The Optimal Lattice Quantizer in Three Dimensions[1]

E. S. Barnes
Department of Pure Mathematics
The University of Adelaide
Adelaide, S.A. 5001, Australia

N. J. A. Sloane[2]
AT&T Labs - Research
Murray Hill, New Jersey 07974, USA

Abstract

The body-centered cubic lattice is shown to have the smallest mean squared error of any lattice quantizer in three dimensions, assuming that the input to the quantizer has a uniform distribution.

1. Introduction

Let Λ be a lattice in real three-dimensional Euclidean space $\mathbb{R}^3$. Around each lattice point $\mathbf{t} member \Lambda$ is its *Voronoi* (or nearest neighbor) region $S(\mathbf{t})$, consisting of all points of the space that are at least as close to $\mathbf{t}$ as to any other lattice point ([1], [2], [5], [14]). The Voronoi regions $S(\mathbf{t})$ are all congruent, and have volume $\sqrt{D}$, where D is the determinant of Λ, i.e. the square of the volume of a fundamental cell of Λ. If Λ is used as a quantizer, for quantizing data that is uniformly distributed over a large region of $\mathbb{R}^3$, its average mean squared error per symbol is given by

$$G = G(\Lambda) = \frac{1}{3} \frac{\int_{S(\mathbf{0})} \boldsymbol{\tau} \cdot \boldsymbol{\tau} d\boldsymbol{\tau}}{D^{5/6}} \tag{1}$$

— see [5]-[7], [9], [15]. (This formula ignores the fact that points near the boundary of the input region have irregular Voronoi regions, and so applies to the case when the number of output levels of the quantizer is very large.) $G(\Lambda)$ is a normalized second moment of $S(\mathbf{0})$, the Voronoi region around the origin, the denominator being determined by the condition that $G(\Lambda)$ should be dimensionless.

It was conjectured by Gersho in [9] that the body-centered cubic lattice D_3^* has the smallest value of $G(\Lambda)$ of any three-dimensional lattice. It is the goal of this paper to establish that conjecture. Furthermore we shall see that there is no other lattice for which $G(\Lambda)$ is even a local minimum. Thus our main result is the following.

Theorem 1. *For any three-dimensional lattice* $\Lambda, G(\Lambda) \geq 19/(192\ 2^{1/3}) = 0.0785433\ldots$, *with equality if and only if* Λ *is equivalent to the body-centered cubic lattice* D_3^*. *Furthermore for no other lattice is* $G(\Lambda)$ *a local minimum.*

Three (of the infinitely many) lattices which compete with the body-centered cubic lattice are the face-centered cubic lattice D_3, for which $G = 2^{-11/3} = 0.0787451\ldots$; the lattice $\sqrt{3}A_2 \oplus \sqrt{5}\mathbb{Z}$, where A_2 is the hexagonal lattice in the plane with minimum norm 2, for which $G = 5^{2/3}/36 = 0.0812227\ldots$; and the cubic lattice $\mathbb{Z}^3$, with $G = 1/12 = 0.0833333\ldots$. However, as we shall see,

---

[1] A slightly different version of this paper appeared in *SIAM J. Algebraic Discrete Methods*, **4** (No. 1, March 1983), 30–41.

[2] Current address: AT&T Shannon Labs, 180 Park Avenue, Florham Park, NJ 07932–0971

these three values of G can all be reduced by perturbing the lattices slightly. The Voronoi regions for D_3^*, D_3, $\sqrt{3}A_2 \oplus \sqrt{5}\mathbb{Z}$, and $\mathbb{Z}^3$ are respectively truncated octahedra, rhombic dodecahedra, hexagonal prisms, and cubes (see [5], [9]).

Finally, it is worth pointing out that our results have wider application than to just uniformly distributed data, because (i) Zador (see [8], [9], [15]) has reduced the problem of finding the minimal quantization error for data with any integrable density function to that of solving the uniformly distributed case, and (ii) the so-called *companding* techniques for quantizing (see [9], [10]) handle nonuniform data by first applying a nonlinear transformation, then a uniform quantizer, and finally the inverse transformation.

The proof of Theorem 1 will be given in Sections II and III. In Section II we first recall some properties of lattices in $\mathbb{R}^3$, in particular the fact that a lattice Λ can be represented by a vector $[\rho_{01}, \rho_{02}, \rho_{03}, \rho_{12}, \rho_{13}, \rho_{23}]$ with six nonnegative real components. We then derive a fundamental formula (Theorem 2) which expresses $G(\Lambda)$ in terms of the ρ_{ij}. In Section III we complete the proof by showing that the only local minimum of this expression for $G(\Lambda)$ occurs when the ρ_{ij} are all equal, which is precisely the case when Λ is a body-centered cubic lattice.

2. A formula for $G(\Lambda)$

In this paper all vectors are column vectors, written for example as $\mathbf{t} = (t_1, t_2, t_3)^{tr}$, where tr denotes transpose. As is customary we shall represent lattices by their associated quadratic forms (see [4], [11]). If a lattice Λ in $\mathbb{R}^3$ is spanned by three vectors $\mathbf{t}^{(1)} = (t_1^1, t_2^1, t_3^1)^{tr}, \ldots, \mathbf{t}^{(3)} = (t_1^3, t_2^3, t_3^3)^{tr}$, then $f(\mathbf{x}) = f(x_1, x_2, x_3) = \mathbf{x}^{tr} A \mathbf{x}$ is a quadratic form associated with Λ, where $A = T^{tr} T$ and T has columns $\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \mathbf{t}^{(3)}$. A typical lattice point can be described in three ways, either by its Euclidean coordinates $\mathbf{t} = (t_1, t_2, t_3)^{tr}$, its $\mathbf{x}$-coordinates $\mathbf{x} = (x_1, x_2, x_3)^{tr}$, where x_1, x_2, x_3 are integers satisfying $\mathbf{t} = T\mathbf{x}$, or by its $\mathbf{y}$-coordinates $\mathbf{y} = (y_1, y_2, y_3)^{tr}$, given by $\mathbf{y} = A\mathbf{x}$. The *norm*, or squared distance from the origin, of this point is

$$\mathbf{t} \cdot \mathbf{t} = \mathbf{t}^{tr}\mathbf{t} = \mathbf{x}^{tr} A \mathbf{x} = f(\mathbf{x}) = f^{-1}(\mathbf{y}) \,, \tag{2}$$

where $f^{-1}(\mathbf{x}) = \mathbf{x}^{tr} A^{-1} \mathbf{x}$ is the inverse form of f.

