Higher genus arithmetic-geometric means

Frazer Jarvis
Department of Pure Mathematics
University of Sheffield
Sheffield S3 7RH
England
a.f.jarvis@shef.ac.uk

(Running title: Higher genus AGMs)

Abstract

We suggest a notion of arithmetic-geometric mean (AGM) of four numbers, and relate the definition to duplication formulae for genus 2 theta functions, enabling us to give a generalisation of Gauss's work on the classical AGM, as described in [5].

Address: Department of Pure Mathematics, University of Sheffield, Sheffield S3 7RH, U.K.

Keywords: Arithmetic-geometric means, theta functions, Siegel modular forms

In this paper, we recall the notion of arithmetic-geometric mean (AGM), due to Gauss, and suggest a definition of an AGM of four numbers. According to [2], this definition was first used by Borchardt [1]. When the four numbers are real and positive, we prove several lemmas demonstrating that the behaviour of our AGM resembles that of the classical AGM. Just as for the classical AGM, however, there is some ambiguity in our definition, arising from the fact that there is no canonical square root at each step. In the classical case, Gauss was able to completely determine all possible limits of the AGM (see Theorem 1.3 below). This paper constitutes our first attempt to generalise Gauss's work to higher genus AGMs, and in the same way as Gauss (see [5]), we analyse the possible values that our 4-variable AGM can take, using the theory of theta functions in genus 2. Most of our analysis generalises to arbitrary genus, and we restrict attention to genus 2 largely for notational convenience.

1 Introduction: the AGM of Gauss

In this section we briefly review the theory of Gauss's AGM of two numbers. Suppose we are given two numbers, a and b. We can form the arithmetic mean $A = \frac{a+b}{2}$ and the geometric mean $B = (ab)^{\frac{1}{2}}$. By the AM-GM inequality, if a and b are real and positive, then $A \geq B$.

Gauss suggested iterating this process. Set $a_0 = a$ and $b_0 = b$, and define inductively sequences (a_n) and (b_n) as follows:

$$a_{n+1} = \frac{a_n + b_n}{2} \qquad \text{and} \qquad b_{n+1} = (a_n b_n)^{\frac{1}{2}}.$$

We refer the reader to [5] for complete details of this operation. Let us first remark that there is an ambiguity in this definition, arising from the choice of sign in the square root. This means that for all initial values, there are

an uncountable number of resulting sequences (a_n) and (b_n). However, as is pointed out in [5], all of these possible sequences converge; for any choice of initial values, the two sequences (a_n) and (b_n) converge to a common limit whatever choices are made for b_n at every step, which we call *a value* of the multi-valued AGM function. Only countably many of these values are non-zero.

1.1 The real and positive case

Suppose that $a_0 = a$ and $b_0 = b$ are real and positive. Then we may choose the positive value of the square root at every stage of the algorithm, to obtain a 'canonical' value $M(a, b)$ for the AGM, which is also real and positive. If $a = b$, then $M(a, b) = a$, so we assume a and b are different, and, without loss of generality, we suppose $a > b$. The arithmetic mean-geometric mean inequality implies that $a_n > b_n$ for all n.

It seems to have already been known to Legendre, and was certainly known to Gauss, that this value $M(a, b)$ is related to the value of a certain elliptic integral:

$$\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} = \frac{\pi}{2M(a, b)}.$$

One of the wonderful properties of the AGM is the speed at which the two sequences converge to the common limit; this enables one to perform extremely quick calculations of elliptic integrals. It is easy to prove that

$$a_{n+1} - b_{n+1} < C(a_n - b_n)^2$$

for some constant C depending on a and b, which gives quadratic convergence.

1.2 The complex case

If, however, the initial values a and b are arbitrary complex numbers, then there is no obvious canonical choice of the square root in the geometric mean. Nonetheless, Gauss was able to completely determine all possible values taken by the AGM: see [5] for a full proof of the theorem we state below.

In fact, there is a 'correct' choice of square root ([5], p.284):

Definition 1.1 Suppose a, $b \in \mathbb{C}^\times$, and $a \neq \pm b$. Then a square root B of ab is the *right choice* if $|A - B| \leq |A + B|$, and, when $|A - B| = |A + B|$, we also have $\mathrm{Im}(B/A) > 0$.

(If a and b are positive and real, this definition coincides with the 'natural' choice of positive square root.) Using this, we can make the following definition ([5], p.287):

Definition 1.2 The *simplest value* $M(a,b)$ of the AGM is defined to be the value taken by the AGM if, at every stage, the right choice of square root is taken.

(Thus this coincides with the previous definition of $M(a,b)$ when a and b are positive and real.) We now state the theorem, due to Gauss ([5], Theorem 2.2):

Theorem 1.3 *Suppose* $a,\ b \in \mathbb{C}^\times$ *satisfy* $a \neq \pm b$ *and* $|a| \geq |b|$. *Then* μ *is a value of the AGM of* a *and* b *if and only if there exist coprime integers* $c \equiv 0 \pmod 4$ *and* $d \equiv 1 \pmod 4$ *such that*

$$\frac{1}{\mu} = \frac{d}{M(a,b)} + \frac{ic}{M(a+b,a-b)}.$$

2 A four variable AGM

2.1 The definition

In the proof of Gauss's Theorem, much use is made of duplication formulae for theta functions. Motivated by the corresponding formulae for (Siegel) theta functions in genus 2 (see below), we propose the following definition:

Definition 2.1 Let a, b, c and d be given, and define

$$\begin{aligned} A &= \frac{1}{4}(a+b+c+d), \\ B &= \frac{1}{2}(\sqrt{ab}+\sqrt{cd}), \\ C &= \frac{1}{2}(\sqrt{ac}+\sqrt{bd}), \\ D &= \frac{1}{2}(\sqrt{ad}+\sqrt{bc}). \end{aligned}$$

Although there appear to be six choices of signs in the square root at every step, these are not independent; all are determined by the four choices $\sqrt{a}$, $\sqrt{b}$, $\sqrt{c}$ and $\sqrt{d}$.

Note also that a permutation of a, b, c and d fixes A and permutes B, C and D, giving a morphism $S_4 \longrightarrow S_3$; the kernel of this morphism is the

Klein 4-group, and indeed any double transposition of a, b, c and d leaves A, B, C and D fixed.

In the same way as above, we may iterate this process to find uncoutably many sequences (a_n), (b_n), (c_n) and (d_n). Exactly as in [5], we may ask which values may occur as the limits of these sequences. The same methods which Gauss used to solve the corresponding problem for the 2-variable AGM (Theorem 1.3) also work in the 4-variable case, as we explain below. The relation with Richelot's theory of abelian integrals (see [10]) may form the subject of future work.

As in the 2-variable AGM, we will treat the cases where all variables are real and positive first, just to indicate some of the basic properties of the function.

2.2 The real and positive case

We suppose a, b, c and d are real and positive. In this case, we suppose that all square roots are taken to be positive.

