On a Sequence Arising in Series for $\pi$

By Morris Newman* and Daniel Shanks

Abstract. In a recent investigation of dihedral quartic fields [6] a rational sequence $\{a_n\}$ was encountered. We show that these $a_n$ are positive integers and that they satisfy surprising congruences modulo a prime $p$. They generate unknown $p$-adic numbers and may therefore be compared with the cubic recurrences in [1], where the corresponding $p$-adic numbers are known completely [2]. Other unsolved problems are presented. The growth of the $a_n$ is examined and a new algorithm for computing $a_n$ is given. An appendix by D. Zagier, which carries the investigation further, is added.

1. Introduction. The sequence $\{a_n\}$ that begins with

$$
\begin{align*}
  a_1 &= 1, & a_2 &= 47, & a_3 &= 2488, & a_4 &= 138799, \\
  a_5 &= 7976456, & a_6 &= 467232200,
\end{align*}
$$

and which is defined below, is encountered in a set of remarkable convergent series for $\pi$. These are (see [6]):

$$
\pi = \frac{1}{\sqrt{N}} \left( -\log |U| - 24 \sum_{n=1}^{\infty} (-1)^n \frac{a_n}{n} U^n \right),
$$

where $N$ is a positive integer and $U = U(N)$ is a real algebraic number determined by $N$. Some of these series are remarkable because of their almost unbelievably rapid rates of convergence.

For example, for $N = 3502$, (2) converges at 79 decimals per term and its leading term, namely

$$
- \frac{1}{\sqrt{3502}} \log U,
$$

differs from $\pi$ by less than $7.37 \cdot 10^{-82}$. In this case,

$$
U = U(3502) = (2defg)^{-6},
$$

where

$$
\begin{align*}
  d &= D + \sqrt{D^2 - 1}, & e &= E + \sqrt{E^2 - 1}, \\
  f &= F + \sqrt{F^2 - 1}, & g &= G + \sqrt{G^2 - 1},
\end{align*}
$$

for the quadratic surds

$$
\begin{align*}
  D &= \frac{1}{2}(1071 + 184\sqrt{34}), & E &= \frac{1}{2}(1553 + 266\sqrt{34}), \\
  F &= 429 + 304\sqrt{2}, & G &= \frac{1}{2}(627 + 442\sqrt{2}).
\end{align*}
$$

In this example, the six $a_n$ in (1) already give $\pi$ correctly to over 500 decimals.
For $N = 2737$, and the more general

$$U = (-1)^N (2\text{defg})^{-6},$$

the quadratic surds

$$D = \frac{1}{2}(621 + 49\sqrt{161}), \quad E = \frac{1}{2}(321 + 25\sqrt{161}),$$

$$F = \frac{1}{4}(393 + 31\sqrt{161}), \quad G = \frac{1}{4}(2529 + 199\sqrt{161}),$$

and (4) unchanged, define its negative value of $U(2737)$. Now (2) converges at only 69 decimals per term. See [6] for other examples of even and odd $N$, and the corresponding positive and negative values of $U$, where (2) also converges very rapidly.

The definition given in [6] of $a_n$ is rather complicated. We have a relation

$$U = V \prod_{n=1}^{\infty} \left(1 + V^n\right)^{24}$$

between our $U = U(N)$ and the number

$$V = V(N) = (-1)^N e^{-\pi\sqrt{N}}.$$

The inversion of (8) gives $V$ as a power series in $U$:

$$V = \sum_{n=1}^{\infty} (-1)^{n-1} c_n U^n$$

that begins with $c_1 = 1$, $c_2 = 24$, $c_3 = 852$, ... Now, in the power series for

$$\log\left(\prod_{n=1}^{\infty} \left(1 + V^n\right)\right) = V + \frac{V^2}{2} + \cdots,$$

substitute (10), and thereby define $a_n$ recursively by

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{a_n}{n} U^n = \log\left(\prod_{n=1}^{\infty} \left(1 + V^n\right)\right).$$

Then, the logarithm of (8) gives us (2).

In [6], only the six coefficients in (1) were given, since they were computed by hand, a tedious operation. (The original $a_n$ so computed contained an error which was discovered when R. Brent kindly attempted to verify (2) for $N = 3502$ to the aforementioned 500 decimals.) Clearly, the $a_n$ are best calculated using a digital computer. The first 100 values of $a_n$ and $c_n$ were so computed in about 8 minutes. The first 50 values of $a_n$ and $c_n$ are given in Tables 1 and 2.

2. Properties of $a_n$. A. We observe that all $a_n$ in Table 1 are positive integers. It was obvious from the recursion above that the $a_n$ are rational but not that they are positive and integral. However, we prove below that

$$24a_n$$

is the coefficient of $x^n$ in

$$\prod_{k=1}^{\infty} \left(1 + x^{2^{k-1}}\right)^{24n},$$

which implies that $a_n$ is a positive integer.

B. We observe that all $a_n$ in Table 1 satisfy

$$a_n$$

is odd if and only if $n$ is a power of 2.
This unexpected result is reminiscent of C. R. Johnson's conjecture for the parity of the number of subgroups of the classical modular group of a given index $N$, see [7]. That conjecture was proved by Stothers and, independently, by A. O. L. Atkin. The present observation (14) is proved below.

C. A striking paradox about this proven (14) for the parity of $a_n$ is this: As presented above, the $c_n$ in (10) would appear to constitute a simpler sequence than our $a_n$ in (12), since its definition is much more direct. Nonetheless, we have been unable to determine the parity of $c_n$. In Table 2 one readily observes that

\[ c_n \text{ is odd only when } n = 8k + 1 \text{ and is odd if } k = 0, 1, 2, 4, 6. \]

But what are these $k$? We do not know, and do not even have a conjecture for the parity of $c_n$.

It is easy to prove (14a) and to compute $c_n$ modulo 2. The parity of $c_n$ appears to be random with increasing $k$ just as is the parity of the unrestricted partition function $p(n)$. (See [8] for the latter.) As for the claim above that we have a paradox here, see Zagier's comment in the appendix.

