Supersingular j-invariants, hypergeometric series, and Atkin's orthogonal polynomials

Kaneko, Masanobu
Faculty of Mathematics, Kyushu University

Zagier, Don
Max-Planck-Institut für Mathematik

出版情報：AMS/IP studies in advanced mathematics. 7, pp.97-126, 1998. American Mathematical Seciety
バージョン：
権利関係：

AMS/IP Studies in Advanced Mathematics
Volume 7, 1998

Supersingular j-invariants, hypergeometric series, and Atkin's orthogonal polynomials

M. Kaneko and D. Zagier

1 Introduction.

An elliptic curve E over a field K of characteristic $p > 0$ is called *supersingular* if the group $E(\bar{K})$ has no p-torsion. This condition depends only on the j-invariant of E and it is well known (cf. §2 for a review) that there are only finitely many supersingular j-invariants in $\bar{\mathbb{F}}_p$. We are interested in computing the polynomial

$$ss_p(j) = \prod_{\substack{E/\bar{\mathbb{F}}_p \\ E \text{ supersingular}}} (j - j(E)) \qquad \in \quad \mathbb{F}_p[j]\,.$$

The polynomial describing supersingularity in terms of the λ-invariant of E (defined by writing E over $\bar{K}$ in Legendre form $y^2 = x(x-1)(x-\lambda)$) has a well-known and simple explicit expression, but a convenient expression for the polynomial expressing the condition of supersingularity directly in terms of the j-invariant (i.e., in terms of a Weierstrass model over K, without numbering the 2-torsion points over $\bar{K}$) is less easy to find. In this (partially expository) paper, we will describe several different ways of constructing canonical polynomials in $\mathbb{Q}[j]$ whose reductions modulo p give $ss_p(j)$. These will be of three kinds:

A. polynomials coming from special modular forms of weight $p-1$,

B. the Atkin orthogonal polynomials, and

C. other orthogonal polynomials coming from hypergeometric series.

In the rest of this introduction, we will describe in more detail these various ways of getting the supersingular polynomials.

A. For any even integer $k > 2$, let M_k denote the space of modular forms of weight k on $\Gamma = PSL(2,\mathbb{Z})$. We can write k uniquely in the form

$$k = 12m + 4\delta + 6\varepsilon \qquad \text{with} \quad m \in \mathbb{Z}_{\geq 0}, \quad \delta \in \{0,1,2\}, \quad \varepsilon \in \{0,1\}\,, \tag{1}$$

and then $\dim M_k = m+1$ and any modular form in M_k can be written uniquely as

$$f(\tau) = \Delta(\tau)^m\, E_4(\tau)^\delta\, E_6(\tau)^\varepsilon\, \widetilde{f}\big(j(\tau)\big) \tag{2}$$

for some polynomial $\widetilde{f}$ of degree $\leq m$ in $j(\tau)$, the coefficient of j^m in $\widetilde{f}$ being equal to the constant term of the Fourier expansion of f. (Here Δ, E_4, E_6 and j have their standard meanings, recalled in §3.) On the other hand, if $k = p-1$ for a prime number $p \geq 5$, then $\deg ss_p = m+\delta+\varepsilon$ and the polynomial $ss_p(j)$ is divisible by $j^\delta(j-1728)^\varepsilon$. We will describe for each k four (three if $k \equiv 2 \bmod 3$) modular forms E_k, F_k, G_k, and H_k of weight k such that if $k = p-1$ then the corresponding polynomial in j,

multiplied by $j^\delta(j-1728)^\varepsilon$, reduces modulo p to the supersingular polynomial. These forms are defined as follows:

E_k is the normalized Eisenstein series of weight k;

G_k is the coefficient of X^k in $(1-3E_4(\tau)X^4+2E_6(\tau)X^6)^{-1/2}$;

H_k is the coefficient of X^k in $(1-3E_4(\tau)X^4+2E_6(\tau)X^6)^{k/2}$;

F_k for $k \not\equiv 2 \pmod 3$ is the unique normalized solution in M_k of the differential equation $\vartheta_{k+2}\vartheta_k F_k = \frac{k(k+2)}{144} E_4 F_k$. Here $\vartheta_k : M_k \to M_{k+2}$ is the derivation $f \mapsto f' - kE_2 f/12$ where $f' = (2\pi i)^{-1} df/d\tau = q\, df/dq$ and $E_2 = \Delta'/\Delta = 1 - 24q - \cdots$ is the "nearly modular" Eisenstein series of weight 2; the existence and uniqueness of F_k will be shown in §3. The first result is then:

Theorem 1 *Let $k = p-1$ where $p \geq 5$ is prime and let f be any of the four modular forms E_k, F_k, G_k, H_k described above. Then the coefficients of the associated polynomial $\widetilde{f}$ are p-integral and*

$$ss_p(j) \equiv \pm j^\delta (j-1728)^\varepsilon \widetilde{f}(j) \pmod p.$$

Of these four descriptions of $ss_p(j)$, the ones in terms of H_k and E_k are well-known, the former being a classical result of Hasse and Deuring and the latter a result apparently first noticed by Deligne (cf. [8]). We will give self-contained and elementary proofs of all four in §§2–3. We will also give alternate descriptions of G_k as the residue at 0 of the $\frac{k+1}{2}$-th power of the Weierstrass $\wp$-function and of F_k as a hypergeometric function. As a numerical example, for $k = 28$ the polynomials $\widetilde{f}(j)$, related to the corresponding modular forms by $f(\tau) = \Delta(\tau)^2 E_4(\tau)\widetilde{f}(j(\tau))$, are given by

$$\begin{aligned}
\widetilde{E}_k(j) &= j^2 - \tfrac{5699870640000}{3392780147} j + \tfrac{1180807372800000}{3392780147}, \\
\widetilde{G}_k(j) &= \tfrac{3304503}{2048} j^2 - \tfrac{8394435}{4} j + 176359680, \\
\widetilde{H}_k(j) &= 6608316 j^2 - 23558895360 j - 1434705592320, \\
\widetilde{F}_k(j) &= \tfrac{391}{72} j^2 - 11424 j + 4644864.
\end{aligned}$$

and with $p = 29$ we indeed find

$$\widetilde{E}_k(j) \equiv \widetilde{F}_k(j) \equiv -\widetilde{G}_k(j) \equiv -\widetilde{H}_k(j) \equiv j^2 + 2j + 21 \equiv ss_p(j)/j \pmod p.$$

B. The second, and even more beautiful, description of the supersingular polynomials was found about ten years ago by Atkin, who was inspired by a paper of Rankin [7] on the zeros of Eisenstein series. However, Atkin's proofs were apparently never published and are not well-known. One main purpose of this paper is to popularize and to provide simpler proofs of his discoveries.

Atkin defines a sequence of polynomials $A_n(j) \in \mathbb{Q}[j]$, one in each degree n, as the orthogonal polynomials with respect to a special scalar product. Recall that, if V is the space of polynomials in one variable over a field K, and $\phi : V \to K$ a linear functional, then one can consider the scalar product on V defined by $(f,g) =$

$\phi(fg)$ and the family—which for generic ϕ exists and unique—of monic polynomials which are mutually orthogonal with respect to it. The study of such polynomials, which we will review briefly in §4, is an old subject and is important in many parts of mathematics. In our context, we take V to be the space of polynomials in j; thinking of j as the modular invariant $j(\tau) = q^{-1} + 744 + \cdots$, we can identify V with the space of holomorphic Γ-invariant functions in the upper half-plane $\mathcal{H}$ which are meromorphic at infinity, i.e. $f\left(\frac{a\tau+b}{c\tau+d}\right) = f(\tau)$ for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ and f has a Laurent series expansion $f(\tau) = \sum_{n \gg -\infty} c_n q^n$.

Theorem 2 (Atkin) *There is a unique functional ϕ on V (up to a scalar multiple) for which all Hecke operators $T_n : V \to V$ ($n \in N$) are self-adjoint with respect to the associated scalar product $(f, g) = \phi(fg)$, and a unique family of monic polynomials $A_n(j)$ of degree $n = 0, 1, 2, \ldots$ which are orthogonal with respect to this scalar product.*

We will prove this theorem, and at the same time give several explicit descriptions of the scalar product, in §5. We mention only one here. Take the weight 12 cusp form $\Delta(\tau) = q - 24q^2 + 252q^3 + \cdots$ (rather than q or j^{-1}) as a local parameter for $\mathcal{H}/\Gamma$ at infinity. Then

$$(f, g) = \text{constant term of } f(\tau)\, g(\tau) \text{ as a Laurent series in } \Delta(\tau)\,. \tag{3}$$

The polynomials A_n can be found by the Gram-Schmidt orthogonalization procedure or by the explicit formulas given in Theorem 4 below. The first few are

$$\begin{aligned}
A_0(j) &= 1\,,\\
A_1(j) &= j - 720\,,\\
A_2(j) &= j^2 - 1640j + 269280\,,\\
A_3(j) &= j^3 - \frac{12576}{5}j^2 + 1526958j - 107765856\,,\\
A_4(j) &= j^4 - 3384j^3 + 3528552j^2 - 1133263680j + 44184000960\,.
\end{aligned}$$

The coefficients of A_n are rational numbers in general, but they are p-integral for primes $p > 2n$. In particular, if n_p ($\approx p/12$) is the degree of the supersingular polynomial ss_p, then A_{n_p} has p-integral coefficients, and we have:

Theorem 3 (Atkin) *Let p be a prime number. Then $ss_p(j) \equiv A_{n_p}(j) \pmod{p}$.*

The form of this theorem is a little surprising: unlike the descriptions in Theorem 1, where the polynomials depended separately on m, δ and ε and we therefore had four polynomials of each degree, here the polynomial depends only on $n = m + \delta + \varepsilon$. Thus a single Atkin polynomial A_n may have to do duty for as many as four different supersingular polynomials, if the four numbers $12n - 13$, $12n - 7$, $12n - 5$ and $12n + 1$ are all prime, e.g. the supersingular polynomial for each of the four prime numbers $p = 23, 29, 31, 37$ is the mod p reduction of the same polynomial $A_3(j)$. For instance,

for $p = 29$ we find $A_3(j) \equiv j^3 + 2j^2 + 21j \pmod{p}$, in accordance with the numerical examples given in **A** above.

Theorem 4 *The polynomials A_n are determined in each of the following ways:*
i) *Recursion relation:*

$$\begin{aligned} A_{n+1}(j) &= \left(j - 24\,\frac{144n^2-29}{(2n+1)(2n-1)}\right) A_n(j) \\ &\quad - 36\,\frac{(12n-13)(12n-7)(12n-5)(12n+1)}{n(n-1)(2n-1)^2}\, A_{n-1}(j) \end{aligned} \tag{4}$$

for $n \geq 2$, with initial values A_0, A_1, A_2 as given above;
ii) *Closed formula:*

$$A_n(j) = \sum_{i=0}^{n} 12^{3i} \left[\sum_{m=0}^{i} (-1)^m \binom{-\frac{1}{12}}{i-m} \binom{-\frac{5}{12}}{i-m} \binom{n+\frac{1}{12}}{m} \binom{n-\frac{7}{12}}{m} \binom{2n-1}{m}^{-1}\right] j^{n-i}$$

iii) *Differential equation:*

$$\begin{aligned} &j^2(j-c)^2\,(n^2 j - 144)\,A_n'''' + j(j-c)\,[6n^2j^2 - 144(36n^2+7)j + c^2/3]\,A_n''' \\ &- [(2n^4-7n^2)j^3 - 48(72n^4-245n^2-30)j^2 - 4c(240n^2+413)j + 320c^2]\,A_n'' \\ &- [(2n^4-n^2)j^2 - 24(72n^4-13n^2-12)j + 2c(192n^2-107)]\,A_n' \\ &+ [n^6 j - 24(18n^4-n^2)]\,A_n = 0 \end{aligned}$$

where $c = 1728$, and A_n is the unique monic polynomial solution of this equation.

We will give a proof of Theorem 3 from the point of view of modular forms theory in §6 and a second proof, from the point of view of the theory of hypergeometric functions and with the recurrence (4) as the definition of the Atkin polynomials, in §7. (Atkin's original proof also used modular forms and hypergeometric functions, but involved higher hypergeometric functions ${}_pF_q$ and was considerably more complicated.) §7 also contains the proof of Theorem 4 and of other explicit formulas for A_n in terms of truncated hypergeometric series.

C. In §8 we will show that the polynomials $\widetilde{F}_k(j)$ attached to the modular forms F_k ($2|k$, $k \not\equiv 2 \pmod 3$) defined in part **A** have beautiful expressions as hypergeometric polynomials and also enjoy properties like those of the Atkin polynomials: not only do their reductions modulo p give the supersingular polynomials, but they are also orthogonal with respect to a suitable scalar product on a space of modular functions on $PSL(2,\mathbb{Z})$. This scalar product does not have the nice property of making the Hecke operators self-adjoint, but is in other respects much simpler than the Atkin scalar product (the scalar product of two monomials in j and $j - 1728$ is given by a very simple formula). Moreover, the polynomials $\widetilde{F}_k$ are given by formulas similar to those in Theorem 4, but rather simpler: they satisfy a recursion of almost exactly the same form as (4) (more precisely, four recursions, one for each of the residue classes modulo 12 which occur), but are given by a much simpler closed formula and satisfy

a much simpler differential equation, of order 2 rather than 4. The detailed statement will be given as Theorem 5 in §8 when we have established more notation.