2.2.1 Easy lemmas

Lemma 2.2 *Suppose $a > b > c > d$. Then $A > B > C > D$.*

Proof. Firstly, note that the classical AM-GM inequality implies that $A > B$. Also,

$$\begin{aligned} B - C &= \frac{1}{2}(\sqrt{b} - \sqrt{c})(\sqrt{a} - \sqrt{d}), \\ B - D &= \frac{1}{2}(\sqrt{b} - \sqrt{d})(\sqrt{a} - \sqrt{c}), \\ C - D &= \frac{1}{2}(\sqrt{c} - \sqrt{d})(\sqrt{a} - \sqrt{b}), \end{aligned}$$

so that if $a > b > c > d$, it follows that $A > B > C > D$. □

In a similar way, we can read off precisely what conditions on a, b, c and d imply that two of A, B, C and D are equal.

We now iterate the process, to find sequences (a_n), (b_n), (c_n) and (d_n).

Lemma 2.3 *The sequences (a_n), (b_n), (c_n) and (d_n) have a common limit, $M(a, b, c, d)$, and convergence is quadratic.*

Proof. Note that the sequence (a_n) (resp. (d_n)) is monotonically decreasing (resp. increasing), and is bounded below (resp. above) by d_0 (resp. a_0). Both these sequences therefore converge.

As both sequences (a_n) and (d_n) converge, we see from the definition of a_{n+1} and d_{n+1} that the sequences $(b_n + c_n)$ and $(\sqrt{b_n c_n})$ converge. But now it follows easily that the sequences $(\sqrt{b_n})$ and $(\sqrt{c_n})$, and therefore the sequences (b_n) and (c_n), converge. Suppose the limits of the sequences (a_n), (b_n), (c_n) and (d_n) are α, β, γ and δ. Certainly $\alpha \geq \beta \geq \gamma \geq \delta$. Then, from the definition of a_{n+1}, we see that $\alpha = \frac{1}{4}[\alpha+\beta+\gamma+\delta]$, so that $\alpha = \beta = \gamma = \delta$. To deduce quadratic convergence, note:

$$\begin{aligned} a_{n+1} - d_{n+1} &= \frac{1}{4}\left[(\sqrt{a_n} - \sqrt{d_n})^2 + (\sqrt{b_n} - \sqrt{c_n})^2\right] \\ &\leq \frac{1}{2}(\sqrt{a_n} - \sqrt{d_n})^2 \\ &= \frac{1}{2}\left[\frac{a_n - d_n}{\sqrt{a_n} + \sqrt{d_n}}\right]^2 \\ &\leq \frac{1}{2}\left[\frac{a_n - d_n}{2\sqrt{d_0}}\right]^2 \\ &= \frac{1}{8d_0}(a_n - d_n)^2. \end{aligned}$$

It follows that there is a constant C such that $a_{n+1} - d_{n+1} \leq C(a_n - d_n)^2$, so all four sequences converge quadratically to the common limit $M(a, b, c, d)$. □

Note that $M(a, a, a, a) = a$ and $M(\lambda a, \lambda b, \lambda c, \lambda d) = \lambda M(a, b, c, d)$.

2.2.2 The inverse map

We sketch the easy verification that, given a quadruple $\{A, B, C, D\}$ with $A > B > C > D$, there exist, in general, two possible preimages $\{a, b, c, d\}$ with $a > b > c > d$ under a single step of the algorithm. This is reflected in the observation that the Richelot isogeny ([8], [10]) between two hyperelliptic curves of genus 2 is derived from a 2-2 correspondence—see [3] for more details.

Observe that

$$\begin{aligned} A + B &= \frac{1}{4}[(\sqrt{a} + \sqrt{b})^2 + (\sqrt{c} + \sqrt{d})^2], \\ C + D &= \frac{1}{2}[(\sqrt{a} + \sqrt{b})(\sqrt{c} + \sqrt{d})]. \end{aligned}$$

It follows that

$$\begin{aligned}\sqrt{a}+\sqrt{b} &= \sqrt{A+B+C+D}+\sqrt{A+B-C-D},\\ \sqrt{c}+\sqrt{d} &= \sqrt{A+B+C+D}-\sqrt{A+B-C-D}.\end{aligned}$$

as certainly $\sqrt{c}+\sqrt{d} < \sqrt{a}+\sqrt{b}$.

In the same way,

$$\begin{aligned}A-B &= \frac{1}{4}[(\sqrt{a}-\sqrt{b})^2+(\sqrt{c}-\sqrt{d})^2],\\ C-D &= \frac{1}{2}[(\sqrt{a}-\sqrt{b})(\sqrt{c}-\sqrt{d})].\end{aligned}$$

(Note that $A-B \geq C-D$ by the AM-GM inequality.) It follows that

$$\begin{aligned}&\{\sqrt{a}-\sqrt{b},\sqrt{c}-\sqrt{d}\}\\ =\ &\{\sqrt{A-B+C-D}+\sqrt{A-B-C+D},\sqrt{A-B+C-D}-\sqrt{A-B-C+D}\}\end{aligned}$$

and so there are two possibilities, depending whether $\sqrt{a}-\sqrt{b}$ is to be taken larger or smaller than $\sqrt{c}-\sqrt{d}$.

This concludes the proof that our 4-variable AGM is a 2-to-1 map, at least when restricted to positive real numbers.

3 Theta functions and the AGM

As is noted in [3] and in [5], Gauss's 2-variable AGM has strong connections with the classical theta functions. Our definition of a 4-variable AGM was in turn motivated by the duplication formulae for theta functions in genus 2. In this section, however, we will work in arbitrary genus, and will later restrict to genus 2 largely for notational reasons. We use the standard notation for Siegel modular forms of genus g; $\mathbb{H}_g$ will denote the Siegel upper-half space of genus g.

3.1 Theta functions

Next, we recall the definition of certain theta functions in genus g.

Let $\mathbf{a}$, $\mathbf{b} \in (\frac{1}{2}\mathbb{Z}/\mathbb{Z})^g$, and define

$$\theta\left[\begin{array}{c}\mathbf{a}\\ \mathbf{b}\end{array}\right](\Omega) = \sum_{\mathbf{n}\in\mathbb{Z}^g} \exp(\pi i^t(\mathbf{n}+\mathbf{a})\Omega(\mathbf{n}+\mathbf{a})+2\pi i^t(\mathbf{n}+\mathbf{a})\mathbf{b}).$$

We will be particularly concerned with those theta series where $\mathbf{a} = 0$. For simplicity, if $\mathbf{b} \in (\frac{1}{2}\mathbb{Z}/\mathbb{Z})^g$, we write

$$\theta_{\mathbf{b}} = \theta \begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix}.$$

Theta series are examples of Siegel modular forms of weight $\frac{1}{2}$ (see [9], p.189).

Our purpose is slightly different: we study the collection of theta functions, and show that they embed certain quotients of $\mathbb{H}_g$ into projective space.

Let $\mathbf{b}_0 = 0, \mathbf{b}_1, \ldots, \mathbf{b}_{2^g-1}$ be some enumeration of the vectors in $(\frac{1}{2}\mathbb{Z}/\mathbb{Z})^g$.