Two lattices Λ and M are equivalent, written $\Lambda \cong M$, if one can be obtained from the other by a rotation and change of scale. Two forms $f(\mathbf{x}) = \mathbf{x}^{tr} A \mathbf{x}$ and $g(\mathbf{x}) = \mathbf{x}^{tr} B \mathbf{x}$ are equivalent if $B = U^{tr} A U$, where U is integral and $\det U = \pm 1$ ([4],[11]). Similarly two lattices are equivalent if they are associated with equivalent forms ([4], [11]).

If Λ is a lattice in $\mathbb{R}^3$, Voronoi (see [1], [11], [13, p. 150]) has shown that Λ has a quadratic form of the shape

$$f(x_1, x_2, x_3) = \rho_{01} x_1^2 + \rho_{02} x_2^2 + \rho_{03} x_3^2 + \rho_{12}(x_1 - x_2)^2 + \rho_{13}(x_1 - x_3)^2 + \rho_{23}(x_2 - x_3)^2$$

associated with it, where the ρ_{ij} are nonnegative. If we define $x_0 = 0, \rho_{ii} = 0$, and $\rho_{ij} = \rho_{ji}$ for $i > j$, this may be written more symmetrically as

$$f(x_1, x_2, x_3) = \frac{1}{2} \sum_{i=0}^{3} \sum_{j=0}^{3} \rho_{ij} (x_i - x_j)^2 \,. \tag{3}$$

Thus Λ is represented by the six nonnegative parameters $[\rho_{01}, \rho_{02}, \rho_{03}, \rho_{12}, \rho_{13}, \rho_{23}]$. In general Λ has 24 such representations, corresponding to the 4! permutations of the subscripts of the ρ_{ij} [1, Lemma 2.1]; multiplying f by a scalar leads to an equivalent lattice. For example, applying the permutation (01), we find that $[\rho_{01}, \rho_{12}, \rho_{13}, \rho_{02}, \rho_{03}, \rho_{23}]$ also represents Λ.

For later reference we mention that the body-centered cubic lattice D_3^* may be represented by the parameters [1,1,1,1,1,1], D_3 by [0,1,1,1,1,0] (more generally with any pair $\rho_{ij} = \rho_{k\ell} = 0$, where i, j and k, ℓ are disjoint subscripts, and all other ρ_{ij} equal), $\mathbb{Z}^3$ by [1,1,1,0,0,0] for example, and $\sqrt{3}A_2 \oplus \sqrt{5}\mathbb{Z}$ by [3,3,5,3,0,0].

Our main result in this section is the following.

Theorem 2. *The average mean squared error of Λ (or the normalized second moment of the Voronoi region $S(\mathbf{0})$) is given by*

$$G(\Lambda) = \frac{D \cdot S_1 + 2S_2 + K}{36 D^{4/3}}\,, \tag{4}$$

where

$$D = \det \Lambda = \sum^{(4)} \rho_{01}\rho_{02}\rho_{03} + \sum^{(3)} \rho_{01}\rho_{23}(\rho_{02} + \rho_{03} + \rho_{12} + \rho_{13})\,,$$

$$\begin{aligned} S_1 &= = \rho_{01} + \rho_{02} + \rho_{03} + \rho_{12} + \rho_{13} + \rho_{23}\,, \\ S_2 &= \rho_{01}\rho_{02}\rho_{13}\rho_{23} + \rho_{01}\rho_{03}\rho_{12}\rho_{23} + \rho_{02}\rho_{03}\rho_{12}\rho_{13}\,, \end{aligned} \tag{5}$$

and

$$K = \sum^{(4)} \rho_{01}\rho_{02}\rho_{03}(\rho_{12} + \rho_{13} + \rho_{23})\,.$$

We are using the standard notation for symmetric functions, so that $\sum^{(4)} \rho_{01}\rho_{02}\rho_{03}$ for example is an abbreviation for $\rho_{01}\rho_{02}\rho_{03} + \rho_{01}\rho_{12}\rho_{13} + \rho_{02}\rho_{12}\rho_{23} + \rho_{03}\rho_{13}\rho_{23}$, and the superscript indicates the number of distinct summands. The formula for $\det \Lambda$ was given in [1], and the quantity K also occurs there. In the proof of Theorem 2 we shall make considerable use of the information about the Voronoi regions of Λ given in [1], and in Section III the proof of Theorem 1 is modeled on the proof of the main result in [1] (although the techniques used are quite different).

Proof of Theorem 2. The matrix A associated with the quadratic form (3) is

$$A = \begin{bmatrix} \rho_{01} + \rho_{12} + \rho_{13} & -\rho_{12} & -\rho_{13} \\ -\rho_{12} & \rho_{02} + \rho_{12} + \rho_{23} & \\ -\rho_{13} & -\rho_{23} & \rho_{03} + \rho_{13} + \rho_{23} \end{bmatrix},$$

and $\det A = \det \Lambda = D$ gives (5). The norm of a vector (see (2)) is best expressed in terms of its $\mathbf{y}$-coordinates, and is given by

$$\begin{aligned} f^{-1}(\mathbf{y}) &= \frac{1}{D}\{(\rho_{12}\rho_{13} + \rho_{12}\rho_{23} + \rho_{13}\rho_{23})y_0^2 + (\rho_{02}\rho_{03} + \rho_{02}\rho_{23} + \rho_{03}\rho_{23})y_1^2 && (6) \\ &\quad +(\rho_{01}\rho_{03} + \rho_{01}\rho_{13} + \rho_{03}\rho_{13})y_2^2 + (\rho_{01}\rho_{02} + \rho_{01}\rho_{12} + \rho_{02}\rho_{12})y_3^2 && (7) \\ &\quad +\rho_{03}\rho_{12}(y_1 + y_2)^2 + \rho_{02}\rho_{13}(y_1 + y_3)^2 + \rho_{01}\rho_{23}(y_2 + y_3)^2\}\,, && (8) \end{aligned}$$

where we have set $y_0 = -y_1 - y_2 - y_3$.