The last two sections of the paper contain a few complementary results. As we already mentioned, the recursion (4) implies that $A_n(j)$ has rational coefficients and is p-integral for $p > 2n$. But the recursion shows only that its denominator divides $\frac{(2n)!(2n-2)!}{2^{2n-1}n!}$, while from numerical examples we find that the denominators are in fact far smaller. (For instance, the denominator of $A_9(j)$ is only 34, and only three of its coefficients are non-integral, and the previous $A_n(j)$ have even fewer non-integral coefficients.) In §9 we will use the closed formula in Theorem 4 to study this phenomenon and some related congruences. Finally, in §10 we will describe an elementary argument relating the properties of the classical modular polynomial $\Phi_p(X,Y) \in \mathbb{Z}[X,Y]$ to the mod p reduction of the polynomial $\widetilde{E}_{p-1}(j)$, and thus the supersingular polynomial. This yields at the same time easy proofs of some properties of supersingular polynomials which were mentioned in the text and a partial answer to a question of E. de Shalit [9].

2 Supersingular elliptic curves

The definition of supersingular elliptic curves over a field of characteristic p was given at the beginning of the paper. We begin by recalling the statement and proof of the standard criterion for deciding whether a given curve over a finite field $\mathbb{F}_q$ ($q = p^r$, p odd) is supersingular or not.

Proposition 1 *Let E be the elliptic curve over $\mathbb{F}_q$ defined by the equation $y^2 = f(x)$ ($f \in \mathbb{F}_q[x]$ of degree 3), and a_p the coefficient of x^{p-1} in $f(x)^{(p-1)/2}$. Then $|E(\mathbb{F}_q)| \equiv 1 - N_{\mathbb{F}_q/\mathbb{F}_p} a_p \pmod{p}$.*

Corollary *E is supersingular if and only if $a_p = 0$.*

Proof For $x \in \mathbb{F}_q$ the number of solutions in $\mathbb{F}_q$ of $y^2 = f(x)$ is equal to $1 + f(x)^{(q-1)/2}$ (namely, to 0, 1, or 2 for $f(x) \notin (\mathbb{F}_q)^2$, $f(x) = 0$, or $f(x) \in (\mathbb{F}_q^\times)^2$, respectively). Counting also the point at infinity, we find

$$|E(\mathbb{F}_q)| = 1 + \sum_{x \in \mathbb{F}_q} \left(1 + f(x)^{\frac{q-1}{2}}\right) \qquad \text{in } \mathbb{F}_q\,.$$

Since the sum over $x \in \mathbb{F}_q$ of x^j equals -1 for $j = q-1$ and 0 for all other j in the range $0 \le j \le 3(q-1)/2$, this gives

$$|E(\mathbb{F}_q)| = 1 - a_q \qquad \text{in } \mathbb{F}_q\,,$$

where a_q denotes the coefficient of x^{q-1} in $f(x)^{(q-1)/2}$. (In particular, a_q belongs to $\mathbb{F}_p$ and not merely to $\mathbb{F}_q$.) But from the expansion

$$f(x)^{\frac{q-1}{2}} = f(x)^{\frac{p-1}{2}(1+p+\cdots+p^{r-1})} = f(x)^{\frac{p-1}{2}} f^{(p)}\left(x^p\right)^{\frac{p-1}{2}} \cdots f^{(p^{r-1})}\left(x^{p^{r-1}}\right)^{\frac{p-1}{2}},$$

where $f^{(p^j)}$ is the polynomial obtained from f by raising all its coefficients to the p^j-th power, we see that $a_q = a_p^{1+p+\cdots+p^{r-1}} = N_{\mathbb{F}_q/\mathbb{F}_p}(a_p)$. This proves the proposition.

To prove the corollary, note that if $a_p = 0$, then $|E(\mathbb{F}_{q^n})| \equiv 1 \not\equiv 0 \pmod{p}$ for all n, so E has no p-torsion over $\bar{\mathbb{F}}_p$. Conversely, if $a_p \neq 0$, then $|E(\mathbb{F}_{q^n})| \equiv 1 - (N_{\mathbb{F}_q/\mathbb{F}_p} a_p)^n$ is divisible by p for n divisible by the order of $N_{\mathbb{F}_q/\mathbb{F}_p}(a_p)$ modulo p, so $E(\bar{\mathbb{F}}_p)$ does contain p-torsion.

□

We now write out the contents of the corollary explicitly in terms of the standard Weierstrass equation. This will give the proof of Theorem 1 in the case $f = H_k$. Suppose that $p \geq 5$. Then any elliptic curve over a field K of characteristic p can be written in a Weierstrass form

$$E: \quad y^2 = x^3 - 3Qx + 2R, \tag{5}$$

where the factors -3 and 2 have been included to coincide with traditional notations. (If Q and R are replaced by the Eisenstein series $E_4(\tau)$ and $E_6(\tau)$ then (5) is an equation of the elliptic curve $\mathbb{C}/(\mathbb{Z}\tau+\mathbb{Z})$ over $\mathbb{C}$ with j-invariant $j(\tau)$.) The j-invariant of E is equal to Q^3/Δ, where $\Delta = (Q^3 - R^2)/1728$. We define a graded homogeneous polynomial $H_{p-1}(Q,R)$ of degree $p-1$ in Q and R (where Q and R have degrees 4 and 6, respectively), the *Hasse polynomial*, as the coefficient of x^{p-1} in $(x^3-3Qx+2R)^{(p-1)/2}$, so that the modular form $H_{p-1}(E_4(\tau), E_6(\tau))$ is the same as the modular form of weight $p-1$ denoted $H_{p-1}(\tau)$ in the introduction. As explained there, we can write this polynomial in the form

$$H_{p-1}(Q,R) = \Delta^m Q^\delta R^\varepsilon \, \widetilde{H}_{p-1}(j)$$

for some polynomial $\widetilde{H}_{p-1} \in \mathbb{Z}[j]$, where m, δ and ε are the numbers defined by (1) with $k = p-1$. They are given explicitly by

$$m = \left[\frac{p}{12}\right], \quad \delta = \begin{cases} 0 & \text{if } p \equiv 1 \pmod 3, \\ 1 & \text{if } p \equiv 2 \pmod 3, \end{cases} \quad \varepsilon = \begin{cases} 0 & \text{if } p \equiv 1 \pmod 4, \\ 1 & \text{if } p \equiv 3 \pmod 4. \end{cases} \tag{6}$$

It now follows from the corollary above that, at least in the case $K \subset \bar{\mathbb{F}}_p$, the curve E is supersingular if and only if $j^\delta(j-1728)^\varepsilon \widetilde{H}_{p-1}(j) = 0$ and hence that

$$ss_p(j) \,|\, j^\delta(j-1728)^\varepsilon \widetilde{H}_{p-1}(j)\,.$$

The fact that the two polynomials agree up to a constant, as claimed in Theorem 1, therefore follows from the well-known formula

$$n_p \,(:= \deg ss_p(j)) \;=\; m + \delta + \varepsilon\,, \tag{7}$$

but to make the paper self-contained we give a direct proof. It suffices to show that the polynomial $j^\delta(j-1728)^\varepsilon \widetilde{H}_{p-1}(j)$ has no multiple roots, since we have already shown that it has the same zeros as ss_p, which is square-free by definition. We first treat the roots 0 and 1728. From the expansion

$$(x^3 - 3Qx + 2R)^{\frac{p-1}{2}} = \left(x^3+2R\right)^{\frac{p-1}{2}} - 3\frac{p-1}{2} Qx \left(x^3+2R\right)^{\frac{p-3}{2}} + O(Q^2)$$

we find

$$H_{p-1}(Q,R) = \begin{cases} \binom{\frac{p-1}{2}}{\frac{p-1}{3}} (2R)^{\frac{p-1}{6}} + O(Q) & \text{if } p \equiv 1 \pmod 3 \\ -3\,\frac{p-1}{2} \binom{\frac{p-3}{2}}{\frac{p-2}{3}} (2R)^{\frac{p-5}{6}}\, Q + O(Q^2) & \text{if } p \equiv 2 \pmod 3 \end{cases}$$

and hence in both cases that $\widetilde{H}_{p-1}(0) \not\equiv 0 \pmod p$. A similar argument works for $j = 1728$. That the other numbers in $\bar{\mathbb{F}}_p$ cannot be multiple zeros of $\widetilde{H}_{p-1}(j)$ follows from the fact that this polynomial satisfies a second-order linear differential equation with polynomial coefficients and with leading coefficient $j(j-1728)$. (We will show in §3 that $\widetilde{H}_{p-1}$ and $\widetilde{F}_{p-1}$ agree modulo p, and the differential equation for the latter is a simple translation of the definition of F_k, given explicitly in §8.) This implies that any common zero in $\bar{\mathbb{F}}_p \setminus \{0, 1728\}$ of $\widetilde{H}_{p-1}$ and its first derivative would be a zero of all higher derivatives and hence have infinite order, which is impossible. (Note that since we are in characteristic p, this argument would fail if $\widetilde{H}_{p-1}$ were a polynomial in j^p, but this is not the case since $\deg \widetilde{H}_{p-1} = m < p$.)

Remark We have proved in particular that there are only finitely many supersingular invariants in $\bar{\mathbb{F}}_p$, as mentioned in the introduction. In fact, one knows that these are the only supersingular j-invariants (i.e., the j-invariant of a supersingular elliptic curve over any field K of characteristic p lies in $\bar{\mathbb{F}}_p$), and also that they all lie in $\mathbb{F}_{p^2}$ (equivalently, $ss_p(j)$ factors into linear and quadratic polynomials in $\mathbb{F}_p[j]$). For proofs we refer the reader to [3] and [1], which are the two basic references for the theory of supersingular elliptic curves. In §10 we will give an elementary proof of the fact that all roots of $ss_p(j)$ lie in $\mathbb{F}_{p^2}$ (using the formula $ss_p(j) = j^\delta(j-1728)^\varepsilon \widetilde{E}_{p-1}(j)$, which will be established in the next section), as well as another proof of the simplicity of the zeros $\neq 0, 1728$ of $\widetilde{H}_{p-1}(j) \pmod p$.

3 Modular forms and supersingular polynomials

Our object now is to prove Theorem 1 of the Introduction. We begin by giving the definitions of the special modular forms E_k, F_k, G_k, and H_k in more detail. For k even and positive we denote by B_k the kth Bernoulli number and by $E_k(\tau)$ the kth Eisenstein series

$$E_k(\tau) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \left(\sum_{d|n} d^{k-1} \right) q^n \qquad (q = e^{2\pi i \tau}).$$

It is a modular form of weight k for $k \geq 4$ and for $k = 2$ is "nearly modular":

$$E_2\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2\, E_2(\tau) + \frac{6}{\pi i}\, c(c\tau+d) \qquad \left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma\right). \tag{8}$$

We also define $\Delta = (E_4^3 - E_6^2)/1728 \in M_{12}$ and $j(\tau) = E_4(\tau)^3/\Delta(\tau)$, the modular invariant. From (8) it follows easily that

$$E_2' = \frac{E_2^2 - E_4}{12}, \quad E_4' = \frac{E_2 E_4 - E_6}{3} \quad E_6' = \frac{E_2 E_6 - E_4^2}{2}, \quad \Delta' = E_2 \Delta \tag{9}$$

and more generally $\vartheta_k(f) = f' - \frac{k}{12}E_2 f \in M_{k+2}$ for any $f \in M_k$, where $'$ denotes differentiation with respect to $2\pi i\tau$. If k is represented as in (1), then any $f \in M_k$ has a zero of multiplicity $\geq \delta$ at $\tau = e^{\pi i/3}$ and a zero of multiplicity $\geq \varepsilon$ at $\tau = e^{\pi i/2}$ and hence is divisible by $E_4^\delta E_6^\varepsilon$, and it follows that f has a representation as in (2) with $\widetilde{f}$ a polynomial of degree at most m.

If $k + 4 \not\equiv 0 \pmod 3$, then every element of M_{k+4} is divisible by E_4, so we have an endomorphism ϕ_k of M_k defined by $\phi_k(f) = E_4^{-1}\vartheta_{k+2}(\vartheta_k(f))$. Since the constant term of $\phi_k(f)$ is $\kappa_k := k(k+2)/144$ times the constant term of f, this map preserves the codimension 1 subspace of cusp forms and induces on the quotient space the map multiplication by κ_k. It follows that κ_k is an eigenvalue of ϕ_k. If we pick some corresponding eigenvector F_k, then the other eigenvectors are the modular forms $\Delta^i F_{k-12i}$ $(1 \leq i \leq m)$ with eigenvalues $\kappa_{k-12i} \neq \kappa_k$ (because $\vartheta_k \circ \Delta^i = \Delta^i \circ \vartheta_{k-12i}$ and $\dim M_k = m + 1$), so F_k is unique up to a normalizing factor, which we will fix later when we give natural formulas for F_k in terms of hypergeometric series. Together with the definitions of H_k and G_k given in §1, this defines the four modular forms E_k, F_k, G_k, $H_k \in M_k$. (The letters E, F an H are meant to suggest Eisenstein series, hypergeometric function, and Hasse invariant, while the remaining letter fills the gap.) We will prove that for any prime number $p \geq 5$

$$\widetilde{E}_{p-1}(j) \equiv \widetilde{F}_{p-1}(j) \equiv (-1)^{\delta+\varepsilon}\,\widetilde{G}_{p-1}(j) \equiv (-1)^{\delta+\varepsilon}\,\widetilde{H}_{p-1}(j) \pmod p \tag{10}$$

and

$$ss_p(j) = (-1)^{\delta+\varepsilon}\, j^\delta\, (j-1728)^\varepsilon\, \widetilde{H}_{p-1}(j) \pmod p, \tag{11}$$

which together form a slightly more precise version of Theorem 1. (Here and in what follows, δ and ε are defined by (6).)

