ISSN 1364-0380 (on line) 1465-3060 (printed)

Geometry & Topology
Volume 6 (2002) 329–353
Published: 25 June 2002

New upper bounds on sphere packings II

Henry Cohn

Microsoft Research, One Microsoft Way
Redmond, WA 98052-6399, USA

Email: `cohn@microsoft.com`

URL: `http://research.microsoft.com/~cohn`

Abstract

We continue the study of the linear programming bounds for sphere packing introduced by Cohn and Elkies. We use theta series to give another proof of the principal theorem, and present some related results and conjectures.

This article is in the arXiv as: `arXiv:math.MG/0110010`

Dedicated to Daniel Lewin (14 May 1970 – 11 September 2001)

AMS Classification numbers Primary: 52C17, 52C07

Secondary: 33C10, 33C45

Keywords: Sphere packing, linear programming bounds, lattice, theta series, Laguerre polynomial, Bessel function

Proposed: Robion Kirby
Seconded: Michael Freedman, Walter Neumann

Received: 5 October 2001
Accepted: 25 May 2002

1 Introduction

In [4], Cohn and Elkies introduce linear programming bounds for the sphere packing problem, and use them to prove new upper bounds on the sphere packing density in low dimensions. These bounds are the best bounds known in dimensions 4 through 36, and seem to be sharp in dimensions 8 and 24, although that has not yet been proved. Here, we continue the study of these bounds, by giving another derivation of the main theorem of [4]. We then prove an optimality theorem of Gorbachev [8], and outline in some conjectures how the proof techniques should apply more generally.

We continue to use the notation of [4]. See the introduction of that paper for background and references on sphere packing.

The main theorem Cohn and Elkies prove is the following:

Theorem 1.1 *Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is a radial, admissible function, is not identically zero, and satisfies the following two conditions:*

(1) *$f(x) \le 0$ for $|x| \ge 1$, and*

(2) *$\widehat{f}(t) \ge 0$ for all t.*

Then the center densities of n–dimensional sphere packings are bounded above by

$$\frac{f(0)}{2^n \widehat{f}(0)}.$$

Here, the Fourier transform is normalized by

$$\widehat{f}(t) = \int_{\mathbb{R}^n} f(x) e^{2\pi i \langle t,x\rangle}\, dx,$$

and *admissibility* means that there is a constant $\varepsilon > 0$ such that both $|f(x)|$ and $|\widehat{f}(x)|$ are bounded above by a constant times $(1+|x|)^{-n-\varepsilon}$. More broadly, we could in fact take f to be any function to which the Poisson summation formula applies: for every lattice $\Lambda \subset \mathbb{R}^n$ and every vector $v \in \mathbb{R}^n$,

$$\sum_{x\in\Lambda} f(x+v) = \frac{1}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} \sum_{t\in\Lambda^*} e^{-2\pi i\langle v,t\rangle} \widehat{f}(t).$$

However, the narrower definition of admissibility is easier to check and seemingly suffices for all natural examples.

Section 2 gives another proof of Theorem 1.1, for $n > 1$. This proof is not as simple as the one in [4], but the method is of interest in its own right, as are some of the intermediate results. Section 3 proves Gorbachev's theorem [8] that certain admissible functions (those constructed in Proposition 6.1 of [4], or independently by Gorbachev) are optimal, among functions whose Fourier transforms have support in a certain ball. Finally, Section 4 discusses the dual linear program, and puts the techniques of Section 3 into a broader context.

Acknowledgements

I thank Richard Askey, Noam Elkies, Pavel Etingof, and George Gasper for their advice about special functions, and Tom Brennan, Harold Diamond, Gerald Folland, Cormac Herley, David Jerison, Greg Kuperberg, Ben Logan, László Lovász, Steve Miller, Amin Shokrollahi, Neil Sloane, Hart Smith, Jeffrey Vaaler, David Vogan, and Michael Weinstein for helpful discussions. I was previously supported by an NSF Graduate Research Fellowship and a summer internship at Lucent Technologies, and currently hold an American Institute of Mathematics Five-year Fellowship.

2 Positivity of theta series coefficients

We will prove Theorem 1.1 using the positivity of the coefficients of the theta series of lattices. For each lattice, the theta series of its dual must have positive coefficients, and these coefficients are some transformation of those for the original lattice. This puts strong constraints on the theta series of a lattice, which we exploit below. For simplicity, we will deal only with the case of lattice packings, but everything in this section applies to all sphere packings, by replacing the theta series of a lattice with the average theta series of a periodic packing (see [5, page 45]). Also, for technical reasons we will deal only with the case $n > 1$, which is not a serious restriction as 1–dimensional sphere packing is trivial.

Unfortunately, carrying this program out rigorously involves dealing with a number of technicalities. If one simply wants an idea of the overall argument, without worrying about rigor, one can follow this plan: Ignore Lemma 2.4 and all references to Cesàro sums, and assume that all Laguerre series converge. Ignore the uniformity of convergence in Lemma 2.6 (in which case the proof becomes far simpler). Ignore the justification of interchanging the sum and integral in Lemma 2.7. Following this plan will of course not lead to a rigorous proof, but it may make the underlying ideas clearer.

Before going further, we need a lemma about Laguerre polynomials. Let L_k^α be the Laguerre polynomial of degree k and parameter $\alpha > -1$. These polynomials are orthogonal with respect to the weight $x^\alpha e^{-x}\,dx$ on $[0,\infty)$.

Lemma 2.1 *For every non-negative integer k, $\alpha > -1$, and $y \in \mathbb{R}$, we have*

$$\frac{(-1)^k}{k!}\frac{d^k}{du^k}\left(u^{-\alpha-1}e^{-y/u}\right) = u^{-\alpha-1-k}e^{-y/u}L_k^\alpha(y/u).$$

Proof This is easily proved by induction, using standard properties of Laguerre polynomials (see Section 6.2 of [1], or Sections 4.17–4.24 of [10]). □

Suppose $\Lambda \subset \mathbb{R}^n$ is a lattice, and define a measure μ on $[0,\infty)$ consisting of a point mass at x for each vector in Λ of norm x, where the norm of v is $\langle v, v\rangle$. The purpose of μ is to allow us to sum over all lattice vectors without having to index the sum in our notation; instead, we simply integrate with respect to μ. Although μ depends on Λ, for simplicity our notation does not make that dependence explicit.

The key positivity property of μ is the following lemma:

Lemma 2.2 *For all $y > 0$ and all non-negative integers k,*

$$\int_0^\infty L_k^{n/2-1}(xy)e^{-xy}\,d\mu(x) \ge 0.$$

Proof The theta series of Λ is given by

$$\Theta_\Lambda(z) = \int_0^\infty e^{i\pi xz}\,d\mu(x),$$

and it follows from the Poisson summation formula that the theta series of the dual lattice Λ^* is given by

$$\Theta_{\Lambda^*}(z) = \operatorname{vol}(\mathbb{R}^n/\Lambda)\left(\frac{i}{z}\right)^{n/2}\Theta_\Lambda\left(-\frac{1}{z}\right).$$

(See equation (19) in [5, page 103].)

It will be more convenient for us to work with the variable y given by $y = -i\pi z$. Let $T(y) = \Theta_\Lambda(z)$, so that

$$T(y) = \int_0^\infty e^{-xy}\,d\mu(x).$$

Then up to a positive factor, the theta series of Λ^* is given by $y^{-n/2}T(\pi^2/y)$.