D. A second, more important paradox concerns $a_n$ modulo 3. We conjectured

\[ a_n \not\equiv 0 \mod 3 \]

for all $n$. While (15) appears simpler than (14), we did not prove it. Every positive integer $n$ has a unique representation

\[ n = 3^k (3m \pm 1) \]

with nonnegative $k$, $m$. A stronger conjecture than (15) is

\[ d_{3^k (3m \pm 1)} \equiv \pm 1 \mod 3. \]

For greater clarity, let us rewrite (17) as follows:

\[ a_{3m+1} \equiv 1 \mod 3, \]
\[ a_{3m-1} \equiv -1 \mod 3, \]
\[ a_{3m} \equiv a_m \mod 3. \]

These are clearly equivalent to (17). We did not prove the simple-looking (18a) and (18b). The more subtle-looking (18c) we did prove; it is a simple corollary of a much more general congruence given in E below.

We did verify (17) up to $a_{143} \equiv -1 \mod 3$ by computer, and we both believed it to be true. After we finished the first version of this paper, we showed the conjecture to D. Zagier, and, as we expected, he proved it. See the appendix.

E. The important general congruence alluded to above, and proved below, is

\[ a_{mp^k} \equiv a_{mp^{k-1}} \mod p^k, \]

valid for every prime $p$ and all positive integers $m$ and $k$. For $k = 1$ this gives us

\[ a_{mp} \equiv a_m \mod p \]

and (18c) is obviously the case $p = 3$.

Congruence (20) is computationally useful. For example, what is $a_{94}$ modulo 94? Since

\[ a_{2 \cdot 47} \equiv a_2 = 47 \mod 47, \]
we have \( a_{94} \equiv 0 \mod 47 \). But also \( a_{94} \equiv 0 \mod 2 \), by (14). Therefore \( a_{94} \equiv 0 \mod 94 \). Similarly, we can evaluate \( a_{2^p} \) modulo \( 2^p \) for any prime \( p \), and in particular we see that, for any prime \( p \),

\[
(21) \quad a_{2^p} \equiv 1 \mod 2^p.
\]

F. The choice \( m = 1 \) in (20) gives us

\[
(22) \quad a_p \equiv a_1 \equiv 1 \mod p,
\]

which we call the Fermat Property. It is a necessary condition for primality. Of course, we ask: Is

\[
(23) \quad a_n \equiv 1 \mod n, \quad n > 1,
\]

a sufficient condition for primality?

We have just seen in (21) that \( n = 2^p \) can never satisfy (23). But consider

\[ a_3 = 2488 = 3 \cdot 829 + 1. \]

Since 829 is prime, we have by (20) that

\[ a_{2487} \equiv a_3 = 1 \mod 829, \]

and similarly

\[ a_{2487} \equiv a_{829} \mod 3. \]

But 829 = 3m + 1, and since (18a) is now true, we also have

\[
(24) \quad a_{2487} \equiv 1 \mod 3.
\]

Then (23) holds for the composite 2487 = 3 \cdot 829. So (23) is not a sufficient condition for primality. Even if it were, it would not be a practical test for primality. The calculation of \( a_n \) modulo \( n \) requires at least \( O(n) \) operations by any algorithm known to us.

G. We return to (19) and specialize in a different direction; \( m = 1 \) gives us

\[
(25) \quad a_p^k \equiv a_p \mod p^k.
\]

Fix \( p \) and consider the sequence

\[
(26) \quad \{a_p^k \mod p^k\}, \quad k = 1, 2, 3, \ldots.
\]

If we write these numbers to the base \( p \), (25) guarantees that each time \( k \) is increased by 1, and we add one more \( p \)-adic digit on the left, all the earlier \( p \)-adic digits on the right remain unchanged. Thus, for each \( p \), the sequence (26) defines a \( p \)-adic number.

For example, for \( p = 2 \), (26) begins (in decimal) as 1, 3, 7, 15, 15, 47, \ldots, and so we have the 2-adic number (reading from right to left)

\[ \ldots 1000101111. \quad \text{(base 2)} \]

Similarly, for \( p = 3 \) and 5, we have

\[ \ldots 0111. \quad \text{(base 3)} \]

\[ \ldots 411. \quad \text{(base 5)} \]

But what are these \( p \)-adic numbers? We do not know. Are they algebraic or transcendental? We do not know. Contrast this ignorance with the situation in I below.
We do have, for every $p$,
\begin{equation}
\label{eq:27}
a_p^2 \equiv 1 + p \mod p^2,
\end{equation}
so the first two $p$-adic digits on the right are both 1. The first 1 follows from the Fermat Property (22) but the second 1 does not follow from the general congruence (19), and again contrasts with the situation in I below. This (27) was first proved by our colleague L. Washington. Our proof below is different.

Perhaps we should note that the sequence
\begin{equation}
\label{eq:28}
\langle a_p^k \rangle, \quad k = 1, 2, 3, \ldots
\end{equation}
defines the same $p$-adic number that (26) does. The latter looks a little simpler since it adds exactly one $p$-adic digit each time.

H. After we discovered (18c), we were inspired to generalize it to (19) because of a recent paper [1] concerning some entirely different sequences; namely, a doubly infinite set of cubic recurrences. It suffices for our discussion here to examine only one of these recurrences. Let
\begin{equation}
\label{eq:29}
A(1) = 1, \quad A(2) = 1, \quad A(3) = 4, \quad A(n + 3) = A(n + 2) + A(n).
\end{equation}
We have [1]
\begin{equation}
\label{eq:30}
A(mp^k) \equiv A(mp^{k-1}) \mod p^k
\end{equation}
just as before. So we also have the Fermat Property and $p$-adic numbers defined by
\begin{equation}
\label{eq:31}
\{ A(p^k) \text{ modulo } p^k \}.
\end{equation}

I. But the $A(n)$ are nonetheless quite different than the $a_n$. First, since
\begin{equation}
\label{eq:32}
A(4) \equiv 1 \mod 4, \quad A(9) \equiv 4 \mod 9,
\end{equation}
(27) does not hold, and the second $p$-adic digit is not invariant. Second, we can identify the $p$-adic numbers (31). For example, for $p = 2$, we now have
\[
\ldots 100101_2 = x \quad (\text{base 2}).
\]
Squaring this, it is easy to show that
\[
x^2 + x + 2 = 0,
\]
and so $x$ is one of the 2-adic numbers
\[
\frac{1}{2} \left( -1 \pm \sqrt{-3} \right).
\]
In fact, for every $p$, (31) is an abelian algebraic integer; see [1], [2].