Equation (11) was already proved in the last section, except for the value of the constant $(-1)^{\delta+\varepsilon}$. Since ss_p is by definition monic, we only have to compute the leading coefficient of $\widetilde{H}_{p-1}$. The leading coefficient of the polynomial $\widetilde{f}(j)$ for any modular form $f \in M_k$ is just the constant term of the Fourier expansion of f, i.e. the limiting value of f as $q \to 0$. Since E_4 and E_6 have the value 1 at $q = 0$, this number for $\widetilde{H}_{p-1}$ is just the coefficient of X^{p-1} in $(1 - 3X^4 + 2X^6)^{(p-1)/2} = (1-X^2)^{p-1}(1+2X^2)^{(p-1)/2}$, and from

$$\begin{aligned}(1-X^2)^{p-1}(1+2X^2)^{\frac{p-1}{2}} &\equiv \frac{1-X^{2p}}{1-X^2}\left[(1+2X^2)^{\frac{p-1}{2}} - 3^{\frac{p-1}{2}} + 3^{\frac{p-1}{2}}\right] \\ &\equiv \left(1-X^{2p}\right)\left(\text{polynomial of degree } p-3\right) + \left(\frac{3}{p}\right)\frac{1-X^{2p}}{1-X^2} \pmod p\end{aligned}$$

we find that this coefficient is congruent to $\left(\frac{3}{p}\right) = (-1)^{\delta+\varepsilon}$ modulo p, as claimed. We remark that the evaluation of the constant in (11) would also follow from equation (10) which we will prove independently, since both ss_p and $\widetilde{E}_{p-1}$ are monic.

We now proceed to equation (10). The congruence $\widetilde{G}_{p-1} \equiv \widetilde{H}_{p-1} \pmod p$ is obvious, since the generating functions of the modular forms G_k and H_k differ by a factor $(1 - 3E_4X^4 + 2E_6X^6)^{p/2} \equiv 1 + O(X^p) \pmod p$. For the congruence between $\widetilde{G}_{p-1}$ and $\widetilde{E}_{p-1}$ we again use a generating function. Consider the elliptic curve over $\mathbb{C}$ with

Weierstrass model (5), where $Q = E_4(\tau)$, $R = E_6(\tau)$. It can be parametrized in a well-known way by the Weierstrass $\wp$-function (here renormalized to keep coefficients rational), namely by $x = P(u)$, $y = -\frac{1}{2}P'(u)$ where

$$P(u) = u^{-2} - \sum_{n\geq 4,\ n \text{ even}} \frac{12^{n/2}B_n}{n(n-2)!}\, E_n(\tau)\, u^{n-2}$$

(B_n = nth Bernoulli number). Therefore the change of variables $X = P(u)^{-1/2} = u + \dots$ gives

$$\begin{aligned} G_k &= \mathrm{Res}_{X=0} \frac{dX}{X^{k+1}\sqrt{1 - 3E_4X^4 + 2E_6X^6}} \\ &= \mathrm{Res}_{u=0}\, P(u)^{\frac{k+1}{2}}\, du \\ &= \text{coeff. of } u^k \text{ in } \left(1 - \sum_{n\geq 4 \text{ even}} \frac{12^{n/2}B_n}{n(n-2)!}\, E_n\, u^n\right)^{\frac{k+1}{2}}. \end{aligned}$$

Now take $k = p-1$ and observe that $B_nE_n/n!$ for $n < p-1$ is a polynomial in E_4 and E_6 with p-integral coefficients, while $pB_{p-1}/(p-1)! \equiv 1 \pmod p$, so

$$\left(1 - \sum_{n\geq 4 \text{ even}} \frac{12^{n/2}B_n}{n(n-2)!}\, E_n\, u^n\right)^{\frac{p}{2}} \equiv 1 + 12^{\frac{p-1}{2}}\, E_{p-1}\, u^{p-1} + \mathrm{O}(u^p) \qquad (\mathrm{mod}\ p).$$

Using $12^{(p-1)/2} \equiv \left(\frac{12}{p}\right) = (-1)^{\delta+\varepsilon}$, we obtain the desired congruence for G_{p-1}.

Finally, we have to consider F_k. This function was defined (up to a constant) as the unique modular form annihilated by the operator $\vartheta_{k+1}\vartheta_k - \kappa_k E_4$. Since for $k = p-1$ the eigenvalues κ_{k-12r} $(0 \leq r \leq k/12)$ of the operator $E_4^{-1}\vartheta_{k+1}\vartheta_k$ remain distinct after reduction modulo p, this characterization remains valid also in characteristic p. (By a modular form of weight k modulo p we mean a polynomial in E_4 and E_6 of the right degree with coefficients in $\mathbb{F}_p$.) But using formulas (9) we find that

$$\left(\vartheta_{k+2}\vartheta_k - \kappa_k\, E_4\right) f = f'' - \frac{k+1}{6}\, E_2\, f' + \frac{k(k+1)}{12}\, E_2'\, f$$

and this certainly vanishes modulo p if $k = p-1$ and f is E_{p-1}, since then the Fourier expansion of f reduces to 1 mod p. The proportionality of E_{p-1} and F_{p-1} modulo p follows. To get the exact constant of proportionality in (10) we still have to normalize F_k, which we have not done. For reasons to be explained in §8, we will do this by

$$\text{constant term of the Fourier expansion of } F_k(\tau) = (-1)^m \binom{\frac{k-5}{6}}{m}, \qquad (12)$$

where m is defined as in equation (1). For $k = p-1$ the right-hand side of (12) is clearly congruent to 1 modulo p. This completes the proof of Theorem 1.

4 Orthogonal polynomials

In this section we review what we will need from the theory of orthogonal polynomials. Let V be the vector space of polynomials in one variable over a field K and $(\ ,\ )$ a scalar product on V of the form $(f,g) = \phi(fg)$ where $\phi : V \to K$ is a linear functional. (For the classical orthogonal polynomials, and also for the ones we shall be considering, K is a subfield of $\mathbb{R}$ and ϕ has the form $\phi(f) = \int_a^b f(X)\,w(X)\,dX$ for some real numbers $a < b$ and some positive function w on (a,b).) Applying the Gram-Schmidt process to the basis $\{X^n\}_{n\geq 0}$ of V, we obtain a unique basis of orthogonal monic polynomials P_n by the recursion

$$P_n(X) = X^n - \sum_{m=0}^{n-1} \frac{(X^n, P_m)}{(P_m, P_m)}\, P_m(X)\,,$$

provided that at each stage the scalar product (P_n, P_n) is not zero. (This condition is satisfied generically and is automatic for ϕ of the special form mentioned above, since then $(f,f) > 0$ for all $f \neq 0$.) The next proposition describes the recursive calculation of the polynomials $P_n(X)$, assuming this non-degeneracy condition.

Proposition 2 i) *The polynomials P_n satisfy a three-term recursion of the form*

$$P_{n+1}(X) = (X - a_n)\,P_n(X) - b_n\,P_{n-1}(X) \qquad (n \geq 1) \tag{13}$$

for some constants a_n, $b_n \in K$, $\quad b_n = \dfrac{(P_n, P_n)}{(P_{n-1}, P_{n-1})} \neq 0$.

ii) *Define a second sequence of polynomials $\{Q_n\}_{n\geq 0}$ in $K[X]$ by the same recurrence as in* i), *but with initial values $Q_0 = 0$, $Q_1(X) = \phi(1)$. Then*

$$\frac{Q_n(X)}{P_n(X)} \;=\; \Phi(X) \,+\, \mathrm{O}(X^{-2n-1}) \;\in\; K[[X^{-1}]] \tag{14}$$

where

$$\Phi(X) = \sum_{n=0}^{\infty} g_n X^{-n-1} \in K[[X^{-1}]]\,, \qquad g_n = (X^n, 1) = \phi(X^n)\,. \tag{15}$$

This property characterizes P_n (assumed to be monic of degree n) and Q_n uniquely.

iii) *Define numbers $\lambda_n \in K$ $(n \geq 1)$ by the continued fraction expansion*

$$g_0 + g_1 x + g_2 x^2 + \cdots \;=\; \cfrac{g_0}{1 - \cfrac{\lambda_1 x}{1 - \cfrac{\lambda_2 x}{1 - \ddots}}} \;\in\; K[[x]]\,. \tag{16}$$

Then all λ_n are non-zero and $a_n = \lambda_{2n} + \lambda_{2n+1}$, $b_n = \lambda_{2n-1}\lambda_{2n}$ for $n \geq 1$.

Proof i) Since the $P_n(X)$ are monic, we have $XP_n = P_{n+1} + a_{nn}P_n + \cdots + a_{n0}P_0$ for some constants $a_{nm} \in K$. The orthogonality of the P_m and the fact that the scalar

product of two polynomials depends only on their product imply that

$$a_{nm}(P_m, P_m) = (XP_n, P_m) = (P_n, XP_m) = \begin{cases} 0 & \text{if } m \le n-2, \\ (P_n, P_n) & \text{if } m = n-1. \end{cases}$$

This proves the assertion (with $a_n = a_{nn}$, $b_n = a_{n\,n-1}$).

ii) From the definitions, (f, g) is just the coefficient of X^{-1} in the Laurent series $f(X)g(X)\Phi(X)$. In particular, the fact that P_n is orthogonal to all monomials of degree $< n$ says that the coefficient of X^{-r-1} in $P_n(X)\Phi(X)$ vanishes for $0 \le r \le n-1$, i.e., that we have

$$P_n(X)\,\Phi(X) = Q_n(X) + O(X^{-n-1}) \in K[X, X^{-1}]] \tag{17}$$

for some polynomial $Q_n(X)$ of degree $n-1$. (Here $K[X, X^{-1}]]$ denotes the ring of Laurent series in X^{-1}, i.e. sums of polynomials in X and power series in X^{-1}, where X is to be thought of as large.) Reversing the argument shows that property (17), which is obviously the same as (14), is equivalent to the orthogonality of P_n with all lower degree polynomials and hence characterizes the (monic) polynomial P_n completely, as asserted. Finally, from (17) and the recursion (13) we find that $Q_{n+1}(X) - (X - a_n)Q_n(X) + b_nQ_{n-1}(X) = O(X^{-n})$, so that this expression, which is a polynomial, must vanish for $n \ge 1$. Hence the Q_n satisfy the same recursion relation as the P_n, with $Q_0 = 0$, $Q_1 = g_0$.

iii) Define another vector space V^* as $K[Y]$ with scalar product given by $(f, g) = \psi(fg)$, where $\psi(Y^n)$ is 0 for n odd and $g_{n/2}$ for n even. Then we get a family of orthogonal polynomials $P_n^*(Y)$ by the same construction as before. Since odd and even polynomials in Y are orthogonal to each other, it is obvious by induction that P_n^* has parity n for all n (i.e., the even- and odd-index polynomials are even and odd, respectively), so the first part of the proposition applied to (V^*, ψ) gives a recursion of the form $P_{n+1}^*(Y) = YP_n^*(Y) - \lambda_n P_{n-1}^*$ for some non-zero constants $\lambda_n = (P_n^*, P_n^*)/(P_{n-1}^*, P_{n-1}^*) \in K$. The argument of ii) gives a sequence of companion polynomials Q_n^* to the P_n^* (of degree one less, and hence of opposite parity) for which the rational functions $Q_n^*(Y)/P_n^*(Y)$ are the best possible approximations at infinity to $\sum g_k Y^{-2k-1}$, and they satisfy the same recursion $Q_{n+1}^* = YQ_n^* - \lambda_n Q_{n-1}^*$ as the P^*'s. From this one gets by induction on n the formula

$$\begin{pmatrix} Q_{n+1}^* & Q_n^* \\ P_{n+1}^* & P_n^* \end{pmatrix} = \begin{pmatrix} g_0 & 0 \\ Y & 1 \end{pmatrix} \begin{pmatrix} Y & 1 \\ -\lambda_1 & 0 \end{pmatrix} \cdots \begin{pmatrix} Y & 1 \\ -\lambda_n & 0 \end{pmatrix}.$$

This translates by a standard calculation into the continued fraction

$$\cfrac{g_0 Y^{-1}}{1 - \cfrac{\lambda_1 Y^{-2}}{1 - \cfrac{\lambda_2 Y^{-2}}{\ddots\, 1 - \lambda_n Y^{-2}}}} = \frac{Q_{n+1}^*(Y)}{P_{n+1}^*(Y)} = \frac{g_0}{Y} + \frac{g_1}{Y^3} + \cdots \frac{g_n}{Y^{2n+1}} + O\Big(\frac{1}{Y^{2n+3}}\Big).$$

Setting $x = Y^{-2}$ and letting n tend to infinity, we get (16). Finally, the recurrence satisfied by the P_n^* implies the recurrence $P_{n+2}^* = (Y^2 - \lambda_n - \lambda_{n+1})P_n^* - \lambda_{n-1}\lambda_n P_{n-2}^*$

for the P_n^* of a given parity. But V can be identified via $X = Y^2$ with the even part of V^*, with compatible scalar products, so $P_{2n}^*(Y) = P_n(Y^2)$. The relation asserted in the proposition between the coefficients a_n and b_n and the numbers λ_n follows. □

Remark From the proof we see that the necessary and sufficient condition for the scalar product defined by ϕ to be non-degenerate (in the sense that $(P_n, P_n) \neq 0$ for all n) is that the power series $\sum g_n x^n$ has a continued fraction expansion as in (16) with g_0 and all λ_n different from 0. We also see that if $(\ ,\ )$ is positive definite then the numbers $b_n = \lambda_{2n-1}\lambda_{2n}$ are always positive, and if $(f,g) = \int_a^b f(X)g(X)w(X)dX$ with $w \geq 0$ and $a \geq 0$ then all λ_n are positive. (The condition $a \geq 0$ is equivalent to the scalar product on V^* being positive definite, and then $\lambda_n = (P_n^*, P_n^*)/(P_{n-1}^*, P_{n-1}^*) > 0$.)