The Voronoi region $S(\mathbf{0})$ is described in [1]. It is a (possibly degenerate) truncated octahedron, with in general 14 faces, given by

$$F_i : 2y_i = \sum_{\ell \neq i}^{(3)} \rho_{i\ell}\,,$$

$$F_{ij} : 2(y_i + y_j) = \sum_{\ell \neq i,j}^{(2)} (\rho_{i\ell} + \rho_{j\ell}) \, ,$$

$$F_{ijk} : 2(y_i + y_j + y_k) = \sum_{\ell \neq i,j,k}^{(1)} (\rho_{i\ell} + \rho_{j\ell} + \rho_{k\ell}) \, ,$$

where all subscripts and summations run from 0 to 3, and the subscripts on F are unordered. There are in general 24 vertices $\mathbf{v}_{ijk}$, where the subscripts are an ordered 3-subset of $\{0, 1, 2, 3\}$. For example the vertex $\mathbf{v}_{123}$ lies at the intersection of the faces F_1, F_{12} and F_{123} (see [1, Eq. (2.4)]). The Voronoi region is sketched in Figure 1.

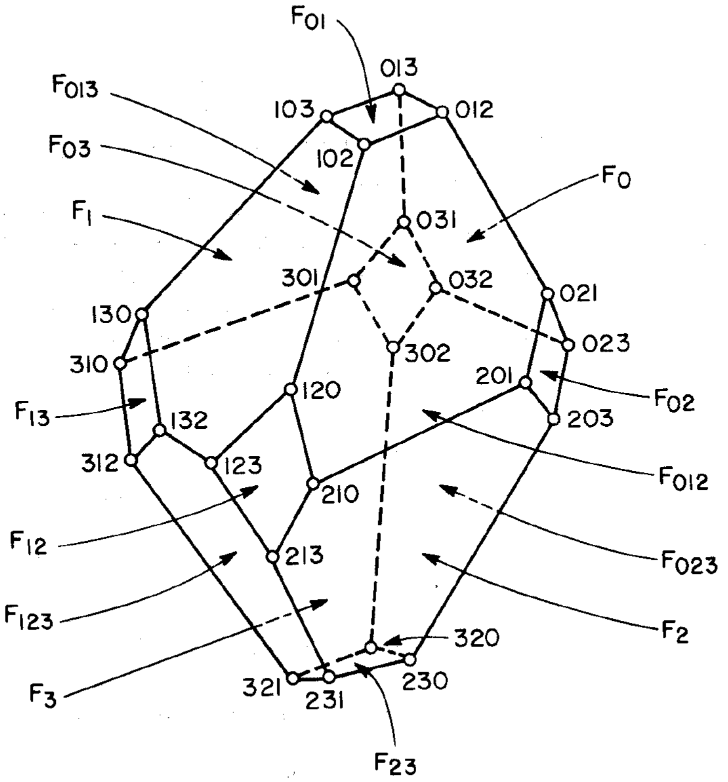

Figure 1: Voronoi region $S(\mathbf{0})$ for a three-dimensional lattice, showing the 14 faces F_i, F_{ij}, F_{ijk} and the 24 vertices $\mathbf{v}_{ijk}$ (only the subscripts are given). The faces $F_i, F_{jk\ell}$ and the faces $F_{ij}, F_{k\ell}$ are parallel, where i, j, k, ℓ is any permutation of 0,1,2,3.

If P is any polyhedron in $\mathbb{R}^3$ we define its unnormalized second moment $U(P)$, its moment of inertia $I(P)$, and its normalized second moment $G(P)$ (all about the origin) by

$$\begin{aligned} U(P) &= \int_P \boldsymbol{\tau} \cdot \boldsymbol{\tau} d\boldsymbol{\tau} \, , \\ I(P) &= \frac{U(P)}{\mathrm{V}olume(P)} \, , \end{aligned}$$

and

$$G(P) = \frac{1}{3} \frac{U(P)}{\text{Volume } (P)^{5/3}} \, .$$

Then the theorem asserts that $G(S(\mathbf{0}))$ is given by (4). To compute $G(S(\mathbf{0}))$ we shall dissect $S(\mathbf{0})$ into 60 tetrahedra, and use the fact that there is an explicit formula (see for example [5]) for the moment of inertia of a tetrahedron. In fact, if T is a tetrahedron with vertices $\mathbf{0}, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ then its barycenter is $\mathbf{q} = (\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3)/4$, and

$$I(T) = \frac{4}{5}\mathbf{q} \cdot \mathbf{q} + \frac{1}{20}(\mathbf{p}_1 \cdot \mathbf{p}_1 + \mathbf{p}_2 \cdot \mathbf{p}_2 + \mathbf{p}_3 \cdot \mathbf{p}_3) \,. \tag{9}$$

If $\mathbf{0}, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ are the $\mathbf{y}$-coordinates of the vertices then T has volume $(6\sqrt{D})^{-1} \det(\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3)$.

We first consider a hexagonal face, say F_1, and divide it into six triangles meeting at the center ($\mathbf{m}_1$) of the face. By joining these triangles to $\mathbf{0}$ we form six tetrahedra. Let us analyze one of these tetrahedra, say the tetrahedron T_{12} with vertices $\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{120}, \mathbf{v}_{123}$. The $\mathbf{y}$-coordinates for its vertices are easily found to be

$$\begin{aligned}
\mathbf{0} &: \quad (0,0,0)\,, \\
\mathbf{m}_1 &: \quad \frac{1}{2}(\rho_{01} + \rho_{12} + \rho_{13}, -\rho_{12}, -\rho_{13})\,, \\
\mathbf{v}_{120} &: \quad \frac{1}{2}(\rho_{01} + \rho_{12} + \rho_{13}, \rho_{02} - \rho_{12} + \rho_{23}, -\rho_{03} - \rho_{13} - \rho_{23})\,, \\
\mathbf{v}_{123} &: \quad \frac{1}{2}(\rho_{01} + \rho_{12} + \rho_{13}, \rho_{02} - \rho_{12} + \rho_{23}, \rho_{03} - \rho_{13} - \rho_{23})\,.
\end{aligned}$$

