DUAL LINEAR PROGRAMMING BOUNDS
FOR SPHERE PACKING VIA MODULAR FORMS

HENRY COHN AND NICHOLAS TRIANTAFILLOU

ABSTRACT. We obtain new restrictions on the linear programming bound for sphere packing, by optimizing over spaces of modular forms to produce feasible points in the dual linear program. In contrast to the situation in dimensions 8 and 24, where the linear programming bound is sharp, we show that it comes nowhere near the best packing densities known in dimensions 12, 16, 20, 28, and 32. More generally, we provide a systematic technique for proving separations of this sort.

1. INTRODUCTION

The sphere packing problem asks for the densest packing of congruent spheres in $\mathbb{R}^d$. In other words, what is the greatest proportion of $\mathbb{R}^d$ that can be covered by congruent balls with disjoint interiors? The case $d = 1$ is trivial, $d = 2$ was solved by Thue [29], and $d = 3$ was solved by Hales [17] with a computer-assisted proof that has since been formally verified [18]. These proofs make essential use of the geometry of packings in $\mathbb{R}^d$ in a way that seems difficult to extend to higher dimensions, and so another approach is needed when d is large. Based on a long history of linear programming bounds in coding theory, Cohn and Elkies [6] developed a linear programming bound for sphere packing. It yields the best upper bounds known for the packing density in high dimensions [12], and Cohn and Elkies conjectured that the linear programming bound is sharp when $d = 8$ or $d = 24$.

In a recent breakthrough, Viazovska [31] proved this conjecture for $d = 8$, and thus showed that the E_8 root lattice yields the densest sphere packing in $\mathbb{R}^8$. Shortly thereafter, Cohn, Kumar, Miller, Radchenko, and Viazovska [10] proved the conjecture for $d = 24$. These are the only two cases beyond $d = 3$ in which the sphere packing problem has been solved.

These advances raise numerous questions. Is it possible that the linear programming bound is sharp in some other dimensions? Could it even be sharp in every dimension? (Surely not, but why not?) What happens in $\mathbb{R}^{16}$, and why does that case seemingly not behave like $\mathbb{R}^8$ and $\mathbb{R}^{24}$? These questions remain mysterious, but in this paper we take some initial steps towards answering them.

The difficulty in analyzing the linear programming bound stems from the use of an auxiliary function, which must satisfy certain inequalities. The quality of the bound depends on the choice of this function, and optimizing the bound amounts to optimizing a functional over the infinite-dimensional space of auxiliary functions. This optimization problem has not been solved exactly except when $d \in \{1, 8, 24\}$.
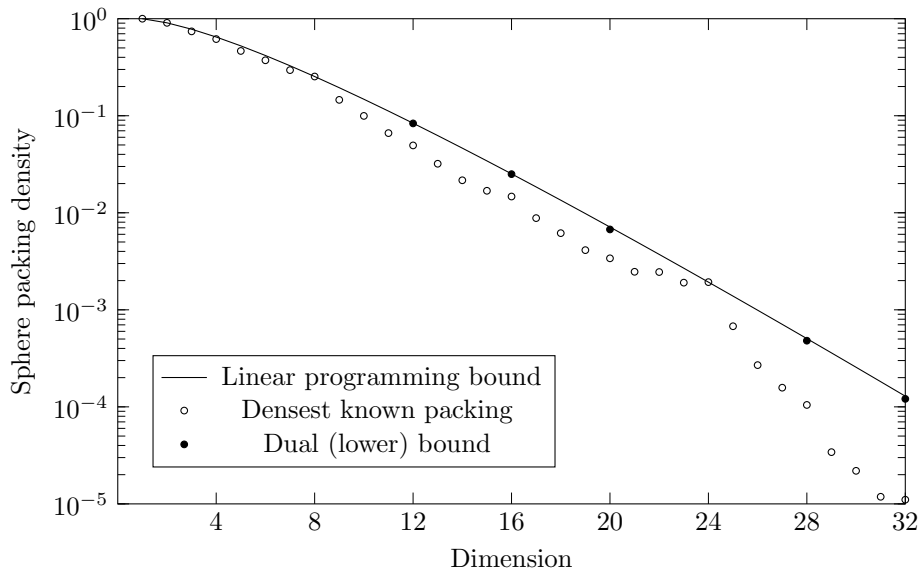
FIGURE 1.1. The upper curve is the linear programming bound computed using the best auxiliary functions currently known, while the white circles are the densest sphere packings currently known (see [13, pp. xix–xx]). Our new obstructions, drawn as black circles, are lower bounds for the linear programming bound. They show that further optimizing the choice of auxiliary function cannot improve the linear programming bound by much.

In other dimensions, we can approximate the true optimum by using a computer to optimize over a finite-dimensional subspace. The resulting auxiliary function always proves some bound for the sphere packing density, and we expect it to be close to the optimal linear programming bound if the subspace is large and generic enough. However, nobody has been able to determine how close it must be. What if these numerical computations are woefully far from the true optimum? If that were the case, then they would shed very little light on the linear programming bound. It is even possible, albeit implausible, that the linear programming bound might be sharp for relatively small values of d that nobody has noticed yet.

As shown in Figure 1.1, the linear programming bound seems to vary smoothly as a function of dimension, and the sharp bounds in 8 and 24 dimensions fit perfectly with the curve as a whole. These observations raise our confidence that the numerical optimization is not in fact misleading. However, there remains a fundamental gap in the theory of the linear programming bound: how can one prove a corresponding lower bound, beyond which no auxiliary function can pass? In optimization terms, such a bound amounts to a *dual linear programming bound*, which controls how good the optimal linear programming bound could be.

In this paper, we show how to compute such a bound when the dimension is a multiple of four, by optimizing over spaces of modular forms. (We expect that other dimensions work similarly, but we have not carried out the modular form calculations in those cases.) Our results for dimensions 12, 16, 20, 28, and 32 are

shown in Figure 1.1 and Table 6.1. The most noteworthy cases are dimensions 12 and 16, where the Coxeter-Todd and Barnes-Wall lattices are widely conjectured to be optimal sphere packings:

Theorem 1.1. *The linear programming bound for the sphere packing density in $\mathbb{R}^{16}$ is greater than 1.7 times the density of the Barnes-Wall lattice, and the bound in $\mathbb{R}^{12}$ is greater than 1.686 times the density of the Coxeter-Todd lattice. In particular, the linear programming bound cannot prove that either lattice is an optimal sphere packing.*

Unsurprisingly, in neither case is the linear programming bound even close to reaching the best density known. The ratios 1.7 and 1.686 are almost certainly not quite optimal, and we expect that they could be improved to 1.712 and 1.694, respectively, which would match the known upper bounds to three decimal places. See Section 7 for further discussion.

Note that even when the linear programming bound is far from sharp, determining its value is of interest in its own right. For example, it can be interpreted as describing an uncertainty principle for the signs of a function and its Fourier transform (see [7]). Thus, it has significance beyond just the topic of sphere packing.

