SIX-DIMENSIONAL SPHERE PACKING AND LINEAR PROGRAMMING

MATTHEW DE COURCY-IRELAND, MARIA DOSTERT, MARYNA VIAZOVSKA

Abstract. We prove that the Cohn–Elkies linear programming bound for sphere packing is not sharp in dimension 6. The proof uses duality and optimization over a space of modular forms, generalizing a construction of Cohn–Triantafillou to the case of odd weight and non-trivial character.

1. Introduction

The method of linear programming gives an upper bound for the density achievable by sphere packings in Euclidean space of any given dimension. In some cases, especially dimensions 8 and 24 as shown in [31, 12], the bound is equal to the density of a known configuration, giving a proof of optimality. In most dimensions, we do not know the optimal packing, and we have only a numerical approximation to the linear programming bound. The optimal density and the linear programming bound are not expected to be equal in general. It seems plausible that equality occurs only in dimensions 1, 2, 8, and 24, which are the values conjectured by Cohn and Elkies [6, Conjecture 7.3], with 2 the last of the expected sharp cases to defy proof. However, it has only recently become possible to rule out the scenario that linear programming gives a sharp bound in all dimensions: it provably fails in dimensions 3, 4, 5, and [9, Table 6.1] for 12 and 16. In this article, we prove that the bound is not sharp in dimension 6. Intuitively, the reason is that the linear programming method applies not only to packings, but also to “fake packings” with impossibly high densities. We construct one such example.

The first obstacle is that the densest packing in dimension 6 is not known. We therefore compare the linear programming bound (LP bound) to a more powerful bound, based on semidefinite programming (SDP) and developed by Cohn, de Laat, and Salmon. We refer to [8, Table 1.1] for bounds in a range of dimensions.

Theorem 1.1 (Cohn, de Laat, Salmon). The sphere packing density in dimension 6 is at most 0.410304.

Theorem 1.2. The LP bound for density in dimension 6 is at least 0.410948.

Corollary 1.3. The linear programming bound in dimension 6 is strictly higher than the density of any packing.
For comparison, the highest density known in dimension 6 is
\[
\frac{\pi^3}{48\sqrt{3}} = 0.372947 \ldots
\]
This is achieved by the $E_6$ root lattice [11, p. 126], which is closely related to the sharp case of $E_8$ in dimension 8. Numerical optimization over a restricted class of functions in the linear program gives an upper bound 0.417674, which is presumably very close to the true value, but in principle might differ considerably from the optimum over the full infinite-dimensional space of candidates. Theorem 1.2 shows that the true value cannot be much lower, and in particular it exceeds both the highest known density (1.1) and the upper bound from Theorem 1.1. The proof sheds some light on why the linear programming bound is not sharp in dimension 6, as it is for dimension 8, despite the close connection between the exceptional structures $E_6$ and $E_8$. On the other hand, it is known that there are infinitely many other packings achieving the same density as $E_6$, but with no particularly close connection to $E_8$. They are obtained by stacking layers of the (presumed) densest 5-dimensional packing [11, p. 144].

The proof of Theorem 1.2 uses duality for linear programs. To describe it further, let us recall more specifically how the LP bound works. It is convenient to work with the “center density” of a packing, which is the density divided by the volume of a ball in $\mathbb{R}^D$ of the corresponding radius. This simplifies various factors of $\pi$, and can be thought of as the number of centers per unit volume. For example, the center density of $E_6$ is the value from (1.1) divided by $\pi^D/(D/2)! = \pi^3/6$, namely
\[
\frac{\pi^3}{48\sqrt{3}} + \frac{\pi^3}{6} = \frac{1}{8\sqrt{3}} = 0.072168 \ldots
\]
Theorem 1.1 implies that all packings have center density at most 0.079398. Theorem 1.2 shows that the LP bound for center density is at least 0.0795223, and numerical optimization leads to 0.08084 as an upper bound. In fact, our number 0.079523 is rounded down from a quadratic irrational. What we really show is that the LP bound for center density in dimension 6 is at least
\[
-277385984684414834701547634832199852580621960702176236773103
+ 5542322057902681856636220216828951751933789602321848882511869
+ 160710053876796738434707078882772389137901213002069082699776
\]
\[
\sqrt{3}
\]
\[
0.079522333845052286373845030218205166528
\]
The fractions in (1.3) work out to $\alpha + \beta\sqrt{3}$ where $\alpha \approx -0.5178$ and $\beta \approx 0.3448$.

**Theorem 1.4** (Cohn–Elkies, [6]). *If an integrable, continuous function $f : \mathbb{R}^D \to \mathbb{R}$ with Fourier transform $\widehat{f}$ is not identically 0 and satisfies $f(x) \leq 0$ for $|x| \geq r$, and $\widehat{f} \geq 0$, then the center density of any sphere packing in $\mathbb{R}^D$ is at most*
\[
\left(\frac{r}{2}\right)^D f(0) \div \widehat{f}(0)
\]
By the Cohn–Elkies linear programming bound, or simply the LP bound, we mean the infimum of (1.5) over all admissible $f$. We have assumed $f$ is continuous and integrable so that $f(0)$ and $\widehat{f}(0)$ are defined in the most straightforward
way. It is also of interest to take the infimum over other classes of functions (for instance, Schwartz), and these variants could also be called “the” LP bound. In fact, the original formulation from [6] assumed that \( f \) is sufficiently smooth and rapidly decaying so that both sides of the Poisson summation formula converge. This was later relaxed in [7, section 9] and [10, Theorem 3.3].

The following theorem is an instance of weak duality for linear programs. It appears in [5] and a version of the same linear program, expressed in terms of pair-correlation functions, was studied by Torquato and Stillinger [30]. Depending on the space of \( f \) one considers for the Cohn–Elkies bound, the dual object \( \mu \) will be one kind of generalized function or another. We suppose \( f \) varies over continuous functions, so that \( \mu \) varies over measures.

**Theorem 1.5.** If \( \mu = \delta_0 + \nu \) is a measure on \( \mathbb{R}^D \) with \( \nu \geq 0 \) supported in \( \{ x \in \mathbb{R}^D : |x| \geq r \} \) and \( \hat{\mu} \geq c\delta_0 \), then the Cohn–Elkies linear programming bound for center density in \( \mathbb{R}^D \) is at least

\[
(1.6) \quad c \left( \frac{r}{2} \right)^D.
\]

**Proof.** The duality at work here is the Plancherel formula. Let \( f \) and \( \mu \) be as in the Cohn–Elkies program and its dual. We can assume it is the same \( r \) for both \( f \) and \( \mu \), by scaling if necessary. Indeed, replacing \( f(x) \) by \( sf(s'x) \) for \( s, s' > 0 \) does not change (1.5). Then \( \langle f, \mu \rangle \leq f(0) \) because \( f \leq 0 \) on the support of \( \mu - \delta_0 \). By Plancherel, since \( \hat{\mu} \geq c\delta_0 \) and \( \hat{f} \geq 0 \),

\[
(1.7) \quad f(0) \geq \langle f, \mu \rangle = \langle \hat{f}, \hat{\mu} \rangle \geq c\hat{f}(0)
\]

so \( f(0)/\hat{f}(0) \geq c \). Since this holds for any \( f \) satisfying the Cohn–Elkies constraints, the LP bound is at least \( c(r/2)^D \). \( \square \)

In other words, the LP bound thinks there is a packing of center density \( c(r/2)^D \), even if there is only a measure imitating such a packing. For comparison, given a lattice \( \Lambda \) with center density \( \rho \) and dual lattice \( \Lambda^* \), the following measure has transform given by Poisson summation:

\[
\mu = \sum_{x \in \Lambda} \delta_x, \quad \hat{\mu} = \rho \sum_{\xi \in \Lambda^*} \delta_{\xi}
\]

The factor \( c \) in this case is exactly \( \rho \), and the result from Theorem 1.5 is the same as one would find by comparing the LP bound to the packing with spheres centered at the points of \( \Lambda \) (here, we scale so that the spheres in the packing have radius 1). In this sense, the \( \mu \) in the “dual” program has a more direct connection to the underlying packings, compared to the Cohn–Elkies function \( f \). In the sharp cases in 8 or 24 dimensions, the packing is related to the zeros of \( f \) and \( \hat{f} \), and \( f \) can be recovered from these zeros by interpolation [12, 31].

Figure 1.1 compares our candidate \( \mu \) to a numerical approximation of the optimal \( f \). The proof of Theorem 1.5 shows that, for \( f \) and \( \mu \) to give optimal bounds, \( \mu - \delta_0 \) should be supported on the zeros of \( f \). That way, there is no loss in the inequality \( f(0) \geq \langle f, \mu \rangle \). Similarly, \( \hat{\mu} \) should be supported on the zeros of \( \hat{f} \) together with the origin. In our example, the support is larger, but aligns well with the initial zeros. The function plotted in Figure 1.1 approximates \( f \) by a polynomial of degree 24 times a Gaussian \( \exp(-\pi x) \), written as a function of squared distance. It was computed in Julia using sum-of-squares techniques implemented by de Laat and
Leijenhorst in [24]. For visibility, we plot only the polynomial factor without the Gaussian, which of course vanishes at the same points as $f$.

To make a $\mu$ satisfying the constraints of Theorem 1.5, Cohn and Triantafillou [9] proposed

$$\mu = \sum_{n=0}^{\infty} a_n \delta_n \sqrt{n}$$

where $a_n$ are the Fourier coefficients of a modular form. They assumed that the dimension $D$ is a multiple of 4, and considered modular forms of weight $D/2$ with trivial character on $\Gamma_0(N)$. Suitable values of $N$ gave results in dimensions 12 and 16, showing that the linear programming bound does not match the densities of the presumed best packings. However, these dimensions were too high to make a favourable comparison with semidefinite methods at the time. The possibility of a better packing remained, which could even match the LP bound, until [8] gave a strong enough SDP bound in dimensions 12 and 16 to show that the earlier bounds from [9] imply non-sharpness in those cases. In dimension 6, we apply a similar strategy, showing decisively that the linear programming bound is not sharp by comparing it to Theorem 1.1. Theorem 1.2 follows from the existence of a suitable modular form. Since 6 is not a multiple of 4, we must work with modular forms of odd weight 3, using non-trivial Dirichlet characters. In particular, we used the quadratic characters $\chi_3$ and $\chi_4$ defined by $-1 \mod 3 \mapsto -1$ and $-1 \mod 4 \mapsto -1$ respectively. For each of these there are spaces of modular forms denoted by $M_k(\Gamma_0(N), \chi)$, which we will discuss further in Sections 2 and 3.

**Proposition 1.6.** There is a modular form

$$g \in M_3(\Gamma_0(48), \chi_3) \oplus M_3(\Gamma_0(48), \chi_4)$$

whose Fourier expansion and that of the transform $\tilde{g}(z) = -i48^{-3/2}z^{-3}g(-1/(48z))$ satisfy

$$g(z) = \sum_{n=0}^{\infty} a_n e^{2\pi i nz}, \quad \tilde{g}(z) = \sum_{n=0}^{\infty} b_n e^{2\pi i nz}$$

where $a_n \geq 0$ and $b_n \geq 0$ for all $n$, $a_0 = 1$, $b_0 \geq 0.6168035$, and $a_n = 0$ for $1 \leq n \leq 6$. Moreover, $a_n = 0$ for all $n \equiv 1 \mod 4$.

Again, this is a rounded version of a more precise statement: the $g$ we construct has $b_0$ equal to the value from (1.3) multiplied by $\frac{768}{343} \sqrt{12}$, via (2.7) below.

The direct sum in Proposition 1.6 is a 44-dimensional space of modular forms, in which we find $g$ by solving a linear program that forces a finite number of the constraints $a_n \geq 0$ and $b_n \geq 0$. To verify the remaining inequalities as $n \to \infty$, we decompose $g$ as a sum of Eisenstein series and cuspforms. The Eisenstein contribution to $a_n$ or $b_n$ grows quadratically with $n$, whereas the cuspidal part has absolute value at most $n^{1+o(1)}$, from which one deduces the inequalities for large $n$. The different rates of growth are illustrated in Figure 1.2. To show the exact equalities $a_n = 0$ for $n \equiv 1 \mod 4$, we must argue more carefully using the action of Hecke operators and quadratic twists. The fact that $a_n = 0$ for $n \equiv 1 \mod 4$ is visible in Figure 1.1, where, so to speak, every fourth spike seems to be missing. By increasing $N$ and using more characters instead of only $\chi_3$ and $\chi_4$, one could try to remove even more spikes and obtain a better match with the zeros of $f$.

There are other approaches to rigorously estimating the LP bound. Li [25] discretizes Theorem 1.4, replacing $\mathbb{R}^D$ by a grid, and using duality for the resulting
Figure 1.1. Top: a numerical approximation to the optimal Cohn–Elkies function $f$ in green, compared to the measure $\mu = \sum_n a_n \delta_{\sqrt{n}}$ in blue, as functions of squared distance $x$. We rescale $\mu$ by $n \mapsto n/7$ so that the first spike coincides with the sign change of $f$ at $x = 1$. Bottom: $\hat{\mu} = (2/\sqrt{N})^k \sum_n b_n \delta_{2\sqrt{n}/N}$ in blue compared to the polynomial part of $\hat{f}$ in green. With $N = 48$, and scaling the Fourier transform reciprocally to $\mu$ by $n \mapsto 7n$, the spikes of $\hat{\mu}$ occur at $x = 7n/12$. For visibility, without affecting the locations of zeros or spikes, the plots show only the polynomials $f(x) \exp(\pi x)$ and $\hat{f}(x) \exp(\pi x)$ without the Gaussian factor, and large values of $a_n$ or $b_n$ are cut off. The first two zeros of both $f$ and $\hat{f}$ align closely with the supports of $\mu$ and $\hat{\mu}$. 
finite-dimensional linear programs rather than Theorem 1.5. This proves that the Cohn–Elkies bound is not sharp in dimensions 3, 4, and 5, and gives bounds in all dimensions $3 \leq D \leq 13$ (ignoring the sharp case $D = 8$). In dimension 6, the bound from [25] is 0.07632412 in terms of center density, whereas one would need roughly 0.0794 to conclude from Theorem 1.1 that the LP bound is not sharp. However, we imagine the only obstacle is computational power (and the current value would already suffice if one knew that $E_6$ is a densest packing, or if the SDP bound could be improved enough). Torquato and Stillinger [30], motivated by hyperuniformity and the idea that disordered sphere packings in high dimensions could be much denser than known examples based on lattices, analyzed a choice of $\mu$ that gives the best result currently known asymptotically as $D \to \infty$. Their result, in terms of density rather than center density, shows that the LP bound is at least $2^{-\tau D + o(D)}$ where $\tau = \frac{3 - 1}{2} \log 2 \approx 0.7786$ and $o(D)$ denotes sublinear terms, meaning that $o(D)/D \to 0$ as $D \to \infty$. Numerical extrapolations from [1] suggest the LP bound is $2^{-\lambda D + o(D)}$ where $\lambda \approx 0.6044$ and the authors speculate on a closed form for $\lambda$. However, the best proven asymptotic
| $D$ | Record density | SDP bound | LP bound | Dual bound | Upper bound |
|-----|----------------|-----------|----------|------------|-------------|
| 1   | 1              | 1         | 1        | 1          |             |
| 2   | 0.906899       | 0.906899  | 0.906899 | 0.906899   | 0.906899    |
| 3   | 0.740480       | 0.770271  | 0.770657 | 0.779747   |             |
| 4   | 0.616850       | 0.636108  | 0.637303 | 0.647705   |             |
| 5   | 0.465257       | 0.512646  | 0.517236 | 0.524981   |             |
| 6   | 0.372947       | 0.410304  | **0.410948** | 0.417674 |
| 7   | 0.295297       | 0.321148  | 0.301191 | 0.327456   |             |
| 8   | 0.253669       |           | 0.253669 |             |             |
| 9   | 0.145774       | 0.191121  | 0.164925 | 0.194556   |             |
| 10  | 0.099615       | 0.143411  | 0.106256 | 0.147954   |             |
| 11  | 0.066238       | 0.106726  | 0.078504 | 0.111691   |             |
| 12  | 0.049454       | 0.079712  | 0.083381 | 0.083776   |             |
| 13  | 0.032014       | 0.060165  | 0.032522 | 0.062482   |             |
| 14  | 0.021624       | 0.045062  | 0.046365 |             |             |
| 15  | 0.016857       | 0.033757  | 0.034249 |             |             |
| 16  | 0.014708       | 0.023995  | 0.025011 | 0.025195   | 0.025195    |
| 24  | 0.001929       | 0.001929  |           |             | 0.001929    |

Table 1.1. The LP bound exceeds the SDP bound in dimensions $D = 3, 4, 5, 6, 12, 16$, shown in grey. The LP bound is sharp for $D = 1, 8, 24$ and conjecturally $D = 2$. Values of the Cohn–de Laat–Salmon SDP bound are taken from [8, Table 1.1], as are the highest known densities and the numerical LP bounds computed in [1]. Li’s dual bound on the LP bound from [25] for $D \leq 13$, and the values for $D = 12, 16$ from Cohn–Triantafillou [9, Table 6.1], have been converted from center density to density. The value from Theorem 1.2 is shown in bold.