The evaluation of these algebraic integers is of much algorithmic interest and is also of much mathematical interest since, e.g., it leads to new ideas in cyclotomy; see [5]. But more to the present investigation, this $p$-adic approach enables one to solve problems about $A(n)$ that were previously intractable, as in [2].

One might hope that the determination of the $p$-adic numbers in (26) would be equally valuable for $a_n$. Presumably, the distinctive property (27) plays a role in their arithmetic characterization. We commend these problems to the reader.

J. If we generalize (31) to
\begin{equation}
\label{eq:33}
\{ A(mp^k) \text{ modulo } p^k \}
\end{equation}
for $p$ fixed, and $m$ any integer, we define a set of $p$-adic numbers. This set is finite, and each of these numbers is either an algebraic conjugate of that for $m = 1$, or is a related abelian integer of a lower degree.

Similarly, in the present investigation,

$$\{a_{mp^e} \mod p^k\},$$

with $m$ a fixed positive integer, defines a $p$-adic number for each $m$ generalizing (26). But we have not seriously examined this set of $p$-adic numbers and know little about it.

K. Let us note some other differences between $A(n)$ and $a_n$. The former sequence is periodic modulo $p$ for every $p$, but the latter is not. The former is a reversible recurrence, and so we have

$$A(0) = 3, \quad A(-1) = 0, \quad A(-2) = -2, \ldots,$$

while $a_n$ is not defined for $n < 1$. The value of $A(n)$ modulo $n$ can be computed in $O(\log n)$ operations. We know of no algorithm that is that efficient for our $a_n$ modulo $n$. We have

$$A(n) = a^n + \beta^n + \gamma^n$$

for known values of $\alpha, \beta, \gamma$ while we know of no explicit formula for $a_n$.

Since $a_n$ and $A(n)$ are so very different, it is all the more surprising that they have, in (19) and (30), an elaborate, important property in common. We call this property the generalized $p$-adic law.

Naturally, one asks: Can one characterize all sequences $a(n)$ that satisfy this law? This may already be known.

Zagier also comments on the comparison of $a_n$ and $A(n)$.

L. We now turn to the growth of the $a_n$. In the analytic function $V(U)$ in (10) the closest singularity to the point $U = 0, \ V = 0$ is the branch point at $U = -\frac{1}{34}, \ V = -e^{-7}$; see [6, Appendix B]. Therefore, the radius of convergence of (10) is $\frac{1}{34}$, and it follows that

$$\lim_{n \to \infty} \frac{c_{n+1}}{c_n} = 64.$$

In the substitution of (10) into (11), the growth of the $a_n$ is dominated by the growth of the $c_n$, and it may be shown that also

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = 64.$$

M. We therefore have the asymptotic formula

$$\log a_n \sim n \log 64,$$

but an asymptotic formula for $a_n$ itself was lacking. We expected that

$$a_n \sim \frac{C}{n^\beta} (64)^n, \quad C, \beta \text{ constants},$$

but we did not prove it.

In the Appendix, Zagier determines that $\beta = \frac{1}{2}$ (as we expected), and that

$$C = \frac{\sqrt{\pi}}{12} \left( \frac{\Gamma(3/4)}{\Gamma(1/4)} \right)^2.$$
Further, he gives two more terms in the asymptotic series, and thereby enables one to estimate \( a_n \) very accurately.

Prior to this work we had already found the inequalities (38) below, and since these are of some interest, we include the derivation.

\[
\frac{1}{3\sqrt{n}}(63.87)^n < 24a_n < (64)^n.
\]

N. Zagier's evaluation of \( C \) suggests the following sequel. This \( C \) is closely related to the famous lemniscate constant, and, in retrospect, some such result should have been expected. In [6], the group \( C(4) \) was basic, and therefore our sequence \( a_n \) is intimately connected with this group. But the lemniscate constant often arises with \( C(4) \); for example, \( Q(\sqrt{-14}) \) has \( C(4) \) as its class group, and, in counting numbers of the form \( u^2 + 14v^2 \), the lemniscate constant enters via the constant \( \beta_{14} \) referred to in [9, Eq. (5)].

Now, in the modular group, one encounters \( \rho = \sqrt[3]{1} \) as well as \( i = \sqrt{-1} \), and therefore \( C(3) \) as well as \( C(4) \), and [6, p. 405] specifically refers to analogous theories for \( C(3) \) and \( C(6) \). So, there may well be other sequences analogous to \( a_n \) that would arise in this way. We have not yet studied this.

In the quadratic form \( 4u^2 + 2uv + 7v^2 \) we do have class number 3, and in counting numbers of \( this \) form one does indeed encounter a constant which contains \( \Gamma(1/6) \) instead of \( \Gamma(1/4) \); see [10, Eq. (5)]. If there are such sequences, one would expect Zagier's calculations to have analogues here.

### 3. Proofs of the Theorems

The function
\[
y = x \prod_{k=1}^{\infty} (1 + x^k)^{24}
\]
defined in (8) (the variable names have been changed) is of importance in the theory of the elliptic modular functions. \( y \) is a Hauptmodul for the congruence subgroup \( \Gamma_0(2) \) of the classical modular group \( \Gamma \), considered as a function of the complex variable \( \tau \), where \( x = \exp(2\pi i \tau) \), \( \text{im} \tau > 0 \). (See [4] for a good general reference on this topic.) However, all that is required here is a formal study of the coefficients of \( y^n \), where \( n \) is an integer. In this connection certain complex integral formulas associated with the inversion of a function of the form \( y = x + b_2x^2 + \cdots \) (or the reversion of a power series of this form) will be used freely. These are classical, and may be found for example in the book by Behnke and Sommer [3].