1.1. **The linear programming bound.**

Before proceeding further, let us review how the linear programming bound works. Recall that a *sphere packing* in $\mathbb{R}^d$ is a disjoint union $\bigcup_{x \in C} B(x, \rho)$ of open unit balls of some fixed radius ρ and centered at the points of some subset C of $\mathbb{R}^d$.

Given a sphere packing $\mathcal{P}$, the *upper density* $\Delta_{\mathcal{P}}$ of $\mathcal{P}$ is defined by

$$\Delta_{\mathcal{P}} = \limsup_{r \to \infty} \frac{\mathrm{vol}(B(x, r) \cap \mathcal{P})}{\mathrm{vol}(B(x, r))}$$

for any $x \in \mathbb{R}^d$ (the upper density does not depend on the choice of x). If the limit exists, and not just the limit superior, then we say that $\mathcal{P}$ has *density* $\Delta_{\mathcal{P}}$. The *sphere packing density in $\mathbb{R}^d$* is

$$\Delta_d = \sup_{\mathcal{P} \subset \mathbb{R}^d} \Delta_{\mathcal{P}},$$

where the supremum is over sphere packings $\mathcal{P}$. We will often renormalize and work with the *upper center density*

$$\delta_{\mathcal{P}} = \frac{\Delta_{\mathcal{P}}}{\mathrm{vol}(B(0, 1))} = \limsup_{r \to \infty} \frac{\#(B(x, r) \cap C)}{\mathrm{vol}(B(x, r))} \cdot \frac{\mathrm{vol}(B(0, \rho))}{\mathrm{vol}(B(0, 1))},$$

which measures the number of center points per unit volume in space if we use spheres of radius $\rho = 1$. Of course the center density has no theoretical advantage over the density, but it is often convenient not have to carry around the factor of $\mathrm{vol}(B(0, 1)) = \pi^{d/2}/(d/2)!$. For example, $\delta_{24} = 1$, while $\Delta_{24} = \pi^{12}/12! = 0.00192957\ldots$.

We normalize the *Fourier transform* $\widehat{f}$ of an integrable function $f \colon \mathbb{R}^d \to \mathbb{R}$ by

$$\widehat{f}(y) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \langle x, y \rangle} \, dx,$$
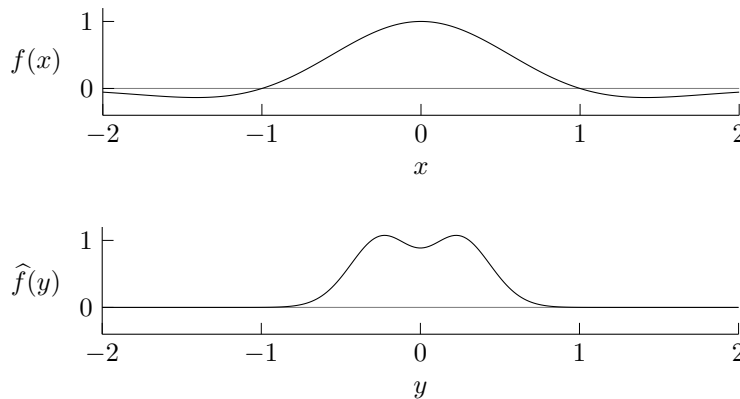
FIGURE 1.2. A sample auxiliary function and its Fourier transform (namely, $f(x) = (1 - x^2)e^{-x^2}$ on $\mathbb{R}^1$, with $r = 1$).

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product on $\mathbb{R}^d$. Cohn and Elkies [6] showed[1] how to use harmonic analysis to bound the sphere packing density as follows:

Theorem 1.2 (Cohn and Elkies [6]). *Let $f \colon \mathbb{R}^d \to \mathbb{R}$ be a continuous, integrable function, such that $\widehat{f}$ is integrable as well and $\widehat{f}$ is real-valued (i.e., f is even). Suppose f and $\widehat{f}$ satisfy the following inequalities for some positive real number r:*

(1) *$f(0) > 0$ and $\widehat{f}(0) > 0$,*
(2) *$f(x) \leq 0$ for $|x| \geq r$, and*
(3) *$\widehat{f}(y) \geq 0$ for all y.*

Then every sphere packing in $\mathbb{R}^d$ has upper center density at most

$$\frac{f(0)}{\widehat{f}(0)} \cdot \left(\frac{r}{2}\right)^d.$$

The *linear programming bound* in $\mathbb{R}^d$ is the infimum of the center density upper bound

$$\frac{f(0)}{\widehat{f}(0)} \cdot \left(\frac{r}{2}\right)^d$$

over all *auxiliary functions* f satisfying the hypotheses of Theorem 1.2. See Figure 1.2 for an example of an auxiliary function, which is far from optimal.

Without loss of generality, we can assume that the auxiliary function f is radial, because we can simply average its rotations about the origin. For a radial function f, we write $f(t)$ with $t \in [0, \infty)$ to denote the common value $f(x)$ with $|x| = t$. If f is radial, then $\widehat{f}$ is radial as well, and

$$\widehat{f}(y) = \frac{2\pi}{|y|^{d/2-1}} \int_0^\infty f(t) J_{d/2-1}(t|y|) t^{d/2} \, dt,$$

where $J_{d/2-1}$ is the Bessel function of the first kind of order $d/2-1$ (see, for example, Theorem 9.10.3 in [1]).

---

[1]Strictly speaking, the paper [6] imposed stronger hypotheses on f, but one can easily remove those hypotheses by mollifying f, using the approach from the first paragraph of Section 4 in [5]. The fact that they could be removed was first observed in [8].

The density bound

$$\frac{f(0)}{\widehat{f}(0)} \cdot \left(\frac{r}{2}\right)^d$$

is invariant under replacing f with $x \mapsto f(\rho x)$ and r with r/ρ for any scaling factor $\rho \in (0, \infty)$. Without loss of generality we can use this invariance to fix $r = 1$, and we can assume $\widehat{f}(0) = 1$ as well. Then the constraints on f from Theorem 1.2 are linear inequalities, and the density bound is also a linear functional of f. Thus, optimizing the choice of f amounts to solving an infinite-dimensional linear optimization problem, which explains the name "linear programming bound." In practice, however, fixing r may not lead to the prettiest answers. For example, Cohn and Elkies found more elegant behavior if one instead fixes $f(0) = \widehat{f}(0)$ and lets r vary (see Section 7 of [6]).

The best choice of f is not known, except when $d \in \{1, 8, 24\}$, and little is known about how good the optimal bound might be. It is not hard to produce upper bounds by numerically optimizing over finite-dimensional spaces of functions, and in most cases these upper bounds seem to be close to the optimal linear programming bound (see [7] for the most extensive calculations so far). However, these computational methods leave open the possibility that other auxiliary functions might prove much better bounds.

What sort of obstructions prevent the linear programming bound from reaching the density of the best sphere packing? In this paper we provide a partial answer, with an algorithm to compute such obstructions via linear programming over spaces of modular forms of weight $d/2$. The algorithm is based on optimizing a summation formula for radial Schwartz functions, which is an analogue of Voronoi summation.