5 The Atkin scalar product and the Atkin polynomials

We gave one definition of Atkin's scalar product in §1. The following result gives several alternate descriptions. Recall that V is the set of Γ-invariant holomorphic functions on $\mathcal{H}$ which grow at most like q^{-N} at infinity for some N, that V coincides with the set of polynomials in j, and that one can take q, j^{-1} or Δ as a local parameter at infinity, where $q = e^{2\pi i\tau}$, $j = j(\tau) = q^{-1} + 744 + 196884q + \cdots$ is the modular invariant, and $\Delta = q - 24q^2 + 252q^3 + \cdots$ is the discriminant function.

Proposition 3 *The following four definitions of a scalar product on V coincide:*

i) (f,g) = *constant term of* fg *as a Laurent series in* Δ ;

ii) (f,g) = *constant term of* fgE_2E_4/E_6 *as a Laurent series in* j^{-1} ;

iii) (f,g) = *constant term of* fgE_2 *as a Laurent series in* q ;

iv) $(f,g) = \frac{6}{\pi}\int_{\pi/3}^{\pi/2} f(e^{i\theta})\, g(e^{i\theta})\, d\theta$.

Corollary *The scalar product $(\ ,\ )$ is positive definite on $V_{\mathbb{R}} = \mathbb{R}[j]$.*

Proof The equivalence of the first three formulas is immediate by writing the constant terms as $1/(2\pi i)$ times the corresponding residues and using the formulas

$$\frac{d\Delta(\tau)}{\Delta(\tau)} = 2\pi i\, E_2(\tau)\, d\tau = E_2(\tau)\,\frac{dq}{q} = -\frac{E_2(\tau)E_4(\tau)}{E_6(\tau)}\,\frac{dj(\tau)}{j(\tau)}\,.$$

For the fourth, we use the global residue formula: Let $\mathcal{F}_a$ denote the standard fundamental domain of Γ, truncated at some height $a > 1$ (i.e., the domain $|x| \leq \frac{1}{2}$, $x^2 + y^2 \geq 1$, $y \leq a$, where $\tau = x + iy$). The integral of $f(\tau)g(\tau)E_2(\tau)d\tau$ over the top edge of this domain equals (f,g) by formula (iii), so the holomorphy of fgE_2 implies that (f,g) is also given by the sum of the integrals over the rest of the boundary, taken with the appropriate signs. The integrals along the vertical edges $x = \pm\frac{1}{2}$ cancel because fgE_2 is periodic of period 1. Replacing τ by $-1/\tau$ on the left half of the bottom edge and noting that f and g are invariant under this transformation, we find that (f,g) equals the integral along the arc from $e^{\pi i/3}$ to $e^{\pi i/2}$

of $[E_2(\tau) - \tau^{-2}E_2(-1/\tau)]f(\tau)g(\tau)d\tau$. But the expression in square brackets equals $-6i/\pi\tau$ by the transformation law (8) of E_2, and this, with $\tau = e^{i\theta}$, gives formula (iv).

The corollary follows immediately from (iv), since $j(e^{i\theta})$ is real for $\theta \in [\pi/3, \pi/2]$ and consequently $\int_{\pi/3}^{\pi/2} f(e^{i\theta})^2\, d\theta > 0$ for $f(\tau)$ any non-zero polynomial in $j(\tau)$ with real (or, a fortiori, rational) coefficients. Note that formula (iv) can be rewritten in the form

$$(f,g) = \int_0^{1728} f(j)\, g(j)\, w(j)\, dj\,, \qquad w(j) = \frac{6}{\pi}\theta'(j)$$

(here and from now on we will commit the standard abuse of notation of using the same letter to denote an element of V thought of as a function of τ, q, j or Δ, indicating by one of these letters the argument intended), where $\theta : [0, 1728] \to [\pi/3, \pi/2]$ is the inverse to the monotone increasing function $\theta \mapsto j(e^{i\theta})$. A graph of the function $w(j)$ is given in Figure 1.

□

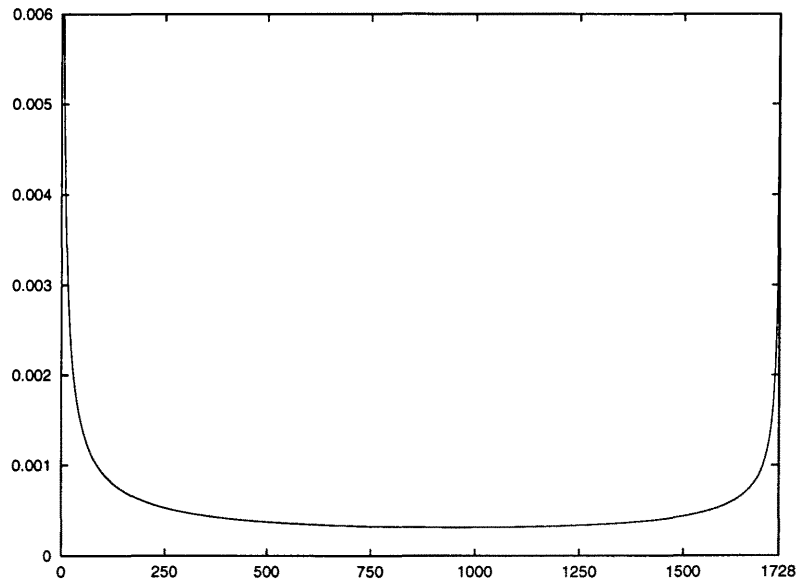

Figure 1.

By the results of the last section we now deduce that (i) there is a unique sequence of monic orthogonal polynomials $A_n(j)$ of degree n; (ii) the scalar product of two monomials j^n and j^m equals g_{n+m}, where g_n is the coefficient of $j(\tau)^{-n-1}$ in

$$\Phi(\tau) = \frac{E_2(\tau)E_4(\tau)}{E_6(\tau)j(\tau)} = q - 24q^2 + 196812q^3 + \cdots = \frac{1}{j(\tau)} + \frac{720}{j(\tau)^2} + \cdots\,;$$

(iii) the A_n are the denominators of the best rational-function approximations to Φ; and (iv) they satisfy a recursion of the form

$$A_{n+1}(j) = \big(j - (\lambda_{2n} + \lambda_{2n+1})\big)\, A_n(j) \;-\; \lambda_{2n-1}\lambda_{2n} A_{n-1}(j) \tag{18}$$

where the λ_n are positive rational numbers defined by the continued fraction expansion of Φ with respect to $1/j$. Computing numerically, we find that the first few values of $g_n = (j^n, 1)$ and λ_n are given by

$$g_0 = 1, \quad g_1 = 720, \quad g_2 = 911520, \quad g_3 = 1301011200, \quad g_4 = 1958042030400,$$
$$\lambda_1 = 720, \quad \lambda_2 = 546, \quad \lambda_3 = 374, \quad \lambda_4 = 475, \quad \lambda_5 = \frac{2001}{5}.$$

In §7, in connection with hypergeometric functions, we will show that

$$\lambda_n = \begin{cases} 720 & \text{if } n = 1, \\ 12\left(6 + \dfrac{(-1)^n}{n-1}\right)\left(6 + \dfrac{(-1)^n}{n}\right) & \text{if } n > 1, \end{cases} \tag{19}$$

giving the explicit recurrence written in part (i) of Theorem 4 of the Introduction. However, this recurrence is not needed for the proofs of the two most important properties of the Atkin polynomials, their Hecke invariance and their relationship to the supersingular polynomials in characteristic p, to which we now turn.

Proof of Theorem 2 For $k \in \mathbb{Z}$ let V_k denote the space of holomorphic functions in $\mathcal{H}$ which transform like modular forms of weight k and have at most exponential growth at infinity (i.e., V_k is the degree k part of the graded ring $\mathbb{C}[E_4, E_6, \Delta^{-1}]$). Note that $V = V_0$ and that V_2 coincides with the set of derivatives of functions in V. We define Hecke operators on V_k by

$$(f|_k T_n)(\tau) = n^{k/2} \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma\backslash\mathcal{M}_n} \frac{1}{(c\tau + d)^k} f\left(\frac{a\tau + b}{c\tau + d}\right) \qquad (n \in \mathbb{N}, \quad f \in V_k),$$

where $\mathcal{M}_n$ denotes the set of 2×2 matrices with integral coefficients and determinant n. (This is not the standard normalization of T_n unless $k = 2$, but will turn out to be a more convenient normalization when we study both positive and negative weights.) Notice that this formula makes sense only for $f \in V_k$, since the expression $(c\tau + d)^{-k} f\left(\frac{a\tau+b}{c\tau+d}\right)$ will not be independent of the choice of representative in $\Gamma\backslash\mathcal{M}_n$ if f is not modular, but the "Hecke operator at infinity"

$$(f|_k T_n^\infty)(\tau) = n^{k/2} \sum_{\substack{ad=n \\ a,d>0}} \sum_{b \,(\mathrm{mod}\ d)} d^{-k} f\left(\frac{a\tau + b}{d}\right)$$

makes sense for any 1-periodic function f and agrees with $|_k T_n$ if $f \in V_k$ because the matrices $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ with $0 \le b < d = \dfrac{n}{a}$ are a set of representatives for $\Gamma\backslash\mathcal{M}_n$. We claim that

$$\mathrm{Res}_\infty\big((f|_k T_n^\infty) \cdot h\big) = \mathrm{Res}_\infty\big(f \cdot (h|_{2-k} T_n^\infty)\big) \qquad (f, h \in \mathbb{C}[q^{-1}, q]]) \tag{20}$$

and

$$(gE_2)|_2 T_n^\infty = (g|_0 T_n) \cdot E_2 \pmod{V_2} \qquad (g \in V_0), \tag{21}$$

where $\mathrm{Res}_\infty(F)$ for a 1-periodic holomorphic function F on $\mathcal{H}$ denotes the residue at infinity of $2\pi i F(\tau)d\tau$, i.e., the constant term of F as a Laurent series in q. Theorem 2 then follows using description (iii) of the Atkin scalar product and the fact that $VV_2 \subseteq V_2$ and that Res_∞ vanishes on V_2:

$$\begin{aligned}(f|_0T_n,\, g) &= \mathrm{Res}_\infty\big((f|_0T_n^\infty)\cdot g\cdot E_2\big) = \mathrm{Res}_\infty\big(f\cdot(gE_2)|_2T_n^\infty\big)\\ &= \mathrm{Res}_\infty\big(f\cdot(g|_0T_n)\cdot E_2\big) = (f,\, g|_0T_n) \qquad (f,\, g\in V)\,.\end{aligned}$$

To prove (20), we note that T_n^∞ acts on Fourier series by

$$\Big(\sum_r A_r\, q^r\Big)|_kT_n^\infty \;=\; n^{k/2}\sum_{ad=n} d^{1-k}\sum_r A_{rd}\, q^{ar}$$

and consequently that

$$\begin{aligned}\mathrm{Res}_\infty\big((f|_kT_n^\infty)\,h\big) &= n^{k/2}\sum_{ad=n}\sum_r d^{1-k}A_{dr}B_{-ar}\\ &= n^{1-k/2}\sum_{ad=n} a^{-1+k}\sum_s B_{as}A_{-ds} = \mathrm{Res}_\infty\big(f\,(h|_{2-k}T_n^\infty)\big)\end{aligned}$$

for $f=\sum A_rq^r$, $h=\sum B_sq^s$. For (21), we use the well-known fact, equivalent to the transformation equation (8) in §3, that $E_2(\tau) = E_2^*(\tau) + \dfrac{3}{\pi y}$, where $y=\Im(\tau)$ and the non-holomorphic function $E_2^*(\tau)$ transforms like a modular form of weight 2. Denoting by V_2^* the space of functions with the last property, and observing that $VV_2^*\subseteq V_2^*$ and that $|_2T_n$ preserves V_2^*, we have

$$(g\,E_2)|_2T_n^\infty - (g|_0T_n)\,E_2 \;\equiv\; \frac{3}{\pi}\,\big((g\,y^{-1})|_2T_n^\infty - (g|_0T_n)\,y^{-1}\big) \pmod{V_2^*}\,.$$

The right-hand side of this formula vanishes by virtue of the calculation

$$\big((g\,y^{-1})|_2T_n^\infty\big)(\tau) = \sum_{\substack{ad=n\\ b\ (\mathrm{mod}\ d)}} \frac{n}{d^2}\,g\Big(\frac{a\tau+b}{d}\Big)\,\Im\Big(\frac{a\tau+b}{d}\Big)^{-1} = y^{-1}\,(g|_0T_n)(\tau)\,,$$

so the left-hand side, which is holomorphic, belongs to V_2 as claimed.

The uniqueness clause in the theorem is easy to prove. Any functional $\phi: V\to\mathbb{C}$ as in Theorem 2 annihilates the polynomials $h_n = j|T_n\cdot 1 - j\cdot 1|T_n$ $(n\geq 2)$ and $h^* = j^2|T_2\cdot j - j^2\cdot j|T_2$. But these polynomials span a codimension 1 subspace of V, since $\deg h_n = n$ and h^* is not a linear combination of the h_n's, so ϕ is unique.

Remark One can prove in the same way the more general adjunction formula

$$(f|_kT_n,\, g) = (f,\, g|_{-k}T_n) \qquad (f\in V_k,\ g\in V_{-k}),$$

where the pairing $(\ ,\ ): V_k\otimes V_{-k}\to\mathbb{C}$ is defined by $(f,g)=\mathrm{Res}_\infty(f\cdot g\cdot E_2)$.