The upper bound remains that of Kabatianskii and Levenshtein [21]: the LP bound is at most $2^{-\kappa D + o(D)}$ where $\kappa \approx 0.5990576$ is the root of an explicit equation. The best known packings have density $2^{-D + o(D)}$. To our knowledge, there is no rigorous analysis of the SDP bound as $D \to \infty$. That would be an important step in showing that there are only finitely many dimensions where the linear programming bound is sharp, and perhaps only $1, 2, 8, 24$.

In Section 2, we review the method of [9], giving an alternative calculation of the main Fourier transform underlying the approach. In Section 3, we describe the optimization space from Proposition 1.6 in more detail. In Section 4, we explain how to find the solution using exact arithmetic in the field $\mathbb{Q}(\sqrt{3})$. In Sections 5 to 7, we estimate at what point the Eisenstein part dominates the cuspidal part, and explain the vanishing $a_n = 0$ for $n \equiv 1 \mod 4$ in Section 6. Supplementary files available at https://arxiv.org/abs/2211.09044 explain these calculations, and the final verification that $a_n, b_n \geq 0$, in more detail. In Section 8, we prove an identity (8.1) between theta series for the lattices $D_6, D_6^*$, and $Z^6$. We had expected these three to be independent elements of the optimization space, and the linear relation counteracting that expectation is specifically 6-dimensional.
2. The method of Cohn–Triantafillou

The key Fourier calculation behind our proof is given in the following proposition. Cohn and Triantafillou proved this in [9, Proposition 2.2] assuming that $k$ is an integer (i.e. the dimension is even), and that $g$ is modular for a congruence group $\Gamma_1(N)$. We give another proof without those assumptions: it is enough for $g$ to be modular for a smaller group given by generators in (2.5), where $N > 0$ could take non-integer values. For $k$ a half-integer and $\text{Im}(z) > 0$, define $(-iz)^{-k}$ by the standard branch satisfying $1^{-k} = 1$. Write $\mathbb{H}$ for the half-plane $\{z \in \mathbb{C} ; \text{Im}(z) > 0\}$.

**Proposition 2.1.** For a function $g : \mathbb{H} \to \mathbb{C}$, $k = D/2$ an integer or half-integer, and $N > 0$, let

$$
(2.1) \quad \bar{g}(z) = (-iz)^{-k}N^{-k/2}g\left(\frac{-1}{Nz}\right).
$$

Suppose $g$ and $\bar{g}$ are both periodic, with Fourier expansions

$$
\begin{align*}
g(z) &= \sum_{n=0}^{\infty} a_n e^{2\piinz} \quad \text{and} \quad \bar{g}(z) = \sum_{n=0}^{\infty} b_n e^{2\piinz}.
\end{align*}
$$

Then

$$
\sum_{n=0}^{\infty} a_n \delta_{\sqrt{n}} \quad \text{and} \quad (2/\sqrt{N})^k \sum_{n=0}^{\infty} b_n \delta_{2\sqrt{n}/N}
$$

are Fourier transforms of each other as tempered distributions on $\mathbb{R}^D$, where $\delta_r$ denotes a spherical delta at radius $r$.

**Proof.** By definition, the claim is that

$$
\sum_{n=0}^{\infty} a_n f(\sqrt{n}) = (2/\sqrt{N})^k \sum_{n=0}^{\infty} b_n \hat{f}\left(2\sqrt{\frac{n}{N}}\right)
$$

for all radial Schwartz functions $f$. By approximation, it is enough to check this for Gaussians $f(x) = e^{\pi i |x|^2 z}$ with $\text{Im}(z) > 0$. The Fourier transform is then

$$
\hat{f}(\xi) = (i/z)^k e^{\pi i |\xi|^2 (1/z)}
$$

On the left,

$$
\sum_{n=0}^{\infty} a_n f(\sqrt{n}) = \sum_{n} a_n e^{\pi inz} = g(z/2).
$$

On the right,

$$
\sum_{n} b_n \hat{f}(2\sqrt{n/N}) = \sum_{n} b_n e^{-4\pi in/(Nz)}(i/z)^k = \bar{g} \left(\frac{-1}{Nz/2}\right)(i/z)^k.
$$

Note that

$$
\bar{g}\left(\frac{-1}{Nz/2}\right) = \left(-i\left(\frac{-1}{Nz/2}\right)^{-k}\right)N^{-k/2}g(z/2) = 2^{-k}N^{k/2}(i/z)^{-k}g(z/2).
$$

So, as required, (right) = $2^{-k}N^{k/2}$(left). \qed

The periodicity of both $g$ and $\bar{g}$ requires $g$ to transform as a modular form, and $\bar{g}$ is an example of an Atkin–Lehner involution applied to $g$. Let

$$
(2.2) \quad w_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}, \quad w_N(z) = \frac{-1}{Nz}
$$
Then, for integer $k$,

$$\tilde{g} = (-i)^{-k}g|w_N$$

where the slash action is given by

$$g| \begin{pmatrix} a & b \\ c & d \end{pmatrix} (z) = (ad - bc)^{k/2}(cz + d)^{-k}g\begin{pmatrix} az + b \\ cz + d \end{pmatrix}$$

and satisfies a chain rule $g|\gamma_1\gamma_2 = (g|\gamma_1)|\gamma_2$. The factor $(-i)^{-k}$, or simply $i^k$, is the correct sign so that both sequences $a_n, b_n$ can be non-negative, as one can see by taking $z = it$ on the imaginary axis in (2.1). The periodicity of $\tilde{g}$ amounts to $g|w_N T = g|w_N$, where $T(z) = z + 1$ is the standard translation. Equivalently, $g|w_N T w_N^{-1} = g$. Thus the periodicity of $g$ and $\tilde{g}$ requires two modular properties for $g$:

$$g, \tilde{g} \text{ both periodic } \iff g = g|T = g|w_N T w_N^{-1}$$

In matrix form, the condition is that $g$ must be modular (of weight $k = D/2$) with respect to the group generated by $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix}$.

If $N > 4$, then there is an infinite-dimensional space of candidates $g$ transforming as per (2.5), since a scaling $z \mapsto \sqrt{N}z$ conjugates these generators into a Hecke triangle group [3, Chapter 4]. However, these are not as explicit as modular forms on $\Gamma_1(N)$. Following [9], we will assume $N$ is an integer and impose modularity under a larger group to obtain a finite-dimensional space that is amenable to computation.

The action of $w_N$ by conjugation is

$$w_N \begin{pmatrix} a & b \\ c & d \end{pmatrix} w_N^{-1} = \begin{pmatrix} d & -c/N \\ -Nb & a \end{pmatrix}$$

In particular, $w_N$ normalizes the congruence subgroups $\Gamma_0(N)$ and $\Gamma_1(N)$. Recall that these are the subgroups of $\text{SL}(2, \mathbb{Z})$ defined by

$$\Gamma_0(N) = \left( \begin{array}{cc} * & * \\ 0 & * \end{array} \right) \mod N, \quad \Gamma_1(N) = \left( \begin{array}{cc} 1 & * \\ 0 & 1 \end{array} \right) \mod N$$

which are clearly normalized by $w_N$, in view of (2.6). It follows that if $g$ is a modular form for $\Gamma_1(N)$, then both $g, \tilde{g}$ are periodic. That was the case considered in [9]. It is convenient for computational purposes because of existing software developed in GP/Pari [2], Sage [29], and Magma [4].

The space of modular forms for $\Gamma_1(N)$ decomposes into spaces for $\Gamma_0(N)$ with various Dirichlet characters modulo $N$ (see, for instance, [22, p. 137]). Recall that, by definition, modular forms with character $\chi$ are holomorphic functions with a growth condition and satisfying the functional equations

$$f\left(\frac{az + b}{cz + d}\right) = \chi(d)(cz + d)^k f(z)$$

whenever $a, b, c, d$ are integers obeying $ad - bc = 1$ and $c$ is divisible by $N$, that is, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$. Such $f$ form a vector space over $\mathbb{C}$ which we denote by $M_k(\Gamma_0(N), \chi)$. Calculations with this space have been implemented in Pari using the command \texttt{mfinit} which takes as input $k, N$, and various possible descriptions of $\chi$. 

Cohn and Triantafillou stated their results in terms of $\Gamma_1(N)$, but used only the trivial character, which was already enough to obtain interesting bounds with $k = 6$ or $k = 8$ for the dimensions they considered. In case $k$ is odd, we must use $\Gamma_1(N)$, rather than $\Gamma_0(N)$ with trivial character, and decompose with respect to characters modulo $N$ satisfying $\chi(-1) = -1$. One must have $\chi(-1) = (-1)^k$ for there to be any non-zero modular forms of character $\chi$ on $\Gamma_0(N)$, in view of the action of $-I: z \mapsto (-z)/(-1) = z$.

Taking the measure $\mu$ in Theorem 1.5 to be a sum of point masses as in Proposition 2.1 led Cohn and Triantafillou to the following, which can be optimized over a finite-dimensional space of modular forms.

**Proposition 2.2** (Cohn–Triantafillou). Suppose $g$ is a modular form of weight $D/2$ for $\Gamma_1(N)$ and the Fourier coefficients of $g = \sum_n a_n q^n$ and $\tilde{g} = \sum_n b_n q^n$ obey the following inequalities, where $q = e^{2\pi i z}$:

\[
\begin{align*}
    a_n &\geq 0, b_n \geq 0 \\
    a_0 &= 1, b_0 > 0 \\
    a_n &= 0 \quad \text{for} \quad 1 \leq n < T
\end{align*}
\]

Then the linear programming bound in $\mathbb{R}^D$ is at least

\[
(2.7) \quad b_0 \left( \frac{2}{\sqrt{N}} \right)^{D/2} \left( \frac{\sqrt{T}}{2} \right)^D
\]

This is a linear program with a finite number of variables, as many as the dimension of the space of modular forms of weight $D/2$ for $\Gamma_1(N)$. However, it involves infinitely many constraints $a_n \geq 0, b_n \geq 0$. In practice, one solves a linear program enforcing only a finite number of these. One must then separate the solution $g$ into a cuspidal part and an Eisenstein part. The remaining inequalities for large $n$ are established by showing that the Eisenstein part is eventually positive, and large enough that the cuspidal part can be neglected.

The cuspidal part can be bounded using Deligne’s theorem on the Ramanujan conjecture for holomorphic forms. For a statement without proof, one can refer to [20, eq. (14.54)]. The latter follows from Deligne’s work on the Riemann Hypothesis for varieties over finite fields [13, Théorème 8.2]. See also [14] and the references therein to earlier work [18, 23] linking the Fourier coefficients and Hecke eigenvalues to the number of points on associated varieties.

**Theorem 2.3** (Deligne). If $k \geq 2$ is an integer, and $f$ is a Hecke cuspform of weight $k$ for $\Gamma_0(N)$ with character $\chi$, normalized so that

\[
f(z) = \sum_{n > 0} c_f(n) e(nz) = 1 \cdot e(n_{\min}z) + \ldots \quad (\epsilon(t) = e^{2\pi it})
\]

with leading coefficient 1, then for all $n$,

\[
|c_f(n)| \leq n^{(k-1)/2} \sigma_0(n)
\]

where $\sigma_0(n)$ is the number of positive divisors of $n$.

The corresponding result for $k = 1$ is also true, by a theorem of Deligne and Serre [15], but that case corresponds to dimension 2, where the LP bound is expected to be sharp, and where both exponents $k - 1$ and $(k - 1)/2$ equal 0. For us, $k = 3$. 
The divisor function \( \sigma_0(n) \) grows slower than any power of \( n \), but can exceed any fixed power of \( \log n \). At its largest, it behaves roughly like
\[
2^{\log n} / \log \log n
\]
as discussed for instance in [17, Theorem 317]. We quote a sharp version of this from [27].

**Theorem 2.4** (Nicolas, Robin). *The maximum*
\[
\max_{n \geq 2} \log \sigma_0(n) \cdot \frac{\log \log n}{(\log 2)(\log n)}
\]
is attained at
\[
n = 2^5 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 = 6983776800
\]
with value \( 1.5379 \ldots \)

This number 1.5379 is larger than the limiting value 1 that one could take as \( n \to \infty \), but in our application, it is helpful if one can show that the Eisenstein part dominates the cuspidal part even for relatively small values of \( n \).

Write \( a_n \) as the sum of the contributions from Eisenstein series and from cusp-forms:
\[
a_n = a_{n,\text{eis}} + a_{n,\text{cusp}}
\]
and similarly for \( b_n \). It can happen that \( a_n \) vanishes along an arithmetic progression (for instance, we will have \( a_n = 0 \) for all \( n \equiv 1 \mod 4 \)). For the remaining \( n \), we will show \( a_n \geq 0 \) by the following inequalities. Using explicit formulas for the Eisenstein series, one can show a bound of the form
\[
a_{n,\text{eis}} \geq \varepsilon n^{k-1}
\]
whereas Deligne’s theorem implies
\[
|a_{n,\text{cusp}}| \leq C n^{(k-1)/2} \sigma_0(n)
\]
where \( \varepsilon > 0 \) and \( C > 0 \) are constants (independent of \( n \)) determined by the choice of variables \( x_j \) in Proposition 2.2. These bounds entail
\[
a_n \geq a_{n,\text{eis}} - |a_{n,\text{cusp}}| \geq \varepsilon n^{k-1} - C n^{(k-1)/2} \sigma_0(n).
\]
This will imply \( a_n \geq 0 \) once \( n \) is large enough that
\[
\sigma_0(n) \leq \frac{\varepsilon}{C} n^{(k-1)/2}
\]
After logarithms, the comparison (2.11) becomes
\[
\log \sigma_0(n) + \log \left( \frac{C}{\varepsilon} \right) \leq \frac{k - 1}{2} \log n
\]
Suppose
\[
\log \sigma_0(n) \leq R \left( \frac{\log n}{\log \log n} \right)
\]
where one can take \( R = 1.5379 \log 2 \) as in Theorem 2.4, or take \( R \) arbitrarily close to \( \log 2 \) by assuming \( n \) is large enough. The criterion (2.12) is satisfied provided that
\[
\frac{R}{\log \log n} + \frac{\log (C/\varepsilon)}{\log n} \leq \frac{k - 1}{2}
\]
which holds for all large enough $n$. Explicitly, if

$$n \geq \max \left( \exp \left( \frac{4 \log(C/\varepsilon)}{k - 1} \right), \exp \exp \left( \frac{4R}{k - 1} \right) \right)$$

then (2.14) is satisfied because each term on the left is at most half the right. For instance, it is enough to take $n \geq \max((C/\varepsilon)^2, 4590)$ when $k = 3$ and $R$ is taken from Theorem 2.4.