The remainder of the paper is organized as follows. In Section 2, we present a general framework for computing dual linear programming bounds. We describe our algorithm in Section 3, and we prove the summation formula underlying the algorithm in Section 4. In Section 5, we expand on the final step of our algorithm by describing a method for checking in finite time that all of the coefficients of the q-expansion of a given modular form are nonnegative. Finally, we present a table of new lower bounds in Section 6, and we conclude with open problems in Section 7.

2. Duality

Computing a bound for the objective function in a linear program is typically straightforward: it just amounts to finding a feasible point in the dual linear program. The difficulty in our case is that the optimization problems are infinite-dimensional. The primal problem is relatively tractable, because the auxiliary functions in Theorem 1.2 are well behaved in practice. We can approximate them with polynomials times Gaussians, and using high-degree polynomials yields excellent results. For example, in $\mathbb{R}^{16}$ the resulting center density bounds seem to converge to

$$0.10705844234092448845891681517141\ldots$$

as the polynomial degree tends to infinity, and we believe this number is the optimal linear programming bound for 16 dimensions, correct to 32 decimal places. Unfortunately, the dual problem is much less tractable. It amounts to optimizing over a space of measures, and we believe the optimal measures will be singular (specifically, supported on a discrete set of radii). In particular, we know of no

simple family of measures we can use to approximate them fruitfully. Instead, the dual problem appears to be quite a bit more subtle.

In Section 4 of [5], Cohn formulated the dual linear program as follows. Here, δ_0 denotes a delta function at the origin, and $\widehat{\mu}$ is the Fourier transform of μ as a tempered distribution.

Proposition 2.1. *Let μ be a tempered distribution on $\mathbb{R}^d$ such that $\mu = \delta_0 + \nu$ with $\nu \geq 0$, $\mathrm{supp}(\nu) \subseteq \{x \in \mathbb{R}^d : |x| \geq r\}$ for some $r > 0$, and $\widehat{\mu} \geq c\delta_0$ for some $c > 0$. Then the linear programming bound in $\mathbb{R}^d$ is at least*

$$c \cdot \left(\frac{r}{2}\right)^d.$$

Sketch of proof. Let $f \colon \mathbb{R}^d \to \mathbb{R}$ be an auxiliary function satisfying the hypotheses of Theorem 1.2, where we use scaling invariance to ensure that the same value of r works for both f and μ. If f and $\widehat{f}$ are rapidly decreasing, then the inequalities on f and μ imply that

$$f(0) \geq \int_{\mathbb{R}^d} f\mu = \int_{\mathbb{R}^d} \widehat{f}\,\widehat{\mu} \geq c\widehat{f}(0),$$

and thus

$$\frac{f(0)}{\widehat{f}(0)} \geq c,$$

as desired. More general auxiliary functions must be mollified, as described in Section 4 of [5], after which the same argument applies to them as well. $\square$

The difficulty in applying this proposition is how to find a plentiful source of distributions μ that could satisfy the hypotheses. One source is Poisson summation for lattices, which says that for any lattice Λ in $\mathbb{R}^d$, the Fourier transform of the distribution

$$\sum_{x \in \Lambda} \delta_x$$

is

$$\frac{1}{\mathrm{vol}(\mathbb{R}^d/\Lambda)} \sum_{y \in \Lambda^*} \delta_y,$$

where Λ^* is the dual lattice. Thus, the hypotheses of Proposition 2.1 are satisfied with $c = 1/\mathrm{vol}(\mathbb{R}^d/\Lambda)$ and $r = \min_{x \in \Lambda \setminus \{0\}} |x|$. The resulting lower bound amounts to proving Theorem 1.2 for lattice packings.

In principle, one could try to improve on individual lattices by using a linear combination of Poisson summation formulas for different lattices (see, for example, the bottom of page 351 in [5]). However, that does not seem fruitful in general. Instead, we use the following analogue of Voronoi summation to produce distributions from modular forms. For definitions related to modular forms, see [16]. In particular, recall that the slash operator is defined as follows: if $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ and $\det M > 0$, then

$$(f|_k M)(z) := (ad - bc)^{k/2}(cz + d)^{-k} f\left(\frac{az + b}{cz + d}\right).$$

Proposition 2.2. *Let $d = 2k$ with $k \in \mathbb{N}$, let $g \in M_k(\Gamma_1(N))$ be a modular form of weight k for the congruence subgroup $\Gamma_1(N)$, let $w_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$, and let*

$$\widetilde{g}(z) = i^k (g|_k w_N)(z) = \frac{i^k}{N^{k/2} z^k} g\left(-\frac{1}{Nz}\right)$$

be i^k times the image of g under the full level N Atkin-Lehner operator (so that $g = i^k \widetilde{g}|_k w_N$ as well). Let the q-expansions of g and $\widetilde{g}$ be

$$g(z) = \sum_{n=0}^{\infty} a_n q^n \qquad and \qquad \widetilde{g}(z) = \sum_{n=0}^{\infty} b_n q^n,$$

where $q = e^{2\pi i z}$. Then for every radial Schwartz function $f\colon \mathbb{R}^d \to \mathbb{C}$,

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

In particular, if δ_r denotes a delta function supported on the sphere of radius r about the origin in $\mathbb{R}^d$, then this proposition says that the tempered distributions

$$\sum_{n=0}^{\infty} a_n \delta_{\sqrt{n}} \qquad \text{and} \qquad \left(\frac{2}{\sqrt{N}}\right)^{d/2} \sum_{n=0}^{\infty} b_n \delta_{2\sqrt{n/N}}$$

are Fourier transforms of each other. Our algorithm will optimize over distributions of this form. The advantage of these distributions is that their supports help enforce the constraint that $\operatorname{supp}(\nu) \subseteq \{x \in \mathbb{R}^d : |x| \geq r\}$ in Proposition 2.1.

For comparison, the techniques in Section 5 of [7] produce what appear to be close numerical approximations to the optimal distributions μ. They have the form

$$\mu = \sum_{n \geq 0} c_n \delta_{r_n}$$

with radii given by $0 = r_0 < r_1 < r_2 < \cdots$ and tending to infinity, coefficients $c_n > 0$, and $\widetilde{\mu} = \mu$. For example, in $\mathbb{R}^{16}$ the first few radii and coefficients are listed in Table 2.1. The only drawback is that the results of these calculations are merely conjectural: we do not know whether such a distribution actually exists.

Our approach in this paper amounts to approximating the optimal μ with a distribution μ' whose existence follows from Proposition 2.2. For comparison, Table 2.1 shows the best μ' we have obtained, which we computed using the parameters $N = 96$ and $T = 20$ in the notation of the next section. This distribution is of the form $\mu' = \sum_{n \geq 0} c'_n \delta_{r'_n}$, with Fourier transform $\widehat{\mu}' = \sum_{n \geq 0} c''_n \delta_{r''_n}$. In the table, we have rescaled the distribution μ' so that $c'_0 = c''_0 = 1$. Note that

$$r_1 \approx r'_1 \approx r''_1 \approx r''_2,$$
$$r_2 \approx r'_2 \approx r'_3 \approx r''_3 \approx r''_4, \text{ and}$$
$$r_3 \approx r'_4 \approx r'_5 \approx r''_5 \approx r''_6,$$

and the sums of the corresponding coefficients are also near each other. The approximation to μ is not yet very close, but one can already see μ roughly emerging from μ'.