We know that $y^{-n/2}T(\pi^2/y)$ is a positive linear combination of functions e^{-cy} with $c \ge 0$, because it is the theta series of a lattice (times a positive constant). Hence, its successive derivatives with respect to y alternate in sign. We have

$$y^{-n/2}T(\pi^2/y) = \int_0^\infty y^{-n/2}e^{-\pi^2 x/y}\,d\mu(x),$$

from which it follows using Lemma 2.1 that

$$\frac{(-1)^k}{k!}\frac{d^k}{dy^k}\left(y^{-n/2}T(\pi^2/y)\right) = \int_0^\infty y^{-n/2-k}e^{-\pi^2 x/y}L_k^{n/2-1}(\pi^2 x/y)\,d\mu(x).$$

(Differentiating under the integral sign, which really denotes a sum, is justified by uniform convergence of the differentiated sum; see Theorem 7.17 of [13].) Now the change of variable $y \leftrightarrow \pi^2/y$ shows us that

$$\int_0^\infty L_k^{n/2-1}(xy)e^{-xy}\,d\mu(x) \ge 0,$$

as desired. □

When we use only the fact that the derivatives of $y^{-n/2}T(\pi^2/y)$ alternate in sign, we do not lose much information—by a theorem of Bernstein (see Section 12 of Chapter IV of [22]), this property characterizes functions of the form

$$\int_0^\infty e^{-xy}\,d\mu(x)$$

for some measure μ on $[0,\infty)$. Also, it is not surprising that the inequalities in Lemma 2.2 occur for all scalings y, because so far our setup is scale-invariant.

If the shortest non-zero vectors in Λ have length 1 (that is, Λ leads to a packing with balls of radius $1/2$), then the center density of the lattice packing given by Λ equals

$$(4\pi)^{-n/2}\lim_{y\to 0+} y^{n/2}T(y).$$

The proof is as follows. The relationship between the theta series of Λ^* and Λ is

$$\frac{T_{\Lambda^*}(y)}{2^n \operatorname{vol}(\mathbb{R}^n/\Lambda)} = (4\pi)^{-n/2}\left(\frac{\pi^2}{y}\right)^{n/2} T_\Lambda\left(\frac{\pi^2}{y}\right).$$

As we let $y \to \infty$, the right hand side becomes the limit above, and the left hand side tends to $1/(2^n \operatorname{vol}(\mathbb{R}^n/\Lambda))$, which is the center density.

Using Lemma 2.2, we can bound the center density. First, we need a definition and a lemma.

Definition 2.3 A function $f\colon [0,\infty) \to \mathbb{R}$ has the α–SILP property ("scale-invariant Laguerre positivity") if the following conditions hold:

(1) f is continuous and for some $\varepsilon > 0$ and $C > 0$, we have
$$|f(x)| \le C(1+|x|)^{-\alpha-1-\varepsilon}$$
for all x, and

(2) for every $y > 0$, the Laguerre series
$$\sum_{j\ge 0} a_j(y) L_j^\alpha(x),$$
for $x \mapsto f(x/y)$ has $a_j(y) \ge 0$ for all j.

Condition (1) is merely a technical restriction; condition (2) is the heart of the matter. Notice that the orthogonality of the Laguerre polynomials implies that
$$a_j(y) = \frac{\int_0^\infty f(x/y) L_j^\alpha(x) x^\alpha e^{-x}\,dx}{\int_0^\infty L_j^\alpha(x)^2 x^\alpha e^{-x}\,dx} = \frac{\int_0^\infty f(x/y) L_j^\alpha(x) x^\alpha e^{-x}\,dx}{\Gamma(j+\alpha+1)/j!}.$$
We make no assumption about convergence for the Laguerre series in Definition 2.3. However, the following analogue of Fejér's theorem on Fourier series holds. It is a simple consequence of results in [20]. We could also make use of [16] to prove a marginally weaker result (which would still suffice for our purposes).

Lemma 2.4 *Let $\alpha \ge 0$, and let $f\colon [0,\infty) \to \mathbb{R}$ be an α–SILP function. Then for all $k > \alpha + 1/2$, the (C,k) Cesàro means*
$$\binom{k+m}{m}^{-1} \sum_{j=0}^{m} \binom{k+m-j}{m-j} a_j(y) L_j^\alpha(x) e^{-x/2}$$
of the partial sums of the series
$$\sum_{j\ge 0} a_j(y) L_j^\alpha(x) e^{-x/2}$$
converge uniformly to $f(x/y)e^{-x/2}$ on $[0,\infty)$, as $m \to \infty$. (Here, $a_j(y)$ is as above.)

Proof We take $y = 1$ for notational simplicity; of course, the same proof holds for each $y > 0$. For a function $g\colon [0,\infty) \to \mathbb{R}$, let $\tilde{g}(x) = g(x)e^{-x/2}$, and let $\sigma_m g(x)$ denote the Cesàro mean
$$\sigma_m g(x) = \binom{k+m}{m}^{-1} \sum_{j=0}^{m} \binom{k+m-j}{m-j} b_j L_j^\alpha(x) e^{-x/2},$$

where g has Laguerre coefficients b_j. Theorem 6.2.1 of [20] says that there exists a constant C such that for all m and all g such that $\tilde{g} \in L^\infty([0,\infty), x^\alpha\, dx)$,

$$||\sigma_m g||_\infty \leq C||\tilde{g}||_\infty,$$

where $||\cdot||_\infty$ denotes the norm on $L^\infty([0,\infty), x^\alpha\, dx)$.

We can then imitate the proof of Theorem 2 in [12]. Let $\varepsilon > 0$. By Theorem 18 of [17], $\tilde{f}$ can be uniformly approximated on $[0,\infty)$ by $\tilde{g}$ with g a polynomial. Choose g so that

$$\left|\left|\tilde{f} - \tilde{g}\right|\right|_\infty < \frac{\varepsilon}{2+2C}.$$

Then

$$||\sigma_m f - \sigma_m g||_\infty < \frac{C\varepsilon}{2+2C}.$$

For sufficiently large m, we have

$$||\sigma_m g - \tilde{g}||_\infty < \frac{\varepsilon}{2},$$

since g is a polynomial. It follows that

$$\left|\left|\sigma_m f - \tilde{f}\right|\right|_\infty < \varepsilon.$$

Thus, $\sigma_m f$ tends uniformly to f as $m \to \infty$. □

Of course, this proof made no use of the positivity of the Laguerre coefficients, and in fact could be carried out with far weaker constraints on the behavior of f at infinity. We stated it in terms of α–SILP functions only because those are the functions to which we will apply it. The requirement that α be non-negative is part of the hypotheses of Theorem 6.2.1 of [20]. Perhaps one could prove an analogue of Lemma 2.4 for $\alpha < 0$, but in terms of sphere packing that would cover only the one-dimensional case.

Theorem 2.5 *Let $n > 1$. Suppose f has the $(n/2-1)$–SILP property, with $f(0) = 1$ and $f(x) \leq 0$ for $x \geq 1$. Then the center density for n–dimensional lattice packings is bounded above by*

$$\frac{\Gamma(n/2)}{2^n \pi^{n/2} \int_0^\infty f(x) x^{n/2-1}\, dx}.$$

As was pointed out above, the same bound holds for all sphere packings, not just lattice packings. One can prove this more general result by replacing the theta series of a lattice with the averaged theta series of a periodic packing in Lemma 2.2, but for simplicity we restrict our attention to lattices.

Proof Without loss of generality, we can assume that our lattice is scaled so as to have packing radius $1/2$ (that is, every non-zero vector has norm at least 1). Define μ, T, $a_k(y)$, etc. as before.

We have

$$f(0) \geq \int_0^\infty f(x)e^{-xy}\,d\mu(x),$$

since all contributions to the integral from $x > 0$ are non-positive.