In practice, one can do better by solving (2.14) numerically. The estimates from Sections 5 and 7 will show that $C = 21.6161$ and $\varepsilon = 8.7536 \times 10^{-6}$ are admissible for $a_n$, while $C = 24.0265$ and $\varepsilon = 0.001358$ are admissible for $b_n$. It follows that $a_n \geq 0$ for $n \geq 5347177639 \approx 5 \times 10^9$, while $b_n \geq 0$ for $n \geq 8126856 \approx 8 \times 10^6$.

The verification that $a_n, b_n \geq 0$ for the remaining values of $n$ can be done on a personal computer running Pari. The calculation takes a matter of hours for $b_n$, or a matter of days for $a_n$. It seemed too memory-intensive to compute all the values of $a_{n, \text{cusp}}$ for $n \leq n_0 = 5347177638$, whereas $a_{n, \text{eis}}$ and $\sigma_0(n)$ are given by elementary formulas in terms of the factors of $n$. We first compute $a_{n, \text{eis}}$ for all $n \leq n_0$, and check whether $a_{n, \text{eis}} - Cn\sigma_0(n) \geq 0$, in which case positivity follows from Theorem 2.3. If not, we add $n$ to a list of indices for which we will compute $a_{n, \text{cusp}}$. In the more difficult case of $a_n$, this list contains 67250 values, culminating in $n = 261777516$. For $b_n$, we only needed 4329 values of $b_{n, \text{cusp}}$, the last being $n = 556738$. The values can be found in the supplementary files checklist-an-67250.txt and checklist-bn-4329.txt at https://arxiv.org/abs/2211.09044. It took a week to make the list for $a_n$, and two days to confirm that $a_{n, \text{eis}} + a_{n, \text{cusp}} \geq 0$ for each of these $n$. No doubt some improvements could be made to our implementation, and in particular it would have been possible to check different ranges in parallel.

The dependence on $k$ in (2.14) was a challenge making our six-dimensional example more difficult in some ways than the higher-dimensional ones from [9]. When $k$ is large, an Eisenstein part of order $n^{k-1}$ will very quickly dominate a cuspidal part of order $n^{(k-1)/2}$ (although of course there could be difficulties for any given $k$ if the ratio $C/\varepsilon$ is too large). When $k$ is only 3, we encountered several examples where $n$ must be quite large before the asymptotics make themselves felt. For instance, with $T = 7$ and $N = 48$ in Proposition 2.2, forcing the inequalities for $n \leq 2000$ might have seemed more than sufficient, but it failed to make $a_n, b_n \geq 0$ for larger $n$. Choosing $x_j'$ according to that smaller linear program, instead of the solution $x_j$ used here, leads to an infeasible candidate where some of the corresponding coefficients $a_n', b_n'$ are negative. Indeed, in that case, $b_{2890}' < 0$. For various pairs $(N, T)$, some of our candidates seemed valid for tens of thousands of coefficients, only to fail at large values of $n$ with a special factorization for which the Eisenstein contribution was negative. The role of the factors is partly explained by Proposition 7.2.

3. Space of modular forms

In this section, we describe the space of candidates $g$ over which we optimize, and indicate how it relates to the presumed best packing $E_6$. We will take $g = \sum_j x_j f_j$ to be a linear combination of 44 modular forms $f_1, \ldots, f_{14}$. The quantity $b_0$, which
The weight relevant to 6-dimensional packing is
\[ k = \frac{D}{2} = 3. \]

The space also contains theta series for the dual lattices, but as we explain in Proposition 2.2, taking
\[ N = 48 = 2^4 \cdot 3 \]
The other subspace for character \( \chi_4 \), induced from the characters mod 3 and mod 4 given by \(-1 \mod 3 \mapsto -1\) and \(-1 \mod 4 \mapsto -1\). In principle, one could do even better using more characters, but these were already enough to prove Theorem 1.2. We choose \( T = 7 \) in Proposition 2.2, that is, we seek a solution of the form
\[ g(z) = 1 + a_7 q^7 + \ldots \quad (q = e^{2\pi i z}) \]
in the direct sum
\[ M_3(\Gamma_0(48), \chi_3) \oplus M_3(\Gamma_0(48), \chi_4). \]
The summand for \( \chi_3 \) contains the theta series of the lattice \( E_6 \), which is presumed to give a sphere packing of maximal density. In terms of the basis \( f_1, \ldots, f_{44} \) that we will use, this theta series is expressed as
\[ \vartheta_{E_6}(z) = 81f_1(z) - 9f_6(z) = 1 + 72q + 270q^2 + 936q^3 + 2160q^4 + \ldots \]
This motivated our choice of characters. From the outset, the summand \( M_3(48, \chi_3) \) in the optimization space gives enough flexibility to at least recover the density of the (presumed) optimal packing. The dual lattice \( E_6^* \), suitably scaled, has theta series
\[ 9f_1(z) - 9f_6(z) = 1 + 0q + 54q^2 + 72q^3 + 4q^4 + 432q^5 + 270q^6 + 0q^7 + 918q^8 + 720q^9 + \ldots \]
The other subspace for character \( \chi_4 \) contains the theta series of the lattices \( D_6 \) and \( Z^6 \), namely
\[ \vartheta_{D_6}(z) = 64f_{23}(z) - 4f_{29}(z) = 1 + 60q + 252q^2 + 544q^3 + 1020q^4 + \ldots \]
\[ \vartheta_{Z_6}(z) = 16f_{23}(z) - 4f_{29}(z) = 1 + 12q + 60q^2 + 160q^3 + 252q^4 + \ldots \]
The space also contains theta series for the dual lattices, but as we explain in Section 8, there is a linear relation between those of \( D_6 \), \( D_6^* \), and \( Z^6 \). We increased the level \( N \) along multiples of 12 so that all these theta series would be available, and tried several values of \( T \) at each level, stopping at the pair \((N = 48, T = 7)\) where the resulting bound was strong enough to prove Corollary 1.3. Theta series are convenient because they automatically satisfy the inequalities from Proposition 2.2.

An important feature of the choice of level \( N = 48 \) is that it allows certain quadratic twists. Recall, for instance from [20, Proposition 14.19], that for a modular form \( f \) and a Dirichlet character \( \chi \), the twist \( f \otimes \chi \) is defined by
\[ f(z) = \sum_n c_f(n)q^n \implies f \otimes \chi(z) = \sum_n c_f(n)\chi(n)q^n \]
Figure 3.1. Dimensions of the spaces of modular forms for each divisor of 48. The subspaces for the character $\chi_3$ are written to the left of each circled divisor, and those for $\chi_4$ to the right. The dimensions are written as $e + c$ where $e$ is the dimension of the Eisenstein subspace, and $c$ the dimension of the cuspidal subspace. At the bottom levels 3 and 4, the theta series of the lattices $E_6$, $E_6^*$, $D_6$, and $D_6^*$ form a basis of Eisenstein series. A new scaling $z \mapsto sz$, where $s$ divides 12 or 16, applies at each level. These scalings, together with quadratic twists of $E_6$ and $E_6^*$ arising at the top level, form a basis of Eisenstein series.

This is again modular, but generally of level higher than $f$. If $f$ has level $N$ with a character of conductor $N^*$, while $\chi$ is a character modulo $t$, then $f \otimes \chi$ has level equal to the least common multiple $\text{lcm}(N, N^*t, t^2)$ and character equal to that of $f$ multiplied by $\chi^2$. In particular, for a quadratic character, $\chi^2 = 1$ so $f \otimes \chi$ and $f$ have the same character. We will take $f$ of character $\chi_3$ at level $N = 3$, 12, or 24 and twist by $\chi_4$. In all these cases, $t = 4$ and $\text{lcm}(N, N^*t, t^2) = 48$, so these twists remain in the optimization space.

Another important aspect of the analysis is scaling. The divisors of $N = 48 = 2^4 \cdot 3$ are 1, 2, 3, 4, 6, 8, 12, 16, 24, 48. There are (non-zero) modular forms for $\chi_3$ at each sub-level 3, 6, 12, 24, 48 and modular forms for $\chi_4$ at 4, 8, 12, 16, 24, 48. Figure 3.1 shows the dimensions of these subspaces. These dimensions can be computed with Pari by finding a basis for each subspace using $\text{mfbasis}$. The spaces for $\chi_3$ and $\chi_4$ are implemented as $\text{mfinit}([48,3,-3])$ and $\text{mfinit}([48,3,-4])$, and similarly for divisors of 48, with $\text{mfinit}(...,[1])$ for the cuspidal space or $\text{mfinit}(...,[3])$ for the Eisenstein space. There are also explicit dimension formulas, which are explained for instance in [28, section 6.3] and implemented in Pari’s command $\text{mfdim}$.

In total, we have a 44-dimensional space of modular forms, with 24-dimensional Eisenstein subspace and 20-dimensional cuspidal subspace. As shown in the chart 3.1, there are 2 independent Eisenstein series at each of the bottom levels 3 and 4. These can be thought of as the theta series of $E_6$ and its dual $E_6^*$ for level 3, or $D_6$ and $D_6^*$ for level 4. Rather than these theta functions, the output of $\text{mfbasis}$ involves four Eisenstein series $f_1, f_6, f_{23}$, and $f_{29}$ which take a simpler form in
Atkin–Lehner transforms, and compute a finite number of Fourier coefficients.

The constant terms involve Dirichlet $L$-values at $1 - k$ for weight $k = 3$. See [16, Theorem 4.5.1].

The rest of the Eisenstein basis is given by scaling $f_1$ and $f_6$ by divisors of $48/3 = 16$, and scaling $f_{23}$ and $f_{29}$ by divisors of $48/4 = 12$, as well as two quadratic twists. Note that twisting by $\chi_4$ increases levels by $4^2$, hence from 3 to 48, as per [20, Proposition 14.19]. The twists are

$$f_{11} = f_1 \otimes \chi_4, \quad f_{11}(z) = \sum_{n=1}^{\infty} q^n \chi_4(n) \sum_{d|n} d^2 \chi_3(n/d)$$

$$f_{12} = f_6 \otimes \chi_4, \quad f_{12}(z) = \sum_{n=1}^{\infty} q^n \chi_4(n) \sum_{d|n} d^2 \chi_3(d)$$

The scalings are indexed by the size of $s$ in $z \mapsto sz$, so that

$$f_2(z) = f_1(2z), \quad f_7(z) = f_6(2z), \quad f_{24}(z) = f_{23}(2z), \quad f_{30}(z) = f_{29}(2z)$$

$$f_{25}(z) = f_{23}(3z), \quad f_{31}(z) = f_{29}(3z)$$

$$f_3(z) = f_1(4z), \quad f_8(z) = f_6(4z), \quad f_{26}(z) = f_{23}(4z), \quad f_{32}(z) = f_{29}(4z)$$

$$f_{27}(z) = f_{23}(6z), \quad f_{33}(z) = f_{29}(6z)$$

$$f_4(z) = f_1(8z), \quad f_9(z) = f_6(8z), \quad f_{28}(z) = f_{23}(12z), \quad f_{34}(z) = f_{29}(12z)$$

$$f_5(z) = f_1(16z) \quad f_{10}(z) = f_6(16z)$$

In the approach of Cohn–Triantafillou, the cuspforms are eventually discarded as a small correction to the Eisenstein series. The method can be applied without knowing them in closed form. All the necessary information can be obtained automatically using Pari’s commands to diagonalize the Hecke operators, calculate Atkin–Lehner transforms, and compute a finite number of Fourier coefficients.

However, our proof of Theorem 1.2 uses a candidate with vanishing coefficients $a_n = 0$ for all $n \equiv 1 \mod 4$. To show that an infinite number of coefficients vanish, we will need to know slightly more about the basis functions. In particular, for the character $\chi_3$, the cuspforms of level 48 can all be expressed as quadratic twists of $\chi_3$ and $\chi_4$:

$$f_1(z) = \sum_{n=1}^{\infty} q^n \sum_{d|n} d^2 \chi_3(n/d) = q + 3q^2 + 9q^3 + 13q^4 + \ldots$$

$$f_6(z) = \frac{1}{2} L(-2, \chi_3) + \sum_{n=1}^{\infty} q^n \sum_{d|n} d^2 \chi_3(d) = -\frac{1}{9} + q - 3q^2 + q^3 + 13q^4 + \ldots$$

$$f_{23}(z) = \sum_{n=1}^{\infty} q^n \sum_{d|n} d^2 \chi_4(n/d) = q + 4q^2 + 8q^3 + 16q^4 + \ldots$$

$$f_{29}(z) = \frac{1}{2} L(-2, \chi_4) + \sum_{n=1}^{\infty} q^n \sum_{d|n} d^2 \chi_4(d) = -\frac{1}{4} + q + q^2 - 8q^3 + q^4 + \ldots$$

The constant terms involve Dirichlet $L$-values at $1 - k$ for weight $k = 3$. See [16, Theorem 4.5.1].
The structure of the basis with respect to scalings is as follows:

\[
\begin{pmatrix}
  f_{20} \\
  f_{21} \\
  f_{22}
\end{pmatrix} = \begin{pmatrix}
  1 & 1 & 0 \\
  3 & 0 & -1 \\
  0 & 2 & -2
\end{pmatrix} \begin{pmatrix}
  f_{13} \otimes \chi_4 \\
  f_{16} \otimes \chi_4 \\
  f_{18} \otimes \chi_4
\end{pmatrix}
\]

The remaining basis elements \( f_{13}, f_{16}, f_{18}, f_{35}, f_{38}, f_{41}, f_{43}, f_{44} \) are handled using the GP/Pari library [2]. Their Fourier expansions begin as follows:

\[
\begin{align*}
  f_{13}(z) &= q - 3q^3 + 2q^7 + 9q^9 - 22q^{13} + \ldots \\
  f_{16}(z) &= 2q + 2q^3 - 12q^7 - 14q^9 + 20q^{13} + \ldots \\
  f_{18}(z) &= 2q - 14q^3 + 32q^5 - 12q^7 - 46q^9 + 32q^{11} + 20q^{13} + \ldots \\
  f_{35}(z) &= 2(q - q^2 - 2q^4 + 3q^6 + 8q^8 - 3q^9 + 2q^{10} - 6q^{12} + 2q^{13} + \ldots) \\
  f_{38}(z) &= -2(q + 2q^2 - 3q^3 - 8q^4 - 2q^5 + 6q^6 + 12q^7 + 8q^8 - 3q^9 - 4q^{10} - \ldots) \\
  f_{41}(z) &= q - 6q^5 + 9q^9 + 10q^{13} + \ldots \\
  f_{43}(z) &= 2(q + 6q^5 - 3q^9 - 14q^{13} + \ldots) \\
  f_{44}(z) &= -6(q^3 + 4q^7 - 12q^{11} + \ldots)
\end{align*}
\]

Some of them have simple expressions as products of Dedekind eta functions, for example

\[
f_{13}(z) = \eta(2z)^3 \eta(6z)^3 = q \prod_{n=1}^{\infty} (1 - q^{2n})^3 (1 - q^{6n})^3
\]

but we do not make any use of such formulas.