Definition 3.1 For $\Omega \in \mathbb{H}_g$, and $n \geq 1$, define the function $\Theta^{(n)}(\Omega)$ by

$$\Theta^{(n)}(\Omega) = [\theta^n_{\mathbf{b}_0}(\Omega) : \theta^n_{\mathbf{b}_1}(\Omega) : \cdots : \theta^n_{\mathbf{b}_{2^g-1}}(\Omega)] \in \mathbb{P}^{2^g-1}(\mathbb{C}).$$

Note that as the functions $\theta_{\mathbf{b}}$ do not vanish simultaneously on $\mathbb{H}_g$, this function does indeed take values in $\mathbb{P}^{2^g-1}(\mathbb{C})$ (see also [6]).

Definition 3.2 We define $\Gamma^{(n)}$ to be the group

$$\{\gamma \in \mathrm{Sp}_{2g}(\mathbb{Z}) | \Theta^{(n)}(\gamma\Omega) = \Theta^{(n)}(\Omega) \text{ for all } \Omega \in \mathbb{H}_g\}.$$

Thus $\Theta^{(n)} : \Gamma^{(n)} \backslash \mathbb{H}_g \hookrightarrow \mathbb{P}^{2^g-1}(\mathbb{C})$.

We first record the following special case of [9], p.190.

Corollary 3.3

$$\theta \begin{bmatrix} -C\mathbf{b} \\ A\mathbf{b} \end{bmatrix} \left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} \Omega \right) =$$
$$\zeta . \det(C\Omega + D)^{\frac{1}{2}} \exp(-\pi i {}^t\mathbf{b}^t AC\mathbf{b}) \theta \begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix} (\Omega)$$

for $\Omega \in \mathbb{H}_g$ *and* $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_{1,2}$*, the group of integral symplectic* $2g \times 2g$*-matrices such that* $\mathrm{diag}({}^tAC)$ *and* $\mathrm{diag}({}^tBD)$ *are even.*

We first explicitly compute the groups $\Gamma^{(i)}$ in the cases $i = 1, 2$.

Theorem 3.4 *We have:*

1. $\Gamma^{(1)} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{Z}) \middle| \begin{array}{l} A \equiv D \equiv I_g \text{ (mod 2)}, C \equiv 0 \text{ (mod 4)} \\ \mathrm{diag}(B) \text{ is even}, \mathrm{diag}(C) \equiv 0 \text{ (mod 8)} \end{array} \right\}$.

2. $\Gamma^{(2)} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{Z}) \middle| \begin{array}{l} A \equiv D \equiv I_g \text{ (mod 2)}, C \equiv 0 \text{ (mod 2)} \\ \mathrm{diag}(B) \text{ is even}, \mathrm{diag}(C) \equiv 0 \text{ (mod 4)} \end{array} \right\}$.

Proof. In order that each theta function $\theta_{\mathbf{b}}$ should be preserved by the action of $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, rather than sent to a theta function with different characteristics, we require from Corollary 3.3 that, for all $\mathbf{b}$,

- $C\mathbf{b} \in \mathbb{Z}^g$,
- $(A - I_g)\mathbf{b} \in \mathbb{Z}^g$,

which imply that $C \equiv 0 \pmod 2$ and $A \equiv I_g \pmod 2$. As $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{Z})$, we also deduce that $D \equiv I_g \pmod 2$.

Then each group $\Gamma^{(n)}$ is contained in the group

$$\left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{Z}) \middle| \begin{array}{l} A \equiv D \equiv I_g \pmod 2, C \equiv 0 \pmod 2 \\ \mathrm{diag}(B) \text{ is even} \end{array} \right\}.$$

We just do the case $n = 2$ which will be most important for us. The case $n = 1$ is similar. Then a matrix $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma^{(2)}$ if and only if there is a constant α such that, for each $\mathbf{b}$,

$$\theta^2_{\mathbf{b}}(\gamma\Omega) = \alpha\theta^2_{\mathbf{b}}(\Omega)$$

for all Ω. By considering $\mathbf{b} = 0$, we see that $\alpha = \zeta^2 . \det(C\Omega + D)$. It follows that, in order that this constant be the same for all $\mathbf{b}$, we need

$$\exp(-2\pi i {}^t\mathbf{b}{}^t AC\mathbf{b}) = 1$$

for all $\mathbf{b}$. This is equivalent to the condition that ${}^t\mathbf{b}{}^t AC\mathbf{b} \in \mathbb{Z}$ for all $\mathbf{b}$. Equivalently, we require that ${}^t\mathbf{x}{}^t AC\mathbf{x} \in 4\mathbb{Z}$ for all $\mathbf{x} \in \mathbb{Z}^g$. This is equivalent to the pair of conditions:

- ${}^t AC \equiv 0 \pmod 2$,
- $\mathrm{diag}({}^t AC) \equiv 0 \pmod 4$.

As $A \equiv I_g \pmod 2$, the first condition is automatic, and then the second is equivalent to $\mathrm{diag}(C) \equiv 0 \pmod 4$, as required. □

(Note that none of the groups $\Gamma^{(n)}$ act freely on $\mathbb{H}_g$, as all of them contain $-I_{2g}$.)

Definition 3.5 Define the following subset of $\mathbb{H}_g$:

$$\mathcal{F}^{(2)} = \left\{ \Omega \in \mathbb{H}_g \,\middle|\, |\det(C\Omega + D)| \geq 1 \text{ for all } \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma^{(2)} \right\}.$$

Note that every point of $\mathbb{H}_g$ may be mapped to some point in $\mathcal{F}^{(2)}$ under the action of $\Gamma^{(2)}$. Indeed, $\mathcal{F}^{(2)}$ contains infinitely many copies of fundamental domains for $\Gamma^{(2)}$ in $\mathbb{H}_g$.

If γ denotes the matrix $\begin{pmatrix} \sqrt{2}I_g & 0 \\ 0 & \frac{1}{\sqrt{2}}I_g \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{R})$, so that $\gamma\Omega = 2\Omega$, then composition of the map $\Theta^{(1)}$ with the isomorphism

$$\begin{array}{rcl} \alpha : \Gamma\backslash\mathbb{H}_g & \xrightarrow{\sim} & \gamma\Gamma\gamma^{-1}\backslash\mathbb{H}_g \\ \Omega & \mapsto & 2\Omega \end{array}$$

induces the following embedding:

$$\Theta^{(1)} \circ \alpha : \Gamma_{2,4}\backslash\mathbb{H}_g \hookrightarrow \mathbb{P}^{2^g-1}(\mathbb{C})$$

where

$$\Gamma_{2,4} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma(2) \,\middle|\, \mathrm{diag}(B) \equiv \mathrm{diag}(C) \equiv 0 \pmod 4 \right\},$$

and $\Gamma(2)$ denotes the principal congruence subgroup of level 2 in $\mathrm{Sp}_{2g}(\mathbb{Z})$ (this gives an explicit verification of [6], p.181, where the composition of $\Theta^{(1)}$ and α is the map denoted $\mathrm{Th}^{(2)}$).