The numbers \( a_n \) are defined by the relationship (12), rewritten as

\[
\log y - \log x = 24 \sum_{n=1}^{\infty} (-1)^{n-1} \frac{a_n}{n} y^n.
\]

Differentiating (39) with respect to \( y \), and then multiplying by \( y \), we have that

\[
1 - \frac{y}{x} \frac{dx}{dy} = 24 \sum_{n=1}^{\infty} (-1)^{n-1} a_n y^n.
\]

Hence for some suitable positive number \( r \), we have that

\[
(-1)^{n-1}24a_n = \frac{1}{2\pi i} \oint_{|y|=r} \left(1 - \frac{y}{x} \frac{dx}{dy}\right) y^{-n-1} dy,
\]
so that, for \( n \geq 1, \)
\[
(-1)^{n-1} 24a_n = \frac{1}{2\pi i} \int_{|x| = r'} \left( \frac{1}{x} \frac{dx}{dy} \right) y^{-n} dy.
\]
This implies that, for some suitable positive number \( r' \),
\[
(-1)^{n-1} 24a_n = \frac{1}{2\pi i} \int_{|x| = r'} x^{-n} dx
= -\frac{1}{2\pi i} \int_{|x| = r'} x^{-n-1} \prod_{k=1}^{\infty} (1 + x^k)^{-24n} dx.
\]
It follows that, for \( n \geq 1, \) \((-1)^n \cdot 24a_n\) is the coefficient of \( x^n \) in the power series expansion of \( \prod_{k=1}^{\infty} (1 + x^k)^{-24n}. \) If we use the fact that
\[
\prod_{k=1}^{\infty} (1 + x^k)^{-1} = \prod_{k=1}^{\infty} (1 - x^{2k-1}),
\]
and replace \( x \) by \(-x\), we obtain (13) and write

**Theorem 1.** The number \( 24a_n \) defined by (39) is the coefficient of \( x^n \) in the infinite product \( \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24n}. \)

This proves immediately that these numbers are positive, but a small additional discussion is required to prove that \( a_n \) is an integer (because of the factor 24).

We set
\[
\prod_{k=1}^{\infty} (1 + x^{2k-1})^{24n} = \sum_{k=0}^{\infty} C_n(k) x^k.
\]
so that
\[
24a_n = C_n(n).
\]
We find by logarithmic differentiation of (41) and known properties of Lambert series that the integers \( C_n(k) \) satisfy the recurrence formula
\[
kC_n(k) = 24n \sum_{s=1}^{k} (-1)^{s-1} \sigma^*(s) C_n(k-s), \quad k \geq 1.
\]
where \( C_n(0) = 1, \) and
\[
\sigma^*(s) = \sum_{d | s, d \text{ odd}} d.
\]
For the choice \( k = n, \) (42) and (43) imply that
\[
a_n = \sum_{s=1}^{n} (-1)^{s-1} \sigma^*(s) C_n(n-s),
\]
which shows at once that \( a_n \) is an integer. That is, we have proved

**Theorem 2.** The numbers \( a_n \) defined by (39) are positive integers.

Our next objective is to prove (14), which states the remarkable fact that \( a_n \) is odd if and only if \( n \) is a power of 2. For this purpose we need to know the parity of the function \( \sigma^*(s), \) defined by (44). We have the following simple lemma, whose proof
we omit:

**Lemma 1.** The function $\sigma^*(s)$ is odd if and only if $s$ is a square, or twice a square.

This lemma and formula (45) imply that

$$a_n \equiv \sum C_n(n - s^2) + \sum C_n(n - 2s^2) \mod 2. \quad (46)$$

In the first summation, $s$ runs over all positive integers such that $s^2 \leq n$, and, in the second summation, $s$ runs over all positive integers such that $2s^2 \leq n$.

First note that

$$(1 + u)^{16} = (1 + u^2)^8 \mod 16,$$

where the congruence means that coefficients of corresponding powers of $u$ are congruent. This readily implies that

$$24a_{2n} \equiv 24a_n \mod 16,$$

$$a_{2n} \equiv a_n \mod 2. \quad (47)$$

Congruence (47) is the special case $p = 2$ of the general congruence (20), to be proved later.

Thus, in order to determine the parity of $a_n$, it is only necessary to choose $n$ odd, which we now do. If we note that

$$\prod_{k=1}^{\infty} (1 + x^{2k-1})^{4k^2n} = \prod_{k=1}^{\infty} (1 + x^{4k^2-2})^{24n} \mod 16,$$

which in turn implies that

$$24a_{2n} \equiv 24a_n \mod 16,$$

$$a_{2n} \equiv a_n \mod 2. \quad (47)$$

But $n$ is odd. Thus the second sum in (48) is empty, and in the first sum $s$ must be odd, implying that $n \equiv 1 \mod 8$. Put $n = 8t + 1$. Then

$$a_{8t+1} = \sum_{s \text{ odd}} C_{8t+1}(8t + 1 - s^2) = \sum C_{8t+1} \left(8 \left(t - \frac{r^2 + r}{2}\right)\right) \mod 2. \quad (49)$$

where $r$ runs over all nonnegative integers such that $\frac{1}{2}(r^2 + r) \leq t$.

We have

$$\sum_{k=0}^{\infty} C_{8t+1}(k)x^k = \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24(8t+1)}$$

$$= \prod_{k=1}^{\infty} (1 + x^{2k-1} - 8k) \mod 16$$

so that

$$\sum_{k=0}^{\infty} C_{8t+1}(8k)x^k = \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24t+3} \mod 2.$$
Thus
\[ \prod_{k=1}^{\infty} (1 + x^{2k-1})^{-3} \cdot \sum_{k=0}^{\infty} C_{8r+1}(8k) x^k \equiv \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24r} \mod 2. \]

Now use the Jacobi identity
\[ \prod_{k=1}^{\infty} (1 - x^k)^3 = \sum_{k=0}^{\infty} (-1)^k (2k + 1) x^{(k^2+k)/2} \]
and the fact that
\[ \prod_{k=1}^{\infty} (1 + x^{2k-1})^{-3} \equiv \prod_{k=1}^{\infty} (1 - x^k)^3 \mod 2. \]

Then
\[ \sum_{k=0}^{\infty} x^{(k^2+k)/2} \sum_{k=0}^{\infty} C_{8r+1}(8k) x^k \equiv \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24r} \mod 2. \]

It follows that
\[ \sum C_{8r+1}(8(t - \frac{1}{2}(r^2 + r))) \]
is congruent modulo 2 to the coefficient of \( x^t \) in \( \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24r} \). But this coefficient is odd if and only if \( t = 0 \) (it is divisible by 24 otherwise, since then the coefficient is 24\( a_t \)). It follows from (49) that \( a_{8r+1} \) is odd if and only if \( t = 0 \).