The transforms of the basis functions can be computed using Pari’s commands \texttt{mfatkininit} and \texttt{mfatkin}. The former outputs various data, including a matrix representing the action of \( z \mapsto -1/(Nz) \) on each basis function, together with a complex floating-point scalar for normalization. We must divide by this factor to compare (2.1), which is normalized so that the transform is an involution, with Pari’s convention by which the transform has coefficients in the same field as the original modular form (despite the factors \( i^k \) and \( N^{k/2} \)). The second command \texttt{mfatkin} can then be applied to individual forms.

The transform preserves the Eisenstein and cuspidal spaces for each character. It is represented by a block diagonal matrix:

\[
W = \begin{pmatrix}
  W_{3,e} & W_{3,c} \\
  W_{4,c} & W_{4,e}
\end{pmatrix}
\]
This is a $44 \times 44$ matrix where the first 12 indices correspond to Eisenstein series for $\chi_3$, the next 10 to cuspforms for $\chi_3$, and then likewise there are 12 and 10 more indices corresponding to Eisenstein series and cuspforms for $\chi_4$. Missing entries denote 0 for brevity. Column by column, $W_{3,e}$ gives the transforms of the Eisenstein series for $\chi_3$ and equals

$$
\frac{1}{4} \begin{pmatrix}
1 & 1/2 \\
32 & 12 & 3/2 & 3/2 \\
-4/3 & -1/6 & -3/2 & \\
-4/3 & -12 & & \\
-256/3 & -32/3 & -4/3 & -1/6 & -12 & -3/2 & -3/16 \\
96 & -768 & & & -4/3 & -12
\end{pmatrix}
$$

and likewise for cuspforms

$$
W_{3,c} = \frac{1}{4} \begin{pmatrix}
4 & 1/2 \\
32 & 12 & 3/2 & 3/2 \\
-4/3 & -1/6 & -3/2 & \\
-4/3 & -12 & & \\
-256/3 & -32/3 & -4/3 & -1/6 & -12 & -3/2 & -3/16 \\
96 & -768 & & & -4/3 & -12
\end{pmatrix}
$$

For $\chi_4$, the normalization is different, with $W_{4,e}$ given by

$$
\frac{1}{\sqrt{12}} \begin{pmatrix}
4 & 1/2 \\
32 & 12 & 3/2 & 3/2 \\
-4/3 & -1/6 & -3/2 & \\
-4/3 & -12 & & \\
-256/3 & -32/3 & -4/3 & -1/6 & -12 & -3/2 & -3/16 \\
96 & -768 & & & -4/3 & -12
\end{pmatrix}
$$

and likewise

$$
W_{4,c} = \frac{1}{\sqrt{12}} \begin{pmatrix}
4 & 1/2 \\
32 & 12 & 3/2 & 3/2 \\
-4/3 & -1/6 & -3/2 & \\
-4/3 & -12 & & \\
-256/3 & -32/3 & -4/3 & -1/6 & -12 & -3/2 & -3/16 \\
96 & -768 & & & -4/3 & -12
\end{pmatrix}
$$

The only basis functions with a constant term are $f_6 = -1/9 + \ldots$, $f_{29} = -1/4 + \ldots$, and their scalings $f_{7,8,9,10}$ and $f_{30,31,32,33,34}$. These are proportional to the transforms of $f_{1,2,3,4,5}$ and $f_{23,24,25,26,27,28}$, up to the factors from $W_{3,e}$ and $W_{4,e}$. The coefficient $b_0$ to be maximized therefore takes the form claimed in (3.1).
4. Choice of coefficients

In this section, we specify the values $x_j$ used to define the candidate $g = \sum_j x_j f_j$. Recall that the goal is to maximize a product $b_0 T^3$ where both $g = \sum_n a_n q^n = 1 + a_T q^2 + \ldots$ and its transform $\tilde{g} = \sum_n b_n q^n$ are required to have non-negative coefficients. We choose $T = 7$.

The values $x_1, \ldots, x_{44}$ were obtained as follows. First, consider a truncated linear program forcing the constraints from Proposition 2.2 for finitely many values of $n$. This relies on Pari’s command `mfcoefs` to obtain the coefficients of $f_1, \ldots, f_{44}$ that appear in these constraints, and `mfatkininit` or `mfatkin` for the transforms $e g_j$. To solve the linear program, we used the software package GLPK [26]. The resulting values for $x_j$ are given in floating-point to double precision, and the constraints are only approximately satisfied. We then guess values of $n$ for which a constraint such as $a_n \geq 0$ or $b_n \geq 0$ seems to hold with equality, modulo numerical imprecision. The vanishing coefficients, together with the normalization $a_0 = 1$, give a system of equations for $x_1, \ldots, x_{44}$ with coefficients in $\mathbb{Q}(\sqrt{N})$, since the Atkin-Lehner transform $g_j$ from (2.1) involves a factor $N^{k/2}$. Recall that $k = 3$ is odd in our case.

In our case, it appeared that $a_n = 0$ whenever $n \equiv 1 \mod 4$. As we will see in Section 6, this follows from just eleven equations:

$$
\begin{align*}
&x_{10} = x_{41} = x_{43} = 0 \\
&x_1 + x_{11} = 0 \\
&x_6 + x_{12} = 0 \\
&x_{23} + x_{29} = 0 \\
&x_{25} - x_{31} = 0 \\
&x_{35} - x_{38} = 0 \\
&x_{13} + x_{20} + 3x_{21} = 0 \\
&x_{16} + x_{20} + 2x_{22} = 0 \\
&x_{18} - x_{21} - 2x_{22} = 0
\end{align*}
$$

In addition to this infinite progression where $a_n = 0$, we impose

\begin{align*}
(4.1) & \quad a_0 = 1 \\
(4.2) & \quad a_n = 0 \quad \text{for } n \in \{2, 3, 4, 6, 8, 10, 11, 12, 22, 26, 32, 38, 60, 64, 88, 90, 92, 106, 164, 1932\} \\
(4.3) & \quad b_n = 0 \quad \text{for } n \in \{1, 2, 3, 4, 7, 8, 9, 10, 13, 14, 36, 82\}
\end{align*}

Together, these amount to $11 + 1 + 20 + 12 = 44$ equations for the 44 variables $x_j$. There is a unique solution, which can be calculated exactly in $\mathbb{Q}(\sqrt{N})$, for instance by working in a polynomial ring modulo $X^2 - N$ and using Pari’s command `matinverseimage`. The fact that $a_{1932} = 0$ illustrates that, even for $n \approx 2000$, the Eisenstein part is not yet dominant over the cuspidal part. Moreover, without checking a high enough number of initial coefficients, one would have fewer than 44 equations and the system would not uniquely specify the solution. In fact, in our first attempt at writing the solution exactly, we had only 43 equations but could determine the one remaining free parameter from the inequalities $a_n, b_n \geq 0$ and the objective to maximize $b_0$. 

Table 4.1. Approximate values of the 44 coefficients $x_i \in \mathbb{Q}(\sqrt{3})$ expressing $g = \sum_i x_i f_i$ in the basis from Section 3, written as $i = 10h + v$ with $h \leq 4$ running horizontally and $v \leq 9$ running vertically. The values $x_{10} = x_{41} = x_{43} = 0$ are exact. Exact values for the other $x_i$ can be found in a supplementary file supplement.txt available at https://arxiv.org/abs/2211.09044. The coefficients $x_1$ through $x_{22}$ correspond to the character $\chi_3$, while $x_{23}$ through $x_{44}$ correspond to $\chi_4$. The Eisenstein indices are $1 \leq i \leq 12$ and $23 \leq i \leq 34$, with the cuspidal part shaded.

| i | 0     | 1     | 2     | 3     | 4     |
|---|-------|-------|-------|-------|-------|
| 0 | 0.0000| 0.1155| -0.0880| -1.0698|
| 1 | 0.5426| -0.5426| 0.1669| -0.3790| 0.0000|
| 2 | -1.2415| 0.0961| -0.0774| -0.0134| -18.0274|
| 3 | -3.3991| -0.6164| -0.0943| 2.8874| 0.0000|
| 4 | -1.7265| 0.0010| 0.0606| -6.7493| -1.6689|
| 5 | -0.8136| 0.4724| -0.3790| -0.4690|
| 6 | -0.0961| 0.0394| 1.4476| -1.3134|
| 7 | -0.2722| -1.0533| 3.5885| -1.5459|
| 8 | 0.5140| 0.0120| -4.8033| -0.4690|
| 9 | 0.4126| 0.08926| 0.0943| -0.8239|

Table 4.1 shows the solution $x$ truncated to 4 digits. The exact values are complicated, for instance

$$x_1 = \frac{1}{2} \cdot 13305835801389714487894701329762261844642710724107326137255$$

\[ \frac{5}{2} \left( 5580210204055442306717732912787404128951045878479406537152 \right) \]

\[ \frac{5}{2} \left( 5933985634895097866727312605651608472538244579636113805857 \right) \sqrt{3} \]

which is approximately 0.5426088049814009651386767534544603187 \ldots Approximate values are enough to check any strict inequality $a_n > 0$ or $b_n > 0$. The exact quadratic values are important in cases where $a_n$ or $b_n$ is 0, which in the first stage of the calculation could have been a very small negative number. However, these zeros have all been accounted for in setting up the system of equations, so we need only the existence of an exact solution rather than the particular values.

The factors of 1932 = 12 · 7 · 23 play a role in the vanishing coefficient $a_{1932} = 0$, the largest from (4.2). We will see in Section 7 that the Eisenstein contribution to $a_n$ is at its smallest relative to $n^2$ when $n = 12n_0$, with $n_0 \equiv 5 \mod 12$ and $n_0$ divisible by many primes $p \equiv 7 \mod 12$.

5. The cuspidal part

In this section, we bound the contributions to $a_n$ and $b_n$ from cuspforms, and prove some observations about the basis functions that will be used to show that $a_n = 0$ for $n \equiv 1 \mod 4$. Both applications rely on a change of basis that is also useful for computational purposes.

Proposition 5.1. The following relations hold between the coefficients of $q^n$ in the basis functions $f_1, \ldots, f_{44}$. Write $[i]_n$ for the coefficient of $q^n$ in $f_i$.

(1) For $n \equiv 1 \mod 4$, $[44]_n = 0$. 

(2) For $n \not\equiv 1 \mod 4$, $[41]_n = 0.$

(3) For $n \equiv 1 \mod 4$, $[35]_n = -[38]_n.$

(4) For $n \equiv 1 \mod 3$, $[22]_n = 0.$

(5) For $n \equiv 2 \mod 3$, $[20]_n = 0.$

All of these patterns can be explained by changing to a basis of Hecke eigenforms. We refer to [19, Chapter 6] for background on Hecke operators. This is one of the aspects of the framework from [9] that changes slightly because we consider modular forms with a non-trivial character. In particular, not all the Hecke operators $T_n$ are self-adjoint, and in fact the adjoint of $T_n$ on a space of modular forms with character $\chi$ is $\chi(n)T_n$ [19, Theorem 6.20]. Nevertheless, the $T_n$ are commuting normal operators, and can therefore be diagonalized simultaneously.

Proposition 5.2. For all $n \geq 1$, the total contributions from cuspforms to $q^n$ in $g$ or $\bar{g}$ are at most:

\begin{align*}
(5.1) \quad |a_{n,\text{cusp}}| &\leq 21.6161 \cdot n\sigma_0(n) \\
(5.2) \quad |b_{n,\text{cusp}}| &\leq 24.0266 \cdot n\sigma_0(n)
\end{align*}

For comparison, note that the cuspidal part involves a large value $x_{42} \approx -18.0274$ from Table 4.1.

To apply Theorem 2.3, we change to a basis of Hecke eigenforms computed from Pari’s commands `mfheckemat` and `mateigen`. The command `mfheckemat` returns a matrix representing a Hecke operator $T_n$ acting on the cuspidal subspace for $\chi_3$ or $\chi_4$, and `mateigen` can then give a basis of eigenvectors of this matrix. The output of `mateigen` in our case involves floating-point approximations to $\sqrt{8}$ and $\sqrt{3}$, which are easy enough to recognize, and can also be justified rigorously by factoring the characteristic polynomial.

It is enough to consider $f_{13}, f_{16}, f_{18}, f_{20}, f_{21}, f_{22}$ for $\chi_3$, and $f_{35}, f_{38}, f_{41}, f_{43}, f_{44}$ for $\chi_4$. The rest of the basis is obtained by scaling these. We divide by the leading coefficient to obtain the following normalized Hecke eigenforms. They come in pairs under complex conjugation, which we denote by $K$. This is an abbreviation for complex conjugation of each coefficient in the $q$-series, not pointwise conjugation as functions of a variable $z$ in the half-plane $\operatorname{Im}(z) > 0$. In the same way, we will write $h = \Re(h) + i \Im(h)$ and $Kh = \Re(h) - i \Im(h)$ coefficient-wise. The basis functions $f_j$ have real coefficients (in fact, integer coefficients).