6 "Modular" proof of Theorem 3

We saw in the last section that the Atkin polynomials are the denominators of the best rational approximations to the function $\Phi = E_2E_4/E_6j$, considered as a function of j near infinity. If $p \geq 5$ is a prime, then we know that the Eisenstein series E_{p-1} and E_{p+1} are congruent (as power series in q and therefore also as power series in $1/j$) to the constant function 1 and to the Eisenstein series E_2, respectively, so we can replace Φ modulo p by the modular form $E_{p+1}E_4/E_{p-1}E_6j$, which has weight 0 and hence is itself a rational function of j. This rational function is then a perfect, and hence certainly a best possible, approximation to itself, so its denominator must be (the mod p reduction of) the corresponding Atkin polynomial. But this denominator is essentially $\widetilde{E}_{p-1}$, and hence essentially the supersingular polynomial for p by the result of §3. We now give the details of this argument.

Let p be a prime ≥ 5 (Theorem 3 is trivial for $p = 2$ or 3), and define m, δ, and ε by (6). They coincide with the m, δ, and ε of (1) for weight $k = p-1$, while the corresponding numbers for weight $p+1$ are $m+\delta+\varepsilon-1$, $2(1-\delta)$, and $1-\varepsilon$, respectively. Equation (2) for the Eisenstein series E_{p-1} and E_{p+1} therefore becomes

$$E_{p-1} = \Delta^m E_4^\delta E_6^\varepsilon \widetilde{E}_{p-1}\,, \qquad E_{p+1} = \Delta^{m+\delta+\varepsilon-1} E_4^{2-2\delta} E_6^{1-\varepsilon} \widetilde{E}_{p+1}\,,$$

so

$$\Phi = \frac{E_2E_4}{E_6j} \equiv \frac{E_{p+1}E_4}{E_{p-1}E_6j} = \frac{\widetilde{E}_{p+1}}{j^\delta(j-1728)^\varepsilon \widetilde{E}_{p-1}} \pmod p\,.$$

The right-hand side here is a rational function whose denominator is $ss_p(j) \pmod p$ by the result of §3, while the numerator is a polynomial of degree $m+\delta+\varepsilon-1 = n_p-1$. On the other hand, Φ is equal to $B_n(j)/A_n(j) + O(j^{-2n-1})$ for any n by equation (14), where A_n is the nth Atkin polynomial and B_n a certain polynomial of degree $n-1$. Taking $n = n_p$ and multiplying A_n and B_n if necessary by a common power of p to make A_{n_p} p-integral and primitive modulo p, we obtain

$$\frac{\widetilde{E}_{p+1}(j)}{ss_p(j)} \equiv \Phi \equiv \frac{\bar{B}_{n_p}}{\bar{A}_{n_p}} + O\big(j^{-2n_p-1}\big) \pmod p$$

where $\bar{A}_{n_p}$ and $\bar{B}_{n_p}$ are polynomials over $\mathbb{F}_p$ of degree $\leq n_p$ and $\leq n_p-1$, respectively. Multiplying this identity by $\bar{A}_{n_p}(j)ss_p(j)$, we find that the expression $\bar{B}_{n_p}ss_p - \bar{A}_{n_p}\widetilde{E}_{p+1}$ is $O(j^{-1})$ as $j \to \infty$ and hence, since it is a polynomial, vanishes. If we show that $\widetilde{E}_{p+1}(j)$ is prime to $ss_p(j)$, then it follows that ss_p divides $\bar{A}_{n_p}(j) \pmod p$ and hence, since A_{n_p} is monic of degree n_p, that A_{n_p} indeed has p-integral coefficients and reduces to ss_p mod p. This coprimality assertion is a consequence (since ss_p has no multiple roots) of the identity

$$\widetilde{E}_{p+1}(j) \equiv -12\,ss_p'(j) + 8\delta\,\frac{ss_p(j)}{j} + 6\varepsilon\frac{ss_p(j)}{j-1728}\,,$$

which in turn follows from the formulas

$$12\vartheta_{p-1}E_{p-1} = 12q\frac{d}{dq}E_{p-1} - (p-1)E_2E_{p-1} \equiv E_2 \equiv E_{p+1} \pmod p,$$

$$(\vartheta_{p-1}E_{p-1})^{\sim} = -j^{\delta}(j-1728)^{\varepsilon}\left(\frac{\delta}{3j} + \frac{\varepsilon}{2(j-1728)} + \frac{d}{dj}\right)\widetilde{E}_{p-1}.$$

This completes the proof.

We remark that the last steps of the proof could have been simplified if we had used the formula (19), which will be proved in the next section, for then it would follow

(i) that the coefficients in (18) are p-integral for $n \le n_p$, and consequently that $A_{n_p}(j)$ is a polynomial with p-integral coefficients and can be reduced modulo p, and

(ii) that the mod p reductions of A_{n_p} and B_{n_p} are coprime (because $A_{n_p}B_{n_p-1} - B_{n_p}A_{n_p-1} = \pm b_1 \cdots b_{n_p-1} \not\equiv 0 \pmod p$).

It then follows by an argument like the one above that the reduction of A_{n_p} divides, and hence, since it is monic, equals, ss_p.

7 Hypergeometric aspects of $A_n(j)$ and $ss_p(j)$

Recall the definition of the classical Gauss hypergeometric series $F = {}_2F_1$:

$$F(a,b;c;x) = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}x^n = \sum_{n=0}^{\infty}\frac{\binom{-a}{n}\binom{-b}{n}}{\binom{-c}{n}}(-x)^n \qquad (|x|<1)$$

where $(a)_n$ denotes the "ascending factorial" $a(a+1)\cdots(a+n-1)$. (The minus signs in the second formula are because Gauss chose to use ascending rather than descending factorials.) If $c = -k$ is a non-positive integer, then $F(a,b;c;x)$ is not defined, since all terms after the kth have infinite coefficients; if instead a or b is a non-positive integer, then $F(a,b;c;x)$ is a polynomial. In this section we will relate the Atkin polynomials to hypergeometric and truncated hypergeometric series, and use this relationship to give a second proof of Theorem 3 and to prove the various formulas for $A_n(j)$ given in Theorem 4 of the Introduction.

We first define four monic polynomials U_n^{ε}, V_n^{δ} of every degree $n \ge 0$ by the formulas

$$\begin{aligned}
j^n\,F\left(\tfrac{1}{12},\tfrac{5}{12};1;\tfrac{1728}{j}\right) &= U_n^0(j) + O(1/j)\\
j^{n-1}\,(j-1728)\,F\left(\tfrac{7}{12},\tfrac{11}{12};1;\tfrac{1728}{j}\right) &= U_n^1(j) + O(1/j)\\
(j-1728)^n\,F\left(\tfrac{1}{12},\tfrac{7}{12};1;\tfrac{1728}{1728-j}\right) &= V_n^0(j) + O(1/j)\\
j\,(j-1728)^{n-1}\,F\left(\tfrac{5}{12},\tfrac{11}{12};1;\tfrac{1728}{1728-j}\right) &= V_n^1(j) + O(1/j)
\end{aligned}$$

as $j \to \infty$ (truncated hypergeometric functions).

We will prove two results about these polynomials: that the Atkin polynomials can be expressed as linear combinations of them, and that their reductions modulo primes give the supersingular polynomials. Together, these give a second proof of the

relation between the Atkin and the supersingular polynomials. But now we will take the recursion relation (4) rather than the orthogonality property with respect to the Atkin scalar product as the defining property of A_n, so that the new proof is simpler or more complicated than the first one we gave depending on which definition of the A_n's one considers to be the more fundamental one.

Proposition 4 *The functions $A_n(j)$ $(n \geq 0)$ defined by the recurrence and initial conditions given in part (i) of Theorem 4 have the following expressions in terms of the polynomials introduced above:*

$$A_n(j) = \sum_{m=0}^{n} (-12)^{3m} \binom{n+\frac{1}{12}}{m} \binom{n-\frac{7}{12}}{m} \binom{2n-1}{m}^{-1} U^0_{n-m}(j),$$

$$A_n(j) = \sum_{m=0}^{n} (-12)^{3m} \binom{n-\frac{5}{12}}{m} \binom{n-\frac{13}{12}}{m} \binom{2n-1}{m}^{-1} U^1_{n-m}(j),$$

$$A_n(j) = \sum_{m=0}^{n} 12^{3m} \binom{n+\frac{1}{12}}{m} \binom{n-\frac{5}{12}}{m} \binom{2n-1}{m}^{-1} V^0_{n-m}(j),$$

$$A_n(j) = \sum_{m=0}^{n} 12^{3m} \binom{n-\frac{7}{12}}{m} \binom{n-\frac{13}{12}}{m} \binom{2n-1}{m}^{-1} V^1_{n-m}(j).$$

Proof We will prove only the first of these formulas (which is the same as the formula given in part (ii) of Theorem 4), the other cases being similar. Denote the expression on the right of this formula by A^0_n. The equality $A^0_n = A_n$ is checked directly for $n \leq 2$, so we must prove the recursion $A^0_{n+1} = (j - a_n)A^0_n - b_n A^0_{n-1}$, where a_n and b_n are the rational functions of n occurring in formula (4). We can rewrite the definition of A^0_n as $A^0_n = \sum_{k=0}^{n} c(n,k) U^0_k$ with

$$c(n,0) = (-12)^{3n} \binom{-5/12}{n} \binom{-13/12}{n} \binom{2n-1}{n}^{-1},$$

$$c(n,k) = c(n,0) \cdot 12^{-3k} \binom{n}{k} \binom{-n}{k} \binom{-5/12}{k}^{-1} \binom{-13/12}{k}^{-1}.$$

Then, noting that $jU^0_k = U^0_{k+1} - 12^{3k+3}\binom{-1/12}{k+1}\binom{-5/12}{k+1}$, we find

$$\begin{aligned} &A^0_{n+1}(j) - (j - a_n)\, A^0_n(j) + b_n\, A^0_{n-1}(j) \\ &= \sum_{k=0}^{n} \left[c(n+1,k) - c(n,k-1) + a_n\, c(n,k) + b_n\, c(n-1,k)\right] U^0_k \\ &\quad + \sum_{k=0}^{n} 12^{3k+3} \binom{-1/12}{k+1} \binom{-5/12}{k+1} c(n,k) \end{aligned} \tag{22}$$

for $n \geq 2$, where we have set $c(n,-1) = 0$ and used that $c(n,n) = 1$, $c(n-1,n) = 0$. We can check directly that the coefficient of U^0_k on the right equals 0 for $k \geq 1$ and

equals $84c(n,0)/(n^2-1)$ for $k=0$. Substituting the value $U_0^0=1$ and the values of a_n, b_n and $c(n,k)$ and writing $k=n-m$, we then find that the right-hand side of (22) equals

$$\frac{12c(n,0)}{n^2-1}\sum_{m=0}^{n+1}\binom{-n+1}{n+1-m}\left[7\binom{n+1}{m}-12(n+1)\binom{n}{m}\right],$$

where the "extra" term $m=n+1$ comes from the multiple of U_0^0 in (22). This last sum is just the coefficient of x^{n+1} in $(1+x)^{1-n}[7(1+x)^{n+1}-12(n+1)(1+x)^n]$ and hence vanishes for $n\geq 2$.

□

Remark The formulas in Proposition 4 can be inverted, e.g., we have

$$U_n^0(j)=\sum_{m=0}^{n}12^{3m}\binom{n+\frac{1}{12}}{m}\binom{n-\frac{7}{12}}{m}\binom{2n-m}{m}^{-1}A_{n-m}(j).$$

Proposition 5 *Let $p\geq 5$ be a prime number and write p as $12n-8\delta-6\varepsilon+1$ with $n\in\mathbb{N}$, $\delta,\varepsilon\in\{0,1\}$. Then*

$$ss_p(j)\equiv U_n^\varepsilon(j)\equiv V_n^\delta(j)\pmod p.$$

Proof Again we treat only the case of U_n^0 in detail, the other cases being similar. So assume that $p\equiv 1\pmod 4$; then we want to show that $ss_p(j)\equiv U_n^0(j)\pmod p$. Write $p=4l+1$. Expanding H_{p-1} by the trinomial theorem, we have

$$\begin{aligned}H_{p-1}&=\text{coefficient of } X^{p-1}\text{ in }\left(1-3E_4X^4+2E_6X^6\right)^{2l}\\&=\sum_{\substack{r,s\geq 0\\2r+3s=2l}}\frac{(2l)!}{r!\,s!\,(2l-r-s)!}\left(-3E_4\right)^r(2E_6)^s\\&=\left(-3E_4\right)^l\sum_{k=0}^{[l/3]}\frac{(2l)!}{(l-3k)!\,(2k)!\,(l+k)!}\left(-\frac{4}{27}\,\frac{j-1728}{j}\right)^k,\end{aligned}$$

where $k=s/2$ and we have used that $E_6^2/E_4^3=(j-1728)/j$. (The calculation with U_n^1 in the case $p\equiv 3\pmod 4$ would be similar but with $s=2k+1$, while to obtain the results for V_n^δ we would instead set $r=3k+\delta$ and then expand in powers of $j/(j-1728)$.) But one checks easily, either directly or by induction on k, that

$$\frac{(2l)!}{(l-3k)!\,(2k)!\,(l+k)!}\left(-\frac{4}{27}\right)^k\equiv\binom{2l}{l}\frac{(\frac{1}{12})_k\,(\frac{5}{12})_k}{k!\,(\frac{1}{2})_k}\pmod p,$$

so using (11) and noting that $[l/3]=m=n-\delta$ and $(-1)^\delta=(-3/p)\equiv 3^{2l}\pmod p$, we find (writing $F_m(a,b;c;x)$ for the hypergeometric series truncated at degree m)

$$ss_p(j)\equiv(-j)^\delta\,\widetilde{H}_{p-1}(j)\equiv(-3)^{3l}\binom{2l}{l}j^n\,F_{[l/3]}\left(\tfrac{1}{12},\tfrac{5}{12};\tfrac{1}{2};1-\tfrac{1728}{j}\right)\pmod p.$$

Now, taking into account that the coefficients of x^k and y^k in $F(\frac{1}{12}, \frac{5}{12}; 1; x)$ and $F(\frac{1}{12}, \frac{5}{12}; \frac{1}{2}; y)$ vanish modulo p if k is in the range $[l/3] < k \leq 2l$ and that both $F(\frac{1}{12}, \frac{5}{12}; 1; x)$ and $F(\frac{1}{12}, \frac{5}{12}; \frac{1}{2}; 1-x)$ satisfy the same second order linear differential equation with polynomial coefficients of degree at most 2, we can conclude that the polynomial on the right-hand of the last formula is a multiple of $U_n^0(j)$. This multiple must then be 1 because the supersingular polynomial is monic.