3.2 Duplication formulae

Now we fix $g = 2$. However, similar results exist for higher genus Siegel theta functions.

Let $\Omega \in \mathbb{H}_g$. We define, for $(n_1, n_2) \in \mathbb{Z}^2$,

$$q_\Omega \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = \exp(\pi i (n_1 \; n_2)\Omega \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}).$$

We define the following theta functions:

$$\begin{aligned}
\theta_{00}(\Omega) &= \sum_{\binom{n_1}{n_2}\in\mathbb{Z}^2} q_\Omega \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}, \\
\theta_{01}(\Omega) &= \sum_{\binom{n_1}{n_2}\in\mathbb{Z}^2} (-1)^{n_2} q_\Omega \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}, \\
\theta_{10}(\Omega) &= \sum_{\binom{n_1}{n_2}\in\mathbb{Z}^2} (-1)^{n_1} q_\Omega \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}, \\
\theta_{11}(\Omega) &= \sum_{\binom{n_1}{n_2}\in\mathbb{Z}^2} (-1)^{n_1+n_2} q_\Omega \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}.
\end{aligned}$$

In the notation of the previous section, these are $\theta_{(0,0)}$, $\theta_{(0,\frac{1}{2})}$, $\theta_{(\frac{1}{2},0)}$ and $\theta_{(\frac{1}{2},\frac{1}{2})}$ respectively.

Definition 3.6 For $\Omega \in \mathbb{H}_2$, and $n \geq 1$, define the function $\Theta^{(n)}(\Omega)$ by

$$\Theta^{(2)}(\Omega) = [\theta_{00}^2(\Omega) : \theta_{01}^2(\Omega) : \theta_{10}^2(\Omega) : \theta_{11}^2(\Omega)] \in \mathbb{P}^3(\mathbb{C}).$$

We first state the duplication formulae, which motivated our definition of the 4-variable AGM.

Theorem 3.7 *Let $\Omega \in \mathbb{H}_2$. Then we have the following duplication formulae:*

$$\begin{aligned}
\Omega_{00}(2\Omega)^2 &= \frac{1}{4}[\theta_{00}(\Omega)^2 + \theta_{01}(\Omega)^2 + \theta_{10}(\Omega)^2 + \theta_{11}(\Omega)^2], \\
\Omega_{01}(2\Omega)^2 &= \frac{1}{2}[\theta_{00}(\Omega)\theta_{01}(\Omega) + \theta_{10}(\Omega)\theta_{11}(\Omega)], \\
\Omega_{10}(2\Omega)^2 &= \frac{1}{2}[\theta_{00}(\Omega)\theta_{10}(\Omega) + \theta_{01}(\Omega)\theta_{11}(\Omega)], \\
\Omega_{11}(2\Omega)^2 &= \frac{1}{2}[\theta_{00}(\Omega)\theta_{11}(\Omega) + \theta_{01}(\Omega)\theta_{10}(\Omega)].
\end{aligned}$$

Proof. Note that

$${}^t(\mathbf{n}+\mathbf{m})\Omega(\mathbf{n}+\mathbf{m}) + {}^t(\mathbf{n}-\mathbf{m})\Omega(\mathbf{n}-\mathbf{m}) = 2({}^t\mathbf{n}\Omega\mathbf{n} + {}^t\mathbf{m}\Omega\mathbf{m}).$$

It follows that

$$q_\Omega(\mathbf{n}+\mathbf{m}).q_\Omega(\mathbf{n}-\mathbf{m}) = q_{2\Omega}(\mathbf{n})q_{2\Omega}(\mathbf{m}).$$

Then

$$\begin{aligned}
& \theta_{00}(\Omega)^2 + \theta_{01}(\Omega)^2 + \theta_{10}(\Omega)^2 + \theta_{11}(\Omega)^2 \\
= & \sum_{\mathbf{m}\in\mathbb{Z}^2} \sum_{\mathbf{n}\in\mathbb{Z}^2} q_\Omega(\mathbf{m}) q_\Omega(\mathbf{n}) [1 + (-1)^{m_2+n_2} + (-1)^{m_1+n_1} + (-1)^{m_1+m_2+n_1+n_2}] \\
= & 4 \sum_{\mathbf{m}\in\mathbb{Z}^2} \sum_{\substack{\mathbf{n}\in\mathbb{Z}^2 \\ \mathbf{m}\equiv\mathbf{n}(\mathrm{mod}2)}} q_\Omega(\mathbf{m}) q_\Omega(\mathbf{n}) \\
= & 4 \sum_{\mathbf{m}'\in\mathbb{Z}^2} \sum_{\mathbf{n}'\in\mathbb{Z}^2} q_\Omega(\mathbf{m}'+\mathbf{n}') q_\Omega(\mathbf{m}'-\mathbf{n}') \\
= & 4 \sum_{\mathbf{m}'\in\mathbb{Z}^2} \sum_{\mathbf{n}'\in\mathbb{Z}^2} q_{2\Omega}(\mathbf{m}') q_{2\Omega}(\mathbf{n}') \\
= & 4\theta_{00}(2\Omega)^2.
\end{aligned}$$

The remaining relations are similar. □

The duplication formulae are, of course, special cases of more general addition formulae for theta functions.

Corollary 3.8 *The set*

$$\{\theta_{00}(2\Omega)^2, \theta_{01}(2\Omega)^2, \theta_{10}(2\Omega)^2, \theta_{11}(2\Omega)^2\}$$

may be derived from the set

$$\{\theta_{00}(\Omega)^2, \theta_{01}(\Omega)^2, \theta_{10}(\Omega)^2, \theta_{11}(\Omega)^2\}$$

by applying the AGM process.

Proof. This is clear from the duplication formulae. □

So doubling of the period matrix $\Omega \in \mathbb{H}_2$ corresponds to a possible application of the AGM.

As an immediate corollary, we have the following lemma (compare [5], Lemma 2.3):

Lemma 3.9 *Let a, b, c and d be in $\mathbb{C}^\times$, and suppose there exists $\Omega \in \mathbb{H}_2$ such that*

$$\Theta^{(2)}(\Omega) = [a : b : c : d].$$

Let

$$\mu = \frac{a}{\theta_{00}^2(\Omega)} = \frac{b}{\theta_{01}^2(\Omega)} = \frac{c}{\theta_{10}^2(\Omega)} = \frac{d}{\theta_{11}^2(\Omega)}.$$

If

$$\begin{aligned} a_n &= \mu\theta_{00}^2(2^n\Omega), \\ b_n &= \mu\theta_{01}^2(2^n\Omega), \\ c_n &= \mu\theta_{10}^2(2^n\Omega), \\ d_n &= \mu\theta_{11}^2(2^n\Omega), \end{aligned}$$

then the sequence of quadruples (a_n, b_n, c_n, d_n) may be derived from the AGM process, and the common limit is given by μ.

