17), and (3.18) become Mennicke’s relations (X), (Y), (P), (Z), (Q), and (R), respectively.

For Lemma 3.5, we consider the following relations:

\[
\begin{align*}
  1 &= L^N, \\
  1 &= [a, r], \\
  1 &= [b, l], \\
  1 &= (ab^{-1}a)^4, \\
  (ab^{-1}a)^2 &= (b^{-1}a)^3, \\
  (ab^{-1}a)^3 &= (b^2a^{-2})^3, \\
  1 &= (lr^{-1}l)^4, \\
  (lr^{-1}l)^2 &= (r^{-1}l)^3, \\
  (lr^{-1}l)^3s(lr^{-1}l) &= s^{-1}, \\
  s^{-1}rs &= r^{25}, \\
  (lr^{-1}l)^2 &= (sr^5lr^{-1}l)^3.
\end{align*}
\]

All notation is as described in (2) and (3c)(i–iii) of the algorithm. Note that \( 1 = L^N \) implies that \( L = al \) and \( R = br \).

**Lemma 3.5.** \( \text{SL}_2(\mathbb{Z}/N) \) has a presentation with generators \( L \) and \( R \), and defining relations (3.19)–(3.29). The relations are fulfilled by \( L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) and \( R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \) in \( \text{SL}_2(\mathbb{Z}/N) \).

Proof. The Chinese Remainder Theorem implies that

\[
\text{SL}_2(\mathbb{Z}/N) \cong \text{SL}_2(\mathbb{Z}/m) \times \text{SL}_2(\mathbb{Z}/e).
\]

It also follows from the Chinese Remainder Theorem that, if \( L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) and \( R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \) in \( \text{SL}_2(\mathbb{Z}/N) \), the \( \text{SL}_2(\mathbb{Z}/m) \) factor is precisely \( \langle a, b \rangle \) and the \( \text{SL}_2(\mathbb{Z}/e) \) factor is precisely \( \langle l, r \rangle \). Therefore, the above relations are satisfied in \( \text{SL}_2(\mathbb{Z}/N) \).

On the other hand, since (3.19) implies (3.4) and (3.13), comparison with Lemmas 3.3 and 3.4 shows that the above presentation is the direct product of \( \text{SL}_2(\mathbb{Z}/m) \) and \( \text{SL}_2(\mathbb{Z}/e) \). The lemma follows.

**Proof of Theorem 3.1.** After steps 1 and 2 of the procedure, we know that the relations

\[
\begin{align*}
  1 &= L^N, \\
  1 &= (LR^{-1}L)^2, \\
  1 &= (R^{-1}L)^3.
\end{align*}
\]
must be satisfied. From Theorem 2.4, we see that if (3.31)–(3.33) and (A) (resp. (B), (C)) are defining relations for $SL_2(\mathbb{Z}/N)/\{\pm I\}$ when $N$ is odd (resp. $N$ is a power of 2, and $e$ and $m$ are greater than 1), then Theorem 3.1 follows. Comparing (A) and Lemma 3.3, with $a = L$ and $b = R$, and comparing (B) and Lemma 3.4, with $l = L$ and $r = R$, the first two cases follow easily, so it remains to check the third.

Comparing (C) and (3.19)–(3.29), we see that it is enough to show that given (3.31)–(3.33) and (3.19)–(3.29), the relations (3.21), (3.25), and (3.26) are redundant. First, (3.31), (3.32), (3.20), and (3.21) give us

\[ 1 = (LR^{-1}L)^4 = (alr^{-1}b^{-1}a)^4 = (ab^{-1}a)^4(lr^{-1}l)^4, \]

which means that (3.22) implies (3.25). Similarly, (3.31), (3.32), (3.33), (3.20), and (3.21) imply

\[ (LR^{-1}L)^2 = (R^{-1}L)^3, \]
\[ (ab^{-1}a)^2(lr^{-1}l)^2 = (b^{-1}a)^3(r^{-1}l)^3, \]

which means that (3.23) implies (3.26). Finally, since (3.32), (3.33), and the braid trick (3.3) imply that $L$ is conjugate to $R^{-1}$, we can eliminate (3.21), since it is implied by (3.20).