□

"Hypergeometric" proof of Theorem 3 The theorem is trivial for $p = 2$ or 3. Otherwise n_p is the same as the number n in the proposition. Applying Proposition 4 to U_n^ε or to V_n^δ with this value of n immediately gives the desired result, since all coefficients except the one for $m = 0$ vanish modulo p. So we actually get *two* proofs.

□

Proof of Theorem 4 i) We need to prove only the explicit formula (19) for the coefficients of the continued fraction expansion of $\Phi = E_2E_4/E_6j$ with respect to $1/j$, since then (4) reduces to equation (18), which we have already proved. From the formulas (9) we immediately find that $\Phi = -d(\log \Delta)/dj$. On the other hand, Δ can be expressed hypergeometrically in terms of j by the formula $\Delta = \frac{1}{j} F(\frac{1}{12}, \frac{5}{12}; 1; \frac{1728}{j})^{12}$. It follows with the aid of Gauss's contiguous relations that $\Phi = F(\frac{13}{12}, \frac{5}{12}; 1; \frac{1728}{j})/jF(\frac{1}{12}, \frac{5}{12}; 1; \frac{1728}{j})$, and (19) now follows from Gauss's well-known formula for the continued fraction expansion of a quotient of contiguous hypergeometric functions, as given in [2].

ii) We have just shown that the Atkin polynomials are indeed the A_n of Proposition 4. The closed formula given in Theorem 4 is then just a rewriting of the first formula of that proposition, as already mentioned.

iii) The closed formula of part (ii) is equivalent to saying that $A_n(j)$ can be obtained by truncating the product

$$F(\tfrac{1}{12}, \tfrac{5}{12}; 1; x)\, F(-n - \tfrac{1}{12}, -n + \tfrac{7}{12}; 1 - 2n; x) \tag{23}$$

at x^n and inverting it (i.e., set $x = 1728/j$ and multiply by j^n). Notice that some truncation is necessary, since the coefficients of the second factor in (23) become infinite from degree $2n$ onwards. In fact we can truncate at x^m for any m between n and $2n-1$, since the coefficient of x^i in (23) vanishes for $n < i < 2n$, as can be seen by letting $\gamma \to 1$ in the following identity, which is a consequence of Gauss's contiguous relation and two formulas of Heine ([4], proved in [5]):

$$\begin{aligned} &F(\alpha, \beta; \gamma; x)\, F(-n - \alpha, -n + 1 - \beta; -2n + 2 - \gamma; x) \\ &+ \delta_n\, x^{2n}(1 - x)\, F(1 - \alpha, 1 - \beta; 2 - \gamma; x)\, F(\alpha + n + 1, \beta + n; \gamma + 2n; x) \\ &= \text{ polynomial of degree } n \qquad \Big(n \geq 1, \quad \delta_n = \frac{\binom{-\alpha}{n+1}\binom{-\beta}{n}\binom{\alpha-\gamma}{n-1}\binom{\beta-\gamma}{n}}{\binom{-\gamma}{2n}\binom{1-\gamma}{2n}\binom{2n}{n}\binom{2n}{n-1}}\Big). \end{aligned}$$

The differential equation for A_n now follows from this truncation argument and the fact that the product of two hypergeometric functions (or of any two functions satisfying linear differential equations of second order) satisfies a fourth order linear differential equation whose coefficients can be calculated by an explicit procedure. For the uniqueness, we observe that if the function $A_n(j)$ is replaced by a polynomial beginning j^d for some integer $d \geq 0$, then the left-hand side of the differential equation in (iii) has

leading term $n^2(n^2-d^2)^2$, so the differential equation can be satisfied only if $d=n$. □

As a corollary to Theorem 4, we have

Proposition 6 (Atkin) *The scalar product of A_n with itself and its special values at $j=0$ and $j=1728$ are given for $n \geq 1$ by*

$$(A_n, A_n) = -12^{6n+1} \frac{(-1/12)_n\,(5/12)_n\,(7/12)_n\,(13/12)_n}{(2n-1)!\,(2n)!},$$

$$A_n(0) = (-12)^{3n+1} \frac{(-1/12)_n\,(5/12)_n}{(2n-1)!},$$

$$A_n(1728) = -12^{3n+1} \frac{(-1/12)_n\,(7/12)_n}{(2n-1)!}.$$

Proof The values at $j=0$ and $j=1728$ can be checked directly from the recursion (4), while the formula for the scalar product follows from the recursion and part (i) of the Proposition 2, §4. The fact that $(A_n, A_n) > 0$ for all n gives a second proof of the fact, already mentioned in §5, that Atkin's scalar product is positive definite. From the explicit formula for (A_n, A_n) and Stirling's formula we see that the length of A_n in the Atkin norm is asymptotically equal to $\sqrt{6/\pi}\,432^n$ as $n \to \infty$.

8 Hypergeometric properties of F_k.

Recall that the modular form $F_k(\tau)$ for $k \not\equiv 2 \pmod 3$ is the unique normalized solution of the second order differential equation

$$\vartheta_{k+2}\,\vartheta_k\,F_k - \frac{k(k+2)}{144}\,E_4\,F_k = 0, \tag{24}$$

the normalization being given as in (12). We introduce the following notations:

$$\nu_0 = \frac{1-2\delta}{3}, \quad \nu_1 = \frac{1-2\varepsilon}{2}, \quad \nu_\infty = \frac{k+1}{6} \qquad (\nu_0+\nu_1+\nu_\infty = 2m+1),$$

$$X_0 = J = \frac{j}{1728}, \quad X_1 = 1-J, \quad X_\infty = -1 \qquad (X_0+X_1+X_\infty = 0),$$

$$Y_0 = E_4^3, \quad Y_1 = -E_6^2, \quad Y_\infty = -1728\,\Delta \qquad (Y_0+Y_1+Y_\infty = 0),$$

where m, δ and ε are associated to k by (1) as usual. The following theorem gives various explicit descriptions of the F_k and their associated polynomials $\widetilde{F}_k(j)$, similar to those given earlier for A_n, $\widetilde{G}_k$, and $\widetilde{H}_k$.

Theorem 5 *Suppose $k \geq 0$, $k \not\equiv 2 \pmod 3$. Then we have:*

i) *Differential equation: $\widetilde{F}_k(j)$ is the unique normalized polynomial solution of*

$$j(j-1728)\,\widetilde{F}_k'' + \{(1-\nu_1)\,j + (1-\nu_0)(j-1728)\}\,\widetilde{F}_k' + m(m-\nu_\infty)\,\widetilde{F}_k = 0.$$

ii) *Closed formulas: Let σ be any permutation of $\{0,1,\infty\}$. Then*

$$\widetilde{F}_k(j) = (sgn(\sigma)\cdot 1728)^m \binom{m-\nu_{\sigma(\infty)}}{m} X_{\sigma(0)}^m \, F\Big(-m,-m+\nu_{\sigma(0)};1-\nu_{\sigma(\infty)};-\frac{X_{\sigma(\infty)}}{X_{\sigma(0)}}\Big)$$

and

$$F_k(\tau) = sgn(\sigma)^m \, E_4^\delta \, E_6^\varepsilon \sum_{l=0}^{m} (-1)^l \binom{m-\nu_{\sigma(0)}}{l} \binom{m-\nu_{\sigma(\infty)}}{m-l} Y_{\sigma(\infty)}^l \, Y_{\sigma(0)}^{m-l} .$$

iii) *Recursion relation: The $\widetilde{F}_k(j)$ satisfy*

$$\begin{aligned}&(m+1)(m-\nu_\infty)(1-\nu_\infty)\,\widetilde{F}_{k+12}\\ &\quad -\nu_\infty\left[(1+\nu_\infty)(1-\nu_\infty)\,j - 1728((1-\nu_0)(\nu_0+\nu_1)+2m(m-\nu_\infty))\right]\widetilde{F}_k\\ &\quad +1728^2(m-\nu_0)(m-\nu_1)(1+\nu_\infty)\,\widetilde{F}_{k-12} = 0 \qquad (k\geq 12).\end{aligned}$$

iv) *"Generating function": For $k\in\mathbb{Z}_{\geq 0}$ and any α denote by $G_{k,\alpha}(\tau)$ the coefficient of X^k in $(1-3E_4(\tau)\,X^4+2E_6(\tau)\,X^6)^\alpha$. Then*

$$F_k(\tau) = (-1)^{m+\delta}\, 2^{-2m-\varepsilon} \binom{2m+\varepsilon}{m} \binom{\frac{1}{6}(k-2)}{m+\varepsilon}^{-1} G_{k,\frac{k-2}{6}}(\tau).$$

Proof i) We can easily transform the equation (24) into the one in terms of j, by using formulas (9) and the relation $F_k = \Delta^m E_4^\delta E_6^\varepsilon \widetilde{F}_k$. The uniqueness follows from the same argument used in the proof of Theorem 4, iii) in §7.

ii) The equation in i) is in fact a hypergeometric differential equation and has a polynomial solution $F(-m,-m+\nu_\infty;1-\nu_0;j/1728)$. We have 6 polynomial solutions out of Kummer's 24 solutions to this equation. By the uniqueness, they differ only by constant factors. Making the expressions symmetric, we obtain the formula in the theorem. The formula for F_k follows immediately.

iii) The first three arguments of the hypergeometric series in the formula for $\widetilde{F}_k(j)$ will change by 1 if we replace k by $k\pm 12$. The recursion is therefore a consequence of Gauss's contiguous relations.

iv) Put $Y_\alpha = (1-3E_4X^4+2E_6X^6)^\alpha$. From the relations

$$\begin{aligned}&Y_\alpha = Y_{\alpha-1}\cdot(1-3E_4X^4+2E_6X^6),\\ &\frac{\partial}{\partial X}Y_\alpha = \alpha\, Y_{\alpha-1}\,(-12E_4X^3+12E_6X^5),\\ &\sum_{k=0}^{\infty}\vartheta_k G_{k,\alpha}\,X^k = \frac{1}{2\pi i}\frac{\partial}{\partial\tau}Y_\alpha - \frac{E_2}{12}\,X\,\frac{\partial}{\partial X}\,Y_\alpha = \alpha\, Y_{\alpha-1}\,(E_6\,X^4 - E_4^2\,X^6)\end{aligned}$$

we obtain respectively

$$G_{k,\alpha} = G_{k,\alpha-1} - 3\,E_4\,G_{k-4,\alpha-1} + 2\,E_6\,G_{k-6,\alpha-1}\,, \tag{25}$$

$$k\,G_{k,\alpha} = -12\,\alpha\,E_4\,G_{k-4,\alpha-1} + 12\,\alpha\,E_6\,G_{k-6,\alpha-1}\,, \tag{26}$$

$$\vartheta_k\,G_{k,\alpha} = \alpha\,E_6\,G_{k-4,\alpha-1} - \alpha\,E_4^2\,G_{k-6,\alpha-1}\,. \tag{27}$$

Solving (25) and (26) for $E_4 G_{k-4,\alpha-1}$ and $E_6 G_{k-6,\alpha-1}$ and substituting the expressions obtained into (27), we get

$$\vartheta_k G_{k,\alpha} = -\frac{1}{\alpha+1}\left(\alpha - \frac{k}{6} + \frac{1}{3}\right)\left(\alpha - \frac{k}{6} + \frac{2}{3}\right) G_{k+2,\alpha+1} + \left(\alpha - \frac{k}{12} + \frac{1}{2}\right) G_{k+2,\alpha}\,.$$

Using this repeatedly we finally obtain

$$\begin{aligned}
&\vartheta_{k+2}\,\vartheta_k\, G_{k,\alpha} - \frac{k(k+2)}{144} E_4\, G_{k,\alpha} \\
&= \left(\alpha - \frac{k}{6} + \frac{1}{3}\right)\left[\frac{1}{(\alpha+1)(\alpha+2)}\left(\alpha - \frac{k}{6} + \frac{2}{3}\right)\left(\alpha - \frac{k}{6} + 1\right)\left(\alpha - \frac{k}{6} + \frac{4}{3}\right) G_{k+4,\alpha+2}\right. \\
&- \frac{1}{\alpha+1}\left\{\left(\alpha - \frac{k}{6} + \frac{2}{3}\right)\left(\alpha - \frac{k}{12} + \frac{4}{3}\right) + \left(\alpha - \frac{k}{6}\right)\left(\alpha - \frac{k}{12} + \frac{1}{2}\right) - \frac{k(k+2)}{144}\right\} G_{k+4,\alpha+1} \\
&\left. + \left(\alpha + \frac{1}{2}\right) G_{k+4,\alpha}\right]\,.
\end{aligned}$$

Since the right-hand side vanishes if $\alpha = (k-2)/6$ we deduce that $G_{k,(k-2)/6}(\tau)$ satisfies the same differential equation as $F_k(\tau)$. The constant term of the Fourier expansion of $G_{k,(k-2)/6}$ equals the coefficient of X^k in $(1-3X^4+2E^6)^{(k-2)/6} = (1-X^2)^{(k-2)/3}(1+2X^2)^{(k-2)/6}$, which in turn equals

$$\sum_{i=0}^{k/2} (-1)^i\, 2^{\frac{k}{2}-i} \binom{\frac{k-2}{3}}{i} \binom{\frac{k-2}{6}}{\frac{k}{2}-i} = (-1)^{\frac{k}{2}} \binom{\frac{k-2}{3}}{\frac{k}{2}} \sum_{i=0}^{k/2} 2^i \binom{\frac{k}{2}}{i} = (-3)^{k/2} \binom{\frac{k-2}{3}}{\frac{k}{2}}\,.$$

This together with (12) gives $F_k(\tau) = c\, G_{k,\frac{k-2}{6}}(\tau)$ with

$$c = (-1)^m \binom{\frac{k-5}{6}}{m} \Big/ (-3)^{\frac{k}{2}} \binom{\frac{k-2}{3}}{\frac{k}{2}} = (-1)^{m+\delta}\, 2^{-2m-\varepsilon} \binom{2m+\varepsilon}{m} \binom{\frac{k-2}{6}}{m+\varepsilon}^{-1}\,.$$

□

Remarks 1. The symmetry in the closed formula ii) is the reason why we chose the normalization (12) of F_k.