Proof. Clearly $(a_0, b_0, c_0, d_0) = (a, b, c, d)$; the previous corollary implies that the sequence (a_n, b_n, c_n, d_n) may be given by the AGM. As $\Omega \in \mathbb{H}_2$, its imaginary part is totally positive. It follows that $q_{2^n\Omega}\begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \longrightarrow 0$ as $n \longrightarrow \infty$, for fixed n_1 and n_2. Then $\theta_{ij}(2^n\Omega) \longrightarrow 1$ as $n \longrightarrow \infty$ for i, $j \in \{0, 1\}$, so that the common limit is visibly μ. □

3.3 Other possible values for the AGM

In the previous section, we observed that $\Theta^{(2)}(2\Omega)$ was a possible answer after applying one step of the AGM process to $\Theta^{(2)}(\Omega)$. In this section, we compute all other possibilities.

The answer $\Theta^{(2)}(2\Omega)$ was obtained by choosing the square root $\theta_{ij}(\Omega)$ of $\theta_{ij}^2(\Omega)$, rather than $-\theta_{ij}(\Omega)$.

Using slightly more general duplication formulae (see, for example, [8], p.38), together with Corollary 3.3 above, we find easily the following:

Proposition 3.10 *Suppose $[a : b : c : d] = \Theta^{(2)}(\Omega)$. Then all possible results of the AGM map also lie in the image of $\Theta^{(2)}$.*

Sketch proof. After scaling, we may suppose without loss of generality that $a = \theta_{00}^2(\Omega)$ etc. Recall that

$$\begin{aligned} A &= \frac{1}{4}(a + b + c + d), \\ B &= \frac{1}{2}(\sqrt{a}\sqrt{b} + \sqrt{c}\sqrt{d}), \\ C &= \frac{1}{2}(\sqrt{a}\sqrt{c} + \sqrt{b}\sqrt{d}), \\ D &= \frac{1}{2}(\sqrt{a}\sqrt{d} + \sqrt{b}\sqrt{c}). \end{aligned}$$

We just consider one case; the others are similar. Suppose we choose the roots $\theta_{00}(\Omega)$, $\theta_{01}(\Omega)$, $\theta_{10}(\Omega)$ and $-\theta_{11}(\Omega)$. Then using the duplication formulae of [8], we see that the values of (A, B, C, D) are given by

$$(\theta_{00}^2(2\Omega), \theta^2 \left[\begin{smallmatrix} (\frac{1}{2}, 0) \\ (0, \frac{1}{2}) \end{smallmatrix} \right] (2\Omega), \theta^2 \left[\begin{smallmatrix} (0, \frac{1}{2}) \\ (\frac{1}{2}, 0) \end{smallmatrix} \right] (2\Omega), -\theta^2 \left[\begin{smallmatrix} (\frac{1}{2}, \frac{1}{2}) \\ (\frac{1}{2}, \frac{1}{2}) \end{smallmatrix} \right] (2\Omega)).$$

Using Corollary 3.3, we see that now

$$[A : B : C : D] = \Theta^{(2)} \left(\left(\begin{pmatrix} I_2 & 0 \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & I_2 \end{pmatrix} (2\Omega) \right).$$

In the same way, all of the possible images are of the form $\Theta^{(2)} \left(\left(\begin{pmatrix} I_2 & 0 \\ C & I_2 \end{pmatrix} (2\Omega) \right)$, where C runs over the set

$$0, \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

as claimed. □

We now make a generalisation of Definitions 1.1 and 1.2.

Definition 3.11 Given a quadruple (a, b, c, d) such that there exists $\Omega \in \mathcal{F}^{(2)}$ with $\Theta^{(2)}(\Omega) = [a : b : c : d]$, we say that $(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d})$ is a *right choice* of square roots of (a, b, c, d) if $\Theta^{(1)}(\Omega) = [\sqrt{a} : \sqrt{b} : \sqrt{c} : \sqrt{d}]$. We say a sequence of quadruples (a_n, b_n, c_n, d_n) derived from the AGM algorithm is a *good sequence* if a right choice of square roots is taken at all but finitely many steps. Finally, a *simplest value* of the AGM is a value of the AGM corresponding to a sequence in which a right choice of square root is taken at each step.

Note that good sequences converge: the analysis above shows that the right choice of square root corresponds to doubling the period matrix—Lemma 3.9 now implies that the sequences converge to a common limit. We say that values of the AGM of good sequences are *good values* of the AGM.

Of course, right choices are only defined up to sign, but note also that in any case the right choice is not uniquely defined in general—if Ω lies on the boundary of $\mathcal{F}^{(2)}$, then there may be several possible right choices of square root. It would be easy to distinguish a unique one by giving an additional criterion, but this does not seem to us to be useful at the moment. The main weakness in our results is that we have so far been unable to translate

this definition of 'right' choice of square roots into one solely involving a, b, c and d. We guess, however, that the right choice of square roots will be the one where $|\sqrt{a}+\sqrt{b}+\sqrt{c}+\sqrt{d}|$ is maximised. We also conjecture (less confidently) that all sequences which converge to a non-zero limit are good sequences.

Lemma 3.12 *For all $\Omega \in \mathbb{H}_2$, there exists N such that $2^n\Omega \in \mathcal{F}^{(2)}$ for all $n \geq N$.*

Proof. By [7], §3, Lemma 1, there are only finitely many pairs of bottom rows (C, D) (up to multiplication by a unimodular matrix) such that $|\det(C\Omega + D)| \leq 1$. Choose N such that $2^N \nmid C$ for every such $C \neq 0$. Then $|\det(C.2^n\Omega + D)| > 1$ for all $n \geq N$, and so $2^n\Omega \in \mathcal{F}^{(2)}$. □

It follows that all sequences obtained by successive period doubling as in Lemma 3.9 are good sequences. The same proof shows that in all such sequences, there is eventually a unique right choice (up to sign) of square roots—eventually $2^n\Omega$ will lie in the interior of $\mathcal{F}^{(2)}$.

3.4 The inverse map

Next, we explain that, given a point in the image of $\Theta^{(2)}$, all of its possible preimages are also in the image of $\Theta^{(2)}$, and we make all these explicit.