For $\chi_3$, let

\[
\begin{align*}
h_1 &= f_{13} \\
h_2 &= \frac{1}{4\sqrt{2}i} \left(-1 + \sqrt{8}i\right)f_{16} + f_{18} \\
h_3 &= \frac{-1}{4\sqrt{2}i} \left(-1 - \sqrt{8}i\right)f_{16} + f_{18} = Kh_2 \\
h_4 &= \frac{-1}{8} \left(-f_{20} - 2f_{21} + f_{22}\right) \\
h_5 &= \frac{3}{16(1 - \sqrt{2}i)} \left(2 - \sqrt{8}i\right)f_{20} + \frac{-2 + \sqrt{8}i}{3}f_{21} + f_{22} \\
h_6 &= \frac{3}{16(1 + \sqrt{2}i)} \left(2 + \sqrt{8}i\right)f_{20} + \frac{-2 - \sqrt{8}i}{3}f_{21} + f_{22} = Kh_5
\end{align*}
\]
and, for $\chi_4$,

\[
\begin{align*}
    h_7 &= \frac{1}{\sqrt{12}i} \left( (1 + \sqrt{3}i) f_{35} + f_{38} \right) \\
    h_8 &= \frac{-1}{\sqrt{12}i} \left( (1 - \sqrt{3}i) f_{35} + f_{38} \right) = Kh_7 \\
    h_9 &= \frac{-1}{\sqrt{12}i} \left( -\sqrt{3}i f_{43} + f_{44} \right) \\
    h_{10} &= \frac{1}{\sqrt{12}i} \left( \sqrt{3}i f_{43} + f_{44} \right) = Kh_9 \\
    h_{11} &= f_{41} - 3 f_{42} \\
    h_{12} &= f_{41} + 3 f_{42}
\end{align*}
\]

(in the last two, $f_{42}(z) = f_{41}(3z)$ is included to diagonalize $T_3$; $f_{41}$ is already an eigenform for $T_p$ with $p$ coprime to 48). The original basis functions are then

\[
\begin{align*}
    f_{13} &= h_1 \\
    f_{16} &= h_2 + h_3 = 2 \text{Re}(h_2) \\
    f_{18} &= (1 + \sqrt{8}i)h_2 + (1 - \sqrt{8}i)h_3 = f_{16} - 2 \sqrt{8} \text{Im}(h_2) \\
    f_{20} &= \frac{8}{7} (h_4 + h_5 + h_6) = \frac{8}{7} (h_4 + 2 \text{Re}(h_5)) \\
    f_{21} &= \frac{24}{7} h_4 + \left( -\frac{4}{7} - \sqrt{8}i \right) h_5 + \left( -\frac{4}{7} + \sqrt{8}i \right) h_6 = \frac{24}{7} h_4 - \frac{8}{7} \text{Re}(h_5) + 2 \sqrt{8} \text{Im}(h_5) \\
    f_{22} &= 4 \sqrt{2}i (h_6 - h_5) = 8 \sqrt{2} \text{Im}(h_5)
\end{align*}
\]

for $\chi_3$, and for $\chi_4$,

\[
\begin{align*}
    f_{35} &= h_7 + h_8 = 2 \text{Re}(h_7) \\
    f_{38} &= (-1 + \sqrt{3}i) h_7 - (1 + \sqrt{3}i) h_8 = -f_{35} - 2 \sqrt{3} \text{Im}(h_7) \\
    f_{41} &= \frac{1}{2} h_{11} + \frac{1}{2} h_{12} \\
    f_{43} &= h_9 + h_{10} = 2 \text{Re}(h_9) \\
    f_{44} &= \sqrt{3}i (h_{10} - h_9) = 2 \sqrt{3} \text{Im}(h_9)
\end{align*}
\]

**Proof of Proposition 5.1.** The calculation relies on [19, Theorem 6.20], which gives the adjoint of a Hecke operator acting on a space of modular forms with character $\chi$:

\[
\langle T_n f, g \rangle = \chi(n) \langle f, T_n g \rangle
\]

In the case that $f = g$ is an eigenform $h$ with eigenvalue $\lambda_n$ for $T_n$, it follows from the conjugate-linearity in the second factor that $\lambda_n = \chi(n) \overline{\lambda_n}$. With $\chi = \chi_4$, this implies that the coefficient of $q^n$ in a Hecke eigenform $h$ is real or imaginary according as $n \equiv \pm 1 \bmod 4$. Likewise for $\chi_3$, the congruence $n \equiv \pm 1 \bmod 3$ determines whether $h$ has real or imaginary coefficient.

If $n \equiv 1 \bmod 3$, then we are in the self-adjoint case so $h_5$ has real coefficients. This explains why $f_{22} = 8 \sqrt{2} \text{Im}(h_5)$ satisfies $[22]_n = 0$ along this progression, as claimed in item (4) of Proposition 5.1. Likewise for item (5): $f_{20}$ is a real combination of $h_4$ and $2 \text{Re}(h_5)$, whereas these would have imaginary coefficients in the skew-adjoint case $n \equiv 2 \bmod 3$, so $[20]_n = 0$. 
For $n \equiv 1 \pmod{4}$, the eigenforms for $\chi_4$ have real coefficients. It follows that

$$ [38]_n = -[35]_n, \quad [44]_n = 0 $$

as per items (1) and (3). If $n \not\equiv 1 \pmod{4}$, then the eigenforms have imaginary coefficients, whereas $f_{41}$ is real, so $[41]_n = 0$ as per item (2).

Proof of Proposition 5.2. We apply Deligne’s bound Theorem 2.3 to each of the eigenforms $h_j$. For the scalings $s = 1, 2, 3, 4$, this gives a multiple of $\frac{n}{2}\sigma_0(\frac{n}{s})$ and we note for simplicity that $\sigma_0(n/s) \leq \sigma_0(n)$. Some convenient absolute values, in view of the change of basis, are

$$ |1 + \sqrt{3}i| = 3, \quad \left|\frac{-4}{7} + \sqrt{8i}\right| = \frac{12(2 + \sqrt{2})}{7}, \quad |\pm 1 + \sqrt{3}i| = 2 $$

The cuspidal coefficients for $\chi_3$ can then be bounded by

$$ ||13|| = |[h_1]_n| \leq n\sigma_0(n) $$

$$ ||16|| = |2\operatorname{Re}(h_2)_n| \leq 2n\sigma_0(n) $$

$$ ||18|| = |(1 + \sqrt{3}i)[h_2]_n| \leq 6n\sigma_0(n) $$

$$ ||20|| = \frac{8}{7}|[h_4]_n + [h_5]_n + [h_6]_n| \leq \frac{24}{7}n\sigma_0(n) $$

$$ ||21|| = \frac{24}{7}|[h_4]_n + (4/7 - \sqrt{8i})[h_5]_n + (4/7 + \sqrt{8i})[h_6]_n| \leq \frac{24}{7}(3 + \sqrt{2})n\sigma_0(n) $$

$$ ||22|| = 8\sqrt{2}|\operatorname{Im}(h_5)|_n \leq 8\sqrt{2}n\sigma_0(n) $$

It follows from (3.11) that $f_{20}$, $f_{21}$, $f_{22}$ contribute only for odd $n$; for even $n$ the factor $\chi_4(n)$ vanishes. For $\chi_4$, we have

$$ [35]_n = |2\operatorname{Re}(h_7)_n| \leq 2n\sigma_0(n) $$

$$ [38]_n = |(-1 + \sqrt{3}i)h_7 - (1 + \sqrt{3}i)h_8| \leq 4n\sigma_0(n) $$

$$ [41]_n \leq n\sigma_0(n) $$

$$ [43]_n \leq 2n\sigma_0(n) $$

$$ [44]_n = 2\sqrt{3}|\operatorname{Im}(h_9)_n| \leq 2\sqrt{3}n\sigma_0(n) $$

If $n$ is divisible by 12, the cuspidal part is

$$ x_{15}[13]_n + x_{14}[13]_{n/2} + x_{15}[13]_{n/4} $$

$$ x_{16}[16]_n + x_{17}[16]_{n/2} $$

$$ x_{18}[18]_n + x_{19}[18]_{n/2} $$

$$ x_{35}[35]_n + x_{36}[35]_{n/2} + x_{37}[35]_{n/4} $$

$$ x_{38}[38]_n + x_{39}[38]_{n/2} + x_{40}[38]_{n/4} $$

$$ x_{41}[41]_n + x_{42}[41]_{n/3} $$

$$ x_{43}[43]_n $$

$$ x_{44}[44]_n $$

If $n$ is not sufficiently divisible by 2 or 3, we remove the columns for $n/2$, $n/4$, or $n/3$. If $\gcd(n, 48) = 1$ or 3, then we add the contribution from the quadratic twists $f_{20}$, $f_{21}$, and $f_{22}$. Case by case, we substitute the values of $x_j$ from Table 4.1 and apply the triangle inequality for an upper bound on $|a_{n,\text{cusp}}|$. The upper bound for $|b_{n,\text{cusp}}|$
Their coefficients are multiplicative: for a normalized Hecke eigenform \( h \),
\[
h(z) = \sum_n \lambda(n)q^n, \gcd(m, n) = 1 \implies \lambda(mn) = \lambda(n)\lambda(m)
\]
as one can recall for instance from [19, eq. (6.83), (6.60)]. This makes it possible to compute \( \lambda(n) \) in terms of the factors of \( n \), which are much smaller than \( n \) itself in many of the cases we will need to check. We will see in Section 7 that the Eisenstein part is smallest when \( n \) has many factors. It is therefore very efficient to compute the cuspidal part via
\[
\lambda(\prod_{p^\ell || n} p^\ell) = \prod_{p^\ell || n} \lambda(p^\ell)
\]
although we will still need to compute cuspidal coefficients at some large primes.

This multiplicative structure gives the following formula for \( a_{n,\text{cusp}} \) in case \( n = 2^a3^b7^c n_0 \) is even. It is expressed using an indicator function \( \mathbb{1}_{a \geq 2} \) which is 0 if \( n \) is not sufficiently divisible by 2, or 1 if the scalings \( f_{15,37,40} \) under \( z \mapsto 4z \) are present. The products are taken over prime powers exactly dividing \( n_0 \). Recall that the Hecke basis involves several conjugate pairs such as \( h_3 = Kh_2 \), which we combine into real and imaginary parts.

\[
a_{n,\text{cusp}} = \left( x_{13}[h_1]_{2^a} + x_{14}[h_1]_{2^{a-1}} + \mathbb{1}_{a \geq 2} x_{15}[h_1]_{2^{a-2}} \right)[h_1]_{3^b} \prod_{p^\ell || n_0} [h_1]_{p^\ell} + 2\Re \left( (x_{16}[h_2]_{2^a} + x_{18}(1+\sqrt{8}i)[h_2]_{2^a} + (x_{17} + x_{19}(1+\sqrt{8}i)[h_2]_{2^{a-1}})[h_2]_{3^b} \prod_{p^\ell || n_0} [h_2]_{p^\ell} \right)
\]

\[
+ 2\Re \left( (x_{35}[h_7]_{2^a} + x_{36}[h_7]_{2^{a-1}} + x_{37} \mathbb{1}_{a \geq 2} [h_7]_{2^{a-2}} + (-1 + \sqrt{3}i) \left( x_{38}[h_7]_{2^a} + x_{39}[h_7]_{2^{a-1}} + x_{40} \mathbb{1}_{a \geq 2} [h_7]_{2^{a-2}} \right)[h_7]_{3^b} \prod_{p^\ell || n_0} [h_7]_{p^\ell} \right)
\]

Even values of \( n \) are the most important for us. When we generate the list of values for which \( a_{n,\text{cusp}} \) needs to be computed, the last odd value is \( n = 315 \). We already know \( a_n \geq 0 \) up to that point since these inequalities are among the constraints of the linear program that was solved in the first place. For \( n \) of this size, it is also easy to compute \( a_{n,\text{cusp}} \) in the original basis. The Hecke basis becomes decisive for larger values of \( n \), where an attempt to compute \( a_{n,\text{cusp}} \) directly in the original basis exhausted our available memory. Using (5.4) instead, the calculation is only a matter of time.

If \( n \) were odd, (5.4) would need to be modified by adding the quadratic twists \( f_{20}, f_{31}, f_{42} \) from (3.11), as well as \( f_{43}, f_{44} \). The latter do not contribute for even \( n \) because they are combinations of Hecke eigenforms for \( \chi_4 \), as in Proposition 5.1 parts (1) and (2).
These lead to pairwise cancellations between the Eisenstein series \( f \).

To prove this, we analyze the contributions from Eisenstein series and cuspforms for

variables:

\[
 x \equiv n f \mod \text{divisors } d \chi \text{ Eisenstein series for } \text{both of the characters } \chi \text{ and } \chi_4. \text{ Each of the four vanishes separately.}
\]

The Eisenstein contribution vanishes because of four linear relations between the variables:

\[
 x_1 + x_{11} = x_6 + x_{12} = x_{23} + x_{29} = x_{25} - x_{31} = 0.
\]

These lead to pairwise cancellations between the Eisenstein series \( f_1, f_{11}, f_6, f_{12}, f_{23}, f_{29}, f_{25}, \text{and } f_{31}. \) Recall that the coefficient of \( q^n \) in each of these is given by

\[
 f_1 : \sum_{d|n} d^2 \chi_3(n/d), \quad f_{11} : \chi_4(n) \sum_{d|n} d^2 \chi_3(n/d)
\]

\[
 f_6 : \sum_{d|n} d^2 \chi_3(d), \quad f_{12} : \chi_4(n) \sum_{d|n} d^2 \chi_3(d)
\]

\[
 f_{23} : \sum_{d|n} d^2 \chi_4(n/d), \quad f_{29} : \sum_{d|n} d^2 \chi_4(d)
\]

\[
 f_{25}(z) = f_{23}(3z), \quad f_{31}(z) = f_{29}(3z)
\]

If \( n \equiv 1 \mod 4 \), then \( \chi_4(n) = 1 \), while (6.2) gives \( x_1 = -x_{11} \) and \( x_6 = -x_{12} \). Thus the contributions to \( q^n \) from \( f_1 \) and \( f_{11} \) cancel out. Likewise for \( f_6 \) and \( f_{12} \). The Eisenstein series for \( \chi_3 \) thus cancel in pairs.

To see the cancellation between \( f_{23} \) and \( f_{29} \), we observe that each term vanishes in the sum over divisors \( d | n \). Indeed, since \( x_{29} = -x_{23} \) from (6.2),

\[
 x_{23} \sum_{d|n} d^2 \chi_4(n/d) + x_{29} \sum_{d|n} d^2 \chi_4(d) = x_{23} \left( \sum_{d|n} d^2 (\chi_4(n/d) - \chi_4(d)) \right) = x_{23} \sum_{d|n} d^2 \cdot 0
\]

Here, \( \chi_4(n/d) = \chi_4(d) \) because \( n \equiv 1 \mod 4 \).

Something similar happens for \( f_{25} \) and \( f_{31} \). These two are rescalings of \( f_{23} \) and \( f_{29} \), which only contribute if \( n \) is divisible by 3. In that case, again assuming \( n \equiv 1 \mod 4 \), we have \( n = 3 \cdot n/3 \) where \( n/3 \equiv -1 \mod 4 \). From (6.2), we have \( x_{25} = x_{31} \). The contribution of this pair is then

\[
 x_{25} \sum_{d|n/3} d^2 \chi_4 \left( \frac{n/3}{d} \right) + x_{31} \sum_{d|n/3} d^2 \chi_4(d) = x_{25} \sum_{d|n/3} d^2 \left( \chi_4 \left( \frac{n/3}{d} \right) + \chi_4(d) \right) = x_{25} \sum_{d|n/3} d^2 \cdot 0
\]
This time, \( \chi_4(n/3d) \) and \( \chi_4(d) \) have the opposite sign because \( n/3 \equiv -1 \mod 4 \).

The cuspsforms for \( \chi_4 \) contribute nothing because of the relations

\[
x_{35} = x_{38}, \quad x_{41} = x_{43} = 0.
\]

There is no contribution from \( f_{41} \) and \( f_{43} \) because \( x_{41} = x_{43} = 0 \), and the rest can be analyzed using Proposition 5.1. For \( n \equiv 1 \mod 4 \), there is no contribution from \( f_{44} \) and the coefficients of \( q^n \) in \( x_{35} \) and \( x_{38} \) are negatives of each other. The latter cancel because \( x_{35} = x_{38} \). If \( n \) is divisible by 3, then there is a further term \( x_{42}\{41\}/3 \) from the scaling \( f_{42}(z) = f_{41}(3z) \). In this case, \( n = 3 \cdot n/3 \) where \( n/3 \equiv 3 \mod 4 \), so \( \{41\}/3 = 0 \) by Proposition 5.1 again.