(The $\mathbf{x}$-coordinates for $\mathbf{m}_1$ are $(\frac{1}{2}, 0, 0)$.) The norms of these vectors can be obtained from (6). It turns out that

$$\mathbf{m}_1 \cdot \mathbf{m}_1 = \frac{1}{4}(\rho_{01} + \rho_{12} + \rho_{13})\,,$$

$$\mathbf{v}_{120} \cdot \mathbf{v}_{120} = \frac{1}{4D}\{D \cdot S_1 - K - 4\lambda_2\lambda_3\}\,, \tag{10}$$

$$\mathbf{v}_{123} \cdot \mathbf{v}_{123} = \frac{1}{4D}\{D \cdot S_1 - K - 4\lambda_1\lambda_3\}\,, \tag{11}$$

where D, S_1 and K have already been defined (see Theorem 1) and

$$\lambda_1 = \rho_{01}\rho_{23}, \lambda_2 = \rho_{02}\rho_{13}, \lambda_3 = \rho_{03}\rho_{12}\,.$$

(Formulas (8) and (9) are given in [1].) To find the moment of inertia $I(T_{12})$ we compute the barycenter $\mathbf{q} = \frac{1}{4}(\mathbf{m}_1 + \mathbf{v}_{120} + \mathbf{v}_{123})$ and use (7), eventually obtaining

$$\begin{aligned}
I(T_{12}) = \frac{1}{40D}\{&D(6\rho_{01} + 3\rho_{02} + \rho_{03} + 6\rho_{12} + 6\rho_{13} + 3\rho_{23}) \\
&-\rho_{01}\rho_{02}\rho_{03}(3\rho_{12} + \rho_{13} + \rho_{23}) - 3\rho_{01}\rho_{02}\rho_{12}(\rho_{13} + \rho_{23}) \\
&-\rho_{01}\rho_{03}(\rho_{12}\rho_{13} + 4\rho_{12}\rho_{23} + \rho_{13}\rho_{23}) - \rho_{02}\rho_{03}(4\rho_{12}\rho_{13} + \rho_{12}\rho_{23} + \rho_{13}\rho_{23}) \\
&\qquad -3\rho_{12}\rho_{13}\rho_{23}(\rho_{01} + \rho_{02} + \rho_{03})\}.
\end{aligned}$$

Also the volume of T_{12} is

$$\frac{1}{24\sqrt{D}}\rho_{03}(\rho_{02} + \rho_{23})(\rho_{01} + \rho_{12} + \rho_{13})\,.$$

The six tetrahedra meeting at the face F_1 are:

Name	Vertices
T_{12}	$(\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{120}, \mathbf{v}_{123})$
$T_{1\bar{3}}$	$(\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{102}, \mathbf{v}_{120})$
T_{10}	$(\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{103}, \mathbf{v}_{102})$
$T_{1\bar{2}}$	$(\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{130}, \mathbf{v}_{103})$
T_{13}	$(\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{132}, \mathbf{v}_{130})$
$T_{1\bar{0}}$	$(\mathbf{0}, \mathbf{m}_1, \mathbf{v}_{123}, \mathbf{v}_{132})$.

We find that $I(T_{1\bar{2}}) = I(T_{12}), I(T_{1\bar{3}}) = I(T_{13}) = \pi_{23} I(T_{12}), I(T_{1\bar{0}}) = I(T_{10}) = \pi_{02} I(T_{12})$, where π_{ab} denotes the transposition (ab) applied to the subscripts of the ρ_{ij} (e.g. $\pi_{02}(\rho_{23}) = \rho_{03}$). Up to this point the calculations may be performed by hand. But to proceed further a computer is desirable. We used the symbolic manipulation program Altran [3]. The contribution to the unnormalized second moment $U(S(\mathbf{0}))$ from all six tetrahedra meeting the face F_1 is

$$\sum^{(6)} \text{Volume } (T_{ij}) I(T_{ij}) = U_1 (\text{say}) \, .$$

The total contribution from all eight hexagonal faces is then

$$U_{hex} = 2(U_1 + \pi_{01} U_1 + \pi_{12} U_1 + \pi_{13} U_1) \, .$$

It remains to consider the quadrilateral faces. It is simplest to divide each of them into two triangles. For example we divide the quadrilateral face F_{12} into the triangles $\mathbf{v}_{123}, \mathbf{v}_{120}, \mathbf{v}_{213}$ and $\mathbf{v}_{120}, \mathbf{v}_{213}, \mathbf{v}_{210}$. Proceeding as before we find that the moments of inertia of the corresponding tetrahedra $\mathbf{0}, \mathbf{v}_{123}, \mathbf{v}_{120}, \mathbf{v}_{213}$ and $\mathbf{0}, \mathbf{v}_{120}, \mathbf{v}_{213}, \mathbf{v}_{210}$ are both equal to

$$\frac{1}{20D} \{ D(3\rho_{01} + 3\rho_{02} + \rho_{03} + \rho_{12} + 3\rho_{13} + 3\rho_{23}) - K - 2\lambda_1 \lambda_3 - 2\lambda_2 \lambda_3 \} \, ,$$

and that both tetrahedra have volume

$$\frac{1}{12\sqrt{D}} \rho_{03} \rho_{12} (\rho_{01} + \rho_{02} + \rho_{13} + \rho_{23}) \, .$$

Twice the product of these two expressions gives the contribution U_{12} to $U(S(\mathbf{0}))$ from the face F_{12}. The total contribution from all six quadrilateral faces is then

$$U_{\text{quad}} = 2(U_{12} + U_{23} + U_{13}) = 2(U_{12} + \pi_{13} U_{12} + \pi_{23} U_{12}) \, .$$

Finally we use Altran to compute $U(S(\mathbf{0})) = U_{\text{hex}} + U_{\text{quad}}$, and $G(S(\mathbf{0})) = U(S(\mathbf{0}))/3D^{5/6}$. After a factor D is removed from the numerator, the result is the right-hand side of (4), which proves Theorem 2. □

3. Minimizing $G(\mathbf{\Lambda})$

We complete the proof of Theorem 1 by establishing the following result.

Theorem 3. *The only local minimum of the right-hand side of Eq. (4) subject to the constraints* $\rho_{ij} \geq 0$ *(for all* i, j*) and* $D \neq 0$ *occurs when all the* ρ_{ij} *are equal.*

Proof. Our method is the one used in [1], namely to exhibit small variations in the ρ_{ij} which will reduce the right-hand side of (4) unless all the ρ_{ij} are equal. For convenience we define

$$\begin{aligned}
&\boldsymbol{\rho} = [\rho_{01}, \rho_{02}, \rho_{03}, \rho_{12}, \rho_{13}, \rho_{23}]\,, \\
&G(\boldsymbol{\rho}) = (D \cdot S_1 + 2S_2 + K)/36D^{4/3}\,, \\
&N = D \cdot S_1 + 2S_2 + K\,.
\end{aligned}$$

The proof will be divided into several steps.