2. Part iv) of the theorem makes it clear why the modular forms H_{p-1}, G_{p-1} and F_{p-1} in Theorem 1 are (up to scalar factors) congruent to one another modulo p (up to scalar factors): they are just the specializations of $G_{p-1,\alpha}$ to the three values $\alpha = -\frac{1}{2}$, $\frac{p-1}{2}$ and $\frac{p-3}{6}$, which are the same modulo p.

3. In the notation of [10], the formula for F_k can be written as

$$F_k = \mathrm{sgn}(\sigma)^m\, E_4^\delta\, E_6^\varepsilon\, H_m(1-\nu_{\sigma(\infty)}, 1-\nu_{\sigma(0)}; Y_{\sigma(\infty)}, Y_{\sigma(0)}),$$

where $H_n(k,l;X,Y)$ is the polynomial $\sum_{r+s=n}(-1)^r \binom{n+k-1}{s}\binom{n+l-1}{r} X^r Y^s$, which satisfies the "hidden symmetry" that it is ± 1-symmetric in the three variables (k,X), (l,Y), (m,Z), where $k+l+m = n-2$, $X+Y+Z=0$. (This polynomial is essentially the Wigner 3J-symbol of quantum mechanics and arose in [10] in connection with the Cohen bracket operation on modular forms.)

We now interpret the $\widetilde{F}_k$ as orthogonal polynomials. Consider the space $W = \mathbb{C}[j^{1/3}, (j-1728)^{1/2}]$ which is identified with the space of holomorphic functions on $\mathcal{H}$ invariant under the commutator subgroup $[\Gamma,\Gamma]$ of $\Gamma = PSL(2,\mathbb{Z})$ and growing at most like a polynomial in q^{-1}. For each residue class r modulo 6, denote by χ^r the character of the cyclic group $\Gamma/[\Gamma,\Gamma]$ determined uniquely by $\chi\left(\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right) \bmod [\Gamma,\Gamma]\right) = e^{\pi i r/3}$, and let $W = \bigoplus_{r \bmod 6} W(r)$ be the corresponding decomposition of W, i.e., $W(r)$ is the χ^r-eigenspace with respect to the action of Γ. We note that if $\delta \in \{0,1,2\}$ and $\varepsilon \in \{0,1\}$ are determined by the congruence $2r \equiv 4\delta + 6\varepsilon \pmod{12}$, then $W(r)$ is identified with $j^{\delta/3}(j-1728)^{\varepsilon/2}\mathbb{C}[j]$. Now we define a scalar product on W by the formula

$$(f,g) = \int_0^{1728} \frac{f(j)\,g(j)\,dj}{j^{1/3}\,(1728-j)^{1/2}}, \qquad (f,g \in W). \tag{28}$$

On each subspace $W(r)$, the scalar product induced by (28) is positive or negative definite on $j^{\delta/3}(j-1728)^{\varepsilon/2}\mathbb{R}[j]$ depending whether $\varepsilon = 0$ or 1. Hence we have six families of monic polynomials $\{f_m^{(r)}\}_{m\geq 0}$, $f_m^{(r)}$ being of degree m, such that the $j^{\delta/3}(j-1728)^{\varepsilon/2} f_m^{(r)}$ are orthogonal with respect to this scalar product. On the other hand, for each even residue class $2r \equiv 4\delta + 6\varepsilon \pmod{12}$ and $m \geq 0$, put

$$\widehat{F}_m^{(r)}(j) = j^m\, F\left(-m, -m+\nu_0; 1-\nu_\infty; \frac{1728}{j}\right)$$

where $\nu_0 = \frac{1}{3}(1-2\delta)$ and $\nu_\infty = \frac{1}{6}(12m+4\delta+6\varepsilon+1)$. If $2r \not\equiv 2 \bmod 3$ and $k = 12m+4\delta+6\varepsilon$, this is nothing but $\widetilde{F}_k(j)$, renormalized to be monic.

Theorem 6 *$f_n^{(r)} = \widehat{F}_n^{(r)}$ for any r mod 6 and $m \geq 0$.*

Proof For fixed r, the $\widehat{F}_m^{(r)}$ satisfy the following recursion which is equivalent to part iii) of Theorem 5 if $2r \not\equiv 2 \pmod 3$:

$$\widehat{F}_{m+1}^{(r)}(j) = \big(j - (\lambda_{2m} + \lambda_{2m+1})\big)\,\widehat{F}_m^{(r)}(j) - \lambda_{2m-1}\lambda_{2m}\,F_{m-1}^{(r)} \qquad (m \geq 1)$$

where

$$\lambda_n = 12\left(6 - (-1)^n \frac{3-6\nu_0}{n-\nu_0-\nu_1}\right)\left(6 - (-1)^n \frac{3-6\nu_0}{n+1-\nu_0-\nu_1}\right)$$

($\nu_1 = (1-2\varepsilon)/2$). By the general theory reviewed in §4, the $\widehat{F}_m^{(r)}$ are orthogonal with respect to the scalar product whose values $g_n = (j^n, 1)$ are given by the continued fraction (16). This continued fraction is equal to $g_0\,F(1, 1-\nu_0; 2-\nu_0-\nu_1; 1728/j)$ by [2]. Hence the $(j^n,1)$ are given by $(j^n,1) = g_0 \cdot 1728^n \binom{\nu_0-1}{n} / \binom{\nu_0+\nu_1-2}{n}$, which is a constant multiple of $\int_0^{1728} j^{n-\nu_0}(1728-j)^{-\nu_1}dj$ (beta function). This latter integral is just $(-1)^\varepsilon$ times the inner product $(j^{\delta/3}(j-1728)^{\varepsilon/2}j^n,\ j^{\delta/3}(j-1728)^{\varepsilon/2}\cdot 1)$ in (28). The theorem follows.

□

Remark Up to a rescaling by 1728, our polynomial is essentially a Jacobi polynomial. These are polynomials $P_n^{(\alpha,\beta)}$ generalizing the Chebyshev polynomials, which have parameters $(\frac{1}{2},\frac{1}{2})$ and come in four types (even and odd, first and second kind), corresponding to the fourfold decomposition of $\mathbb{C}[x^{1/2}, (1-x)^{1/2}]$, whereas our parameters are $(\frac{1}{3},\frac{1}{2})$ and we have six families.

9 The denominators of the Atkin polynomials

From the relation (4) we get by induction that the denominator of $A_{n+1}(j)$ is at most $(2n+1)!(2n)!/2^{2n}n!$, but this is far too large, e.g., the denominator of $A_8(j)$ is only 5 rather than 1380566997810000, and indeed only two of its coefficients, those of j^7 and j^2, have any denominator at all. Similarly, if we fix a prime $p > 3$, then (4) would lead one to expect that p would occur in the denominator of all A_n with $n > p/2$, but instead we find, for instance, that for $p = 11$ all of the A_n with $n \le 150$ are p-integral unless n is congruent to 1 or 6 modulo 11 or n belongs to one of the intervals [62,70] or [123,130]. Finally, if we look at the values of $pA_n(j)$ modulo p for the first values of n for which p occurs in the denominator, then we find a very precise pattern. For instance, for the first such value $n = \frac{1}{2}(p+1)$ we find experimentally

$$p\,A_n(j) \;\equiv\; 0j^n + 84j^{n-1} + 37800j^{n-2} + \cdots + c_r\,j^{n-r-1} + \cdots \qquad (\mathrm{mod}\ p)$$

with coefficients $c_r \in \mathbb{Z}$ independent of p, and with a little numerical work we discover the empirical formula

$$c_r = 12\,(8r+7)\,\frac{(6r+1)!}{(3r)!\,r!\,(r+1)!^2} \qquad (r \ge 0)$$

from which in turn it follows that all of the coefficients of $A_{\frac{1}{2}(p+1)}(j)$ with $r \ge \frac{p-1}{6}$ are actually p-integral. In this section we will describe these phenomena in a little more detail and provide some explanations.

We first generalize the above congruence for $pA_{\frac{1}{2}(p+1)}(j)$. Recall from the discussion of orthogonal polynomials in §4 that in terms of the variable $Y = \sqrt{j}$ the Atkin polynomials are the even members of a sequence of monic polynomials $A_n^*(Y)$ satisfying the recurrence $A_{n+1}^*(Y) = YA_n^*(Y) - \lambda_n A_{n-1}^*(Y)$ $(n \ge 1)$ with λ_n as in (19). Moreover, A_n^* is a polynomial of the same parity as n, so we can write $A_n^*(Y) = Y^n a_n(1/Y^2)$ where $a_n(t)$ is a polynomial of degree $\le n/2$. In terms of the a_n the recursion becomes

$$a_{n+1}(t) = a_n(t) - \lambda_n\,t\,a_{n-1}(t) \qquad (n \ge 1), \tag{29}$$

and the relation to the Atkin polynomials is $A_n(j) = j^n a_{2n}(1/j)$. From the recursion it follows that $a_n(t)$ has p-integral coefficients for $n \le p$, and looking at numerical examples we find empirically that

$$\begin{aligned} a_p(t) &\equiv \Phi_0(t) \quad \mathrm{mod}\ (p,\,t^p)\,, \\ p\,a_n(t) &\equiv \Phi_{n-p}(t) \quad \mathrm{mod}\ (p,\,t^p) \qquad (p < n < 2p), \end{aligned} \tag{30}$$

where the $\Phi_n(t)$ are certain power series independent of p, the first few being

$$\begin{aligned} \Phi_0(t) &= 1 + 120\,t + 83160\,t^2 + 81681600\,t^3 + 93699005400\,t^4 + \cdots\,, \\ \Phi_1(t) &= 84\,t\,\big(1 + 450\,t + 394680\,t^2 + 429557700\,t^3 + \cdots\big)\,, \\ \Phi_2(t) &= 27720\,t^2\,\big(1 + 944\,t + 1054170\,t^2 + 1297994880\,t^3 + \cdots\big)\,, \\ \Phi_3(t) &= 13693680\,t^3\,\big(1 + 1335\,t + 1757970\,t^2 + 2386445040\,t^3 + \cdots\big)\,. \end{aligned} \tag{31}$$

Comparing the recursion (29) and the congruence (30), and using formula (19) for the λ_n, we find that these power series, if they exist at all, must satisfy the recursion

$$\Phi_{n+1}(t) = \Phi_n(t) - \lambda_n^* \, t \, \Phi_{n-1}(t) \qquad (n \geq 1) \tag{32}$$

with

$$\lambda_n^* = \begin{cases} 84 & \text{if } n = 1, \\ 12\left(6 - \dfrac{(-1)^n}{n-1}\right)\left(6 - \dfrac{(-1)^n}{n}\right) & \text{if } n > 1. \end{cases}$$

Moreover, by inspection of the first few coefficients one finds the formulas

$$\Phi_0(t) = \sum_{r=0}^{\infty} \frac{(6r)!}{(3r)!\, r!^3}\, t^r, \qquad \Phi_1(t) = 12 \sum_{r=0}^{\infty} \frac{(8r+7)\,(6r+1)!}{(3r)!\, r!\, (r+1)!^2}\, t^{r+1}. \tag{33}$$

So far, this is all only experimental. To check that it is true, we first find the solution of the recursion (32) with initial conditions (33). We first observe that the formulas (33) are equivalent to

$$\Phi_0(t) = F\left(\tfrac{1}{12}, \tfrac{5}{12}; 1; 1728t\right)^2, \quad \Phi_1(t) = 84tF\left(\tfrac{1}{12}, \tfrac{5}{12}; 1; 1728t\right)F\left(\tfrac{5}{12}, \tfrac{13}{12}; 2; 1728t\right).$$

(To prove this, verify that in each claimed equality both sides have the the same first few terms and satisfy the same third order differential equation.) Now by induction on n and the continued fraction formulas of Gauss already used in the proof of part i) of Theorem 4 we find that the general solution of (32) is given by

$$\Phi_n(t) \;=\; c_n\, t^n\, F\left(\tfrac{1}{12}, \tfrac{5}{12}; 1; 1728t\right) F\left(\left[\tfrac{n}{2}\right] + \tfrac{5}{12}, \left[\tfrac{n+1}{2}\right] + \tfrac{1}{12}; n+1; 1728t\right) \tag{34}$$

with constants $c_n \in \mathbb{Z}$ given by $c_0 = 1$ and

$$c_n = \lambda_1^* \ldots \lambda_n^* = (-1728)^n\, n \binom{\left[\frac{n}{2}\right] - \frac{1}{12}}{n} \binom{\left[\frac{n+1}{2}\right] - \frac{5}{12}}{n} = \frac{(6n+1)!/(6n+(-1)^n)}{(n-1)!(2n)!(3n)!}$$

for $n \geq 1$. In particular, the power series $\Phi_n(t)$ has integral coefficients and is divisible by t^n for all n, properties which were visible in the examples (31) but are not at all obvious from the recursion.