Summarizing, we have proved

**Theorem 3.** The number \( a_n \) is odd if and only if \( n \) is a power of 2.

Our next objective is to prove (19). If \( p \) is a prime and \( k \) a positive integer, then
\[ (1 + u)^{pt} \equiv (1 + u^p)^{p^{k-1}} \mod p^k, \]
where once again the congruence is understood to hold for corresponding powers of \( u \). It follows that if \( m \) is any positive integer,
\[ (1 + u)^{mpt} \equiv (1 + u^p)^{m^{p^{k-1}}} \mod p^k. \]

Formula (50) now implies that
\[ \prod_{s=1}^{\infty} (1 + x^{2s-1})^{24mp^s} \equiv \prod_{s=1}^{\infty} (1 + x^{2ps-p})^{24mp^{k-1}} \mod p^{k+\delta}, \]
where
\[ \delta = \begin{cases} 3, & p = 2, \\ 1, & p = 3, \\ 0, & p > 3. \end{cases} \]

Comparing coefficients of \( x^{mp^k} \) on both sides of (51), we find that
\[ 24a_{mp^k} \equiv 24a_{mp^{k-1}} \mod p^{k+\delta}, \]
so that, for all primes \( p \),
\[ a_{mp^k} \equiv a_{mp^{k-1}} \mod p^k. \]
That is, we have proved

**Theorem 4.** Let $p$ be a prime, $m$, $k$ positive integers. Then

\[(52) \quad a_{mp^k} \equiv a_{mp^{k-1}} \mod p^k.\]

We now go on to formula (27), which reads

\[a_{p^2} \equiv 1 + p \mod p^2, \quad p \text{ prime.}\]

Since (52) implies that

\[a_{p^2} \equiv a_p \mod p^2,
\]

it is sufficient to prove that

\[a_p \equiv 1 + p \mod p^2, \quad p \text{ prime.}\]

We may assume that $p > 3$, since the cases $p = 2, 3$ may be verified directly. We have

\[(1 + u)^p = 1 + u^p + \sum_{r=1}^{p-1} \binom{p}{r} u^r \equiv 1 + u^p + p \sum_{r=1}^{p-1} \frac{(-1)^{r-1}}{r} u^r \mod p^2,\]

so that

\[\frac{(1 + u)^p}{1 + u^p} \equiv 1 + p \sum_{r=1}^{p-1} \frac{(-1)^{r-1}}{r} \frac{u^r}{1 + u^p} \mod p^2.\]

Now choose $u = x^{2k-1}$, product for $k = 1, 2, 3, \ldots$, and raise both sides to the 24th power. We get

\[
\prod_{k=1}^{\infty} \frac{(1 + x^{2k-1})^{24p}}{(1 + x^{2kp-p})^{24}} = 1 + 24p \sum_{1 \leq r \leq p-1 \atop k \geq 1} \frac{(-1)^{r-1}}{r} \frac{x^{r(2k-1)}}{1 + x^{p(2k-1)}} \mod p^2,
\]

\[
\prod_{k=1}^{\infty} (1 + x^{2k-1})^{24p} \equiv \prod_{k=1}^{\infty} (1 + x^{2kp-p})^{24} \cdot S \mod p^2,
\]

where

\[S = 1 + 24p \sum_{1 \leq r \leq p-1 \atop k \geq 1} \frac{(-1)^{r-1}}{r} \frac{x^{r(2k-1)}}{1 + x^{p(2k-1)}}.\]

Comparing coefficients of $x^p$, we find that

\[24a_p \equiv 24 + 24p \mod p^2,
\]

so that

\[a_p \equiv 1 + p \mod p^2.
\]

We state this result as L. Washington's

**Theorem 5.** Let $p$ be a prime. Then

\[a_{p^2} \equiv a_p \equiv 1 + p \mod p^2.
\]
We note that these congruences may be strengthened, if desired. A slightly more involved proof along the same lines will show for example that

\[(53) \quad a_{p^k} \equiv a_{p^k + p^k + 1} \mod p^{k+1}.\]

However, it does not seem possible to determine \(a_{p^k}\) modulo \(p^k\) precisely, except for small values of \(k\).

We now turn to the inequalities of (38). Theorem 1 implies that \(24a_n\) is equal to

\[(54) \quad \sum \binom{24n}{n_1} \binom{24n}{n_3} \binom{24n}{n_5} \cdots n_1 + 3n_3 + 5n_5 + \cdots = n, \quad n_i \geq 0.\]

Since \(n_1 = n, n_3 = n_5 = \cdots = 0\) is a permissible choice, we find that

\[(55) \quad 24a_n \geq \binom{24n}{n}.\]

A simple application of Stirling's formula gives

\[24a_n > \frac{1}{3\sqrt{n}} \binom{24}{23}^n > \frac{1}{3\sqrt{n}} (63.87)^n.\]

proving the lower bound.

For the upper bound, we have that if \(r\) is any number such that \(0 < r < 1\), then

\[24a_n = \frac{1}{2\pi i} \int_{|x|=r} g(x)^n \frac{dx}{x}.\]

where

\[g(x) = \frac{1}{x} \prod_{k=1}^{\infty} (1 + x^{2k-1})^{24}.\]

It follows that

\[(56) \quad 24a_n \leq g(r)^n.\]

Now the function \(g(x)\) is an entire modular function on the congruence subgroup \(\Gamma_0(4)\) of \(\Gamma\), considered as a function of the complex variable \(\tau\), where \(x = \exp(2\pi i \tau)\), and \(\text{im} \tau > 0\). It is easy to show by the transformation formulae for \(g(x)\) that

\[g(e^{-\tau}) = 64.\]

Choosing \(r = e^{-\pi}\) in (56) gives

\[24a_n < 64^n.\]

which is the desired upper bound.