Fix $\Omega \in \mathbb{H}_2$, and consider the point $\Theta^{(2)}(2\Omega)$. Let

$$\begin{aligned} A &= \theta_{00}^2(2\Omega), \\ B &= \theta_{01}^2(2\Omega), \\ C &= \theta_{10}^2(2\Omega), \\ D &= \theta_{11}^2(2\Omega). \end{aligned}$$

Proposition 3.13 *If (a, b, c, d) is a quadruple mapping to (A, B, C, D) under the AGM process, then $[a : b : c : d]$ is in the image of $\Theta^{(2)}$.*

Proof. Note that

$$\begin{aligned}
A+B+C+D &= \left[\frac{\theta_{00}(\Omega)+\theta_{01}(\Omega)+\theta_{10}(\Omega)+\theta_{11}(\Omega)}{2}\right]^2,\\
A+B-C-D &= \left[\frac{\theta_{00}(\Omega)+\theta_{01}(\Omega)-\theta_{10}(\Omega)-\theta_{11}(\Omega)}{2}\right]^2,\\
A-B+C-D &= \left[\frac{\theta_{00}(\Omega)-\theta_{01}(\Omega)+\theta_{10}(\Omega)-\theta_{11}(\Omega)}{2}\right]^2,\\
A-B-C+D &= \left[\frac{\theta_{00}(\Omega)-\theta_{01}(\Omega)-\theta_{10}(\Omega)+\theta_{11}(\Omega)}{2}\right]^2,
\end{aligned}$$

using the duplication formulae. Considering all possible square roots of these, and using the same method as §2.2, we can write down all possibilities for $(\sqrt{a},\sqrt{b},\sqrt{c},\sqrt{d})$. Just as in the real case, we find that there are essentially two possibilities, together with their permutations under the Klein 4-group, making eight possible preimages in total. It is easy to check that they are given by $\Theta^{(2)}(\Omega+B)$ where B runs over the set

$$0,\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&\frac{1}{2}\\\frac{1}{2}&0\end{pmatrix},\begin{pmatrix}0&\frac{1}{2}\\\frac{1}{2}&1\end{pmatrix},\begin{pmatrix}1&\frac{1}{2}\\\frac{1}{2}&0\end{pmatrix},\begin{pmatrix}1&\frac{1}{2}\\\frac{1}{2}&1\end{pmatrix},$$

as claimed. □

3.5 Theta functions and moduli spaces

Given initial values, a, b, c and d, we would like to interpret these as values of the theta functions as in the previous section. That is, we would like to find $\Omega\in\mathbb{H}_2$ such that

$$\Theta^{(2)}(\Omega)=[a:b:c:d].$$

Recall that the map $\Theta^{(2)}$ induces an embedding

$$\Theta^{(2)}:\Gamma^{(2)}\backslash\mathbb{H}_2\hookrightarrow\mathbb{P}^3(\mathbb{C})$$

embedding the quotient $\Gamma^{(2)}\backslash\mathbb{H}_2$ as an open subset of projective space. We may compactify the quotient using the Satake compactification (see [4]). The precise details of this procedure are rather complicated, and we content ourselves with a few observations.

The compactification is of the form $\Gamma^{(2)}\backslash\mathbb{H}_2^*$, for some analytic space $\mathbb{H}_2^*$, formed by adding to the boundary of $\mathbb{H}_2$ some copies of the usual compactified upper half complex plane, $\mathbb{H}_1^*$. The action of $\Gamma^{(2)}$ extends to the boundary

components, and the resulting quotient $\Gamma^{(2)}\backslash\mathbb{H}_2^*$ looks like $\Gamma^{(2)}\backslash\mathbb{H}_2$ with a finite number of modular curves (i.e., isomorphic to quotients of $\mathbb{H}_1^*$ by subgroups of finite index in $\mathrm{SL}_2(\mathbb{Z})$). In particular, the boundary components have codimension 2 in the compactification.

Then $\Theta^{(2)}$ extends to the compactification $\overline{\Gamma^{(2)}\backslash\mathbb{H}_2}$, and $\Theta^{(2)}(\overline{\Gamma^{(2)}\backslash\mathbb{H}_2})$ must be all of $\mathbb{P}^3(\mathbb{C})$.

It follows that the image of $\Theta^{(2)}$ on the open variety $\Gamma^{(2)}\backslash\mathbb{H}_2$ is all of $\mathbb{P}^3(\mathbb{C})$ apart from a finite number of curves, so the complement of the image of $\Theta^{(2)}$ in $\mathbb{P}^3(\mathbb{C})$ has codimension 2.

3.6 Values of the 4-variable AGM

In this section we prove our main theorem, which should be viewed as a partial generalisation of Gauss's Theorem 1.3.

Theorem 3.14 *For almost all quadruples (a,b,c,d), there exists $\Omega \in \mathbb{H}_2$ such that the good values of the AGM are precisely the values of the following set:*

$$\left\{ \frac{a}{\theta_{00}^2(M(\Omega))} \middle| \, M \in \Gamma^{(2)} \right\}.$$

Further, the simplest values of the AGM are those of maximum modulus.

Proof. In the previous section, we observed that there is an embedding

$$\Theta^{(2)} : \Gamma^{(2)}\backslash\mathbb{H}_2 \hookrightarrow \mathbb{P}^3(\mathbb{C}),$$

embedding the quotient $\Gamma^{(2)}\backslash\mathbb{H}_2$ as an open subset of projective space and extending to the compactification.

Given any quadruple (a,b,c,d), we get a point $[a:b:c:d] \in \mathbb{P}^3(\mathbb{C})$. We know that $\Theta^{(2)}(\Gamma^{(2)}\backslash\mathbb{H}_2^*) = \mathbb{P}^3(\mathbb{C})$, and that the image of $\Gamma^{(2)}\backslash\mathbb{H}_2^*$ has complement in $\mathbb{P}^3(\mathbb{C})$ of codimension 2. Thus, unless $[a:b:c:d]$ happens to lie in this set of codimension 2, there is some point $\Omega \in \mathbb{H}_2$ such that $\Theta^{(2)}(\Omega) = [a:b:c:d]$. We suppose that (a,b,c,d) is not in this exceptional set (by Proposition 3.10, nor are any of its iterates under the AGM process).

Because of our calculation of the monodromy of $\Theta^{(2)}$, we know that $\Theta^{(2)}(M(\Omega)) = [a:b:c:d]$ for all $M \in \Gamma^{(2)}$. By Lemma 3.9, it follows that all of the values $\frac{a}{\theta_{00}^2(M(\Omega))}$ are good values of the AGM as M runs over $\Gamma^{(2)}$.

Conversely, suppose we are given a good value μ of the AGM. Then eventually all steps, all after the Nth say, arise by taking a 'right' choice

for square roots. Then μ is a simplest value of the AGM of the quadruple (a_N, b_N, c_N, d_N). There is $\Omega_N \in \mathbb{H}_2$ which corresponds to the quadruple (a_N, b_N, c_N, d_N); using Proposition 3.13 repeatedly, we find a sequence $\Omega_{N-1}, \ldots, \Omega_0 \in \mathbb{H}_2$ such that $\Theta^{(2)}(\Omega_i) = [a_i : b_i : c_i : d_i]$ and such that

$$(a_n, b_n, c_n, d_n) = \mu(\theta_{00}^2(2^n\Omega_0), \theta_{01}^2(2^n\Omega_0), \theta_{10}^2(2^n\Omega_0), \theta_{11}^2(2^n\Omega)).$$

It follows that μ is obtained as $\frac{a}{\theta_{00}^2(\Omega_0)}$ for some Ω_0; as $\Theta^{(2)}(\Omega_0) = [a_0 : b_0 : c_0 : d_0]$, we see also that $\Theta^{(2)}(\Omega_0) = \Theta^{(2)}(\Omega)$, so that $\Omega_0 = M(\Omega)$ for some M, as required. This completes the proof that the two sets are equal.