TABLE 2.1. Radii and coefficients for dual distributions in $\mathbb{R}^{16}$.

n	r_n	c_n
0	0	1
1	$1.7393272583625204\ldots$	$8431.71627140\ldots$
2	$2.2346642069957498\ldots$	$292026.09352080\ldots$
3	$2.6462005756471079\ldots$	$3111809.14450639\ldots$
n	r'_n	c'_n
0	0	1
1	$1.7385384653461733\ldots$	$8360.61230142\ldots$
2	$2.1990965401230488\ldots$	$4240.44226222\ldots$
3	$2.2331930934327142\ldots$	$282582.90774253\ldots$
4	$2.6366241274825130\ldots$	$2419678.28385080\ldots$
5	$2.6651290005171109\ldots$	$584982.54962505\ldots$
n	r''_n	c''_n
0	0	1
1	$1.6604472109700065\ldots$	$133.02471778\ldots$
2	$1.7414917267847931\ldots$	$8321.61159562\ldots$
3	$2.2277237020673214\ldots$	$245869.54859549\ldots$
4	$2.2887685306282807\ldots$	$50042.27252495\ldots$
5	$2.6253975605696717\ldots$	$1578408.61282183\ldots$
6	$2.6773906784567302\ldots$	$1610965.69273527\ldots$

3. AN ALGORITHM FOR DUAL LINEAR PROGRAMMING BOUNDS

Proposition 2.2 allows modular forms for the congruence subgroup $\Gamma_1(N)$, but for simplicity we will restrict our attention to those for the larger group $\Gamma_0(N)$ (equivalently, to modular forms for $\Gamma_1(N)$ that have trivial Nebentypus). There is some loss of generality, but this case serves as an attractive proving ground for the general theory, and it should suffice when the dimension d is a multiple of 4.

Specifically, let $k = d/2$ be an even integer, and let $M_k(\Gamma_0(N))$ be the space of modular forms of weight k for $\Gamma_0(N)$. Recall that this space has a basis consisting of modular forms with rational coefficients in their q-expansions (see, for example, Corollary 12.3.12 in [15]). Furthermore, the Atkin-Lehner involution on $M_k(\Gamma_0(N))$ preserves the property of having rational coefficients (see Lemma 3.5.3 in [24]).

In practice, to simplify Section 5 we also assume that N is not divisible by 16^2, 9^2, or p^2 for any prime $p > 3$, but this assumption is not essential.

We would like to find a modular form $g = \sum_{n \geq 0} a_n q^n$ in $M_k(\Gamma_0(N))$ with the following properties for some T, where we set $\widetilde{g} = i^k g|_k w_N = \sum_{n \geq 0} b_n q^n$:

(1) $a_0 = 1$ and $b_0 > 0$,
(2) $a_n \geq 0$ and $b_n \geq 0$ for all $n \geq 0$, and
(3) $a_n = 0$ for $1 \leq n < T$.

Then we use the distribution

$$\mu = \sum_{n \geq 0} a_n \delta_{\sqrt{n}}$$

in Proposition 2.1. By Proposition 2.2, we have $c = (2/\sqrt{N})^{d/2}b_0$ and $r = \sqrt{T}$ in the notation of Proposition 2.1. Thus, we obtain a lower bound of

$$b_0 \left(\frac{2}{\sqrt{N}}\right)^{d/2} \left(\frac{\sqrt{T}}{2}\right)^d$$

for the linear programming bound in $\mathbb{R}^d$, and we wish to choose g so as to maximize this bound. We will do so by linear programming, with one caveat: all our calculations will consider only the terms up to q^M in the q-series for some fixed M, and at the end we must check that the inequalities are not violated beyond that point.

Let $g^1, \ldots, g^{\dim M_k(\Gamma_0(N))}$ be a basis of $M_k(\Gamma_0(N))$ with rational q-series coefficients, and let $\widetilde{g}^j = i^k g^j|_k w_N$ be i^k times the image of g^j under the full level N Atkin-Lehner involution. We write the q-expansions of the modular forms g^j and $\widetilde{g}^j$ as

$$g^j = \sum_{n=0}^{\infty} a_n^j q^n \qquad \text{and} \qquad \widetilde{g}^j = \sum_{n=0}^{\infty} b_n^j q^n,$$

and we fix integers T and M with $1 \leq T < \dim M_k(\Gamma_0(N)) < M$. These bases and q-series can all be computed algorithmically (see, for example, [28]).

Now we write $g = \sum_j x_j g^j$ with respect to our basis, and we optimize over the choice of coefficients x_j by solving the following linear program:

$$\begin{aligned}
\text{maximize} \quad & \sum_j x_j b_0^j \\
\text{subject to} \quad & 1 = \sum_j x_j a_0^j, \\
& 0 = \sum_j x_j a_n^j \text{ for } 1 \leq n < T, \\
& 0 \leq \sum_j x_j a_n^j \text{ for } T \leq n \leq M, \text{ and} \\
& 0 \leq \sum_j x_j b_n^j \text{ for } 1 \leq n \leq M.
\end{aligned}$$

These inequalities encode all the desired properties of f and g, except that we examine only the terms up to q^M in the q-series.

We hope that if M is large enough, then all the terms beyond q^M will have nonnegative coefficients automatically, and we attempt to use asymptotic bounds to confirm that all of the coefficients of g and $\widetilde{g}$ are nonnegative (see Section 5). If this verification fails, we can increase M and attempt the optimization problem again. In practice, $M = 2 \cdot \dim M_k(\Gamma_0(N))$ typically seems to be sufficient for the algorithm to succeed, and it works for all the numerical results we report in this paper.

To find the best possible bounds, we run the method for several values of N and T. Larger values of N typically yield better results, but not always. It seems difficult to predict the best values for T in general, although they also tend to increase as N increases. See Section 6 for the results of this method applied to the spaces $M_k(\Gamma_0(N))$ of modular forms of weight $k \in \{6, 8, 10, 14, 16\}$ and level $N = 24$ or 96.