Now write $a_n(t) = \sum_{i=0}^{[n/2]} \alpha(n,i)\, t^i$. The congruences (30) say that $p\alpha(p+n,i)$ is congruent modulo p to the coefficient of t^i in $\Phi_n(t)$ for $0 \leq n, i < p$. Because of the recursions, it suffices to prove them for n even (which anyway is the case we are interested in). By part ii) of Theorem 4, we have

$$\alpha(2n,i) = 12^{3i} \sum_{m=0}^{i} (-1)^m \binom{-\frac{1}{12}}{i-m} \binom{-\frac{5}{12}}{i-m} \binom{n+\frac{1}{12}}{m} \binom{n-\frac{7}{12}}{m} \binom{2n-1}{m}^{-1}.$$

Let $2n = p + 2h + 1$ with $h \geq 0$ fixed and $p > 2h+1$. The binomial coefficient $\binom{2n-1}{m}$ is prime to p for $m \leq 2h$ and divisible by p exactly once for $m \geq 2h+1$ (note that $m \leq i < p$), and in the latter case we have the congruence

$$\begin{aligned} \frac{1}{p}\binom{2n-1}{m} &= \frac{(p+2h)\cdots(p+1)(p-1)\ldots(p-m+2h+1)}{m!} \\ &\equiv \frac{(-1)^{m-1}\,(2h)!\,(m-2h-1)!}{m!} \equiv \frac{(-1)^{m-1}}{m}\binom{m-1}{2h}^{-1} \pmod{p}. \end{aligned}$$

Hence $p\,\alpha(p+2h+1,i)$ is p-integral and is congruent modulo p to

$$-12^{3i}\sum_{2h+1\le m\le i}\binom{-\frac{1}{12}}{i-m}\binom{-\frac{5}{12}}{i-m}\cdot\binom{h+\frac{7}{12}}{m}\binom{h-\frac{1}{12}}{m}m\binom{m-1}{2h},$$

and using (34) one checks that this agrees with the coefficient of t^i in $\Phi_{2h+1}(t)$.

This completes the proof of (30), which give congruences modulo p for the $a_n(t)$ in the range $p\le n<2p$. Going further, one finds further congruences of the same sort. For instance,

$$\begin{aligned} a_{2p}(t) &\equiv \tfrac{1}{2}\,\Psi_0(t) \mod (p,\,t^p)\,,\\ p\,a_n(t) &\equiv \tfrac{1}{2}\,\Phi_{n-2p}(t) \mod (p,\,t^p) \qquad (2p<n<3p), \end{aligned}$$

where the $\Psi_n(t)$ $(n\ge 0)$ are another sequence of power series with properties similar to those of the Φ_n (they satisfy a simple recursion and factor as products of one fixed and one variable hypergeometric series), the first few being

$$\begin{aligned} \Psi_0(t) &= 1-24\,t-17928\,t^2-18117312\,t^3+\cdots\,,\\ \Psi_1(t) &= -60\,t\,(1+522\,t+471288\,t^2+519169620\,t^3+\cdots)\,,\\ \Psi_2(t) &= -32760\,t^2\,(1+896\,t+984042\,t^2+1201855008\,t^3+\cdots)\,. \end{aligned}$$

More generally, for $s<p$ the first p coefficients of $a_{sp}(t)$ and $p\,a_{sp+n}(t)$ $(0<n<p)$ are congruent modulo p to the coefficients of $\gamma_s\Phi_n(t)$ if s is even and to $\gamma_s\Psi_n(t)$ if s is odd, where $\gamma_0=1$, $\gamma_1=\frac{1}{2}$, $\gamma_2=-\frac{1265}{3}$, $\gamma_3=-\frac{1647}{4}$, ... are certain constants. Then starting at $n=p^2$ we get higher powers of p in the denominators and a new sequence of congruences. There are also nice congruences for the power series $\Phi_n(t)$ and $\Psi_n(t)$ modulo p and powers of p which the reader may want to experiment with, but we are getting carried away from our main theme and will stop here.

10 Supersingular polynomials and the modular polynomial

In this section, which is a bit disjoint from the rest of the paper, we tie up some loose ends by giving direct proofs of two facts about supersingular polynomials which appeared in earlier sections. The argument given here is essentially identical with one given in a paper of Koike ([6], p. 136 and 169), but the presentation there is considerably less elementary and makes use of difficult results of Ihara, whereas the discussion here is entirely self-contained. Since the present paper has a partially expository character, and the argument is short and quite pretty, it seemed worth including it here.

We will work with modular forms and functions considered as elements in the ring $\mathbb{Q}((q))$ of Laurent series in q, which we identify with the ring $\mathbb{Q}((j^{-1}))$ of Laurent series in $j^{-1}=j(\tau)^{-1}$. In particular, we define a Laurent series $\varphi_p(j)\in\mathbb{Z}((j^{-1}))$ with leading coefficient $744\,j^{p-1}$ by

$$\varphi_p(j(\tau))=\frac{j(\tau)^p-j(p\tau)}{p}\,. \tag{35}$$

Let $\Phi_p(X,Y) \in \mathbb{Z}[X,Y]$ be the modular polynomial relating the j-invariants of two p-isogenous elliptic curves. The famous (and easily proved) congruence of Kronecker says that

$$\Phi_p(X,Y) = (X^p - Y)(X - Y^p) + p\,R_p(X,Y) \tag{36}$$

for some $R_p(X,Y) \in \mathbb{Z}[X,Y]$. We are interested in the mod p reduction of the polynomial $H_p(X) = R_p(X,X^p)$. Substituting $X = j(\tau)$, $Y = j(p\tau)$ into (36) we find

$$0 = \frac{1}{p}\,\Phi_p\big(j(\tau), j(p\tau)\big) = \varphi_p\big(j(\tau)\big)\big(j(\tau) - j(p\tau)^p\big) + R_p\big(j(\tau),\, j(p\tau)\big)$$

and hence, reducing modulo p and using $j(p\tau) \equiv j(\tau)^p$ (here and from now on $\equiv$ denotes congruence of Laurent series modulo p),

$$\varphi_p(j) \equiv \frac{H_p(j)}{j^{p^2} - j}\,. \tag{37}$$

In particular, the mod p reduction of φ_p is the Laurent series expansion of a rational function all of whose poles are simple and are contained in $\mathbb{F}_{p^2}$.

On the other hand, differentiating (35) gives

$$\varphi_p'\big(j(\tau)\big) = j(\tau)^{p-1} - \frac{j'(p\tau)}{j'(\tau)} \equiv j(\tau)^{p-1} - j'(\tau)^{p-1}\,,$$

where $'$ applied to a function of τ means $(2\pi i)^{-1}d/d\tau$. But from

$$1 \equiv E_{p-1}(\tau) = \Delta^m E_4^\delta E_6^\varepsilon \widetilde{E}_{p-1}(j) \equiv \Delta^m E_4^\delta E_6^\varepsilon j^{-\delta}(j - 1728)^{-\varepsilon} S_p(j),$$

where $S_p(j) \in \mathbb{F}_p[j]$ is *defined* by $S_p(j) \equiv j^\delta (j - 1728)^\varepsilon \widetilde{E}_{p-1}(j)$, we get

$$j'(\tau)^{p-1} = \left(-\frac{E_6(\tau)}{E_4(\tau)}\,j(\tau)\right)^{12m+4\delta+6\varepsilon} \equiv \frac{j(\tau)^{8m+4\delta+4\varepsilon}\,(j(\tau) - 1728)^{6m+2\delta+4\varepsilon}}{S_p(j(\tau))^2}$$

(for $p \geq 5$). Hence the Laurent series $\varphi_p'(j)$ satisfies the congruence

$$\varphi_p'(j) \equiv j^{8m+4\delta+4\varepsilon}\Big(j^{4m+2\varepsilon} - \frac{(j - 1728)^{6m+2\delta+4\varepsilon}}{S_p(j)^2}\Big)\,. \tag{38}$$

Comparing equations (37) and (38), we deduce that:

(a) all zeros of $S_p(j)$ except for possibly 0 and 1728 are simple;

(b) all zeros of $S_p(j)$ lie in $\mathbb{F}_{p^2}$;

(c) the polynomial $H_p(j)$ (mod p) vanishes at $j = 0$ and 1728 and at all values of $j \in \mathbb{F}_{p^2}$ which are *not* roots of $S_p(j)$;

(d) if j is a root of $S_p(j)$, then $H_p(j) \equiv -j^{8m+4\delta+4\varepsilon}(j - 1728)^{6m+2\delta+4\varepsilon}/S_p'(j)^2$.

Statement (a) was needed in §3, where a different proof was given. Statement (b), once $S_p(j)$ has been identified with $ss_p(j)$ as was done in §3, is the fact that all supersingular invariants lie in $\mathbb{F}_{p^2}$, which was mentioned without proof in §2. Statement (c) can be found in the paper [9] by E. de Shalit, and statement (d) is closely related to a question raised there. Namely, de Shalit wrote down the conjectural formula, closely related to previous work of Oesterlé and himself on the p-adic period pairing on $X_0(p)$,

$$H_p(j) = \begin{cases} 0 & \text{if } j \in \mathbb{F}_{p^2} \text{ is not supersingular,} \\ \dfrac{(-1)^{\varepsilon}\, j^{\frac{2\delta(p+1)}{3}}\, (j-1728)^{\frac{\varepsilon(p+1)}{2}}}{ss_p'(j)^{p+1}} & \text{if } j \text{ is supersingular,} \end{cases} \tag{39}$$

and proved this formula completely for $p \equiv 1 \pmod 4$ and up to a possible ambiguity of sign if $p \equiv 3 \pmod 4$, his proof being non-elementary. Comparing with (d), we see that formula (39) is true if and only if

$$ss_p(j) = 0 \quad \Rightarrow \quad ss_p'(j)^{p-1} = (-1)^{\varepsilon-1}\, j^{\frac{2}{3}(\delta-1)(p-1)}\, (j-1728)^{\frac{1}{2}(\varepsilon-1)(p-1)}\,. \tag{40}$$

For instance, if $p \equiv -1 \pmod{12}$, then (40) says simply that $ss_p'(j)$ belongs to $\mathbb{F}_p$ for all supersingular j. It would be nice to have an elementary proof of this simple statement. A sketch of an argument why equation (40) (or at least its 12th power) should be true was shown to one of the authors by Faltings. We also mention that properties (c) and (d) almost characterize the polynomial $H_p(j)$, since they give its values at p^2 arguments and H_p has degree p^2+p+1 if $p \nmid 744$. A complete (though not very elegant) determination of H_p then follows from (37) and the additional formulas $\varphi_p^{(n)}(0) \equiv 0$ $(1 \le n \le 8m+2\delta+4\varepsilon)$ and $\varphi_p^{(n)}(1728) \equiv (n-1)!(-1728)^{n-1}$ $(1 \le n \le 6m+2\delta+2\varepsilon)$, which follow immediately from (38).

References

[1] M. Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörpern*, Abh. Math. Sem. Hamburg, **14** (1941), 197-272.

[2] C.F. Gauss, *Disquisitiones generales circa seriem infinitam* $1 + \frac{\alpha\beta}{1\cdot\gamma}\, x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1\cdot2\cdot\gamma(\gamma+1)}\, xx + \frac{\alpha(\alpha+1)(\alpha+1)\beta(\beta+1)(\beta+2)}{1\cdot2\cdot3\cdot\gamma(\gamma+1)(\gamma+2)}\, x^3 + \text{etc.}$, Pars prior, (1812), 125-162, Werke III.

[3] H. Hasse, *Existenz separabler zyklischer unverzweigter Erweiterungskörper vom Primzahlgrade p über elliptischen Funktionenkörpern der Charakteristik p*, J. Reine Angew. Math., **172** (1934), 77-85; Math. Abhandlungen, **2**, 161-169.

[4] E. Heine, *Auszug eines Schreibens über Kettenbrüche von Herrn E. Heine an den Herausgeber*, J. Reine Angew. Math., **53** (1857), 284-285.

[5] E. Heine, *Über die Zähler und Nennner der Näherungswerthe von Kettenbrüchen*, J. Reine Angew. Math., **57** (1860), 231-247.

[6] M. Koike, *Congruences between modular forms and functions and applications to the conjecture of Atkin*, J. Fac. Sci. Univ. Tokyo, **20** (1973), 129-169.

[7] R.A. Rankin, *The zeros of Eisenstein series*, Publications of the Ramanujan Institute, **1** (1969), 137-144.

[8] J-P. Serre, *Congruences et formes modulaires (d'après H.P.F. Swinnerton-Dyer)* , Sém. Bourbaki, **416** (1971/72), 74-88; or Œuvres III 74-88.

[9] E. de Shalit, *Kronecker's polynomial, supersingular elliptic curves, and p-adic periods of modular curves*, Contemporary Math., **165** (1994), 135-148.

[10] D. Zagier, *Modular forms and differential operators* , Proc. Indian Acad. Sci. (Math. Sci.), **104** (1994), 57-75.