Finally, the cuspsforms for \( \chi_3 \) contribute

\[
x_{13}[13]_n + x_{16}[16]_n + x_{18}[18]_n + x_{20}[20]_n + x_{21}[21]_n + x_{22}[22]_n
\]

Recall from (3.11) that \( f_{20}, f_{21}, f_{22} \) are linear combinations of \( f_{13} \otimes \chi_4, f_{16} \otimes \chi_4, f_{18} \otimes \chi_4 \). The sum becomes

\[
x_{13}[13]_n + x_{16}[16]_n + x_{18}[18]_n + \\
\chi_4(n)(x_{20}(13]_n + 16]_n) + x_{21}(3[13]_n - [18]_n) + x_{22}(2[16]_n - 2[18]_n))
\]

For \( n \equiv 1 \mod 4 \), \( \chi_4(n) = 1 \) and we can collect terms as follows

\[
[13]_n(x_{13} + x_{20} + 3x_{21}) + [16]_n(x_{16} + x_{20} + 2x_{22}) + [18]_n(x_{18} - x_{21} - 2x_{22})
\]

This is \( 0 + 0 + 0 \) because of the linear relations

\[
x_{13} + x_{20} + 3x_{21} = x_{16} + x_{20} + 2x_{22} = x_{18} - x_{21} - 2x_{22} = 0.
\]

7. The Eisenstein part

In this section, we show that the Eisenstein series contribute at least a positive multiple of \( n^2 \) to \( a_n \) and \( b_n \), except for the progression we have already noted where \( a_n = 0 \).

**Proposition 7.1.** For \( n \equiv 1 \mod 4 \), \( a_{n,eis} = 0 \). For \( n \not\equiv 1 \mod 4 \),

\[
a_{n,eis} > \varepsilon n^2, \quad \varepsilon = 8.753 \times 10^{-6}
\]

For all \( n \geq 1 \),

\[
b_{n,eis} > \varepsilon^* n^2, \quad \varepsilon^* = 1.358 \times 10^{-3}
\]

The Eisenstein part is dictated by the factors of \( n = 2^a \cdot 3^b \cdot n_0 \), leading to several cases over which we take the minimum to obtain the numerical values in Proposition 7.1. Table 7.1 gives admissible values for \( \varepsilon \) and \( \varepsilon^* \) depending on the exponents \( a \) and \( b \).

The Eisenstein contribution is expressed in terms of four basic functions:

\[
\sigma_3^{+}(n) = \sum_{d|n} d^2 \chi_3(n/d), \quad \sigma_3^{-}(n) = \sum_{d|n} d^2 \chi_3(d),
\]

\[
\sigma_4^{+}(n) = \sum_{d|n} d^2 \chi_4(n/d), \quad \sigma_4^{-}(n) = \sum_{d|n} d^2 \chi_4(d)
\]

In this notation, the basis functions from (3.5)–(3.8) are \( f_1(z) = \sum_n \sigma_3^{+}(n)q^n \), \( f_6(z) = -1/9 + \sum_n \sigma_3^{-}(n)q^n \), \( f_{23}(z) = \sum_n \sigma_4^{+}(n)q^n \), and \( f_{29}(z) = -1/4 + \sum_n \sigma_4^{-}(n)q^n \). The remaining Eisenstein series are scalings or quadratic twists of these four.
Table 7.1. Admissible values in the inequalities $a_{n,\text{eis}} \geq \varepsilon n^2$ and $b_{n,\text{eis}} \geq \varepsilon^* n^2$ for $n = 2^a3^bn_0$ where $n_0$ is not divisible by 2 or 3. The case $a_{n,\text{eis}} = 0$ for $n \equiv 1 \mod 4$ is treated separately. The worst cases, shown in red, are $a = 2, b \geq 1$ for $a_n$ and $a = 1, b = 0$ for $b_n$.

| $a$ | $b$ | $\varepsilon$ | $\varepsilon^*$ |
|-----|-----|----------------|-----------------|
| 0   | 0   | 0.6468         | 0.1612          |
| 1   | 0   | 0.003742       | 0.001358        |
| 2   | 0   | 0.0008134      | 0.03902         |
| 3   | 0   | 0.0008264      | 0.2038          |
| $\geq 4$ | 0 | 0.0002649 | 0.01181 |

From their definition as sums over divisors, the four functions $\sigma = \sigma_{3,4}^\pm$ are multiplicative:

\[(7.3)\] \[\gcd(m,n) = 1 \implies \sigma(mn) = \sigma(m)\sigma(n)\]

Their values on prime powers can be evaluated as geometric series. Assuming $\chi(p) = \pm 1$, where $p$ is prime,

\[(7.4)\] \[\sigma^-(p^r) = \frac{(p^2\chi(p))^{r+1} - 1}{p^2\chi(p) - 1}\]

\[(7.5)\] \[\sigma^+(p^r) = \frac{p^{2e+2} - \chi(p)^{r-1}}{p^2 - \chi(p)}\]

\[(7.6)\] \[\sigma^\pm(p^r) = \chi(p)^r \sigma^\pm(p^e)\]

and likewise

\[(7.7)\] \[\sigma^-(n) = \chi(n)\sigma^+(n)\]

as long as $n$ has no factors for which $\chi$ vanishes. However, there are special cases when $\chi(p) = 0$ and the sums reduce to a single term: $\sigma^+(p^r) = p^{2e}$, $\sigma^-(p^r) = 1$. In particular,

\[(7.8)\] \[\sigma_3^+(3^e) = 3^{2e}, \sigma_4^+(2^e) = 2^{2e}, \quad \sigma_3^-(3^e) = 1 = \sigma_4^-(2^e)\]

The other values follow the pattern of (7.5) with $\chi_3(2) = \chi_4(3) = -1$:

\[(7.9)\] \[\sigma_3^+(2^e) = \frac{4^{e+1} + (-1)^e}{5}, \quad \sigma_4^+(3^e) = \frac{9^{e+1} + (-1)^e}{10}\]

For composite $n$, the values $\sigma(n)$ are then given by products. It follows that the Eisenstein coefficients are of order $n^2$, except that $\sigma^-(n)$ can be smaller when $n$ is highly divisible by 2 or 3.

To analyze this further, write $n = 2^a3^b n_0$ where $\gcd(n_0, 48) = 1$. In any of the four cases $\sigma = \sigma_{3,4}^\pm$,

\[(7.10)\] \[\sigma(n) = \sigma(2^a)\sigma(3^b)\sigma(n_0).\]
The factors $\sigma(n^e)$ and $\sigma(n^f)$ can be evaluated with (7.6), (7.8), and (7.9). The Eisenstein contribution $a_{n,eis}$ or $b_{n,eis}$ is a sum of up to 22 such terms. We must consider the factorization $n = 2^a3^bn_0$ to see which terms arise, but in any case (7.7) allows us to write the sum as

$$a_{n,eis} = \sigma^+_3(n_0)X_3 + \sigma^+_4(n_0)X_4, \quad b_{n,eis} = \sigma^+_3(n_0)Y_3 + \sigma^+_4(n_0)Y_4$$

where $X_3$ and $X_4$ are linear combinations of the variables $x_j$ (or $y_j$ for the transform, with the corresponding linear combinations $Y_3$ and $Y_4$). We factor out $\sigma^+_3(n_0)$:

$$a_{n,eis} = \sigma^+_3(n_0) \left( X_3 + \frac{\sigma^+_4(n_0)}{\sigma^+_3(n_0)} X_4 \right)$$

Depending on the sign of $X_4$ (or $Y_4$ for the transform), we use either an upper or a lower bound for the ratio of $\sigma^+_3(n_0)$ and $\sigma^+_4(n_0)$. Bounding the ratio is more efficient than considering the worst case for each separately, and for some $n$, it is important to be careful. See (7.21) below for the most delicate case.

**Proposition 7.2.** For $n_0$ not divisible by 2 or 3,

$$\prod_{p \equiv \pm 5 \mod 12} \frac{p^2 - 1}{p^2 + 1} \leq \frac{\sigma^+_3(n_0)}{\sigma^+_4(n_0)} \leq \prod_{p \equiv \pm 5 \mod 12} \frac{p^2 + 1}{p^2 - 1},$$

where the products are over all primes congruent to 5 or 7 mod 12. Numerically,

$$0.94999 \leq \frac{\sigma^+_4(n_0)}{\sigma^+_3(n_0)} \leq 1.06966$$

Separately, still for $n_0$ not divisible by 2 or 3, the two functions satisfy

$$\prod_{p \equiv \pm 3 \mod 4, p \neq \pm 3} (1 - p^{-2}) \leq \frac{\sigma^+_3(n_0)^2}{n_0} \leq \prod_{p \equiv \pm 1 \mod 4} (1 + p^{-2}) < 1.0336$$

$$0.9631 < \prod_{p \equiv \pm 3 \mod 4, p \neq \pm 3} (1 - p^{-2}) \leq \frac{\sigma^+_4(n_0)^2}{n_0} \leq \prod_{p \equiv \pm 1 \mod 4} (1 + p^{-2}) < 1.0545$$

**Proof.** Let $p^e$ be the prime powers dividing $n_0$ exactly. By the geometric series evaluating $\sigma(p^e)$, the ratio in (7.13) is

$$\frac{\sigma^+_4(n_0)}{\sigma^+_3(n_0)} = \prod_{p^e || n_0} \left( \frac{p^{2e+2} - \chi_4(p)e^{-1}}{p^{2e} - \chi_4(p)} \right) \left( \frac{p^2 - \chi_3(p)}{p^{2e+2} - \chi_3(p)e^{-1}} \right)$$

If $\chi_4(p) = \chi_3(p)$, then the corresponding factor is 1. Only the primes congruent to 5 or 7 modulo 12 contribute, so that $\chi_4(p)$ and $\chi_3(p)$ have opposite sign. If $e$ is odd, then $\chi_4(p)^{-1} = \chi_3(p)^{e-1}$ in either case. It follows that

$$\frac{\sigma^+_4(n_0)}{\sigma^+_3(n_0)} = \prod_{p \equiv \pm 5, e \text{ odd}} \frac{p^2 + 1}{p^2 - 1} \prod_{p \equiv \pm 7, e \text{ odd}} \frac{p^2 - 1}{p^2 + 1} \prod_{p \equiv \pm 5, e \text{ even}} \frac{p^2 + 1}{p^2 - 1} \prod_{p \equiv \pm 7, e \text{ even}} \frac{p^2 - 1}{p^2 + 1} \prod_{p \equiv \pm 1, e \text{ even}} \frac{p^2 - 1}{p^2 + 1} \prod_{p \equiv \pm 7, e \text{ even}} \frac{p^2 - 1}{p^2 + 1}$$

For $p \equiv 5$, we have $(p^2 + 1)/(p^2 - 1) > 1$, whereas $(p^2 - 1)/(p^2 + 1) < 1$ for $p \equiv 7$ mod 12. These inequalities are reversed for the terms $(p^{2e+2} \pm 1)/(p^{2e+2} \mp 1)$, so if present, those extra factors bring the product farther from the maximum or minimum. The extreme cases are where $n_0$ is a product of many primes all congruent
to 5, for the upper bound, or to 7 for the lower bound, with multiplicities $e = 1$ throughout.

A similar argument gives (7.15) and (7.16) in parallel. We have

$$\frac{\sigma^+(n_0)}{n_0^2} = \prod_{p' \mid n_0} \frac{p^2\chi(p) - \chi(p)p^{-2e}}{p^2\chi(p) - 1}$$

The factor for primes with $\chi(p) = 1$ satisfies

$$1 \leq \frac{p^2 - p^{-2e}}{p^2 - 1} \leq \frac{p^2 - p^{-2}}{p^2 - 1} = 1 + p^{-2}$$

and for $\chi(p) = -1$,

$$1 \geq \frac{p^2 + (-1)p^{-2e}}{p^2 + 1} \geq 1 - p^{-2}$$

The upper and lower bounds (7.15) and (7.16) follow by extending the products to all $p$, regardless of whether they divide $n_0$. \hfill $\square$

We can now prove Proposition 7.1. We describe the calculations in detail only for some cases, and simply indicate which terms to include for the others.

7.1. Values $n$ coprime to 48: $a = 0$, $b = 0$, $n = n_0$. To begin, let us estimate the Eisenstein contribution to $q^n$ assuming that $n$ has no factors in common with the level $N = 48$. Then we need only consider $f_1$, $f_6$, $f_{11}$, $f_{12}$, $f_{23}$, and $f_{29}$, without any further scalings. Their contribution to $a_n$ is

$$(x_1 + x_{11}\chi_4(n))\sigma_3^+(n) + (x_6 + x_{12}\chi_4(n))\sigma_3^-(n) + x_{23}\sigma_4^+(n) + x_{29}\sigma_4^-(n)$$

and similarly for $b_n$, if one changes the vector of coefficients $x$ to $y = Wy$. Collecting terms gives

$$(x_1 + x_{11}\chi_4(n) + x_6\chi_3(n) + x_{12}\chi_3(n)\chi_4(n))\sigma_3^+(n) + (x_{23} + x_{29}\chi_4(n))\sigma_4^+(n)$$

which is of the form (7.11) with

$$X_3 = x_1 + x_{11}\chi_4(n) + x_6\chi_3(n) + x_{12}\chi_3(n)\chi_4(n)$$

$$X_4 = x_{23} + x_{29}\chi_4(n)$$

If $\chi_4(n) = 1$, then $X_3 = X_4 = 0$ because $x_1 + x_{11} = x_6 + x_{12} = x_{23} + x_{29} = 0$, so we see once again that $a_{n,\text{eis}} = 0$ for $n \equiv 1 \mod 4$. If $\chi_4(n) = -1$, then $X_4 = x_{23} - x_{29} \approx -0.18866 < 0$, so we use the upper bound from (7.14) to see how much could be subtracted. It follows that $a_{n,\text{eis}} > 0.6468n^2$ if $\chi_3(n) = 1$, or $1.0093n^2$ if $\chi_3(n) = -1$.

Similar estimates apply to $b_n$, but without the cancellation in case $\chi_4(n) = 1$. In the four cases according to $\chi_3(n) = \pm 1$ and $\chi_4(n) = \pm 1$,

$$y_1 + y_{11} + y_6 + y_{12} \approx -0.1032 < 0, \quad y_{23} + y_{29} \approx 0.6783$$

$$y_1 + y_{11} - y_6 - y_{12} \approx -0.4734 < 0$$

$$y_1 - y_{11} - y_6 + y_{12} \approx 0.4649 > 0, \quad y_{23} - y_{29} \approx 0.6205$$

$$y_1 - y_{11} + y_6 - y_{12} \approx 0.11168 > 0$$

This time, $Y_4 = y_{23} \pm y_{29} > 0$, so we use the lower bound from (7.14). It follows that $b_{n,\text{eis}}/n^2$ is at least $0.5103$ if $n \equiv 1 \mod 12$, $0.1612$ if $n \equiv 5$, $0.9942$ if $n \equiv 11 \mod 12$, or $0.6611$ if $n \equiv 7 \mod 12$ (listed from top to bottom in the cases above).
7.2. Odd multiples of 3: $\gcd(n, 48) = 3$, $a = 0$, $b \geq 1$, $n_0 = n/3^b$. This case offers all the same terms as for $\gcd(n, 48) = 1$, plus two more corresponding to $f_{25}(z) = f_{23}(3z)$ and $f_{31}(z) = f_{29}(3z)$. The two new terms, with extra factors $\sigma_4^\pm(3^b)$ and $\sigma_3^\pm(3^{b-1})$, lead to

$$X_3 = 9^b x_1 + 9^b \chi_4(n)x_1 + x_6 \chi_3(n_0) + x_{12} \chi_3(n_0) \chi_4(n)$$

$$X_4 = \frac{9^{b+1} + (-1)^b}{10} x_{23} + \frac{(-1)^b}{10} x_{29} \chi_4(n_0) + \frac{9^b + (-1)^{b-1}}{10} x_{31} \chi_4(n_0)$$

The sign $\chi_4(n) = (-1)^b \chi_4(n_0)$ can occur as $\chi_4(n)$ directly from the twists $f_{11}$ and $f_{12}$, or as $\chi_4(n_0)$ when we collect terms using (7.7) to convert $\sigma_4^-$ to $\sigma_4^+$. If $\chi_4(n) = 1$, then $(-1)^{b-1} \chi_4(n_0) = -1$, so

$$X_3 = 9^b (x_1 + x_{11}) + \chi_3(n_0)(x_6 + x_{12})$$

$$X_4 = \frac{1}{10} ((9^{b+1} + (-1)^b)(x_{23} + x_{29}) + (9^b + (-1)^{b-1})(x_{25} - x_{31}))$$

The relations $x_1 + x_{11} = x_6 + x_{12} = x_{23} + x_{29} = x_{25} - x_{31}$ entail $X_3 = X_4 = 0$, giving another explanation of the vanishing of $a_{n, eis}$ when $n \equiv 1 \mod 4$.