Let $k > (n-1)/2$, and

$$\sigma_m f(x) = \binom{k+m}{m}^{-1} \sum_{j=0}^{m} \binom{k+m-j}{m-j} a_j(y) L_j^{n/2-1}(xy)e^{-xy/2}.$$

Then

$$\int_0^\infty \sigma_m f(x)e^{-xy/2}\,d\mu(x) \geq a_0(y)\int_0^\infty e^{-xy}\,d\mu(x) = a_0(y)T(y),$$

since by Lemma 2.2 all the terms in $\sigma_m f(x)$ with $j > 0$ contribute a non-negative amount. Since $\sigma_m f(x)$ converges uniformly to $f(x)e^{-xy/2}$ as $m \to \infty$ by Lemma 2.4 (and because constant functions are integrable with respect to $e^{-xy/2}\,d\mu(x)$), we have

$$\lim_{m\to\infty}\int_0^\infty \sigma_m f(x)e^{-xy/2}\,d\mu(x) = \int_0^\infty f(x)e^{-xy}\,d\mu(x).$$

It follows that

$$\int_0^\infty f(x)e^{-xy}\,d\mu(x) \geq a_0(y)T(y),$$

and hence

$$f(0) \geq a_0(y)T(y).$$

Thus, the center density is bounded above by

$$\lim_{y\to 0+} \frac{y^{n/2}f(0)}{(4\pi)^{n/2}a_0(y)}.$$

We can evaluate that limit, since

$$a_0(y) = \frac{\int_0^\infty f(x/y)x^{n/2-1}e^{-x}\,dx}{\Gamma(n/2)} = \frac{y^{n/2}\int_0^\infty f(u)u^{n/2-1}e^{-yu}\,du}{\Gamma(n/2)},$$

and $\int_0^\infty f(u)u^{n/2-1}e^{-yu}\,du$ converges to $\int_0^\infty f(u)u^{n/2-1}\,du$ as $y \to 0+$, by dominated convergence. Applying this formula leads to the bound in the theorem statement. □

Theorem 2.5 amounts to essentially the same bound as Theorem 1.1, although that is not immediately obvious. The key is Proposition 2.8, which tells us that there is essentially only one α–SILP function for each α, in the sense that every α–SILP function is a positive combination of scalings of this function. First, we need two technical lemmas.

Lemma 2.6 *For $\alpha > -1/2$ and $x \in [0,\infty)$,*

$$\lim_{k\to\infty} k^{-\alpha} L_k^\alpha(x/k) e^{-x/k} = x^{-\alpha/2} J_\alpha(2\sqrt{x}),$$

and convergence is uniform over $[0,\infty)$.

Note that uniform convergence is false for $\alpha = -1/2$, because $k^{-\alpha}L_k^\alpha(x/k)e^{-x/k}$ tends to 0 as $x \to \infty$ but the right side does not. Since we take $\alpha = n/2 - 1$ in dimension n, the only case this rules out is the trivial 1–dimensional case, and that is hardly a problem since it is already ruled out by Theorem 2.5 (via Lemma 2.4).

Proof Pointwise convergence is known (see 10.12 (36) in [7, page 191]), but the statements the author knows of in the literature omit the $e^{-x/k}$ factor that makes the convergence uniform.

We consider two cases. In the first, $x \ge k^{1+\delta}$ with $\delta > 0$ fixed as $k \to \infty$. Then $x^{-\alpha/2}J_\alpha(2\sqrt{x})$ tends uniformly to 0 as $k \to \infty$, and we just need to verify that $k^{-\alpha}L_k^\alpha(x/k)e^{-x/k}$ does as well. For that, we use Theorem 8.91.2 from [19]. It implies that for $a > 0$

$$\max_{x \ge a} \left| e^{-x/2} L_k^\alpha(x) \right| = O(k^C),$$

where $C = \max(-1/3, \alpha/2 - 1/4)$. It follows that $k^{-\alpha}L_k^\alpha(x/k)e^{-x/k}$ tends uniformly to 0 as $k \to \infty$ with $x \ge k^{1+\delta}$.

Thus, we need only deal with the case of $x \le k^{1+\delta}$. We start with (4.19.3) from [10] (which holds for all $\alpha > -1$, not just $\alpha > 1$ as inadvertently stated in [10]), which says that

$$L_k^\alpha(x) = \frac{e^x x^{-\alpha/2}}{k!} \int_0^\infty t^{k+\alpha/2} J_\alpha(2\sqrt{xt}) e^{-t}\, dt.$$

Thus,

$$\begin{aligned} k^{-\alpha}L_k^\alpha(x/k)e^{-x/k} &= \frac{x^{-\alpha/2}k^{k+1}}{k!} \int_0^\infty t^{\alpha/2} J_\alpha(2\sqrt{xt}) e^{k(\log t - t)}\, dt \\ &= (1+o(1))e^k \sqrt{\frac{k}{2\pi}} \int_0^\infty (t/x)^{\alpha/2} J_\alpha(2\sqrt{xt}) e^{k(\log t - t)}\, dt. \end{aligned}$$

The exponent $\log t - t$ is maximized at $t = 1$, so we can use the Laplace method to estimate this integral (see Chapter 4 of [3]). In the following calculations, all constants implicit in big-O terms are independent of x.

Let $\varepsilon > 0$ be small (ε will be a function of k). Our integral nearly equals that over the interval $[1-\varepsilon, 1+\varepsilon]$, since for any $C < 1/2$ we have $\log t - t < -1 - C\varepsilon^2$ outside $[1-\varepsilon, 1+\varepsilon]$ for sufficiently small ε, and hence

$$\left| \int_0^\infty (t/x)^{\alpha/2} J_\alpha(2\sqrt{xt}) e^{k(\log t - t)}\, dt - \int_{1-\varepsilon}^{1+\varepsilon} (t/x)^{\alpha/2} J_\alpha(2\sqrt{xt}) e^{k(\log t - t)}\, dt \right|$$

is bounded by

$$e^{-(k-1)(1+C\varepsilon^2)} \int_0^\infty t^{\alpha/2} \left| \frac{J_\alpha(2\sqrt{xt})}{x^{\alpha/2}} \right| e^{\log t - t}\, dt = O\left(e^{-k(1+C\varepsilon^2)}\right).$$

Thus, we just need to estimate

$$\int_{1-\varepsilon}^{1+\varepsilon} (t/x)^{\alpha/2} J_\alpha(2\sqrt{xt}) e^{k(\log t - t)}\, dt.$$

We would like to approximate it with

$$x^{-\alpha/2} J_\alpha(2\sqrt{x}) \int_{1-\varepsilon}^{1+\varepsilon} e^{k(\log t - t)}\, dt.$$

The difference between these integrals is bounded by a constant times the product of ε, the maximum of the t–derivative of $(t/x)^{\alpha/2} J_\alpha(2\sqrt{xt})$ over $t \in [1-\varepsilon, 1+\varepsilon]$, and

$$\int_{1-\varepsilon}^{1+\varepsilon} e^{k(\log t - t)}\, dt.$$

We have

$$\frac{\partial}{\partial t}(t^{\alpha/2} J_\alpha(2\sqrt{xt})) = \frac{\alpha}{2} t^{\alpha/2-1} J_\alpha(2\sqrt{xt}) + \left(-J_{\alpha+1}(2\sqrt{xt}) + \frac{\alpha J_\alpha(2\sqrt{xt})}{2\sqrt{xt}} \right) \frac{t^{\alpha/2} x}{\sqrt{xt}}.$$

For x near 0, $x^{-\alpha/2} \partial(t^{\alpha/2} J_\alpha(2\sqrt{xt}))/\partial t$ remains bounded; for x away from 0 it is at most $O(x^{1/4-\alpha/2})$, which is at most $O(x^{1/2-\delta})$ if δ is small enough relative to α (which we can assume). Because $x \le k^{1+\delta}$, we have $x^{1/2-\delta} \le k^{1/2-\delta/2}$.