The assertion that simplest values are those of maximal modulus follows from the definition: by Corollary 3.3, $\frac{a}{\theta_{00}^2(\Omega)}$ is maximal if $|\det(C\Omega + D)| \geq 1$ for all $M \in \Gamma^{(2)}$, i.e., if $\Omega \in \mathcal{F}^{(2)}$. But the period doubling sequences, as in Lemma 3.9, starting from such Ω necessarily give rise to simplest values, as $2^n\Omega \in \mathcal{F}^{(2)}$ for all n, so the right choice is made at every step. $\square$

We now compute the values of ζ_M^2 for $M \in \Gamma^{(2)}$. For this, we use the set of generators given in the appendix, together with the calculations in [9], p.194. Mumford shows that the map $M \mapsto \zeta_M^2$ from $\Gamma_{1,2}$ is multiplicative and valued in fourth roots of unity. He further proves:

$$\begin{aligned} \zeta \begin{pmatrix} A & 0 \\ 0 & {}^tA^{-1} \end{pmatrix}^2 &= \det(A), \\ \zeta \begin{pmatrix} I_2 & B \\ 0 & I_2 \end{pmatrix}^2 &= 1. \end{aligned}$$

As $\begin{pmatrix} I_2 & 0 \\ C & I_2 \end{pmatrix}$ is conjugate to $\begin{pmatrix} I_2 & -C \\ 0 & I_2 \end{pmatrix}$, it follows that

$$\zeta \begin{pmatrix} I_2 & 0 \\ C & I_2 \end{pmatrix}^2 = 1.$$

Note also that

$$\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix},$$

and it follows that $\zeta_{X^{(i)}}^2 = -1$ for all i. In particular, $\zeta^2(M)$ takes values only in ± 1 for $M \in \Gamma^{(2)}$.

Appendix: More about $\Gamma^{(2)}$

The main aim of this appendix is to give a set of generators for the group

$$\Gamma^{(2)} = \left\{ \left(\begin{array}{cc} A & B \\ C & D \end{array} \right) \middle| \begin{array}{l} A \equiv D \equiv I_g \pmod 2, C \equiv 0 \pmod 2 \\ \mathrm{diag}(B) \equiv 0 \pmod 2, \mathrm{diag}(C) \equiv 0 \pmod 4 \end{array} \right\}$$

introduced earlier. We also derive a useful corollary (Lemma A.4).

Theorem A.1 *A set of generators for $\Gamma^{(2)}$ is given by*

- $\left(\begin{array}{cc} A & 0 \\ 0 & {}^tA^{-1} \end{array} \right)$ *for $A \equiv I_g$ (mod 2) integral and unimodular,*
- $\left(\begin{array}{cc} I_g & B \\ 0 & I_g \end{array} \right)$ *for B integral symmetric and with even diagonal,*
- $\left(\begin{array}{cc} I_g & 0 \\ 2C & I_g \end{array} \right)$ *for C integral symmetric and with even diagonal,*
- *The matrices $X^{(i)}$ $(i = 1, \ldots, g)$, defined so that*

$$\begin{aligned} X^{(i)}_{ii} &= 3, \\ X^{(i)}_{i,i+g} &= 2, \\ X^{(i)}_{i+g,i} &= 4, \\ X^{(i)}_{i+g,i+g} &= 3, \\ X^{(i)}_{jk} &= \delta_{jk} \qquad \textit{otherwise.} \end{aligned}$$

The proof of this theorem mimics that of [9], Proposition A1. First we have the easy Lemma:

Lemma A.2 *Let $(m, n) \in \mathbb{Z}^2$, not both 0. Under the transformations*

$$\begin{array}{lrcl} \mathbf{A} & (x, y) & \mapsto & (-x, -y), \\ \mathbf{B} & (x, y) & \mapsto & (x + 2y, y), \\ \mathbf{C} & (x, y) & \mapsto & (x, y + 4x), \\ \mathbf{X} & (x, y) & \mapsto & (3x + 2y, 4x + 3y), \end{array}$$

and their inverses, (m, n) may be mapped to one of

$$(0, d), (d, 0), (d, d), (d, 2d),$$

where d denotes the highest common factor (m, n).

Proof. Without loss of generality, $d = 1$. Then each operation also gives coprime integers.

- If $|n| > 2|m|$, apply $\mathbf{C}$ or its inverse; then $|m|$ remains the same, and $|n|$ decreases unless $m = 0$, which happens at the point $(0, \pm 1)$ (if the sign is negative, apply $\mathbf{A}$).
- If $|m| < |n| < 2|m|$, apply $\mathbf{X}$ or its inverse; then both $|m|$ and $|n|$ decrease.
- If $|m| > |n|$, apply $\mathbf{B}$ or its inverse; then $|n|$ remains the same, and $|m|$ decreases unless $n = 0$, which happens at the point $(\pm 1, 0)$ (if the sign is negative, apply $\mathbf{A}$).

Repeat this operation until one reaches $(0, 1)$, $(1, 0)$, or a point with $|m| = |n|$ or $|n| = 2|m|$. In these latter cases, apply $\mathbf{A}$ to ensure that m is positive; then one must be at $(1, 1)$, $(1, -1)$, $(1, 2)$ or $(1, -2)$. Applying $\mathbf{X}$ to $(1, -1)$ gives $(1, 1)$; applying $\mathbf{C}$ to $(1, -2)$ gives $(1, 2)$. □

Proof of Theorem. We use induction on g. Let $\mathbf{e}_1$ be defined to be $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

Let $\gamma \in \Gamma^{(2)}$, and let $\begin{pmatrix} a_1 \\ \vdots \\ a_g \\ c_1 \\ \vdots \\ c_g \end{pmatrix} = \gamma \mathbf{e}_1$. Note that a_1 is odd, every other component is even, and c_1 is divisible by 4.

Premultiplying γ by elements of our generating set allow us to perform row operations on this vector of the following sorts:

$\mathbf{A}_i$	$a_i, c_i \mapsto -a_i, -c_i,$	all other a_j, c_j unaffected;
$\mathbf{B}_i$	$a_i, c_i \mapsto a_i \pm 2c_i, c_i,$	all other a_j, c_j unaffected;
$\mathbf{C}_i$	$a_i, c_i \mapsto a_i, c_i \pm 4a_i,$	all other a_j, c_j unaffected;
$\mathbf{B}_{ij}$	$a_i, a_j, c_i, c_j \mapsto a_i \pm c_j, a_j \pm c_i, c_i, c_j$	all other a_k, c_k unaffected;
$\mathbf{C}_{ij}$	$a_i, a_j, c_i, c_j \mapsto a_i, a_j, c_i \pm 2a_j, c_j \pm 2a_i$	all other a_k, c_k unaffected;
$\mathbf{X}_i$	$a_i, c_i \mapsto 3a_i \pm 2c_i, \pm 4a_i + 3c_i,$	all other a_j, c_j unaffected

coming from special cases of the transformations of type **A, B, C** and **X**.