Step 3.1. $D_3, \sqrt{3}A_2 \oplus \sqrt{5}\mathbb{Z}$ and $\mathbb{Z}^3$ are not local minima. In fact, with ϵ small and positive, when

$$\begin{aligned}
\boldsymbol{\rho} &= \frac{[\in, 1, 1, 1, 1, \in], \qquad G(\boldsymbol{\rho}) = 1}{2^{11/3}(1 - \frac{4\in^3}{81} + \cdots)\,;} \\
\boldsymbol{\rho} &= \frac{[3, 3, 5, 3, \in, 0], \qquad G(\boldsymbol{\rho}) = 5^{2/3}}{36(1 - \frac{209\in^2}{3^6 \cdot 5^2} + \cdots)\,;} \\
\boldsymbol{\rho} &= [1, 1, 1, \in, 0, 0], \qquad G(\boldsymbol{\rho}) = \frac{1}{12}(1 - \frac{2\in^2}{9} + \cdots)\,.
\end{aligned}$$

So in each case a small variation in $\boldsymbol{\rho}$ will reduce $G(\boldsymbol{\rho})$.

Step 3.2. *D_3^* is a local minimum.* Since the ρ_{ij} are homogeneous coordinates for Λ, the effect of any variation of the ρ_{ij} on G is the same as the effect of a variation in which one of the ρ_{ij}, say ρ_{01}, is held constant. Temporarily setting $\rho_{02} = x_2, \rho_{03} = x_3, \ldots, \rho_{23} = x_6$, we find that the first partial derivatives $\frac{\partial G}{\partial x_i}$ $(i = 2, \ldots, 6)$ vanish at [1,1,1,1,1,1], while the matrix of second partial derivatives, $(\frac{\partial^2 G}{partial x_i \partial x_j})$ $(i, j = 2, \ldots, 6)$, is equal to a constant times

$$\begin{bmatrix}
20 & -1 & -1 & -16 & -1 \\
-1 & 20 & -16 & -1 & -1 \\
-1 & -16 & 20 & -1 & -1 \\
-16 & -1 & -1 & 20 & -1 \\
-1 & -1 & -1 & -1 & 20
\end{bmatrix}.$$

Since the eigenvalues of this matrix are positive, the matrix itself is positive definite, and [1,1,1,1,1,1] is a local minimum of G.

It remains to show that there is no other local minimum. From now on we assume that $\bar{\boldsymbol{\rho}} = [\bar{\rho}_{01}, \ldots, \bar{\rho}_{23}]$ is a local minimum, and eventually deduce that all the $\bar{\rho}_{ij}$ must be equal.

Step 3.3. *Not more than two $\bar{\rho}_{ij}$ may be zero.* There are essentially only two cases in which three or more of the $\bar{\rho}_{ij}$ may be zero while D is nonzero, namely (a) $\bar{\rho}_{01} = \bar{\rho}_{02} = \bar{\rho}_{12} = 0$ and (b) $\bar{\rho}_{01} = \bar{\rho}_{02} = \bar{\rho}_{23} = 0$. (a) Suppose $\bar{\boldsymbol{\rho}} = [0, 0, 1, 0, y, z]$ with $y > 0, z > 0$. At $\bar{\boldsymbol{\rho}}$ we must have

$$\frac{\partial G}{\partial \rho_{03}} = \frac{\partial G}{\partial \rho_{13}} = \frac{\partial G}{\partial \rho_{23}} = 0.$$

Now $\frac{\partial G}{\partial \rho_{13}} - \frac{\partial G}{\partial \rho_{23}} = yz(y - z)(y + z + 1)$ at $\bar{\boldsymbol{\rho}}$, so $y = z$. Then $\frac{\partial G}{\partial \rho_{03}} = -2(y - 1)y^4$, so $y = z = 1$, and therefore $\bar{\boldsymbol{\rho}} = [0, 0, 1, 0, 1, 1]$. But this is $\mathbb{Z}^3$, which we have already seen is not a local minimum. Case (b) is almost identical and is omitted.

Step 3.4. *Some variations of the ρ_{ij} that fix D.* We shall generally use δR to denote the first order variation in a function $R(\boldsymbol{\rho})$ resulting from small variations $\delta\rho_{ij}$. Let V_0 denote the following variation of the ρ_{ij}:

$$\begin{aligned}
&\rho_{01} \to \rho_{01}\,, \\
&\rho_{02} \to \rho_{02} - \epsilon\rho_{01}\,, \\
&\rho_{03} \to \rho_{03} + \epsilon\rho_{01}\,, \\
&\rho_{12} \to \rho_{12} + \epsilon(\rho_{01} + \rho_{12} + \rho_{13})\,, \\
&\rho_{13} \to \rho_{13} - \epsilon(\rho_{01} + \rho_{12} + \rho_{13})\,, \\
&\rho_{23} \to \rho_{23} + \epsilon(\rho_{12} - \rho_{13})\,,
\end{aligned}$$

where ϵ is small. When applying V_0, we must be careful to ensure that the ρ_{ij} remain nonnegative. For example, we may not apply V_0 with ϵ positive if $\rho_{02=0}$ and $\rho_{01>0}$, since the new value of ρ_{02} would be negative. The variation V_0 has the useful property that it fixes D to the first order in ϵ. To see this, we note that

$$\frac{\partial D}{\partial \rho_{01}} = \sigma_0 + \sigma_1 + \lambda_2 + \lambda_3,$$

etc., where

$$\sigma_i = \rho_{jk}\rho_{j\ell} + \rho_{jk}\rho_{k\ell} + \rho_{j\ell}\rho_{k\ell},$$

i, j, k, ℓ being a permutation of 0,1,2,3 (see [1, p. 297]). Then the fact that

$$\delta D = \sum^{(6)} \frac{\partial D}{\partial \rho_{ij}} \delta\rho_{ij} = 0$$

is an immediate consequence of the identity

$$\sigma_2\rho_{12} + \lambda_2(\rho_{01} + \rho_{12}) = \sigma_3\rho_{13} + \lambda_3(\rho_{01} + \rho_{13}).$$

Thus the denominator of G is fixed by V_0 to the first order. The numerator is increased by

$$\delta(N) = -2 \in J_0\,, \tag{12}$$

where

$$\begin{aligned}
J_0 \;=\; & \rho_{03}\rho_{12}\{(\rho_{01} + \rho_{13})^2 - \rho_{12}(\rho_{01} + \rho_{13})\} \\
& \quad -\rho_{02}\rho_{13}\{(\rho_{01} + \rho_{12})^2 - \rho_{13}(\rho_{01} + \rho_{12})\} + \rho_{01}\rho_{12}\rho_{13}(\rho_{13} - \rho_{12}).
\end{aligned} \tag{13}$$

Although this formula (and others such as (4), (12) and (19)) could have been obtained by hand, it was actually derived with the aid of the interactive symbolic manipulation program Macsyma [12]. Nevertheless the computer did not produce (11) in its present form. Considerable manipulation by hand is almost always required to transform the computer's output into the most appropriate form. This is especially true of (4), (12) and (19).