Thus,

$$\int_{1-\varepsilon}^{1+\varepsilon} (t/x)^{\alpha/2} J_\alpha(2\sqrt{xt}) e^{k(\log t - t)}\, dt$$

equals

$$\left(x^{-\alpha/2} J_\alpha(2\sqrt{x}) + O\left(\varepsilon k^{1/2-\delta/2}\right) \right) \int_{1-\varepsilon}^{1+\varepsilon} e^{k(\log t - t)}\, dt.$$

If we expand $\log t - t = -1 - (t-1)^2/2 + O((t-1)^3)$, we find that

$$\int_{1-\varepsilon}^{1+\varepsilon} e^{k(\log t - t)}\, dt = (1+o(1))e^{-k}\sqrt{\frac{2\pi}{k}},$$

as long as $k\varepsilon^2 \to \infty$, so that the interval we are integrating over is much wider than the standard deviation of the Gaussian we are using to approximate the integrand.

So far, we know that as long as $k\varepsilon^2 \to \infty$, we have

$$k^{-\alpha} L_k^\alpha(x/k)e^{-x/k} = (1+o(1))x^{-\alpha/2}J_\alpha(2\sqrt{x}) + O\left(\sqrt{k}e^{-kC\varepsilon^2}\right) + O\left(\varepsilon k^{1/2-\delta/2}\right).$$

Now if we take $\varepsilon = k^{-\beta}$ with $(1-\delta)/2 < \beta < 1/2$, we find that

$$k^{-\alpha} L_k^\alpha(x/k)e^{-x/k} = x^{-\alpha/2}J_\alpha(2\sqrt{x}) + o(1),$$

as desired. □

Lemma 2.7 *For $\alpha > -1/2$, if $f\colon [0,\infty) \to \mathbb{R}$ is continuous and satisfies*

$$|f(x)| \le C(1+|x|)^{-\alpha-1-\varepsilon}$$

for some $C > 0$ and $\varepsilon > 0$, then

$$\sum_{k\ge 0} t^k \int_0^\infty f(x/y)L_k^\alpha(x)x^\alpha e^{-x}\, dx = (1-t)^{-\alpha-1}\int_0^\infty f(x/y)x^\alpha e^{-x/(1-t)}\, dx$$

whenever $|t| < 1/3$.

Proof We would like to convert this sum to

$$\int_0^\infty \sum_{k\ge 0} f(x/y)L_k^\alpha(x)x^\alpha e^{-x}t^k\, dx$$

and apply the generating function

$$\sum_{k\ge 0} L_k^\alpha(x)t^k = (1-t)^{-\alpha-1}e^{-xt/(1-t)}$$

((6.2.4) from [1]). To do so, we must justify interchanging the limit with the sum.

Let

$$g(t) = (1-t)^{-\alpha-1}e^{-xt/(1-t)} = (1-t)^{-\alpha-1}e^{x}e^{-x/(1-t)}.$$

Then the Lagrange form of the remainder in Taylor's theorem implies

$$g(t) = \sum_{k=0}^{m-1} L_k^\alpha(x)t^k + \frac{g^{(m)}(s)}{m!}t^m$$

for some s satisfying $|s| \leq |t|$. By Lemma 2.1,

$$\frac{g^{(m)}(s)}{m!} = \pm e^x (1-s)^{-\alpha-1-m} e^{-x/(1-s)} L_m^\alpha(x/(1-s)).$$

It follows from Lemma 2.6 that

$$\left| e^{-x/(1-s)} L_m^\alpha(x/(1-s)) \right| \leq C' m^\alpha$$

for some constant $C' > 0$ (depending on α). Thus,

$$\left| (1-t)^{-\alpha-1} \int_0^\infty f(x/y) x^\alpha e^{-x/(1-t)}\, dx - \sum_{k=0}^{m-1} t^k \int_0^\infty f(x/y) L_k^\alpha(x) x^\alpha e^{-x}\, dx \right|$$

is bounded above by

$$C' \left(\int_0^\infty f(x/y) x^\alpha \, dx \right) (1-s)^{-\alpha-1} m^\alpha \left(\frac{t}{1-s} \right)^m. \tag{2.1}$$

The integral in (2.1) is finite because of the bound on $|f|$ in the lemma statement. Because $|t| < 1/3$ and $|s| \leq |t|$, we have

$$\left| \frac{t}{1-s} \right| < \frac{1}{2},$$

and hence (2.1) tends to 0 as $m \to \infty$. □

Proposition 2.8 *Let $\alpha > -1/2$, and suppose $f\colon [0,\infty) \to \mathbb{R}$ is continuous, and satisfies $|f(x)| \leq C(1+|x|)^{-\alpha-1-\varepsilon}$ for some $C > 0$ and $\varepsilon > 0$. Then f has the α–SILP property iff*

$$f(x) = \int_0^\infty (xy)^{-\alpha/2} J_\alpha(2\sqrt{xy})\, dg(y)$$

for some weakly increasing function g.

Note that one can compute directly the Laguerre coefficients of the scalings of $x^{-\alpha/2} J_\alpha(2\sqrt{x})$ and verify that they are positive (see Example 3 in Section 4.24 of [10]). Proposition 2.8 tells us that this function is essentially the only α–SILP function.

Proof We know that f has the α–SILP property iff for every $y > 0$,

$$\sum_{k \geq 0} t^k \int_0^\infty f(x/y) L_k^\alpha(x) x^\alpha e^{-x}\, dx$$

has non-negative coefficients as a power series in t. By Lemma 2.7, we can write this function (for small t) as

$$(1-t)^{-\alpha-1}\int_0^\infty f(x/y)x^\alpha e^{-x/(1-t)}\,dx,$$

which is a positive constant (a power of y) times

$$(1-t)^{-\alpha-1}\int_0^\infty f(x)x^\alpha e^{-xy/(1-t)}\,dx.$$

Define $\tilde{f}$ to be the Laplace transform of $x\mapsto x^\alpha f(x)$. Then f has the α–SILP property iff

$$(1-t)^{-\alpha-1}\tilde{f}(y/(1-t))$$

has non-negative coefficients as a power series in t. We can rescale t by a factor of y and pull out a power of y to see that this happens iff

$$(1/y-t)^{-\alpha-1}\tilde{f}(1/(1/y-t))$$

has non-negative coefficients. That happens for all $y>0$ iff the function $u\mapsto u^{-\alpha-1}\tilde{f}(1/u)$ has successive derivatives alternating in sign (the function is non-negative, its derivative non-positive, its second derivative non-negative, etc.). By Bernstein's theorem (Theorem 12b of Chapter IV of [22, page 161]), this holds iff it is the Laplace transform of a positive measure.

Thus, we have shown that f has the α–SILP property iff there is a weakly increasing function g such that for $u>0$,

$$u^{-\alpha-1}\int_0^\infty f(x)x^\alpha e^{-x/u}\,dx=\int_0^\infty e^{-yu}\,dg(y).$$

To finish proving the proposition, we can work as follows. We know that

$$\int_0^\infty f(x)x^\alpha e^{-xu}\,dx=u^{-\alpha-1}\int_0^\infty e^{-y/u}\,dg(y).$$

We can now apply the following general theorem for inverting a Laplace transform: if

$$\phi(u)=\int_0^\infty \psi(x)e^{-xu}\,dx,$$

then

$$\psi(x)=\lim_{k\to\infty}\frac{(-1)^k}{k!}\phi^{(k)}\left(\frac{k}{x}\right)\left(\frac{k}{x}\right)^{k+1}$$

wherever ψ is continuous. (See Corollary 6a.2 of Chapter VII in [22, page 289].)