Summarizing, we have proved

**Theorem 6.** The number \(a_n\) satisfies the inequalities

\[\frac{1}{3\sqrt{n}} (63.87)^n < 24a_n < 64^n.\]

4. Computation. The first dozen or so coefficients \(a_n\) were initially computed using the complicated formula (40). After Theorem 1 was discovered, recurrence formula (43) was used. The coefficients \(\sigma^*(s)\) are small and easily computed, and (43) is convenient and simple to implement. The practical programming problems that arise are consequences of the fact that the \(a_n\) become large. This is best handled by
computing them modulo a sufficient number of large primes, and then using the Chinese Remainder Theorem to recover their exact values.

The coefficients $c_n$ were computed by means of a general program that reverts a power series $y = x + \cdots$. This program computes the coefficients of the powers of $y$ and then solves a triangular system of equations to determine the desired coefficients in the reverted power series $x = y + \cdots$. Once again, residual arithmetic must be used, since the coefficients $c_n$ also become large.

The computation of $a_n$ modulo $m$, where some prime factors of $m$ are small, is awkward (if not impossible) using formula (43), because of the necessity of the division there. The alternative here is to generate $u = \sum_{k=0}^{n} x^k$ modulo $24^n$ and then to form $u^{24^n}$ by successive squarings modulo $24^n$. This is time-consuming and becomes impractical if $n$ is only moderately large; say $n = 1000$.

We note that multiprecision computation (rather than modular computation) would be even more time-consuming. In any case there is very little point in calculating the exact value of $a_{1000}$, say, since it is a number of some 1800 decimal digits.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$a_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>2488</td>
</tr>
<tr>
<td>4</td>
<td>138799</td>
</tr>
<tr>
<td>5</td>
<td>7976456</td>
</tr>
<tr>
<td>6</td>
<td>467232200</td>
</tr>
<tr>
<td>7</td>
<td>2773634840</td>
</tr>
<tr>
<td>8</td>
<td>166280372125</td>
</tr>
<tr>
<td>9</td>
<td>10044242737480</td>
</tr>
<tr>
<td>10</td>
<td>6103747246289272</td>
</tr>
<tr>
<td>11</td>
<td>372725787615063808</td>
</tr>
<tr>
<td>12</td>
<td>22852444771010647496</td>
</tr>
<tr>
<td>13</td>
<td>14058860261076389244</td>
</tr>
<tr>
<td>14</td>
<td>8764106172969340032192</td>
</tr>
<tr>
<td>15</td>
<td>53859190340823180403932608</td>
</tr>
<tr>
<td>16</td>
<td>3325772467042429391572615</td>
</tr>
<tr>
<td>17</td>
<td>206553777945446382091995000</td>
</tr>
<tr>
<td>18</td>
<td>1285207805396218649480480977880</td>
</tr>
<tr>
<td>19</td>
<td>80666108867535131490921525593280</td>
</tr>
<tr>
<td>20</td>
<td>49955437325652565560187887772024</td>
</tr>
</tbody>
</table>

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Table 2. $c_n$, $n = 1(1)50$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$c_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>46</td>
<td>46</td>
</tr>
</tbody>
</table>

**Appendix**

By D. Zagier

Asymptotics and Congruence Properties of the $a_n$

In this appendix we prove an asymptotic formula and a congruence modulo 3 for the numbers $a_n$, assuming various more or less well-known facts from the theory of modular forms whose proofs can be found in standard textbooks on modular and elliptic functions (e.g. Lang's or Weil's).

Let $\tau$ denote a variable in the upper half-plane, $q = e^{2\pi i \tau}$, and $U(\tau) = q\Gamma(1 + q^n)^24$ ($q$ and $U$ were denoted by $V$ and $U$ in Section 1 and by $x$ and $y$ in Section 3). Then $U(\tau) = \Delta(2\tau)/\Delta(\tau)$, where $\Delta(\tau) = q\Gamma(1 - q^n)^24$ is the usual...
discriminant function, so \( U \) is a nowhere vanishing modular function on \( \Gamma_0(2) \) and its logarithmic derivative

\[
\frac{f'(\tau)}{f(\tau)} = \frac{1}{2\pi i} \frac{U''(\tau)}{U(\tau)} = 1 + 24 \sum_{n=1}^{\infty} \sigma^*(n) q^n \quad (\sigma^* \text{ as in } (44))
\]

is a modular form of weight 2 on \( \Gamma_0(2) \). The definition of \( a_n \) can be expressed as

\[
\frac{1}{f(\tau)} = 1 + 24 \sum_{n=1}^{\infty} (-1)^n a_n U(\tau)^n,
\]

an identity valid in a neighborhood of \( \tau = i\infty \) (it cannot be valid for all \( \tau \) for which the series converges, since \( U \) is \( \Gamma_0(2) \)-invariant and \( f \) is not). From the formula for the number of zeros of a modular form, we see that \( f(\tau) \) vanishes only at points \( \tau \) which are \( \Gamma_0(2) \)-equivalent to \( \tau_0 = (1 + i)/2 \) (that \( f \) does vanish at \( \tau_0 \) can be seen by applying the transformation equation of \( f \) to \((\frac{1}{2} \quad -\frac{1}{2}) \in \Gamma_0(2) \), and \( 1 \) then shows that \( \tau \to U(\tau) \) is locally biholomorphic except at these points. Hence the only singularity in \( (2) \) occurs at \( U = U(\tau_0) = -1/64 \), so to obtain the asymptotics of the \( a_n \) we must look at the Taylor series expansions of \( f \) and \( U \) near \( \tau_0 \). In view of \( (1) \) and the equation \( f(\tau_0) = 0 \), it will suffice for this to compute the derivatives \( f^{(\nu)}(\tau_0) \) for \( \nu \geq 1 \).