The expression J_0 is linear in ρ_{02} and ρ_{03}, does not involve ρ_{23}, and goes into $-J_0$ under π_{23}. Other variations of the ρ_{ij} can be obtained from V_0 by applying suitable permutations of the subscripts. We shall require the transformations

$$V_1, V_2, V_3, \qquad\qquad V_4, \qquad\qquad V_5, \qquad V_6,$$

which are obtained by applying the permutations

$$\pi_{12}, \pi_{01}, \pi_{02}, \pi_{02}\pi_{01} = (012), (021), (132),$$

respectively to V_0. Under $V_i (i = 1, \ldots, 6)$ we have $\delta(N) = -2 \in J_i$, where J_i is obtained from J_0 by applying the permutation that produced V_i from V_0. To be quite explicit we write out V_1 and J_1 in full:

$$\begin{array}{ll} V_1 : & \rho_{01} \to \rho_{01} - \in \rho_{02}\,, \\ & \rho_{02} \to \rho_{02}\,, \\ & \rho_{03} \to \rho_{03} + \in \rho_{02}\,, \\ & \rho_{12} \to \rho_{12} + \in (\rho_{02} + \rho_{12} + \rho_{23})\,, \\ & \rho_{13} \to \rho_{13} + \in (\rho_{12} - \rho_{23})\,, \\ & \rho_{23} \to \rho_{23} - \in (\rho_{02} + \rho_{12} + \rho_{23})\,, \end{array}$$

$$\begin{aligned} J_1 &= \rho_{03}\rho_{12}\{(\rho_{02} + \rho_{23})^2 - \rho_{12}(\rho_{02} + \rho_{23})\} \\ &\qquad -\rho_{01}\rho_{23}\{(\rho_{02} + \rho_{12})^2 - \rho_{23}(\rho_{02} + \rho_{12})\} + \rho_{02}\rho_{12}\rho_{23}(\rho_{23} - \rho_{12}). \end{aligned}$$

Step 3.5. *Two $\bar{\rho}_{ij}$ cannot be simultaneously zero.* Again there are essentially only two cases: (a) $\bar{\rho}_{01} = \bar{\rho}_{23} = 0$, with disjoint subscripts, or (b) $\bar{\rho}_{01} = \bar{\rho}_{02} = 0$, with overlapping subscripts.

Case (a). We assume $\bar{\rho}_{01} = \bar{\rho}_{23} = 0$, with the remaining $\bar{\rho}_{ij>0}$. V_1 is a valid variation if $\in < 0$, and then

$$\delta(N) = (-2 \in)\bar{\rho}_{02}\bar{\rho}_{03}\bar{\rho}_{12}(\bar{\rho}_{02} - \bar{\rho}_{12})\,.$$

V_4 is also valid if $\epsilon < 0$, and

$$\delta(N) = (-2 \in)\bar{\rho}_{02}\bar{\rho}_{12}\bar{\rho}_{12}\bar{\rho}_{13}(\bar{\rho}_{12} - \bar{\rho}_{02})\,.$$

Since these variations have opposite sign, $\bar{\mathbf{p}}$ is not a local minimum unless $\bar{\rho}_{02} = \bar{\rho}_{12}$. Applying the permutations (23) and (03)(12) (which leave the assumption $\bar{\rho}_{01} = \bar{\rho}_{23} = 0$ invariant) we obtain the further necessary conditions $\bar{\rho}_{03} = \bar{\rho}_{13}$ and $\bar{\rho}_{13} = \bar{\rho}_{12}$. Thus, $\bar{\mathbf{p}}$ is a multiple of $[0, 1, 1, 1, 1, 0]$, which we have seen is not a local minimum.

Case (b). With $\bar{\rho}_{01} = \bar{\rho}_{02} = 0$ and the other $\bar{\rho}_{ij>0}$, we may apply the variation V_0 with an ϵ of either sign, and so $\delta(N) = -2 \in J_0 = -2 \in \bar{\rho}_{03}\bar{\rho}_{12}\bar{\rho}_{13}(\bar{\rho}_{13} - \bar{\rho}_{12}) = 0$, hence $\bar{\rho}_{12} = \bar{\rho}_{13}$. Similarly, V_1 leads to $\bar{\rho}_{12} = \bar{\rho}_{23}$. Then $\frac{\partial G}{\partial \rho_{03}} = 0$ gives $\bar{\rho}_{12=3/5}$, so $\bar{\mathbf{p}}$ is a multiple of $[0, 0, 1, 3/5, 3/5, 3/5]$. But this is $\sqrt{3}\ A_2 \oplus \sqrt{5}\mathbb{Z}$, which is also not a local minimum.