We can apply this to our equation, and differentiate under the integral sign (justified since the differentiated integrals converge uniformly as u ranges over any compact subset of $(0,\infty)$; see Theorem 14 of Chapter 10 in [23, page 358]). Using Lemma 2.1, it follows that

$$x^\alpha f(x) = \lim_{k\to\infty} \int_0^\infty \left(\frac{k}{x}\right)^{-\alpha} L_k^\alpha\left(\frac{xy}{k}\right) e^{-xy/k}\, dg(y).$$

To finish the proof, we apply Lemma 2.6, but we need to check that passage to the limit under the integral sign is justified. Because of the uniform convergence, it is justified as long as constant functions are integrable with respect to dg. However, that is true, for the following reason. By definition, g satisfies

$$u^{-\alpha-1}\int_0^\infty f(x)x^\alpha e^{-x/u}\, dx = \int_0^\infty e^{-yu}\, dg(y),$$

which is equivalent to

$$\int_0^\infty f(ux)x^\alpha e^{-x}\, dx = \int_0^\infty e^{-yu}\, dg(y).$$

When we let $u \to 0+$, the left side converges to

$$f(0)\int_0^\infty x^\alpha e^{-x}\, dx$$

(by the dominated convergence theorem: recall that f is bounded and continuous), so the right side converges as $u \to 0+$. By monotone convergence, we see that constant functions are integrable with respect to dg, which is what we need. □

Corollary 2.9 *For integers $n>1$, a function $f\colon [0,\infty)\to\mathbb{R}$ has the $(n/2-1)$-SILP property iff the function from $\mathbb{R}^n$ to $\mathbb{R}$ given by $x\mapsto f(|x|^2)$ is continuous, satisfies*

$$|f(|x|^2)| \le C(1+|x|)^{-n-\varepsilon}$$

for some $C>0$ and $\varepsilon>0$, and is the Fourier transform of a non-negative distribution.

Corollary 2.9 follows from combining Proposition 2.8 with Theorem 9.10.3 of [1] (see Proposition 2.1 of [4]), after some changes of variables. Using Corollary 2.9, one can check with some simple manipulations that for $n>1$, Theorem 2.5 implies Theorem 1.1 for lattice packings (and, as pointed out above, the general case can be proved similarly). It is seemingly more general, because it does not constrain the Fourier transform at infinity. However, the additional generality does not seem useful, and one could likely generalize the proof in [4] to use a version of Poisson summation with fewer hypotheses (for example, see Theorem D.4.1 in [1]).

Corollary 2.10 *For $\alpha > -1/2$, the product of two α–SILP functions is always an α–SILP function.*

Corollary 2.10 follows immediately from Corollary 2.9 when $\alpha = n/2 - 1$ with $n \in \mathbb{Z}$, and can be proved for arbitrary α using Proposition 2.8 together with 13.46 (3) of [21] or (7) from Section 3 of [18]. It seems surprisingly difficult to prove directly from the definition of a SILP function: it would follow trivially if the product of two Laguerre polynomials were a positive combination of Laguerre polynomials, but that is not the case. In fact, the coefficients of such a product alternate in sign; that is, the polynomials $(-1)^k L_k^\alpha$ have the property that the set of positive combinations of them is closed under multiplication.

3 Optimality of Bessel functions

Let j_ν denote the first positive root of J_ν. According to Proposition 6.1 of [4], the function $f\colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \frac{J_{n/2}(j_{n/2}|x|)^2}{(1-|x|^2)|x|^n} \tag{3.1}$$

satisfies the hypotheses of Theorem 1.1, and leads to the upper bound

$$\frac{j_{n/2}^n}{(n/2)!^2 4^n}$$

for the densities of n–dimensional sphere packings. The Fourier transform $\widehat{f}$ has support in the ball of radius $j_{n/2}/\pi$ about the origin. We will show that among all such functions, f proves the best sphere packing bound. This is analogous to a theorem of Sidel'nikov [15] for the case of error-correcting codes and spherical codes. It was first proved in the setting of sphere packings by Gorbachev [8]. Our proof will be based on the same identity as Gorbachev's, but the proof of the identity appears to be new.

For notational simplicity, we view f and $\widehat{f}$ as functions on $[0,\infty)$; that is, $f(r)$ will denote the common value of f on all vectors of length r. Let $\nu = n/2-1$, and let $\lambda_1 < \lambda_2 < \cdots$ be the positive roots of $J_{\nu+1}(x)$ (equivalently, the positive roots of $-\nu J_\nu(x) + xJ'_\nu(x)$; see equation (4) in Section 3.2 of [21]). Define $B_r(x)$ to be the closed ball of radius r about x.

Our main technical tool is the following identity due to Ben Ghanem and Frappier (the $p=0$ case of Lemma 4 in [2]), who state it with weaker technical hypotheses and a different proof.

Theorem 3.1 (Ben Ghanem and Frappier [2]) *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a radial Schwartz function. If* $\operatorname{supp}(\widehat{f}\,) \subseteq B_r(0)$, *then*

$$\widehat{f}(0) = \frac{(n/2)!2^n}{\pi^{n/2}r^n} f(0) + \sum_{m=1}^{\infty} \frac{4\lambda_m^{n-2}}{(n/2-1)!\pi^{n/2}r^n J_{n/2-1}(\lambda_m)^2} f\left(\frac{\lambda_m}{\pi r}\right).$$

We will postpone the proof of Theorem 3.1 until we have developed several lemmas. First, however, we deduce the desired optimality:

Corollary 3.2 (Gorbachev [8]) *Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is a radial, admissible function, is not identically zero, and satisfies the following three conditions:*

(1) $f(x) \le 0$ *for* $|x| \ge 1$,

(2) $\widehat{f}(t) \ge 0$ *for all* t, *and*

(3) $\operatorname{supp}(\widehat{f}\,) \subseteq B_{j_{n/2}/\pi}(0)$.

Then

$$\frac{\pi^{n/2}}{(n/2)!2^n} \cdot \frac{f(0)}{\widehat{f}(0)} \ge \frac{j_{n/2}^n}{(n/2)!^2 4^n}.$$

Proof of Corollary 3.2 Let $r = j_{n/2}/\pi$. If f were a Schwartz function, then Theorem 3.1 would imply that

$$\widehat{f}(0) \le \frac{(n/2)!2^n}{\pi^{n/2}(j_{n/2}/\pi)^n} f(0),$$

since $\lambda_m/(\pi r) \ge 1$ for $m \ge 1$. For more general functions f, the series

$$\frac{(n/2)!2^n}{\pi^{n/2}r^n} f(0) + \sum_{m=1}^{\infty} \frac{4\lambda_m^{n-2}}{(n/2-1)!\pi^{n/2}r^n J_{n/2-1}(\lambda_m)^2} f\left(\frac{\lambda_m}{\pi r}\right)$$

at least still converges, since the terms are $O(m^{-1-\varepsilon})$ for some $\varepsilon > 0$ (namely, the ε from the definition of admissibility); to verify this, note that λ_m grows linearly with m, and that $J_\nu(z)^2 + J_{\nu+1}(z)^2 \sim 2/(\pi z)$ (see Section 7.21 of [21, page 200]), so $J_\nu(\lambda_m)^2 \sim 2/(\pi\lambda_m)$. However, we must verify that it converges to $\widehat{f}(0)$.