Now the derivative of a modular form is not a modular form, but, if \( F \) is a modular form of weight \( k \) on a subgroup \( \Gamma \) of \( SL(2, \mathbb{Z}) \), then \( F' - (\pi ik/6)E_2F \) is a modular form of weight \( k+2 \) on \( \Gamma \), where \( E_2 = 1 - 24 \sum_{n \geq 1} (\Sigma_{d|n} d) q^n \) is the usual "Eisenstein series of weight 2 on \( SL(2, \mathbb{Z}) \)" (not actually a modular form), related to \( f \) by \( f(\tau) = 2E_2(2\tau) - E_2(\tau) \). Applying this fact \( \nu \) times and using the identity \( E_2^\nu = (\pi i/6)(E_2^\nu - E_4) \), where \( E_4 = 1 + 240 \sum_{n \geq 1} (\Sigma_{d|n} d^3) q^n \) is the Eisenstein series of weight 4 on \( SL(2, \mathbb{Z}) \), we find by induction that the function

\[
\sum_{\mu=0}^{\nu} \binom{\nu}{\mu} \frac{\Gamma(k+\mu)}{\Gamma(k+\nu)} \left( -\frac{\pi i}{6} E_2 \right)^{\nu-\mu} F^{(\mu)}
\]

is a modular form of weight \( k+2\nu \) on \( \Gamma \). We apply this to \( F = f, \Gamma = \Gamma_0(2), k = 2 \). All modular forms on \( \Gamma_0(2) \) are polynomials in \( f \) and \( E_4 \) (this follows easily from the formulas for the dimensions of the spaces of modular forms of given weight), so we can identify \( (3) \) by computing the first few terms of its \( q \)-expansion; we find

\[
f' - \frac{\pi i}{3} E_2 f = -\frac{\pi i}{3} (2f^2 - E_4),
\]

\[
f'' - \pi i E_2 f' - \frac{\pi^2}{6} E_2^2 f = -\frac{\pi^2}{6} f E_4,
\]

\[
f''' - 2\pi i E_2 f'' - \pi^2 E_2^2 f' + \frac{\pi^3 i}{9} E_2^3 f = \frac{\pi^3 i}{9} f^2 (4f^2 - 3E_4),
\]

etc. At \( \tau = \tau_0 = (1 + i)/2 \) we have \( f = 0, E_2 = 6/\pi \) and \( E_4 = -12\alpha^4 \), where

\[
\alpha = \frac{1}{2\sqrt{\pi}} \frac{\Gamma(1/4)}{\Gamma(3/4)} = 0.834626841678 \ldots
\]
(this follows from the well-known $E_2(i) = 3/\pi$ and $E_4(i) = 3\alpha^4$ together with the transformation properties of $E_2$ and $E_4$ under $SL(2, \mathbb{Z})$). Hence we find inductively from the above formulas the values

$$f'(\tau_0) = -4\pi i\alpha^4, \quad f''(\tau_0) = 24\pi\alpha^4, \quad f'''(\tau_0) = 144\pi i\alpha^4$$

and, continuing in the same way,

$$f^{(iv)}(\tau_0) = -960\pi\alpha^4, \quad f^{(v)}(\tau_0) = -7200\pi i\alpha^4 - 96\pi^3 i\alpha^{12}.$$

Using (1), we obtain the Taylor expansions

$$f(\tau_0 + ie) = 4\pi\alpha^4(\varepsilon - 3\varepsilon^2 + 6\varepsilon^3 - 10\varepsilon^4 + (15 + \pi^4\alpha^8/5)\varepsilon^5 + \cdots)$$

and

$$U(\tau_0 + ie) = -\frac{1}{64} e^{-4\pi^2\alpha^4(\varepsilon^2 - 2\varepsilon + 3\varepsilon^3 - 4\varepsilon^4 + (5 + \pi^4\alpha^8/3)\varepsilon^5 + \cdots)}.$$

The second of these expresses $\sqrt{1 + 64U}$ as a power series in $\varepsilon$ with leading term $2\pi\alpha^2\varepsilon$; inverting this power series and substituting the result into the Taylor expansion of $f$, we can write $1/f$ as a Laurent series in $(1 + 64U)^{1/2}$:

$$\frac{1}{f(\tau)} = \frac{1}{2\alpha^2} (1 + 64U)^{-1/2} + \frac{1}{2\pi\alpha^4} \frac{3 - \pi^2\alpha^4}{8\pi^2\alpha^6} (1 + 64U)^{1/2} + \frac{1}{4\pi^2\alpha^8} (1 + 64U) + \frac{15 + 9\pi^2\alpha^4 - 4\pi^4\alpha^8}{96\pi^4\alpha^{10}} (1 + 64U)^{3/2} + \cdots.$$

Comparing this with (2) gives

$$a_n = \frac{64^n}{24 \cdot 2^{2n}(2n)} \left( \frac{1}{2\alpha^2} - \frac{3 - \pi^2\alpha^4}{8\pi^2\alpha^6} \frac{1}{2n - 1} + \frac{15 + 9\pi^2\alpha^4 - 4\pi^4\alpha^8}{96\pi^4\alpha^{10}} \frac{3}{(2n - 1)(2n - 3)} + \cdots \right)$$

$$= \frac{64^n}{48\alpha^2 \sqrt{n}} \left( 1 - \frac{3}{8\pi^2\alpha^4} n^{-1} + \left( \frac{15}{64\pi^4\alpha^8} - \frac{1}{128} \right) n^{-2} + \cdots \right).$$

We have proved

**Theorem.** The sequence $a_n$ has an asymptotic expansion of the form

$$a_n = C \frac{64^n}{\sqrt{n}} \left( 1 - \frac{\alpha_1}{n} + \frac{\alpha_2}{n^2} + \cdots \right),$$

with

$$C = \frac{\sqrt{\pi}}{12} \frac{\Gamma(3/4)^2}{\Gamma(1/4)^2} = 0.0168732651505 \cdots,$$

$$\alpha_1 = \frac{6\Gamma(3/4)^4}{\Gamma(1/4)^4} = 0.07830067 \cdots, \quad \alpha_2 = \frac{60\Gamma(3/4)^8}{\Gamma(1/4)^8} - \frac{1}{128} = 0.002405668 \cdots.$$
We give two numerical examples.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$a_n$</th>
<th>$C \frac{64^n}{\sqrt{n}} \left(1 - \frac{\alpha_1}{n} + \frac{\alpha_2}{n^2}\right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>$4.853249476 \times 10^{87}$</td>
<td>$4.853249382 \times 10^{87}$</td>
</tr>
<tr>
<td>100</td>
<td>$6.996107097 \times 10^{177}$</td>
<td>$6.996107081 \times 10^{177}$</td>
</tr>
</tbody>
</table>