For a concrete illustration of the method, consider the case $d = 16$ and $N = 4$. One can show that the space $M_8(\Gamma_0(4))$ is five-dimensional, with the following basis. Let

$$E_8(z) = 1 + 480 \sum_{n=1}^{\infty} \sigma_7(n) q^n$$

be the Eisenstein series of weight 8 for $\mathrm{SL}_2(\mathbb{Z})$ (not to be confused with the E_8 root lattice), and let f be the newform of weight 8 for $\Gamma_0(2)$ defined by

$$f(z) = q \prod_{n=1}^{\infty} (1 - q^n)^8 (1 - q^{2n})^8.$$

Then $M_g(\Gamma_0(4))$ has the basis $g^1, \ldots, g^5$, where $g^1(z) = E_8(z)$, $g^2(z) = 16E_8(2z)$, $g^3(z) = 256E_8(4z)$, $g^4(z) = f(z)$, and $g^5(z) = 16f(2z)$. The Atkin-Lehner involution acts by $\widetilde{g}^1 = g^3$, $\widetilde{g}^2 = g^2$, $\widetilde{g}^3 = g^1$, $\widetilde{g}^4 = g^5$, and $\widetilde{g}^5 = g^4$. Using this information, we can write down the linear program explicitly and solve it. As usual, the trickiest part is identifying the right choice of T, while we can simply take M large enough (e.g., $M = 10$ is more than sufficient).

For $T = 2$, solving the linear program yields the modular form

$$\sum_j x_j g^j = \frac{1}{17} g^1 + \frac{1}{17} g^2 - \frac{480}{17} g^4$$
$$= 1 + 4320q^2 + 61440q^3 + 522720q^4 + \cdots,$$

which is the theta series of the Barnes-Wall lattice. Similarly, for $T = 4$ we obtain

$$\sum_j x_j g^j = \frac{1}{272} g^2 + \frac{1}{272} g^3 - \frac{30}{17} g^5$$
$$= 1 + 4320q^4 + 61440q^6 + 522720q^8 + \cdots,$$

which is the same modular form with q replaced by q^2 and which yields the same bound. For these two values of T, the space $M_8(\Gamma_0(4))$ is incapable of separating the linear programming bound from the center density 0.0625 of the Barnes-Wall lattice. However, for $T = 3$ we obtain

$$\sum_j x_j g^j = \frac{1}{136} g^1 - \frac{121}{2176} g^2 + \frac{1}{136} g^3 - \frac{60}{17} g^4 - \frac{60}{17} g^5$$
$$= 1 + 7680q^3 + 4320q^4 + 276480q^5 + \cdots,$$

which yields an improved center density lower bound of $3^8/2^{16} = 0.100112\ldots$, more than 60% greater than the center density of the Barnes-Wall lattice. In fact, this modular form has been studied before: it is the *extremal theta series* in 16 dimensions (see equation (47) in [13, p. 190]).

It is tempting to conjecture that the extremal theta series should exactly match the optimal linear programming bound. This conjecture would be a beautiful analogue of the behavior in 8 and 24 dimensions. In those cases the optimal lattices have determinant 1 and minimal norm 2 or 4, respectively. The extremal theta series in 16 dimensions behaves like the theta series of a lattice of determinant 1 and minimal norm 3, exactly interpolating between 8 and 24 dimensions. Presumably no such lattice exists, but the linear programming bound could match the density of a hypothetical lattice.

That is a good approximation in this case, but the answer turns out to be more subtle: in Section 6, we obtain a better lower bound using $N = 96$. Instead of minimal norm 3, the improved lower bound is 3.022. For comparison, we believe the true linear programming bound amounts to a minimal norm of

$$3.025259311682882063282086557901196\ldots,$$

but we are unable to conjecture an exact formula for this number.

4. Poisson summation analogues from modular forms

The main result of this section is Proposition 2.2, which yields a summation formula from a modular form. Summation formulas of this sort are well known to number theorists, and essentially equivalent to the functional equation for the L-function. We record the details here and sketch a proof for the convenience of the reader. (One can also prove such a formula using the density of complex Gaussians among radial Schwartz functions, along the lines of Section 6 in [25] or Section 2.3 in [11].)

Proposition 2.2 is essentially a version of Voronoi summation. Our proof will follow the approach used in standard proofs of Voronoi summation (for example, as in Section 10.2.5 of [4] or Section 2 of [22]). The key idea comes from the classical observation that the usual Poisson summation formula is a consequence of the functional equation of the Riemann zeta function. Similarly, Proposition 2.2 follows from the functional equation relating the L-functions associated to a modular form and its Atkin-Lehner dual.

In what follows, we use the notation established in Proposition 2.2. To state the functional equation, we first define the *L-function*

$$L(s, g) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

when $\mathrm{Re}(s) > k$, and the *completed L-function*

$$\Lambda(s, g) = N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, g).$$

The functional equation relating $\Lambda(s, g)$ and $\Lambda(s, \widetilde{g})$ is classical, dating back to Hecke [19]. It says that the L-functions can be analytically continued so that

$$\Lambda(s, g) + \frac{a_0}{s} + \frac{b_0}{k - s}$$

is entire and bounded in every vertical strip, and we have the functional equation

$$\Lambda(s, g) = \Lambda(k - s, \widetilde{g}),$$

or equivalently

(4.1) $$L\left(k - \frac{s}{2}, g\right) = N^{(s-k)/2}(2\pi)^{k-s}\frac{\Gamma(s/2)}{\Gamma(k - s/2)}L\left(\frac{s}{2}, \tilde{g}\right).$$

See, for example, Theorem 1 in [23, p. I-5].

Sketch of proof of Proposition 2.2. For a radial Schwartz function f on $\mathbb{R}^d$, let

$$S = \sum_{n \geq 1} a_n f(\sqrt{n}).$$

By Mellin inversion,

$$a_n f(\sqrt{n}) = \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=\sigma} \frac{a_n}{n^{s/2}} \mathcal{M}f(s)\, ds$$

for any $\sigma > 0$, where the *Mellin transform* $\mathcal{M}f$ is defined by

$$\mathcal{M}f(s) = \int_0^{\infty} f(x)x^s \frac{dx}{x}.$$

In particular, for $\sigma = d + \varepsilon$ with $\varepsilon > 0$,

$$S = \frac{1}{2\pi i} \sum_{n=1}^{\infty} \int_{\mathrm{Re}(s)=d+\varepsilon} \frac{a_n}{n^{s/2}} \mathcal{M}f(s) \, ds$$

$$= \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=d+\varepsilon} L\left(\frac{s}{2}, g\right) \mathcal{M}f(s) \, ds,$$

where switching the sum and integral is permitted because of the uniform convergence of the sum defining the L-function.

The integrand $L(s/2, g)\mathcal{M}f(s)$ is negligible when s has large imaginary part. To see why, note that by a stationary phase argument the Mellin transform $\mathcal{M}f(s)$ is rapidly decaying as $\mathrm{Im}(s)$ grows, while $L(s/2, g)$ grows at most polynomially in $\mathrm{Im}(s)$ by the Phragmén-Lindelöf principle. Thus, we can shift the contour of integration to the left, as long as we account for poles.