We need to smooth $\widehat{f}$ without increasing its support. Let i_δ denote any nonnegative, smooth function of integral 1 with support in the ball of radius δ about the origin. Let $f_\varepsilon(x) = f(x(1-\varepsilon))\widehat{\imath}_{r\varepsilon/2}(x)$, where $r = j_{n/2}/\pi$. This is a Schwartz function whose Fourier transform has support in the ball of radius $r(1-\varepsilon/2)$, so Theorem 3.1 applies to f_ε. As $\varepsilon \to 0+$, the functions f_ε and $\widehat{f}_\varepsilon$ converge pointwise to f and $\widehat{f}$, respectively. Since $|\widehat{\imath}_{r\varepsilon/2}| \le 1$ everywhere, dominated convergence lets us interchange the limit as $\varepsilon \to 0+$ with the sum over m to conclude that

$$\widehat{f}(0) = \frac{(n/2)!2^n}{\pi^{n/2}r^n} f(0) + \sum_{m=1}^{\infty} \frac{4\lambda_m^{n-2}}{(n/2-1)!\pi^{n/2}r^n J_{n/2-1}(\lambda_m)^2} f\left(\frac{\lambda_m}{\pi r}\right),$$

and we finish the proof as before. □

Lemma 3.3 *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a radial Schwartz function. If* $\operatorname{supp}(\widehat{f}) \subseteq B_r(0)$*, then for $u \in [0,1)$,*

$$2\pi \widehat{f}(ru)r^{\nu+2} = \frac{2\Gamma(\nu+2)}{(\pi r)^\nu} f(0) + \sum_{m=1}^{\infty} \frac{2(\lambda_m/(2\pi r))^\nu f(\lambda_m/(2\pi r))}{J_\nu(\lambda_m)^2} \frac{J_\nu(\lambda_m u)}{u^\nu}.$$

The same holds even if $\widehat{f}$ is not smooth at radius r (but is left continuous at radius r, and still smooth at all smaller radii), as long as the values of f in the sum decrease faster than any power of $1/m$ as $m \to \infty$.

Note that if f is a Schwartz function, then the condition on the decay of the values of f automatically holds.

Proof Because $\operatorname{supp}(\widehat{f}) \subseteq B_r(0)$, we have

$$x^\nu f(x) = \int_0^1 g(u)u^{\nu+1} J_\nu(2\pi rux)\, du,$$

where $g(u) = 2\pi \widehat{f}(ru)r^{\nu+2}$ (see Theorem 9.10.3 of [1], or Proposition 2.1 of [4]). We begin by expanding $g(u)u^\nu$ into a Dini series. For a quick introduction to Dini series, see [10, page 130]. Unfortunately, for a technical reason that reference does not cover the case we need here (see footnote 33 on page 130). For a more thorough reference, which covers everything we need, see Sections 18.3–18.35 of [21]. In Watson's notation, we are dealing with the case $H+\nu = 0$ (see page 597 of [21]). Convergence of the Dini series to $g(u)u^\nu$ for $u \in (0,1)$ follows from standard results (see pages 601–602 of [21]), and at $u = 0$ it follows from continuity of g at 0 and uniform convergence of the Dini series (which itself follows from the decay of $f(\lambda_m/(\pi r))$).

The Dini series expansion of $g(u)u^\nu$ is

$$g(u)u^\nu = 2(\nu+1)u^\nu \int_0^1 t^{\nu+1} g(t) t^\nu \, dt + \sum_{m=1}^{\infty} b_m J_\nu(\lambda_m u),$$

where

$$\begin{aligned} b_m &= \frac{2\lambda_m^2}{(\lambda_m^2-\nu^2)J_\nu(\lambda_m)^2+\lambda_m^2 J_\nu'(\lambda_m)^2}\int_0^1 tg(t)t^\nu J_\nu(\lambda_m t)\,dt \\ &= \frac{2\lambda_m^2(\lambda_m/(2\pi r))^\nu}{(\lambda_m^2-\nu^2)J_\nu(\lambda_m)^2+\lambda_m^2 J_\nu'(\lambda_m)^2} f(\lambda_m/(2\pi r)). \end{aligned}$$

Note also that

$$\lim_{x\to 0}\int_0^1 g(u)u^{\nu+1}J_\nu(2\pi rux)/x^\nu\,du = \int_0^1 g(u)u^{\nu+1}\frac{(\pi ru)^\nu}{\Gamma(\nu+1)}\,du,$$

since as $x \to 0$,

$$\frac{J_\nu(x)}{x^\nu} \to \frac{1}{2^\nu\Gamma(\nu+1)},$$

so

$$\int_0^1 t^{\nu+1}g(t)t^\nu\,dt = f(0)\Gamma(\nu+1)/(\pi r)^\nu.$$

Furthermore, $\lambda_m J_\nu'(\lambda_m) = \nu J_\nu(\lambda_m)$, so

$$(\lambda_m^2-\nu^2)J_\nu(\lambda_m)^2+\lambda_m^2 J_\nu'(\lambda_m)^2 = \lambda_m^2 J_\nu(\lambda_m)^2.$$

Thus,

$$g(u) = \frac{2\Gamma(\nu+2)}{(\pi r)^\nu}f(0) + \sum_{m=1}^\infty \frac{2(\lambda_m/(2\pi r))^\nu f(\lambda_m/(2\pi r))}{J_\nu(\lambda_m)^2}\frac{J_\nu(\lambda_m u)}{u^\nu},$$

as desired. □

Lemma 3.4 *Let f be a function from $[0,\infty)$ to $\mathbb{R}$. The function $x \mapsto f(|x|)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is the Fourier transform of a compactly support distribution iff f extends to an even, entire function on $\mathbb{C}$ that satisfies*

$$|f(z)| \le C(1+|z|)^k e^{C'|\operatorname{Im} z|}$$

for some C, C', and k.

Proof This lemma is essentially a special case of the Paley-Wiener-Schwartz theorem (Theorem 7.3.1 in [9]). The only difference is that the general theorem is not restricted to radial functions, and characterizes Fourier transforms of compactly supported distributions as entire functions g of n complex variables satisfying

$$|g(z_1,\dots,z_n)| \le C\left(1+\sqrt{|z_1|^2+\dots+|z_n|^2}\right)^k e^{C'\sqrt{(\operatorname{Im} z_1)^2+\dots+(\operatorname{Im} z_n)^2}}. \quad (3.2)$$

The only subtlety in deriving the lemma from the general theorem is in showing that if f satisfies the hypotheses above, then the function g defined by

$$g(z_1,\dots,z_n) = f\left(\sqrt{z_1^2+\cdots+z_n^2}\right)$$

satisfies (3.2). To do that, the elementary inequality

$$\left|\operatorname{Im}\sqrt{z_1^2+\cdots+z_n^2}\right| \le \sqrt{(\operatorname{Im} z_1)^2+\cdots+(\operatorname{Im} z_n)^2}$$

can be used. To prove that inequality, one can use induction to reduce to the $n=2$ case, and prove that case by direct manipulation of both sides. □

Now we are ready to prove Theorem 3.1. Notice that it says that to determine the integral of f, we need only half as many values as we need to reconstruct the whole function via Lemma 3.3. This phenomenon is analogous to Gauss-Jacobi quadrature (see Theorem 14.2.1 of [6]). The proof given below is in fact modeled after the proof of Gauss-Jacobi quadrature, although carrying it out rigorously is more involved.

Proof of Theorem 3.1 Let $\varepsilon > 0$, and define $\tilde h\colon [-1,1] \to \mathbb{R}$ by

$$\tilde h(u) = \frac{2\Gamma(\nu+2)}{(\pi(r/2+\varepsilon))^\nu} f(0) + \sum_{m=1}^{\infty} \frac{2\left(\frac{\lambda_m}{2\pi(r/2+\varepsilon)}\right)^\nu f\left(\frac{\lambda_m}{2\pi(r/2+\varepsilon)}\right)}{J_\nu(\lambda_m)^2} \frac{J_\nu(\lambda_m u)}{u^\nu}.$$

(The functions $J_\nu(\lambda_m u)/u^\nu$ are even, so this is no different from defining $\tilde h$ on $[0,1]$.) Since f is a Schwartz function, the values of f in the series above decrease quickly enough that it defines a C^∞ function on $(-1,1)$. Define $\widehat h$ by

$$2\pi \widehat h((r/2+\varepsilon)u)(r/2+\varepsilon)^{\nu+2} = \begin{cases} \tilde h(u) & \text{if } |u| \le 1, \text{ and} \\ 0 & \text{otherwise,} \end{cases}$$

and define h to be the Fourier transform of $\widehat h$. Then $\operatorname{supp}(\widehat h) \subseteq B_{r/2+\varepsilon}(0)$. By Lemma 3.3, combined with uniqueness for Dini series (which follows from orthogonality), we have

$$h\left(\frac{\lambda_m}{2\pi(r/2+\varepsilon)}\right) = f\left(\frac{\lambda_m}{2\pi(r/2+\varepsilon)}\right)$$

for all m, and $h(0) = f(0)$. (Note that $\widehat h$ may not be smooth at radius $r/2+\varepsilon$, but that does not violate the hypotheses of Lemma 3.3.)