If $\chi_4(n) = -1$, then instead

$$X_3 = 9^b (x_1 - x_{11}) + \chi_3(n_0)(x_6 - x_{12})$$

$$X_4 = \frac{1}{10} ((9^{b+1} + (-1)^b)(x_{23} - x_{29}) + (9^b + (-1)^{b-1})(x_{25} + x_{31}))$$

with $x_1 - x_{11} > 0$. The expected size is $n^2 = 9^b n_0^2$, where $b \geq 1$, so we write

$$X_3 \geq 9^b (x_1 - x_{11}) - |x_6 - x_{12}|/9$$

$$X_4 \leq -\frac{1}{10} 9^b (|9x_{23} - 9x_{29} + x_{25} + x_{31}| + |x_{29} + x_{25} + x_{31}|/9)$$

We take the upper bound from (7.14) in case $X_4$ is negative, check that the total is positive, and multiply with the lower bound from (7.15). The result is that $a_{n, eis} > 0.7414 \cdot 9^b n_0^2 = 0.7414 n^2$.

The Eisenstein contribution to $b_n$ is similar, with $y = Wx$ in place of $x$, but $y_1 + y_{11} < 0$ (in fact, $y_1 = 0$, but that plays no role in this case). The positivity comes from $Y_4$ instead of $Y_3$. If $\chi_4(n) = 1$, then

$$Y_3 = 9^b (y_1 + y_{11}) + \chi_3(n_0)(y_6 + y_{12}) \leq -9^b (|y_1 + y_{11}| + |y_6 + y_{12}|/9)$$

$$Y_4 = \frac{1}{10} ((9^{b+1} + (-1)^b)(y_{23} + y_{29}) + (9^b + (-1)^{b-1})(y_{25} - y_{31}))$$

$$\geq \frac{1}{10} 9^b \left( 9y_{23} + 9y_{29} + y_{25} - y_{31} - \frac{|y_{23} + y_{29} - y_{25} + y_{31}|}{9} \right)$$

These estimates for $Y_3$ and $Y_4$ show, first of all, that $Y_4 > 0$, and then, using the lower bound from (7.14), that $Y_3 + \sigma_3^+(n_0)/\sigma_3(n_0) Y_4 > 0.2924$. It follows that $b_{n, eis} > 0.2757n^2$ in this case, using the lower bound from (7.15). If $\chi_4(n) = -1$, then positive contributions from both $Y_3$ and $Y_4$ lead to an even better bound.
### 7.3. Multiples of 2

When $n$ is even, $\chi_4(n) = 0$ so the twists $f_{11}$ and $f_{12}$ disappear. The Eisenstein contribution to $a_n$ is

(7.17)

$$
\begin{align*}
\text{a}_{n,\text{eis}} &= x_1\sigma_3^+(n) + x_6\sigma_3^-(n) + x_{23}\sigma_4^+(n) + x_{29}\sigma_4^-(n) \\
x_2\sigma_3^+(n/2) + x_7\sigma_3^-(n/2) + x_{24}\sigma_4^+(n/2) + x_{30}\sigma_4^-(n/2) \\
x_3\sigma_3^+(n/4) + x_8\sigma_3^-(n/4) + x_{26}\sigma_4^+(n/4) + x_{32}\sigma_4^-(n/4) \\
x_4\sigma_3^+(n/8) + x_9\sigma_3^-(n/8) + x_{25}\sigma_4^+(n/8) + x_{31}\sigma_4^-(n/8) \\
x_5\sigma_3^+(n/16) + x_{10}\sigma_3^-(n/16) + x_{27}\sigma_4^+(n/16) + x_{33}\sigma_4^-(n/16) \\
+ x_{28}\sigma_4^+(n/12) + x_{34}\sigma_4^-(n/12)
\end{align*}
$$

with possible contributions from the scales $n$, $n/2$, $n/4$, $n/8$, $n/16$, $n/3$, $n/6$, and $n/12$ if $n$ is sufficiently divisible by 2 and 3. The same formula can be used equally well if $n$ does not have enough factors of 2 or 3, provided we change the vector $a_n$ to reset the corresponding entries $x_j$ to 0. Likewise, the Eisenstein contribution to $b_n$ has the same form with $x$ replaced by $y = Wx$ where $W$ is the matrix from (3.12), and perhaps with some $y_j$ changed to 0 depending on the factors of $n$.

Write $n = 2^a3^b n_0$ where $\gcd(n_0, 48) = 1$. Each scale $n/s$ has the same form up to changing the values $a, b$. Collecting terms gives (7.11) with

(7.18)

$$
X_3 = 9^b X_3^+ + (-1)^a \chi_3(n_0) X_3^-
$$

where

$$
\begin{align*}
X_3^+ &= x_1\sigma_3^+(2^a) + x_2\sigma_3^+(2^{a-1}) + x_3\sigma_3^+(2^{a-2}) + x_4\sigma_3^+(2^{a-3}) + x_5\sigma_3^+(2^{a-4}) \\
X_3^- &= x_6\sigma_3^+(2^a) - x_7\sigma_3^+(2^{a-1}) + x_8\sigma_3^+(2^{a-2}) - x_9\sigma_3^+(2^{a-3}) + x_{10}\sigma_3^+(2^{a-4})
\end{align*}
$$

and

(7.19)

$$
X_4 = 4^a X_4^+ + (-1)^b \chi_4(n_0) X_4^-
$$

where

$$
\begin{align*}
X_4^+ &= \sigma_4^+(3^b) \left( x_{23} + \frac{x_{24}}{4} + \frac{x_{26}}{4^2} \right) + \sigma_4^+(3^{b-1}) \left( x_{25} + \frac{x_{27}}{4} + \frac{x_{28}}{4^2} \right) \\
X_4^- &= \sigma_4^+(3^b)(x_{29} + x_{30} + x_{32}) - \sigma_4^+(3^{b-1})(x_{31} + x_{33} + x_{34})
\end{align*}
$$
It can be convenient to factor out the expected size $n^2 = 4^a \cdot 9^b \cdot n_0^2$. Substituting the geometric series from (7.8) gives

\begin{equation}
(7.20) \quad a_{n,\text{eis}} = \frac{\sigma_3^+ (n_0) 4^n g^b}{5} \left( 4x_1 + x_2 + \frac{x_3}{4} + \frac{x_4}{16} + \frac{x_5}{64} + \frac{(-1)^a}{4^a} \left( x_1 - x_2 + x_3 - x_4 + x_5 \right) \right.
\end{equation}

\begin{equation}
+ \frac{(-1)^a}{4^a} \left( x_6 - x_7 + \frac{x_8}{4} - \frac{x_9}{16} + \frac{x_{10}}{64} \right) + \frac{(-1)^a}{4^a} \left( x_6 + x_7 + x_8 + x_9 + x_{10} \right) \left) \right)
\end{equation}

\begin{equation}
+ \frac{\sigma_3^+ (n_0) 4^n g^b}{10} \left( 9(x_{23} + x_{24}/4 + x_{26}/16) + x_{25} + \frac{x_{27}}{4} + \frac{x_{28}}{4^2} + \frac{(-1)^b}{9^b} \left( x_{23} + x_{24}/4 + \frac{x_{26}}{16} - x_{25} - \frac{x_{27}}{4} - \frac{x_{28}}{16} \right) \right.
\end{equation}

\begin{equation}
+ \frac{(-1)^b}{9^b} \left( \chi_4(n_0) \right) \left( 9(x_{29} + x_{30} + x_{32}) - x_{31} - x_{33} - x_{34} \right)
\end{equation}

\begin{equation}
\left( \frac{(-1)^b}{9^b} \chi_4(n_0) \left( x_{29} + x_{30} + x_{32} + x_{31} + x_{33} + x_{34} \right) \right)
\end{equation}

assuming $n$ is divisible by 48 so that all terms are present. We describe below which terms to remove, depending on the factors of $n$. The worst cases are when $\gcd(n, 48)$ is 12 or 2, for $a_{n,\text{eis}}$ and $b_{n,\text{eis}}$ respectively.

7.4. **The case** $\gcd(n, 48) = 48$. In this case, $a \geq 4$ and $b \geq 1$, so the terms with factors $4^{-a}$ or $9^{-b}$ are relatively small. Write the subsums appearing in (7.20) as

\begin{align*}
S_1 &= 4x_1 + x_2 + \frac{x_3}{4} + \frac{x_4}{4^2} + \frac{x_5}{4^3} \\
S_2 &= 9(x_{23} + x_{24}/4 + x_{26}/4^2) + x_{25} + \frac{x_{27}}{4} + \frac{x_{28}}{4^2} \\
S_3 &= 4x_6 - x_7 + \frac{x_8}{4} - \frac{x_9}{4^2} + \frac{x_{10}}{4^3} \\
S_4 &= x_{23} + x_{24}/4 + x_{26}/4^2 - x_{25} - \frac{x_{27}}{4} - \frac{x_{28}}{4^2} \\
S_5 &= x_1 - x_2 + x_3 - x_4 + x_5 \\
S_6 &= 9(x_{29} + x_{30} + x_{32}) - x_{31} - x_{33} - x_{34} \\
S_7 &= x_6 + x_7 + x_8 + x_9 + x_{10} \\
S_8 &= x_{29} + x_{30} + x_{32} + x_{31} + x_{33} + x_{34}
\end{align*}

In this notation,

\begin{align*}
X_3 &= \frac{1}{5} \left( 4^a g^b S_1 + 4^a (-1)^a \chi_3(n_0) S_3 + 9^b (-1)^a S_5 + \chi_3(n_0) S_7 \right) \\
X_4 &= \frac{1}{10} \left( 4^a g^b S_2 + 4^a (-1)^b S_4 + 9^b (-1)^b \chi_4(n_0) S_6 + \chi_4(n_0) S_8 \right)
\end{align*}

For $a_{n,\text{eis}}$, there is a positive contribution from $S_2$. Since $a \geq 4$ and $b \geq 1$,

\begin{equation}
X_4 \geq \frac{1}{10} 4^a g^b \left( S_2 - \frac{|S_1|}{9} - \frac{|S_6|}{4^4} - \frac{|S_8|}{9 \times 4^4} \right) > 0.02785 \cdot 4^a g^b
\end{equation}
while
\[ X_4 \leq \frac{1}{5} 4^n 9^b \left( |S_1| + \frac{|S_2|}{9} + \frac{|S_5|}{4^1} + \frac{|S_7|}{9 \times 4^1} \right) < -0.0091177 \cdot 4^n 9^b \]

Since \( X_4 > 0 \), we use the lower bound (7.14) and conclude that
\[ a_{n,\text{eis}} \geq \sigma_3^+ (n_0) \left( X_3 + \frac{\sigma_3^+(n_0)}{\sigma_3^+(n_0)} X_4 \right) > 0.01635 \cdot n_0^2 4^n 9^b = 0.01635 n^2 \]

We have \( b_{n,\text{eis}} > 0.1008 n^2 \) by a similar use of the triangle inequality, with positive contributions from both \( Y_3 \) and \( Y_4 \) leading to a better bound than for \( a_{n,\text{eis}} \).

7.5. \( a \geq 4, \ b = 0 \): the case \( \gcd(n, 48) = 16 \). We discard the terms involving \( x_{25,27,28,31,33,34} \) from (7.20) since \( n \) is not divisible by 3. The sums \( S_3 \) and \( S_4 \), which previously were negligible because of the factor \( 9^{-b} \), should now be regarded as part of the main term. We consider the cases \( a = 4 \) and \( a \geq 5 \) separately, because a naive application of the triangle inequality using \( 4^{-a} \leq 4^{-4} \) would lead to a negative lower bound. In this way, we find that \( a_{n,\text{eis}} \geq 0.0003723 n^2 \) if \( a = 4 \), or \( 0.0002649 n^2 \) if \( a \geq 5 \). Table 7.1 records only the minimum of these two for simplicity, so that the same bound holds for all \( a \geq 4 \).

7.6. \( a = 3, \ b \geq 1 \): the case \( \gcd(n, 48) = 24 \). We discard the terms involving \( x_{5} \) and \( x_{10} \), and we can substitute \((-1)^a = -1 \). Terms that were neglected because of a factor \( 4^{-a} \) can now be treated exactly.

7.7. \( a = 3, \ b = 0 \): the case \( \gcd(n, 48) = 8 \). We discard the terms corresponding to multiples of 16, namely \( x_{5,10} \), as in the case \( a = 3, \ b \geq 1 \), as well as \( x_{25,27,28,31,33,34} \) which correspond to multiples of 3.

7.8. \( a = 2, \ b \geq 1 \): the case \( \gcd(n, 48) = 12 \). We discard \( x_{4,5,9,10} \) and set \((-1)^a = 1 \). Let us detail only the case \( b = 1 \), which is the worst one for \( a_{n,\text{eis}} \). For \( b \geq 2 \), the triangle inequality gives a better bound \( \varepsilon = 1.0003 \times 10^{-3} \) using \( 9^{-b} \leq 9^{-2} \) to neglect terms. For \( b_n \), we obtain \( \varepsilon^* = 0.01363 \) in case \( b = 1 \), and a better bound \( \varepsilon^* = 0.02608 \) in case \( b \geq 2 \). It is important to distinguish those cases since a reckless use of the triangle inequality would give a negative bound for \( b = 1 \).

Taking then the case \( a = 2, \ b = 1 \) for \( a_{n,\text{eis}} \), we have
\[ X_3 = S_1 + \chi_3(n_0) S_3, \quad X_4 = S_2 + \chi_4(n_0) S_4 \]
where
\[ S_1 = 117 x_1 + 27 x_2 + 9 x_3 \]
\[ S_3 = 13 x_6 - 3 x_7 + x_8 \]
\[ S_2 = 8(16 x_23 + 4 x_24 + x_20) + 16 x_25 + 4 x_27 + x_28 \]
\[ S_4 = -8(x_29 + x_30 + x_32) + x_31 + x_33 + x_34 \]
Substituting the values for \( x_j \), we find that \( S_1 \) and \( S_4 \) are negative, while \( S_2 \) and \( S_3 \) are positive. The least favourable signs for proving a lower bound are therefore \( \chi_3(n_0) = -1 \) and \( \chi_4(n_0) = 1 \), in which case we need \( X_4 = S_2 + S_4 \) to beat \( X_3 = S_1 - S_3 \). Numerically, \( S_1 - S_3 \approx -0.7110, S_2 + S_4 \approx 0.7498 \), and their ratio
\[
\frac{S_3 - S_1}{S_2 + S_4} \approx 0.9482
\]
(7.21)
is only slightly less than 0.9499 from (7.14). We would not be able to prove $a_{n,eis} > 0$

in this case using only the worse ratio 0.931869 obtained from separate upper and lower bounds on $\sigma_4^+(n_0) \text{ and } \sigma_3^+(n_0)$. Proposition 7.2 gives

$$a_{n,eis} = \sigma_3^+(n_0) \left( X_3 + \frac{\sigma_4^+(n_0)}{\sigma_3^+(n_0)} X_4 \right) > 8.7536 \times 10^{-6} \ n^2$$

remembering to divide by 144 from $n^2 = 144n_0^2$.