Step 3.6. *No single $\bar{\rho}_{ij}$ may be zero.* We may assume $\bar{\rho}_{01=0}$, the other $\bar{\rho}_{ij>0}$. Using V_0 we obtain $J_{0=0}$, or $\bar{\rho}_{12} = \bar{\rho}_{13}$, and similarly $\bar{\rho}_{02} = \bar{\rho}_{03}$ from V_2. But $\bar{\mathbf{p}} = [0, \bar{\rho}_{02,}\bar{\rho}_{02}, \bar{\rho}_{12}, \bar{\rho}_{12}, \bar{\rho}_{23}]$ is not a local minimum. For if we evaluate G at $[\epsilon, v, v, 1, 1, z]$ we find $G = G_{0-\epsilon}G_1$+higher order terms, where

$$\frac{G_1 = (v+1)z^2 + (v-1)^2(v+3z+1)}{108 \cdot 2^{1/3}v^{4/3}(v+2z+1)^{4/3} > 0.}$$

Step 3.7. *The path lemma and its consequences.* We may now assume that all $\bar{\rho}_{ij}$ are greater than zero. Then all the variations $V_0, \ldots, V_6$ may be used without restriction. Certainly we must have $E \triangleq (F_3 + F_4) = 0$ at $\bar{\boldsymbol{\rho}}$. But E may be written as

$$\begin{aligned} E &= \rho_{12}(\rho_{01} + \rho_{02})(\rho_{13} + \rho_{23})(\rho_{01} + \rho_{02} - \rho_{13} - \rho_{23}) \\ &\qquad +\rho_{01}\rho_{13}(\rho_{02} + \rho_{23})(\rho_{01} - \rho_{13}) + \rho_{02}\rho_{23}(\rho_{01} + \rho_{13})(\rho_{02} - \rho_{23}). \end{aligned} \tag{14}$$

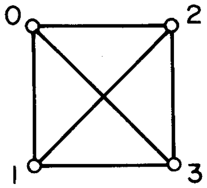

Figure 2:

Therefore if $\bar{\rho}_{01} \geq \bar{\rho}_{13}$ we must have $\bar{\rho}_{02} \leq \bar{\rho}_{23}$, while $\bar{\rho}_{01} \leq \bar{\rho}_{13}$ implies $\bar{\rho}_{02} \geq \bar{\rho}_{13}$. We express this in words by saying that, of the two paths 0-1-3 and 0-2-3 in Figure 2, one must rise and the other must fall (where rise means $\leq$, and fall means $\geq$). This holds between any pair of nodes, and so we may deduce:

Lemma 4. *(The path lemma). In Figure 2, of the two paths i-k-j and i-ℓ-j between any pair of nodes i,j, one must rise and the other must fall.*

Let

$$\left\{ \begin{array}{c} \bar{\rho}_{13} \\ \bar{\rho}_{02} \end{array} \right\} \geq \left\{ \begin{array}{c} \bar{\rho}_{12} \\ \bar{\rho}_{03} \end{array} \right\}$$

be an abbreviation for the inequalities $\bar{\rho}_{13} \geq \bar{\rho}_{12}, \bar{\rho}_{13} \geq \bar{\rho}_{03}, \bar{\rho}_{02} \geq \bar{\rho}_{12}$, and $\bar{\rho}_{02} \geq \bar{\rho}_{03}$.

Lemma 5. *Without loss of generality we may assume that*

$$\left\{ \begin{array}{c} \bar{\rho}_{13} \\ \bar{\rho}_{02} \end{array} \right\} \geq \left\{ \begin{array}{c} \bar{\rho}_{12} \\ \bar{\rho}_{03} \end{array} \right\} \geq \left\{ \begin{array}{c} \bar{\rho}_{23} \\ \bar{\rho}_{01} \end{array} \right\} .$$

Proof. Without loss of generality, no $\bar{\rho}_{ij}$ is larger than $\bar{\rho}_{13}$. By applying Lemma 1 to the paths between nodes 0&1, 0&3, 1&2, and 2&3 we deduce that

$$\left\{ \begin{array}{c} \bar{\rho}_{13} \\ \bar{\rho}_{02} \end{array} \right\} \geq \left\{ \begin{array}{c} \bar{\rho}_{12} \\ \bar{\rho}_{03} \end{array} \right\} \quad \text{and} \quad \left\{ \begin{array}{c} \bar{\rho}_{13} \\ \bar{\rho}_{02} \end{array} \right\} \geq \left\{ \begin{array}{c} \bar{\rho}_{23} \\ \bar{\rho}_{01} \end{array} \right\} .$$

Consideration of the remaining pairs 0&2 and 1&3, and applying the transposition (13) if necessary, then leads to the desired conclusion.

Lemma 6. *By suitably labeling the $\bar{\rho}_{ij}$ we may assume that either*

$$\bar{\rho}_{13} \geq \bar{\rho}_{02} \geq \bar{\rho}_{12} \geq \bar{\rho}_{03} \geq \bar{\rho}_{23} \geq \bar{\rho}_{01}$$

or

$$\bar{\rho}_{13} \geq \bar{\rho}_{02} \geq \bar{\rho}_{03} \geq \bar{\rho}_{12} \geq \bar{\rho}_{23} \geq \bar{\rho}_{01}.$$

Proof. We may assume that $\bar{\rho}_{13}$ is the largest $\bar{\rho}_{ij}$, and then by Lemma 2 the second largest is $\bar{\rho}_{02}$, corresponding to the edge in Figure 2 that does not meet 13. Using (02)(13) we may assume that $\bar{\rho}_{23} \geq \bar{\rho}_{01}$, and finally there are two possibilities for the ranking of $\bar{\rho}_{12}$ and $\bar{\rho}_{03}$. □

Step 3.8. *Two $\bar{\rho}_{ij}$ cannot be equal unless Λ is equivalent to D_3^*.* The subscripts either overlap or are disjoint.

Case (a). If they overlap, we may assume that $\bar{\rho}_{01} = \bar{\rho}_{13} = 1$. Let $\bar{\boldsymbol{\rho}} = [1, v, w, x, 1, z]$. Then $E = 0$ becomes $(z - v)(vxz + vx + 2vz + xz + x) = 0$, so $v = z; F_6 = v(v - 1)(v + 4w + 1) = 0$

implies $v = 1; F_2 = 0$ implies $wx + w - 2x = 0; F_1 = 0$ implies $wx - 2w + x = 0$; and hence $v = w = x = z = 1$, which is D_3^*.

If the equal ρ_{ij} have disjoint subscripts, we shall invoke Lemma 3, and therefore there are three possibilities to consider.

Case (b), $\bar{\rho}_{01} = \bar{\rho}_{23}$. Let $\bar{\mathbf{p}} = [1, v, w, x, y, 1]$, where by Lemma 3 we have $y \geq v \geq x \geq w \geq 1$. If any of these are equal, we can apply case (a), and so we may assume that $y > v > x > w > 1$. By solving $J_3 = 0$ we express v in terms of x and y, and substituting this into $J_4 = 0$ leads to the equation

$$(y-1)((y-x)(xy+y+x)+y+x)(x^2y^2+3xy^2+2x^2y+2y^2-2y+x^2-x) = 0 \tag{15}$$

(found by Macsyma). However, all three factors in (13) are visibly positive, so this case cannot occur.