Now let χ_R denote the characteristic function of a ball of radius R about the origin, so that

$$\widehat{\chi}_R(x) = J_{n/2}(2\pi R|x|)(R/|x|)^{n/2}.$$

The entire function $f-h$ has roots wherever $\widehat{\chi}_{r/2+\varepsilon}$ does, and $\widehat{\chi}_{r/2+\varepsilon}$ has only single roots, so the quotient $g = (f-h)/\widehat{\chi}_{r/2+\varepsilon}$ is entire.

We would like to conclude that g is the Fourier transform of a compactly supported distribution. By Lemma 3.4, this requires bounds for g, and it is not obvious that dividing by a Bessel function does not ruin the bounds. We prove this in two steps. First, Lemma 1 of [11] implies (after rescaling variables) that

$$|J_{n/2}(z)/z^{n/2}| \geq \frac{c_1 e^{c_2|\operatorname{Im} z|}}{(1+|z|)^{c_3}}$$

whenever $|\operatorname{Im} z| \geq c_4$, for some constants c_1, c_2, c_3, c_4, with $c_1 > 0$ of course. That means that dividing by it does not mess up our bounds when the absolute value of the imaginary part is at least c_4. The second step is to deal with points near the real axis. Consider a box with sides on the lines with imaginary part $\pm c_4$ and real part $\pm(k\pi + (\pi n+1)/4)$, where k is a positive integer. By the maximum principle, the maximum of g over the interior of the box must occur on the sides. We know that g satisfies the bound we want on the top and bottom, and g is even, so we only need to estimate g on the right side.

For z in the right half-plane, we have

$$\begin{aligned} J_{n/2}(z) = \sqrt{\frac{2}{\pi z}} \Bigg(&\cos\left(z - \frac{\pi n+1}{4}\right)(1 + O(1/z^2)) \\ &+ \sin\left(z - \frac{\pi n+1}{4}\right)(O(1/z))\Bigg) \end{aligned}$$

(see (1) in Section 7.21 of [21]). When z has real part $k\pi$, we have $\cos(z) = (-1)^k \cosh(\operatorname{Im} z)$, which has absolute value at least 1. Thus, on the right side of the box, the cosine factor is always at least 1. The sine factor is bounded, because $\operatorname{Im} z$ is bounded, so we see that on the right side of the box $J_{n/2}(z)/z^{n/2}$ is never smaller than a power of $1/|z|$.

When we combine these estimates, it follows from Lemma 3.4 that g is the Fourier transform of a distribution with compact support. Furthermore, the Titchmarsh-Lions theorem (see Theorem 4.3.3 in [9]) implies that the convex hull of the support of $f-h$ equals the Minkowski sum of those of $\widehat{g}$ and $\chi_{r/2+\varepsilon}$, so $\operatorname{supp}(\widehat{g}) \subseteq B_{r/2-\varepsilon}(0)$.

Let i_δ denote any non-negative, smooth function of integral 1 with support in the ball of radius δ about the origin. We have

$$f\widehat{\imath}_\delta - h\widehat{\imath}_\delta = (\widehat{\chi}_{r/2+\varepsilon}\widehat{\imath}_\delta)g.$$

Now both sides are integrable functions (note that this is not obviously true of either h or $g\widehat{\chi}_{r/2+\varepsilon}$, which is why we had to multiply by $\widehat{\imath}_\delta$), and we find that

$$(\widehat{f} * i_\delta)(0) - (\widehat{h} * i_\delta)(0) = \int (\chi_{r/2+\varepsilon} * i_\delta)\widehat{g}.$$

Because $\mathrm{supp}(\widehat{g}) \subseteq B_{r/2-\varepsilon}(0)$, if we take $\delta < \varepsilon$ we have

$$\int (\chi_{r/2+\varepsilon} * i_\delta)\widehat{g} = \int \widehat{g} = g(0) = 0,$$

where $g(0) = 0$ because $f(0) = h(0)$. Thus.

$$(\widehat{f} * i_\delta)(0) = (\widehat{h} * i_\delta)(0).$$

If we let $\delta \to 0+$, we find that $\widehat{f}(0) = \widehat{h}(0)$, because both $\widehat{f}$ and $\widehat{h}$ are continuous near 0. It follows from the way $\widehat{h}$ was defined that $\widehat{f}(0)$ equals

$$\frac{(n/2)!2^n}{\pi^{n/2}(r+2\varepsilon)^n} f(0) + \sum_{m=1}^{\infty} \frac{4\lambda_m^{n-2}}{(n/2-1)!\pi^{n/2}(r+2\varepsilon)^n J_{n/2-1}(\lambda_m)^2} f\left(\frac{\lambda_m}{\pi(r+2\varepsilon)}\right).$$

Now sending $\varepsilon \to 0+$ proves the desired result, by dominated convergence. □

4 The dual program

It is natural to view choosing the optimal function f in Theorem 1.1 as solving an infinite-dimensional linear programming problem: if we fix $\widehat{f}(0) = 1$, then we are trying to minimize the linear functional $f(0)$ of f, subject to linear inequalities on f. The technicalities are slightly subtle; for example, it is not immediately clear what the right space of functions to consider is (admissibility might be too ad hoc). It seems likely that Schwartz functions suffice. One can come arbitrarily close to the optimum with functions f such that f and $\widehat{f}$ are smooth and rapidly decreasing, where we say $g\colon \mathbb{R}^n \to \mathbb{R}$ is rapidly decreasing if $g(x) = O((1+|x|)^{-k})$ for every $k > 0$: given any f that satisfies the hypotheses of Theorem 1.1, let

$$f_\varepsilon(x) = ((f * i_\varepsilon * i_\varepsilon)\widehat{\imath}_\varepsilon^{\,2})((1+2\varepsilon)x),$$

where i_ε is any non-negative, smooth function of integral 1 with support in $B_\varepsilon(0)$. Then f_ε has the desired properties, still obeys the required inequalities, and satisfies

$$\lim_{\varepsilon\to 0}\frac{f_\varepsilon(0)}{\widehat{f}_\varepsilon(0)}=\frac{f(0)}{\widehat{f}(0)}.$$

Presumably Schwartz functions also come arbitrarily close, but one would have to worry about making the derivatives rapidly decreasing as well. Despite the fact that rapidly decreasing functions come close to the optimal bounds, it is not clear whether they reach them. For example, even for $n=1$, where one can write down several explicit functions that solve the sphere packing problem (see Sections 3 and 5 of [4]), these functions are not rapidly decreasing.

In this context, it is natural to study the dual linear program, to prove bounds on how good the sphere packing bounds produced by Theorem 1.1 can be. The results of Section 3 amount to doing exactly this, for a restricted linear program in which we limit the support of $\widehat{f}$. Unfortunately, in the unrestricted case the dual program seems no easier to solve in general than the primal program is. However, it leads to several intriguing open problems.

One formulation of the dual program is as follows: find the largest c such that there is a tempered distribution g on $\mathbb{R}^n$ satisfying

(1) $g=\delta+h$ with $h\ge 0$,

(2) $\operatorname{supp}(h)\subseteq\{x:|x|\ge 1\}$, and

(3) $\widehat{g}\ge c\delta$.