As a second application of the modular form description of the $a_n$, we prove the congruence properties (18a, b) of the numbers $a_n \pmod{3}$. These can be written in the form

$$na_n \equiv \begin{cases} 0 \pmod{3} & \text{if } 3 | n, \\ 1 \pmod{3} & \text{if } 3 \nmid n, \end{cases}$$

or

$$\sum_{n=1}^{\infty} (-1)^{n-1} n a_n U^n \equiv \frac{U(1 - U)}{1 + U^3} \pmod{3}.$$ 

On the other hand, differentiating (2) and substituting (1), we see that

$$f(\tau)^3 \sum_{n=1}^{\infty} (-1)^{n-1} n a_n U(\tau)^n = \frac{1}{48\pi i} f'(\tau) = \sum_{n=1}^{\infty} n\sigma^*(n) q^n.$$ 

Since $f \equiv 1 \pmod{3}$, we have to prove that

$$\frac{U(1 - U)}{1 + U^3} \equiv \sum_{n=1}^{\infty} n\sigma^*(n) q^n \pmod{3}.$$ 

From the description of modular forms on $\Gamma_0(2)$ as polynomials in $f$ and $E_4$ it follows that the modular function $U$ must be related to $E_4/f^2$ by a fractional linear transformation; comparing the first few Fourier coefficients we find

$$\frac{E_4}{f^2} = \frac{1 + 256U}{1 + 64U}, \quad U = \frac{1}{64} \frac{E_4 - f^2}{4f^2 - E_4} = \frac{\phi}{f^2 - 64\phi},$$

where

$$\phi = \frac{1}{192} \left( E_4 - f^2 \right) = q + 8q^2 + 28q^3 + \cdots = \sum_{n \geq 1} b(n) q^n, \quad \text{say,}$$

a modular form of weight 4 on $\Gamma_0(2)$. Since $E_4$ and $f^2$ are congruent to 1 (mod 48), it is clear that $4\phi$ has integral coefficients, so that the numbers $b(n)$ are 3-integral, which is all we will need; actually, the $b(n)$ themselves are integral, as one can see from the identity $\phi = U(f^2 - 64\phi)$ or from the formula

$$\phi = \left( \sum_{n > 0 \atop n \text{ odd}} q^{n^2/8} \right)^8.$$
From $U = \phi/(f^2 - 64\phi)$ we obtain

$$
\frac{U(1 - U)}{1 + U^3} = \frac{\phi(f^2 - 64\phi)(f^2 - 65\phi)}{(f^2 - 64\phi)^3 + \phi^3}
\equiv \frac{\phi(f^2 - \phi)(f^2 + \phi)}{f^6} = \frac{\phi}{f^2} - \left(\frac{\phi}{f^2}\right)^3 \pmod{3}.
$$

Since $f \equiv 1 \pmod{3}$, the $q$-expansion of the right-hand side of this is congruent to $\phi - \phi^3$ or $\sum (b(n) - b(n/3))q^n \pmod{3}$ (with the usual convention $b(n/3) = 0$ if $3 \nmid n$), so the congruence we have to prove is

$$
(n\sigma^*(n) \equiv b(n) - b(n/3) \pmod{3}).
$$

The form $E_4(2\tau) = 1 + 240\sum_{n \geq 1} \sigma_3(n)q^{2n}$ is a modular form of weight 4 on $\Gamma_0(2)$ and hence a linear combination of $f^2$ and $E_4$ or of $E_4$ and $\phi$. Comparing two Fourier coefficients gives $E_4(2\tau) = E_4 - 240\phi$ or

$$
\phi(\tau) = \frac{1}{240}(E_4(\tau) - E_4(2\tau)), \quad b(n) = \sigma_3(n) - \sigma_3(n/2).
$$

Clearly $\sigma_3(n) \equiv \sigma_3(n/3) \pmod{3}$ if $3 \mid n$, so (4) is true in this case. On the other hand, $\sigma_3(n) \equiv \sigma_1(n) \equiv \sum_{d \mid n} d \pmod{3}$ since $d^3$ and $d$ are congruent, and, combining the divisors $d$ and $n/d$, we see that $\sigma_1(n) \equiv 0 \pmod{3}$ if $n \equiv -1 \pmod{3}$ or equivalently $\sigma_3(n) \equiv n\sigma_1(n) \pmod{3}$ if $n \equiv 0 \pmod{3}$. Hence for $3 \nmid n$ we have

$$
\sigma_3(n) - \sigma_3(n/2) = n(\sigma_1(n) - 2\sigma_1(n/2)) = n\sigma^*(n) \pmod{3}
$$

as required.

Having proved the formula for $a_n \pmod{3}$ we offer a conjectural formula for $a_n \pmod{5}$:

$$
a_n \equiv \begin{cases} 
a_{n/5} & \text{if } 5 \mid n, \\
0 & \text{if } n = 5k + \delta, 0 < \delta < 5, k \text{ odd}, \\
\delta\left(\frac{2r}{r}\right)^3 & \text{if } n = 10r + \delta, 0 < \delta < 5. 
\end{cases}
$$

It is true up to $n = 100$.

Finally, we make a remark about the nature of the numbers $a_n$. Equation (2) suggests that the natural generalization of this sequence is the sequence $(a_n)$ defined by a generating function of the form $F = \sum a_n u^n$, where $u$ is a Hauptmodul for some group $\Gamma$ of genus 0 (e.g. $\Gamma = SL_2(Z)$, $u = j^{-1}$, $\Gamma = \Gamma_0(2)$, $u = U$, or $\Gamma = \Gamma_0(2) \cup \Gamma_0(2) (\sqrt{2} - \sqrt{2})$, $u = 1/(U + 2^{12}/U)$) and $F$ a meromorphic modular form of some weight $k$ on $\Gamma$. This definition includes both the $a_n$ (with $k = -2$) and the sequence $(A(n))$ mentioned several times in the paper (since these satisfy a recursion with constant coefficients and hence $\sum A(n)U^n$ is a rational function of $U$ and therefore a modular form of weight $k = 0$), which may explain their parallel properties. The sequence $(c_n)$ defined by (10) of the paper has no such interpretation, which may explain why it apparently does not have such nice arithmetic properties.