We first explain that there is a sequence of generators $\delta_1, \ldots, \delta_N$ such that

$$\delta_N \ldots \delta_1 \gamma \mathbf{e}_1 = \mathbf{e}_1.$$

Step 1 Let $d = (a_1, c_1)$, which is odd as a_1 is. By the Lemma, applying the transformations $\mathbf{A}_1$, $\mathbf{B}_1$, $\mathbf{C}_1$ and $\mathbf{X}_1$ changes (a_1, c_1) to $(d, 0)$ (note that $4|c_1$, and this must remain the same after every step, so that $(d, 0)$ is the only possible answer).

Step 2 If $a_1 \nmid c_i$ for some $i > 1$, suppose $d' = \mathrm{hcf}(a_1, c_i)$. Then apply repeatedly $\mathbf{B}_{1i}$ and $\mathbf{C}_{1i}$ to change (a_1, c_i) to $(d', 0)$.

Step 3 Repeat so that $a_1|c_i$ for all $i \geq 1$; repeat Step 1 until c_1 is again 0. Now a_1 divides all c_i. If $a_1 \nmid a_i$, apply $\mathbf{C}_{1i}$ so that the new value of c_1 is $2a_i$. Repeat Step 1. Eventually a_1 divides all a_i and all c_i. As matrices in $\Gamma^{(2)}$ have determinant 1, the highest common factor of entries in any column is 1. Thus $a_1 = \pm 1$. Applying $\mathbf{A}_1$ allows us to take $a_1 = 1$.

Step 4 Make all c_i (for $i > 1$) zero by suitable application of $\mathbf{C}_{1i}$. Then make $c_1 = 0$ by applying $\mathbf{C}_1$.

Step 5 Now $a_1 = 1$ and all other a_i are even. The $g \times g$ matrix A which is the identity, except for its first column, defined to be the vector of a_i's is unimodular, integral and congruent to I_g modulo 2. The inverse of the corresponding element of our generating set maps $\begin{pmatrix} a_1 \\ \vdots \\ a_g \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ back to $\mathbf{e}_1$.

After performing these steps, we arrive at a sequence $\delta_1, \ldots, \delta_N$ with the required property. Now define $\widetilde{\gamma} = \delta_N \ldots \delta_1 \gamma$, so that $\widetilde{\gamma}\mathbf{e}_1 = \mathbf{e}_1$. Define $\mathbf{e}_{g+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Consider $\begin{pmatrix} b_1 \\ \vdots \\ b_g \\ d_1 \\ d_2 \\ \vdots \\ d_g \end{pmatrix} = \widetilde{\gamma}\mathbf{e}_{g+1}$. As $\widetilde{\gamma}$ is symplectic and fixes $\mathbf{e}_1$ we see that $d_1 = 1$. Further, $\widetilde{\gamma} \in \Gamma^{(2)}$, so $2|b_1, d_2, \ldots, d_g$.

Step 6 Make $b_2, \ldots, b_g$ vanish by applying $\mathbf{B}_{1i}$ suitably, and then make b_1 vanish by applying $\mathbf{B}_1$. (Note that these operations fix $\mathbf{e}_1$.)

Step 7 Now form the $g \times g$ matrix D, whose first column is made up of the vector of d_i's, but is otherwise the identity. Let $A = {}^tD^{-1}$. Then A is integral, unimodular and congruent to I_g modulo 2. The corresponding element of the generating set fixes $\mathbf{e}_1$ and maps $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ d_1 \\ d_2 \\ \vdots \\ d_g \end{pmatrix}$ back to $\mathbf{e}_{g+1}$.

Then these steps give transformations $\delta_{N+1}, \ldots, \delta_M$ such that the matrix $\delta_M \ldots \delta_{N+1}\delta_N \ldots \delta_1\gamma$ fixes both $\mathbf{e}_1$ and $\mathbf{e}_{g+1}$. Then restricting the action of this matrix to the complement of the space spanned by $\mathbf{e}_1$ and $\mathbf{e}_{g+1}$ gives a symplectic matrix of dimension $2(g-1)$, and inductively, we see that our putative generating set does indeed generate $\Gamma^{(2)}$. □

Remark A.3 We note that there seems to be an error in Step 7 of the proof of Proposition A1 of [9]; he uses there a transformation which does not fix $\mathbf{e}_1$. It suffices to add to his generating set those elements of our generating set corresponding to unimodular integral matrices A which are congruent to I_g modulo 4, in order that his Step 7 may be made to work. However, the later results in Appendix A of [9] remain valid, as do the applications of these results in the main part of the text.

Lemma A.4 *If $C \equiv 0$ (mod 2), $\mathrm{diag}(C) \equiv 0$ (mod 4) and $D \equiv I_g$ (mod 2) such that the highest common factor of each row of the $2g \times g$ matrix $(C|D)$ is 1, then there exists a matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma^{(2)}$.*

Sketch. This is really a corollary of the previous theorem; we illustrate the principle in the case $g = 1$ and leave the general case to the reader. So suppose $c \equiv 0 \pmod 4$ and d is odd. Then consider the vector (c, d). As in the proof of the theorem, we find a sequence of generators (acting on the right, this time) mapping (c, d) back to $(0, 1)$ (as c and d are coprime). The inverse of this sequence gives a matrix in $\Gamma^{(2)}$ sending $(0, 1)$ to (c, d); this matrix then has the required bottom row. More generally, we use induction on g, as at the end of the previous proof. □

Acknowledgements

We thank Stefan Kühnlein, Nicole Nossem, Nick Shepherd-Barron and John Wilson for helpful remarks during the preparation of this paper.

References

[1] C.W.Borchardt, Gesammelte Werke, Berlin (1888)

[2] J.M.Borwein, P.Borwein, Pi and the AGM, Wiley, New York (1987)

[3] J.-B.Bost, J.-P.Mestre, "Moyenne arithméto-géométrique et périodes des courbes de genre 1 et 2", Gazette des Mathématiciens 38 (1988) 36–64

[4] C.-L.Chai, "Siegel moduli schemes and compactifications over $\mathbb{C}$", in *Arithmetic Geometry*, G.Cornell and J.Silverman (eds.), Springer (1986)

[5] D.A.Cox, "The arithmetic-geometric mean of Gauss", l'Enseignement Mathématique 30 (1984) 275–330

[6] I.Dolgachev, D.Ortland, Point sets in projective spaces and theta functions, Astérisque 165 (1988)

[7] H.Klingen, Introductory Lectures on Siegel Modular Forms, Cambridge University Press (1990)

[8] L.Königsberger, "Ueber die Transformation der Abelschen Functionen erster Ordnung", Crelle 64 (1865)

[9] D.Mumford, Tata lectures on Theta I, II, Birkhäuser (1983, 1984)

[10] F.J.Richelot, "De transformatione integralium Abelianorum primi ordinis commentatio", Crelle 16 (1837)