It is not hard to check that $\mathcal{M}f(s)$ has a possible pole at $s = 0$ with residue $f(0)$, $L(s/2, g)$ has a possible pole at $s = d$ with residue

$$2 \left(\frac{2\pi}{\sqrt{N}}\right)^{d/2} \frac{1}{\Gamma(d/2)} b_0,$$

and $L(0, g) = -a_0$, since the pole of $\Gamma(s)$ at $s = 0$ cancels the pole of $\Lambda(s, g)$ at $s = 0$. Thus,

$$S = -a_0 f(0) + 2b_0 \left(\frac{2\pi}{\sqrt{N}}\right)^{d/2} \frac{1}{\Gamma(d/2)} \mathcal{M}f(d) + \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=-\varepsilon} L\left(\frac{s}{2}, g\right) \mathcal{M}f(s) \, ds.$$

Setting

$$T = \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=-\varepsilon} L\left(\frac{s}{2}, g\right) \mathcal{M}f(s) \, ds$$

and applying the identity $\widehat{f}(0) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \mathcal{M}f(d)$, we see that

(4.2) $$a_0 f(0) + S = \left(\frac{2}{\sqrt{N}}\right)^{d/2} b_0 \widehat{f}(0) + T.$$

Changing variables from s to $d - s$ and applying the functional equation (4.1) yields

$$T = \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=d+\varepsilon} L\left(\frac{d-s}{2}, g\right) \mathcal{M}f(d-s) \, ds$$

$$= \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=d+\varepsilon} N^{s/2-d/4} (2\pi)^{d/2-s} \frac{\Gamma(s/2)}{\Gamma((d-s)/2)} L\left(\frac{s}{2}, \widetilde{g}\right) \mathcal{M}f(d-s) \, ds.$$

Now we use the identity

$$\mathcal{M}\widehat{f}(s) = \frac{\pi^{d/2-s} \Gamma(s/2)}{\Gamma((d-s)/2)} \mathcal{M}f(d-s)$$

(see Theorem 5.9 in [21]). Making this substitution, we find that

$$T = \left(\frac{2}{\sqrt{N}}\right)^{d/2} \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=d+\varepsilon} \left(\frac{4}{N}\right)^{-s/2} L\left(\frac{s}{2}, \widetilde{g}\right) \mathcal{M}\widehat{f}(s) \, ds.$$

Replacing the L-function with its defining sum, switching the sum and integral as above, and applying Mellin inversion again (reversing the steps from the start of

the proof), we see that

$$
\begin{aligned}
T &= \left(\frac{2}{\sqrt{N}}\right)^{d/2} \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=d+\varepsilon} \sum_{n=1}^{\infty} \frac{b_n}{(4n/N)^{s/2}} \mathcal{M}\widehat{f}(s)\, ds \\
&= \left(\frac{2}{\sqrt{N}}\right)^{d/2} \sum_{n=1}^{\infty} \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=d+\varepsilon} \frac{b_n}{(4n/N)^{s/2}} \mathcal{M}\widehat{f}(s)\, ds \\
&= \left(\frac{2}{\sqrt{N}}\right)^{d/2} \sum_{n=1}^{\infty} b_n \widehat{f}\left(\frac{2\sqrt{n}}{\sqrt{N}}\right).
\end{aligned}
$$

Hence, (4.2) implies that

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

as desired. □

5. Checking positivity of modular form coefficients

In this section, we explain how we check whether a modular form of weight k for $\Gamma_0(N)$ has nonnegative coefficients in its q-series. This method uses only standard techniques from the theory of modular forms, but we describe them here for the benefit of readers in discrete geometry. The key idea is that Eisenstein series typically make the dominant contribution asymptotically, which reduces the problem to a finite calculation if the Eisenstein contribution is positive.

As mentioned above, we assume for simplicity that N is not divisible by 16^2, 9^2, or p^2 for any prime $p > 3$. This assumption guarantees that all the characters in this section are real. Furthermore, we assume that $k \geq 3$, because the Eisenstein series for weight 2 must be obtained using different formulas (the formulas that work for $k \geq 3$ no longer converge when $k = 2$).

To verify that $g = \sum_{n=0}^{\infty} a_n q^n$ has $a_n \geq 0$ for all n, we write g as $g_e + g_c$, where $g_e = \sum_{n=0}^{\infty} e_n q^n$ is a linear combination of Eisenstein series and $g_c = \sum_{n=0}^{\infty} c_n q^n$ is cuspidal, and we attempt to carry out the following steps:

(1) Use Weil bounds to show that $|c_n| \leq C_g n^{k/2}$ for some explicit constant C_g.
(2) Use explicit formulas for Eisenstein series to show that $e_n \geq r_g n^{k-1}$ for some explicit constant $r_g > 0$.
(3) Compare the Eisenstein part and the cuspidal part to produce a bound Q such that $a_n > 0$ for $n > Q$.
(4) Explicitly compute the coefficients a_n of g to check that $a_n \geq 0$ for $n \leq Q$.

The first step is straightforward, given some powerful machinery. Deligne's proof of the Weil conjectures [14] implies that, independent of weight, if $h = \sum_{n=1}^{\infty} c_n q^n$ is a cuspidal Hecke eigenform normalized so that $c_{n'} = 1$ for the minimal n' with $c_{n'} \neq 0$, then $|c_n| \leq \sigma_0(n) n^{(k-1)/2} \leq n^{k/2}$. Let $B_k(N)$ be the set of such eigenforms, which are a basis for the cuspidal part of $M_k(\Gamma_0(N))$. (Note that the elements of $B_k(N)$ typically do not have rational coefficients. Instead, we must work over a larger number field.) If

$$
g_c = \sum_{n=1}^{\infty} c_n q^n = \sum_{h \in B_k(N)} x_h h
$$

with coefficients $x_h \in \mathbb{C}$, then

$$|c_n| \leq n^{k/2} \sum_{h \in B_k(N)} |x_h|.$$

Thus, step (1) holds with $C_g = \sum_{h \in B_k(N)} |x_h|$.

For the second step, we need to write down the Eisenstein series explicitly. We can describe them in terms of primitive Dirichlet characters ϕ of conductor u and natural numbers t such that $u^2 t \mid N$ (where $a \mid b$ means a divides b). Thanks to our divisibility hypotheses on N, it follows that $u \mid 24$, and therefore ϕ must be a real character; in other words, it takes on only the values ± 1. Then the Eisenstein series in $M_k(\Gamma_0(N))$ all have the form

$$E_t^{\phi} = \frac{\delta(\phi)}{2} L(1 - k, \phi) + \sum_{\substack{n \geq 1, \\ t \mid n}} \phi(n/t) \sigma_{k-1}(n/t) q^n,$$

where $\sigma_\ell(m) = \sum_{d \mid m} d^\ell$, $L(s, \phi)$ is the L-function of ϕ, and

$$\delta(\phi) = \begin{cases} 1 & \text{if } \phi \text{ is the trivial character of conductor 1, and} \\ 0 & \text{otherwise.} \end{cases}$$

See, for example, Theorem 4.5.2 in [16].