For hand calculations, and for further study, we note the following relations which occur in $SL_2(\mathbb{Z}/N)$:

\[ (SL_2) \quad Z = (LR^{-1}L)^2 = (R^{-1}L)^3, \quad 1 = Z^2, \]
\[ (level) \quad 1 = L^N = R^N, \]
\[ (ab \equiv 0 \pmod{N}) \quad 1 = [L^a, R^b], \]
\[ (ab \equiv -1 \pmod{N}) \quad (L^aR^b)^3 = Z, \]
\[ (ab \equiv -2 \pmod{N}) \quad (L^aR^b)^2 = Z. \]

It has been verified by coset enumeration that the relations $(SL_2)$, (level), and $(ab \equiv 0 \pmod{N})$ are defining relations when $N | 360$. This means that if the level $N$ divides 360, the congruence test reduces to checking that the relations $(ab \equiv 0 \pmod{N})$ are satisfied.

**Acknowledgements**

The author would like to thank J. H. Conway and the referee for many helpful comments and suggestions.

**Appendix A. An arithmetic congruence test**

In this appendix, we present an arithmetic and “invariant” congruence test which uses the Ihara modular group $SL_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
We begin by quoting the following result (Theorem A.1) of J. Mennicke [10]. (Note that Mennicke’s Schur multiplier calculation and subsequent argument require the repairs described in F.R. Beyl [3, §5], but the main result still holds.) Let $N$ be an integer, let $p$ be a prime not dividing $N$, let $R_N$ be the kernel in $\text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)$ resulting from reduction mod $N$, and let $Q_N$ be the normal closure of $L^N$ in $\text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)$.

**Theorem A.1.** $R_N = Q_N$. 

Let $\Gamma$ be a modular subgroup of level $N$ and index $m$ in $\text{SL}_2(\mathbb{Z})$. Consider the commutative diagram in Figure A.1.

![Commutative diagram for Theorem A.2](image)

**Figure A.1.** Commutative diagram for Theorem A.2

Here, $S_m$ is the symmetric group on $m$ objects (the cosets of $\Gamma$ in $\text{SL}_2(\mathbb{Z})$), $r$ is reduction mod $N$, $i$ is inclusion, and $\rho$ is the permutation representation of $\text{SL}_2(\mathbb{Z})$ induced by $\Gamma$. Note that $f_2$ exists if and only if $\Gamma$ is a congruence subgroup, and that such an $f_2$ is uniquely determined.

The setup in Figure A.1 provides us with an invariant congruence test.

**Theorem A.2.** In the notation of Figure A.1, a map $f_1$ exists if and only if $f_2$ exists. In other words, $\Gamma$ is congruence if and only if $\rho$ can be factored through inclusion in $\text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)$.

**Proof.** If $f_2$ exists, let $f_1 = f_2r$. Conversely, if $f_1$ exists, since $L^N$ is in the kernel of $\rho$, $L^N$ must be in the kernel of $f_1$, so in fact, $f_1$ is well defined on $\text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)$.

$$\text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)/Q_N = \text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)/R_N \cong \text{SL}_2(\mathbb{Z}/N),$$

which means that $f_1$ defines an appropriate map $f_2$. 

**Corollary A.3.** In Figure A.1, $f_1$ is determined uniquely if it exists.

One curious feature of Theorem A.2 is that if we know a given family of modular subgroups all have levels relatively prime to $p$, then we can handle all of them in a uniform manner. This is the principle behind Behr and Mennicke’s presentation of $\text{SL}_2(\mathbb{Z}/N)$ for $N$ odd, as these cases can be handled in $\text{SL}_2\left(\mathbb{Z}\left[\frac{1}{p}\right]\right)$.

We also note that if we fix the level $N$, then we can choose any $p$ not dividing $N$ to use in Theorem A.2. This leads to the following idea: For a given family of modular subgroups of level $N$, it seems plausible that one might be able to reduce the extensibility of $\rho$ to the question of whether there exists a $p$ which satisfies certain congruences mod $N$. Dirichlet’s theorem might then be used to find a $p$ which satisfies those congruences.
References


Department of Mathematics, Princeton University, Princeton, New Jersey 08544
E-mail address: timhsu@math.princeton.edu

Current address: Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109
E-mail address: timhsu@math.lsa.umich.edu

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use