Some of the coefficients really are this small: it is not only a failure of our bounds.

For example, with $n = 12n_0$ where $n_0 = 7 \cdot 59, 7 \cdot 71, 7 \cdot 83, \text{ and } 7 \cdot 107$, we have

$$\frac{a_{n,eis}}{n^2} \times 10^5 \approx 6.2643, \quad 6.2649, \quad 6.2652, \quad 6.2656$$

Primes congruent to 7 mod 12 act to minimize (7.14), while primes congruent to 11
are neutral and can be included to arrange $n_0 \equiv 5$ so that the signs align with the worst case for $X_3$ and $X_4$. Taking a smoother $n$ of this form decreases the coefficient further:

$$n = 544236 = 12 \cdot 7 \cdot 11 \cdot 19 \cdot 31 \implies \frac{a_{n,eis}}{n^2} \approx 2.382567 \times 10^{-5}$$

7.9. $a = 2, b = 0$: the case $\gcd(n, 48) = 4$. We have the same terms as for $n = 12n_0$, but without $\chi_{23,27,28,31,33,34}$ which correspond to multiples of 3.

7.10. $a = 1, b \geq 1$: the case $\gcd(n, 48) = 6$. We discard $x_{3,4,5,8,9,10,26,28,32,34}$ because $n$ is not divisible by 4.

7.11. $a = 1, b = 0$: the case $\gcd(n, 48) = 2$. This is the worst case for $b_{n,eis}$.

$$X_3 = S_1 + \chi_3(n_0) S_3, \quad X_4 = S_2 + \chi_4(n_0) S_4$$

$$S_1 = 3x_1 + x_2 \quad S_2 = 4x_{23} + x_{24}$$

$$S_3 = -3x_6 + x_7 \quad S_4 = x_{29} + x_{30}$$

For the values $y_j$, the signs are $S_1 < 0, S_3 > 0, S_2 > 0, S_4 < 0$, so the combinations $S_1 - S_3 \text{ and } S_2 + S_4$ lead to the smallest value when $n_0 \equiv 5 \pmod{12}$:

$$b_{n,eis} > 0.005432 \cdot n_0^2 = 0.001358 \cdot n^2$$

since $n^2 = 4n_0^2$ introduces a final factor of 4. For comparison, some relatively small values $b_{n,eis} / n^2 \approx 0.001791777 \text{ and } 0.0016229 \text{ occur if } n = 2 \cdot 7 \cdot 19 \cdot 53 \text{ or } n = 2 \cdot 7 \cdot 11 \cdot 19 \cdot 31$, which are products of primes congruent to 7 mod 12 with one extra factor 11 or 53 fixing the congruence class $n_0 \equiv 5 \mod{12}$.

8. A relation between $D_6, D_6^*$, and $\mathbb{Z}^6$

The choice of optimization space, using the character $\chi_4$, was motivated in part
by the fact that it contains the theta series of some familiar lattices: $D_6, D_6^*$, and $\mathbb{Z}^6$. Here, $D_6$ is the lattice of integer vectors with even coordinate sum, and $D_6^*$ is its dual (see [11, p. 117]). In the course of proving Theorem 1.2, we noticed a special feature of dimension 6, which is the following linear relation between these theta series:

$$(8.1) \quad \vartheta_{D_6}(z) + 4\vartheta_{D_6^*}(4z) = 5\vartheta_{\mathbb{Z}^6}(2z)$$

A linear relation of this form between $D_n, D_n^*$, and $\mathbb{Z}^n$ seems specific to $n = 6$ (with $n = 1, 2$ as somehow degenerate cases where $D_n$ is congruent to $\mathbb{Z}^n$). Matching the first two coefficients would suggest an identity between, on the one hand,
$\vartheta_{D_6}(z) + (n-2)\vartheta_{D_4}(sz)$ for some scaling $s$, and $(n-1)\vartheta_{Z^n}(2z)$ on the other. For most values of $n$ and $s$, the higher coefficients do not match.

One strategy, knowing that these theta series are modular with the same properties, is to simply check enough coefficients in the $q$-series of both sides. The identity (8.1) is also easy to see from the expressions of the three theta functions in terms of

\begin{align*}
\vartheta_2(z) &= \sum_{m=-\infty}^{\infty} e^{\pi i (m+1/2)^2 z}, \\
\vartheta_3(z) &= \sum_{m=-\infty}^{\infty} e^{\pi i m^2 z}, \\
\vartheta_4(z) &= \sum_{m=-\infty}^{\infty} (-1)^m e^{\pi i m^2 z}
\end{align*}

In terms of these,

\begin{align*}
\vartheta_{Z^n} &= \vartheta_4^n, \\
\vartheta_{D_6} &= \vartheta_3^n + \vartheta_4^n, \\
\vartheta_{D_4^n} &= \vartheta_2^n + \vartheta_3^n
\end{align*}

We will use the following identities, taken from [11, p.104]

\begin{align*}
\theta_2^2 + \theta_4^2 &= \theta_3^2 \\
\theta_3(z) + \theta_4(z) &= 2\theta_3(4z) \\
\theta_3(z) - \theta_4(z) &= 2\theta_2(4z) \\
\theta_3(z)^2 - \theta_4(z)^2 &= 2\theta_2(2z)^2 \\
\theta_3(z)^2 + \theta_4(z)^2 &= 2\theta_4(2z)^2 \\
\theta_3(z)\theta_4(z) &= \theta_4(2z)^2
\end{align*}

The first of these is Jacobi’s identity on sums of four squares. The others are related to the special cases $n = 1, 2$ where $D_n$ is congruent to $Z^n$.

We now take $n = 6$. To express $\vartheta_{Z^n}(2z)$, take $z \rightarrow 2z$ in (8.3) and multiply by $\theta_3(2z)^2$:

\begin{equation}
\vartheta_{Z^n}(2z) = \vartheta_3(2z)^6 = \vartheta_2(2z)^4\vartheta_3(2z)^2 + \vartheta_4(2z)^4\vartheta_3(2z)^2
\end{equation}

To see $D_6$ and $D_4^n$ on the right, we substitute (8.6), (8.7), (8.8) for each term $\theta_j(2z)$. This counteracts the scaling $z \rightarrow 2z$ from the beginning:

\begin{equation}
\vartheta_{Z^n}(2z) = \left(\frac{\theta_3(z)^2 - \theta_4(z)^2}{2}\right)^2 \left(\frac{\theta_3(z)^2 + \theta_4(z)^2}{2}\right)^2 \theta_3(z)^2\theta_4(z)^2 \left(\frac{\theta_3(z)^2 + \theta_4(z)^2}{2}\right)
\end{equation}

Multiplying these terms leads to

\begin{equation}
\vartheta_{Z^n}(2z) = \frac{\theta_3(z)^6 + \theta_4(z)^6 + 3\theta_3(z)^4\theta_4(z)^2 + 3\theta_3(z)^2\theta_4(z)^4}{8}
\end{equation}

The 5 in our target identity comes from a binomial expansion, using (8.5) and (8.4):

\begin{equation}
\vartheta_{D_6}(4z) = \vartheta_2(4z)^6 + \vartheta_4(4z)^6 = \left(\frac{\theta_3(z) - \theta_4(z)}{2}\right)^6 + \left(\frac{\theta_3(z) + \theta_4(z)}{2}\right)^6 = \theta_3(z)^6 + \theta_4(z)^6 + 15\theta_3(z)^4\theta_4(z)^2 + 15\theta_3(z)^2\theta_4(z)^4
\end{equation}

We solve this for the cross-terms:

\begin{equation}
3\theta_3(z)^4\theta_4(z)^2 + 3\theta_3(z)^2\theta_4(z)^4 = \frac{2^5\vartheta_{D_6}(4z) - \theta_3(z)^6 - \theta_4(z)^6}{5}
\end{equation}
Substituting this in (8.9) gives
\[
\vartheta_{2^6}(2z) = \frac{1}{8} \left( \theta_3(z)^6 + \theta_4(z)^6 + \frac{2^5 \vartheta_{D^*_6}(4z) - \theta_3(z)^6 - \theta_4(z)^6}{5} \right)
\]
\[
= \frac{1}{5} \left( \frac{\theta_3(z)^6 + \theta_4(z)^6}{2} + \frac{2^5 \vartheta_{D^*_6}(4z)}{8} \right)
\]
\[
= \frac{1}{5} (\vartheta_{D^*_6}(z) + 4\vartheta_{D^*_6}(4z))
\]
as required.

ACKNOWLEDGEMENTS

It is a pleasure to thank both Henry Cohn and Henri Cohen for their helpful advice. Thanks also to Rupert Li, Henry Cohn, David de Laat, and Andrew Salmon for sharing results from earlier versions of their articles [25] and [8]. Thanks again to David de Laat for advice on computing the Cohn–Elkies function with [24]. Thanks to Claire Burrin for suggesting reference [3]. Thanks to the referees for their comments and efforts reading the paper.

FUNDING

Viazovska’s research was supported by Swiss National Science Foundation project 184927. Dostert was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

REFERENCES

[1] N. Afkhami-Jeddi, H. Cohn, T. Hartman, D. de Laat, A. Tajdini. High-dimensional sphere packing and the modular bootstrap, Journal of High Energy Physics 12 (2020) 66, arXiv:2006.02560
[2] C. Batut, K. Belabas, D. Benardi, H. Cohen, and M. Olivier, “User’s Guide to PARI-GP”, 1998, see also https://pari.math.u-bordeaux.fr/
[3] B. C. Berndt and M. I. Knopp, Hecke’s Theory of Modular Forms and Dirichlet Series, (2008) World Scientific Publishing
[4] W. Bosma, John Cannon, and Catherine Playoust, The Magma algebra system. I. The user language, J. Symbolic Comput., 24 (1997), 235—265.
[5] H. Cohn, New upper bounds on sphere packings II, Geom. Topol. 6 (2002), 329–353. MR1914571 doi:10.2140/gt.2002.6.329 arXiv:math/0110010
[6] H. Cohn and N. Elkies, New upper bounds on sphere packings I, Ann. of Math. (2) 157 (2003), no. 2, 689–714.
[7] H. Cohn and A. Kumar, Universally optimal distribution of points on spheres, J. Amer. Math. Soc. 20 (2007), no. 1, 99–148. MR2257398 doi:10.1090/S0894-0347-06-00546-7 arXiv:math/0607446
[8] H. Cohn, D. de Laat, and A. Salmon, Three-point bounds for sphere packing, 2022, arXiv:2206.15373 [math.MG]
[9] H. Cohn and N. Triantafillou, Dual linear programming bounds for sphere packing via modular forms, Math. Comp. 91 (2022), 491-508 arXiv:1909.04772
[10] H. Cohn and Y. Zhao, Sphere packing bounds via spherical codes, Duke Math. J. 163 (2014), no. 10, 1965–2002. MR3229046 doi:10.1215/00127094-2738857 arXiv:1212.5966
[11] J. H. Conway and N. J. A. Sloane, Sphere Packings, Lattices and Groups, Volume 290, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York (1998)
[12] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, The sphere packing problem in dimension 24, Ann. of Math. (2) 185 (2017), 1017–1033.
[13] P. Deligne, La conjecture de Weil, I, Inst. Hautes Études Sci. Publ. Math. 43 (1974), 273-307
36 MATTHEW DE COURCY-IRELAND, MARIA DOSTERT, MARYNA VIAZOVSKA

[14] P. Deligne, Formes modulaires et représentations ℓ-adiques, Séminaire Bourbaki: vol. 1968/69, exposés 347-363, Séminaire Bourbaki, no. 11 (1971), Exposé no. 355, 34 p.

[15] P. Deligne and J.-P. Serre, Formes modulaires de poids 1, Annales scientifiques de l’École Normale Supérieure, Série 4, Tome 7 (1974) no. 4, pp. 507-530.

[16] F. Diamond and J. Shurman, A First Course in Modular Forms, Graduate Texts in Mathematics 228, Springer, New York, (2005)

[17] G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, (1960) fourth edition

[18] Y. Ihara, Hecke polynomials as congruence zeta functions in elliptic modular case, Ann. of Math., S.2 85 1967, p. 267-295

[19] H. Iwaniec, Topics in Classical Automorphic Forms, American Mathematical Society, Graduate Studies in Mathematics, Volume 17. Providence, Rhode Island (1997)

[20] H. Iwaniec and E. Kowalski, Analytic Number Theory, American Mathematical Society, Colloquium Publications, Volume 53. Providence, Rhode Island (2004)

[21] G. A. Kabatjanski and V. I. Levenshtein, Bounds for packings on a sphere and in space (Russian), Problemy Peredači Informacii 14 (1978), no. 1, 3–25; English translation in Problems of Information Transmission 14 (1978), no. 1, 1–17. MR0514023

[22] N. Koblitz, Introduction to elliptic curves and modular forms, Graduate texts in mathematics 97, Springer, New York (1993) second edition

[23] M. Kuga and G. Shimura, On the zeta function of a fibre variety whose fibres are abelian varieties, Ann. of Math., S.2 82 1965, p.478–539

[24] D. de Laat and N. Leijenhorst, Solving clustered low-rank semidefinite programs arising from polynomial optimization, Preprint (2022), arXiv:2202.12077

[25] R. Li, Dual bounds for sphere packing linear programs via discrete reductions. 2022, arXiv:2206.09876 [math.MG]

[26] A. O. Makhorin, GNU Linear Programming Kit, https://www.gnu.org/software/glpk/, version 5.0 released December 16, 2020

[27] J. Nicolas and G. Robin, Majorations Explicites Pour le Nombre de Diviseurs de N, Canad. Math. Bull. Vol. 26 (4), 1983, doi:10.4153/CMB-1983-078-5

[28] W. A. Stein, Modular Forms: A Computational Approach, American Mathematical Society, Graduate Studies in Mathematics, Volume 79. Providence, Rhode Island (2007) with an appendix by P. E. Gunnells

[29] W. A. Stein et al, Sage Mathematics Software, The Sage Development Team, 2022, http://www.sagemath.org.

[30] S. Torquato and F. H. Stillinger, New conjectural lower bounds on the optimal density of sphere packings, Experiment. Math. 15 (2006), no. 3, 307–331. MR2206469 http://projecteuclid.org/euclid.em/1175789761 arXiv:math/0508381

[31] M. S. Viazovska, The sphere packing problem in dimension 8, Ann. of Math. (2) 185 (2017), 991–1015.

Institute of Mathematics, EPFL, CH-1015 Lausanne, Switzerland and Department of Mathematics, Stockholm University
Email address: matthew.decourcy-ireland@math.su.se

Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden
Email address: maria.dostert@gmail.com

Institute of Mathematics, EPFL, CH-1015 Lausanne, Switzerland
Email address: viazovska@gmail.com