Since the Eisenstein series span the Eisenstein part of $M_k(\Gamma_0(N))$, there exist constants y_t^{ϕ} such that

$$g_e = \sum_{t, \phi} y_t^{\phi} E_t^{\phi}$$

$$= e_0 + \sum_{t, \phi} \sum_{\substack{n \geq 1, \\ t \mid n}} y_t^{\phi} \phi(n/t) \sigma_{k-1}(n/t)$$

$$= e_0 + \sum_{n=1}^{\infty} \sum_{\substack{t \mid N, \\ t \mid n}} \left(\sum_{\phi} y_t^{\phi} \phi(n/t) \right) \sigma_{k-1}(n/t).$$

It is straightforward to check that whenever $t \mid n$,

$$\frac{\sigma_{k-1}(n)}{\sigma_{k-1}(t)} \leq \sigma_{k-1}(n/t) \leq \frac{\sigma_{k-1}(n)}{t^{k-1}}.$$

This implies that if we set

$$r_g(t, n) = \begin{cases} \frac{1}{t^{k-1}} & \text{if } \sum_{\phi} y_t^{\phi} \phi(n/t) < 0, \text{ and} \\ \frac{1}{\sigma_{k-1}(t)} & \text{if } \sum_{\phi} y_t^{\phi} \phi(n/t) \geq 0, \end{cases}$$

$$r_g(n) = \sum_{\substack{t \mid N, \\ t \mid n}} \left(\sum_{\phi} y_t^{\phi} \phi(n/t) \right) r_g(t, n),$$

and

$$r_g = \min_{n \geq 1} r_g(n) = \min_{1 \leq n \leq N} r_g(n),$$

then

$$e_n \geq \sigma_{k-1}(n) r_g \geq n^{k-1} r_g.$$

This completes step (2), provided that r_g is positive. If it is not positive, then our test will be inconclusive, since we are unable to certify that even the Eisenstein part is nonnegative.

Combining the results of the previous two steps, we find that

$$a_n \geq n^{k-1} r_g - n^{k/2} C_g.$$

Since $k > 2$, this inequality provides an easily computed bound $Q = \lfloor (C_g/r_g)^{2/(k-2)} \rfloor$ such that $a_n > 0$ for all $n > Q$. Because of the large gap between n^{k-1} and $n^{k/2}$, the bound Q is typically relatively small. Finally, to certify that the coefficients of g are all nonnegative, we explicitly compute the coefficients a_n for $n \leq Q$.

This method will not always work, without more careful estimates. For example, it fails if a_n is not eventually positive. That can occur in practice: in the example from Section 3 with $d = 16$, $N = 4$, and $T = 2$, the optimal modular form is

$$g = 1 + 4320q^2 + 61440q^3 + 522720q^4 + 2211840q^5 + 8960640q^6 + \cdots,$$

which has eventually positive coefficients, but

$$\widetilde{g} = 16 + 69120q^4 + 983040q^6 + 8363520q^8 + 35389440q^{10} + \cdots,$$

which does not. Thus, proving that $\widetilde{g}$ has nonnegative coefficients requires a little more care. However, we have not observed this phenomenon for the best choices of T in any of the cases we have examined. If it were to occur, it could be handled by distinguishing between the values of $r_g(n)$ for different residue classes of n modulo N, and showing that the cuspidal contribution vanishes whenever $r_g(n) = 0$.

6. Numerical results

Table 6.1 shows our numerical results. We used the SageMath computer algebra system [26] for our calculations, with one exception: we used Magma [2] to compute bases for modular forms and the action of the Atkin-Lehner involution. This combination works conveniently, because SageMath has an interface for calling Magma code.

To produce rigorous results, we used exact rational arithmetic, and we proved nonnegativity of coefficients using the techniques of Section 5. For calculations with forms of level 24, we directly solved the linear program over $\mathbb{Q}$; for level 96, we instead used floating point arithmetic to obtain an approximate solution, which we then used to obtain a rational solution and prove its correctness and optimality. All the numbers in the table are rounded correctly: lower bounds are rounded down, and upper bounds are rounded up.

7. Open problems

Our new lower bounds in Table 6.1 come fairly close to the known upper bounds, but they do not agree to many decimal places. We believe that the upper bounds agree with the true linear programming bound, aside from rounding the last decimal place up, while the lower bounds could be further improved. One difficulty in doing so is that modular forms are inherently quantized: in the summation formula

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

TABLE 6.1. Center density bounds in dimensions 8 through 32. The upper bound is the linear programming bound, computed using the best auxiliary function currently known [7], while the dual bound is based on the given values of N and T, and the record packing is the densest packing currently known [13]. In dimensions 12 and 16, we include both $N = 96$ and $N = 24$ for comparison.

Dimension	Record packing	Dual bound	Upper bound	N	T
8	0.0625	0.0625	0.0625	1	1
12	0.037037	0.062446	0.062742	96	9
		0.059781		24	4
16	0.0625	0.106284	0.107059	96	20
		0.103948		24	6
20	0.131537	0.260996	0.276169	24	9
24	1	1	1	1	2
28	1	4.591741	4.828588	24	9
32	2.565784	28.086665	29.942182	24	12

there is no possibility to perturb the radii $\sqrt{n}$ or $2\sqrt{n/N}$ slightly, and so one must do the best one can using only radii of these forms. In particular, closely matching the upper bound may require N to be very large, perhaps on the order of 10^{10} if we wish to match ten digits, and dealing with such large N is not practical. Any feasible method that could close the gap between the primal and dual bounds to within a factor of $1 + 10^{-10}$ would be a significant advance, and modular forms might not be the right tool for this purpose. For comparison, [30] and [27] obtain dual linear programming bounds in high dimensions using an entirely different approach.

Another topic we leave open is computations in dimensions that are not divisible by 4. We see no theoretical obstacle to such an extension: one must simply use modular forms of odd weight (for dimensions divisible by 2 but not 4) or half-integral weight (for odd dimensions), and replace $\Gamma_0(N)$ with $\Gamma_1(N)$ so that such forms exist. However, we have not implemented these computations. We have also not explored the uncertainty principle introduced in [3] and further studied in [7], for which one could again prove dual bounds using modular forms.

One intriguing possibility that may be nearly within reach is proving that there exists a dimension in which the linear programming bound is not sharp. All dimensions except 1, 2, 8, and 24 seem to have this property, but so far no proof is known. Three dimensions would be a natural target, because we know the optimal packing density, and thus it would suffice to prove any dual bound greater than this density. In higher dimensions, it would require an improvement on the linear programming bound. The only such bound currently known is Theorem 1.4 from de Laat, Oliveira, and Vallentin's paper [20], which is a refinement of the linear programming bound that seems to give a small numerical improvement in dimensions 3, 4, 5, 6, 7, and 9 (see Table 1 in [20]) and presumably higher dimensions as well, aside from 24. Any dual bound greater than this improved upper bound would suffice to show that the linear programming bound is not sharp. Conversely, it would be interesting to prove dual bounds for the theorem of de Laat, Oliveira, and Vallentin itself.