Here δ is a delta function at the origin, and inequalities between distributions mean that applying both sides to non-negative functions preserves this inequality. For g satisfying (1)–(3) above, and any radial function f satisfying the hypotheses of Theorem 1.1 such that f and $\widehat{f}$ are rapidly decreasing, we have

$$f(0)\ge\int_{\mathbb{R}^n} fg=\int_{\mathbb{R}^n}\widehat{f}\,\widehat{g}\ge c\widehat{f}(0).$$

Here, we use the fact that one can apply a non-negative tempered distribution to any rapidly decreasing function, because non-negative tempered distributions are exactly measures μ such that

$$\int_{\mathbb{R}^n}\frac{d\mu(x)}{(1+|x|)^k}<\infty$$

for some k (see Theorem VII in Chapter 7, Section 4 of [14, page 242]). Thus $f(0)/\widehat{f}(0)\ge c$. The duality theorem of linear programming suggests that there is no gap between the smallest $f(0)/\widehat{f}(0)$ and largest c, but it is not clear how to prove it in this infinite-dimensional setting.

Given any lattice Λ with minimum non-zero vector length 1, summing over Λ defines a tempered distribution that clearly satisfies properties (1) and (2), and Poisson summation implies that it has property (3) as well. As is the case for the functions f, we can rotationally symmetrize g, so that g and $\widehat{g}$ are positive linear combinations of spherical delta functions, where we define a spherical delta function δ_r on $\mathbb{R}^n$ to be a distribution with support on the sphere of radius r about the origin, such that integrating any function times δ_r gives the average of that function over the sphere. One would expect that the optimal radial g should always be a linear combination of spherical delta functions, but it is not clear how to prove it. Aside from the origin, g and $\widehat{g}$ should be supported on the zeros of the optimal f and $\widehat{f}$, respectively, but why must these zeros even occur at a discrete set of radii?

Open Question 4.1 Consider tempered distributions g such that g and $\widehat{g}$ are linear combinations of spherical delta functions. Is every such distribution in the span of the rotationally symmetrized Poisson summation distributions?

It seems very unlikely that the answer to Question 4.1 is yes. Any counterexample would be of interest, since the optimal distributions g in most dimensions (not 1, 2, 8, or 24) are probably counterexamples.

One interesting case is 72 dimensions. It is an open question whether there exists an "extremal lattice of Type II" in $\mathbb{R}^{72}$, in other words, an even unimodular lattice in $\mathbb{R}^{72}$ with minimal non-zero norm at least 8 (see [5, page 194] for more details). Such a lattice might be as extraordinary as E_8 or the Leech lattice. Unfortunately, it seems unlikely that one exists. However, its existence cannot be ruled out by Theorem 1.1. The simplest way to see that is in light of Section 2. A proof that the lattice did not exist would amount to a proof that its theta series could not exist. However, although the extremal lattice may not exist, there is a modular form that would be its theta series if it did exist (see [5, page 195]). In fact, the modular form comes from a distribution g as above, because it is a polynomial in the theta series of E_8 and the Leech lattice, and therefore comes from a g that is the corresponding linear combination of Poisson summation for direct sums of E_8 and the Leech lattice. If Θ_n denotes the theta series of E_8, the Leech lattice, and the hypothetical 72–dimensional lattice for $n = 8, 24, 72$, respectively, then

$$\Theta_{72} = \frac{79}{1080}\Theta_{24}^3 + \frac{1183}{720}\Theta_{24}^2\Theta_8^3 - \frac{91}{180}\Theta_{24}\Theta_8^6 - \frac{91}{432}\Theta_8^9.$$

Despite the minus signs, all the coefficients of Θ_{72} are non-negative.

The most elegant form of the dual program comes from a rescaling analogous to that in Theorem 3.2 of [4]. Define a *relaxed lattice* to be a tempered distribution g such that g and $\widehat{g}$ are of the form

$$\sum_{i \ge 0} a_i \delta_{r_i}$$

with $a_i \ge 0$ for all i (not all 0), and $0 = r_0 < r_1 < r_2 < \cdots$. Call a relaxed lattice g *self-dual* if $\widehat{g} = g$. How large can r_1 be?

Conjecture 4.2 In every dimension, the largest possible value of r_1 in a self-dual relaxed lattice equals the smallest value of r possible in Theorem 3.2 of [4].

One might imagine that the self-duality in Conjecture 4.2 would follow from some sort of symmetry of the linear programming problem, but that is not clear. If this conjecture is true, it would explain the otherwise remarkable fact that the minimal values of r in Proposition 7.1 and Theorem 3.2 of [4] always seem to agree (see Conjecture 7.2 in that paper).

References

[1] **G Andrews**, **R Askey**, **R Roy**, *Special Functions*, Cambridge University Press (1999)

[2] **R Ben Ghanem**, **C Frappier**, *Explicit quadrature formulae for entire functions of exponential type*, J. Approx. Theory 92 (1998) 267–279

[3] **N G de Bruijn**, *Asymptotic Methods in Analysis*, Dover Publications, Inc. (1981)

[4] **H Cohn**, **N Elkies**, *New upper bounds on sphere packings I*, to appear in Annals of Mathematics, `arXiv:math.MG/0110009`

[5] **J H Conway**, **N J A Sloane**, *Sphere Packings, Lattices and Groups*, third edition, Springer-Verlag (1999)

[6] **P J Davis**, *Interpolation and Approximation*, Blaisdell Publishing Company (1963)

[7] **A Erdélyi**, **W Magnus**, **F Oberhettinger**, **F G Tricomi**, *Higher Transcendental Functions, Volume II*, based, in part, on notes left by Harry Bateman, McGraw-Hill (1953)

[8] **D V Gorbachev**, *Extremal problem for entire functions of exponential spherical type, connected with the Levenshtein bound on the sphere packing density in* $\mathbb{R}^n$ (Russian), Izvestiya of the Tula State University. Ser. Mathematics. Mechanics. Informatics. 6 (2000) 71–78

[9] **L Hörmander**, *The Analysis of Linear Partial Differential Operators I*, second edition, Springer-Verlag (1990)

[10] **N N Lebedev**, *Special Functions and Their Applications*, Dover Publications, Inc. (1972)

[11] **N Levinson**, *On a problem of Pólya*, Amer. J. Math. 58 (1936) 791–798

[12] **E Poiani**, *Mean Cesàro summability of Laguerre and Hermite series*, Trans. Amer. Math. Soc. 173 (1972) 1–31

[13] **W Rudin**, *Principles of Mathematical Analysis*, third edition, McGraw-Hill (1976)

[14] **L Schwartz**, *Théorie des Distributions*, Hermann (1966)

[15] **V M Sidel'nikov**, *Extremal polynomials used in bounds of code volume*, Problems Inform. Transmission 16 (1980) 174–186

[16] **K Stempak**, *Almost everywhere summability of Laguerre series*, Studia Mathematica 100 (1991) 129–147

[17] **M H Stone**, *The generalized Weierstrass approximation theorem*, Math. Mag. 21 (1948) 167–184 and 237–254

[18] **G Szegő**, *Über gewisse Potenzreihen mit lauter positiven Koeffizienten*, Math. Zeitschr. 37 (1933) 674–688

[19] **G Szegő**, *Orthogonal Polynomials*, fourth edition, AMS Colloquium Publications, Volume 23, American Mathematical Society (1975)

[20] **S Thangavelu**, *Lectures on Hermite and Laguerre Expansions*, Princeton University Press, Mathematical Notes, Volume 42 (1993)

[21] **G N Watson**, *A Treatise on the Theory of Bessel Functions*, second edition, Cambridge University Press (1944)

[22] **D V Widder**, *The Laplace Transform*, Princeton University Press (1941)

[23] **D V Widder**, *Advanced Calculus*, second edition, Dover Publications, Inc. (1989)