References

[1] G. E. Andrews, R. Askey, and R. Roy, *Special Functions*, Encyclopedia of Mathematics and its Applications **71**, Cambridge University Press, Cambridge, 1999. doi:10.1017/CBO9781107325937 MR1688958

[2] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system I: the user language*, J. Symbolic Comput. **24** (1997), no. 3–4, 235–265. doi:10.1006/jsco.1996.0125 MR1484478

[3] J. Bourgain, L. Clozel, and J.-P. Kahane, *Principe d'Heisenberg et fonctions positives*, Ann. Inst. Fourier (Grenoble) **60** (2010), no. 4, 1215–1232. doi:10.5802/aif.2552 MR2722239

[4] H. Cohen, *Number Theory. Volume II: Analytic and Modern Tools*, Graduate Texts in Mathematics **240**, Springer, New York, 2007. doi:10.1007/978-0-387-49894-2 MR2312338

[5] H. Cohn, *New upper bounds on sphere packings II*, Geom. Topol. **6** (2002), 329–353. arXiv:math/0110010 doi:10.2140/gt.2002.6.329 MR1914571

[6] H. Cohn and N. Elkies, *New upper bounds on sphere packings I*, Ann. of Math. (2) **157** (2003), no. 2, 689—714. arXiv:math/0110009 doi:10.4007/annals.2003.157.689 MR1973059

[7] H. Cohn and F. Gonçalves, *An optimal uncertainty principle in twelve dimensions via modular forms*, Invent. Math. **217** (2019), no. 3, 799–831. arXiv:1712.04438 doi:10.1007/s00222-019-00875-4 MR3989254

[8] H. Cohn and A. Kumar, *Universally optimal distribution of points on spheres*, J. Amer. Math. Soc. **20** (2007), no. 1, 99–148. arXiv:math/0607446 doi:10.1090/S0894-0347-06-00546-7 MR2257398

[9] H. Cohn and A. Kumar, *Optimality and uniqueness of the Leech lattice among lattices*, Ann. of Math. (2) **170** (2009), no. 3, 1003–1050. arXiv:math/0403263 doi:10.4007/annals.2009.170.1003 MR2600869

[10] 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), no. 3, 1017–1033. arXiv:1603.06518 doi:10.4007/annals.2017.185.3.8 MR3664817

[11] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *Universal optimality of the E_8 and Leech lattices and interpolation formulas*, preprint, 2019. arXiv:1902.05438

[12] H. Cohn and Y. Zhao, *Sphere packing bounds via spherical codes*, Duke Math. J. **163** (2014), no. 10, 1965–2002. arXiv:1212.5966 doi:10.1215/00127094-2738857 MR3229046

[13] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, third edition, Grundlehren der Mathematischen Wissenschaften **290**, Springer-Verlag, New York, 1999. doi:10.1007/978-1-4757-6568-7 MR1662447

[14] P. Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. **43** (1974), 273–307. `http://www.numdam.org/item?id=PMIHES_1974__43__273_0` MR0340258

[15] F. Diamond and J. Im, *Modular forms and modular curves*, in V. K. Murty, ed., *Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994)*, pp. 39–133, CMS Conf. Proc. **17**, Amer. Math. Soc., Providence, RI, 1995. MR1357209

[16] F. Diamond and J. Shurman, *A First Course in Modular Forms*, Graduate Texts in Mathematics **228**, Springer-Verlag, New York, 2005. doi:10.1007/978-0-387-27226-9 MR2112196

[17] T. C. Hales, *A proof of the Kepler conjecture*, Ann. of Math. (2) **162** (2005), no. 3, 1065–1185. doi:10.4007/annals.2005.162.1065 MR2179728

[18] T. Hales, M. Adams, G. Bauer, T. D. Dang, J. Harrison, L. T. Hoang, C. Kaliszyk, V. Magron, S. McLaughlin, T. T. Nguyen, Q. T. Nguyen, T. Nipkow, S. Obua, J. Pleso, J. Rute, A. Solovyev, T. H. A. Ta, N. T. Tran, T. D. Trieu, J. Urban, K. Vu, and R. Zumkeller, *A formal proof of the Kepler conjecture*, Forum Math. Pi **5** (2017), e2, 29 pp. arXiv:1501.02155 doi:10.1017/fmp.2017.1 MR3659768

[19] E. Hecke, *Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung*, Math. Ann. **112** (1936), no. 1, 664–699. doi:10.1007/BF01565437 MR1513069

[20] D. de Laat, F. M. de Oliveira Filho, and F. Vallentin, *Upper bounds for packings of spheres of several radii*, Forum Math. Sigma **2** (2014), e23, 42 pp. arXiv:1206.2608 doi:10.1017/fms.2014.24 MR3264261

[21] E. H. Lieb and M. Loss, *Analysis*, second edition, Graduate Studies in Mathematics **14**, American Mathematical Society, Providence, RI, 2001. doi:10.1090/gsm/014 MR1817225

[22] S. D. Miller and W. Schmid, *Summation formulas, from Poisson and Voronoi to the present*, in *Noncommutative Harmonic Analysis*, Progr. Math. **220**, 419–440, Birkhäuser Boston, Boston, MA, 2004. arXiv:math/0304187 doi:10.1007/978-0-8176-8204-0_15 MR2036579

[23] A. Ogg, *Modular Forms and Dirichlet Series*, W. A. Benjamin, Inc., New York-Amsterdam, 1969. MR0256993

[24] M. Ohta, *On the p-adic Eichler-Shimura isomorphism for Λ-adic cusp forms*, J. Reine Angew. Math. **463** (1995), 49–98. doi:10.1515/crll.1995.463.49 MR1332907

[25] D. Radchenko and M. Viazovska, *Fourier interpolation on the real line*, Publ. Math. Inst. Hautes Études Sci. **129** (2019), 51–81. arXiv:1701.00265 doi:10.1007/s10240-018-0101-z MR3949027

[26] *SageMath, the Sage Mathematics Software System (Version 8.8)*, The Sage Developers, 2019, `http://www.sagemath.org`.

[27] A. Scardicchio, F. H. Stillinger, and S. Torquato, *Estimates of the optimal density of sphere packings in high dimensions*, J. Math. Phys. **49** (2008), no. 4, 043301, 15 pp. arXiv:0705.1482 doi:10.1063/1.2897027 MR2412293

[28] W. Stein, *Modular Forms, a Computational Approach*, Graduate Studies in Mathematics **79**, American Mathematical Society, Providence, RI, 2007 doi:10.1090/gsm/079 MR2289048

[29] A. Thue, *Om nogle geometrisk-taltheoretiske Theoremer*, Forhandlingerne ved de Skandinaviske Naturforskeres **14** (1892), 352–353.

[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. arXiv:math/0508381 `https://projecteuclid.org/euclid.em/1175789761` MR2264469

[31] M. S. Viazovska, *The sphere packing problem in dimension* 8, Ann. of Math. (2) **185** (2017), no. 3, 991–1015. arXiv:1603.04246 doi:10.4007/annals.2017.185.3.7 MR3664816

Microsoft Research New England, One Memorial Drive, Cambridge, MA 02142
Email address: `cohn@microsoft.com`

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139
Current address: Department of Mathematics, University of Georgia, Athens, GA 30602
Email address: `nicholas.triantafillou@gmail.com`