Case (c), $\bar{\rho}_{02} = \bar{\rho}_{13}$. Let $\bar{\mathbf{p}} = [u, 1, w, x, 1, z]$, where by Lemma 3 and case (a) we may assume $1 > x > w > z > u > 0$. Applying the permutation (032) to Eq. (13) [or alternatively eliminating w and z from $J_0 = 0$, $J_3 = 0$ and $J_2 = 0$] leads us to

$$-(1-x)(ux(x-u)+u(1-u)+x^2+x) \times (u^2x^2+3ux^2+2u^2x+2x^2-2x+u^2-u) = 0. \tag{16}$$

The first two factors in (14) are visibly positive, so

$$u^2x^2+3ux^2+2u^2x+2x^2-2x+u^2-u = 0. \tag{17}$$

Similarly applying the transposition (03) to Eq. (13) [or eliminating w and z from $J_0 = 0$, $J_3 = 0$ and $J_4 = 0$] leads us to

$$u^2x^2+2ux^2+3u^2x+x^2-x+2u^2-2u = 0. \tag{18}$$

Then (15)$-$(16) gives $(x-u)(ux+x+u-1) = 0$, hence $x = (1-u)/(1+u)$, and from (16) we obtain $u = 0$ or $u = 1$, a contradiction.

Case (d), $\bar{\rho}_{03} = \bar{\rho}_{12}$. Let $\bar{\mathbf{p}} = [u, v, 1, 1, y, z]$, where we may assume $y > v > 1 > z > u > 0$. Applying (0321) to (13) [or eliminating y and z from $J_2 = 0$, $J_5 = 0$, $J_0 = 0$] leads us to

$$(v-1)(uv(v-u)+(v^2-u^2)+v+u)(u^2v^2+u(3v^2-1)+2u^2v+2v(v-1)+u^2) = 0,$$

which is impossible since all three factors are positive.

Step 3.8. *The remaining cases*

$$\bar{\rho}_{13} > \bar{\rho}_{02} > \bar{\rho}_{12} > \bar{\rho}_{03} > \bar{\rho}_{23} > \bar{\rho}_{01} > 0 \tag{19}$$

and

$$(18)\bar{\rho}_{13} > \bar{\rho}_{02} > \bar{\rho}_{03} > \bar{\rho}_{12} > \bar{\rho}_{23} > \bar{\rho}_{01} > 0 \tag{20}$$

are impossible. By Lemma 3 these *are* the only remaining cases. We set $\bar{\boldsymbol{\rho}} = [u, v, w, x, y, 1]$ where $y > x > 1 > u > 0$. By showing that this is impossible we rule out both (17) and (18). As in case (b) of the previous step we solve $F_3 = 0$ for v and substitute into $F_4 = 0$. The numerator of the resulting expression is equal to $-u(y-u)$ times

$$\begin{aligned} &\{(y^4x^4 - y^3x^5) + (2y^4x^3u - 2y^3x^4u) + (3y^4x^3 - 3y^2x^5) \\ &+ \ (6y^4x^2u - 2y^3x^3u^2 - 3y^2x^4u) + (y^3x^3 - y^2x^4) \\ &+ \ (3y^4x^2 - 3yx^5) + (6y^4xu - 6y^3x^2u^2) + (6y^3x^2u - 4y^2x^3u^2) + (8y^2x^3u - 6y^3xu^2) \\ &+ \ (2y^3x^2 - 2yx^4) + (y^4x - x^5) + (9y^2x^2u - x^2u^3 - 3y^2x^2u^2) + (7y^3xu - 2y^2u^3 - 3yxu^3) \\ &\qquad +(2y^4u - y^2xu^2) + (y^3x^3u + 2yx^4u + yx^3u^2 + x^4u + 5yx^3u + x^3u^2)\}. \end{aligned} \tag{21}$$

The terms in parentheses are all visibly positive, and so $\delta(N)$ cannot vanish, a contradiction.

This completes the proof of Theorem 3 and therefore of Theorem 1. $\square$

Acknowledgements. Some (but not all!) of the extensive algebraic manipulations needed for the proof were carried out using Altran [3] and Macsyma [12]. We are very grateful to the M.I.T. Laboratory for Computer Science for allowing us to use Macsyma. E. S. B. also wishes to thank the Mathematics Department of Ohio State University at Columbus and Bell Laboratories for their hospitality and support during the course of this work.

References

[1] E. S. Barnes, *The covering of space by spheres*, Canad. J. Math., **8** (1956), pp. 293-304.

[2] E. S. Barnes and T. J. Dickson, *Extreme coverings of n-space by spheres*, J. Austral. Math. Soc., **7** (1967), pp. 115-127; **8** (1968), pp. 638-640.

[3] W. S. Brown, *ALTRAN User's Manual*, Bell Laboratories, Murray Hill, N.J., 4th edition, 1977.

[4] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, New York, 1971.

[5] J. H. Conway and N. J. A. Sloane, *Voronoi regions of lattices, second moments of polytopes, and quantization*, IEEE Trans. Information Theory, **IT-28** (1982), pp. 211-226.

[6] J. H. Conway and N. J. A. Sloane, *Fast quantizing and decoding algorithms for lattice quantizers and codes*, IEEE Trans. Inform. Theory, IT-28 (1982), pp. 227-232.

[7] J. H. Conway and N. J. A. Sloane, *Fast 4- and 8-dimensional quantizers and decoders*, National Telecommunications Record - 1981, IEEE Press, N.Y., Vol. 3, pp. F4.2.1–F4.2.4.

[8] N. C. Gallagher, Jr., and J. A. Bucklew, *Some recent developments in quantization theory*, Proc. 12th Annual Symposium on System Theory, Virginia Beach, Virginia, May 19-20, 1980.

[9] A. Gersho, *Asymptotically optimal block quantization*, IEEE Trans. Information Theory, **IT-25** (1979), pp. 373-380.

[10] A. Gersho, *A structural approach to vector quantization*, National Telecommunications Record - 1981, IEEE Press, N.Y., Vol. 3, pp. F4.1.1–F4.1.5.

[11] C. G. Lekkerkerker, *Geometry of numbers*, Wolters-Noordhoff Publishing, Groningen, and North-Holland Publishing Co., Amsterdam, 1969.

[12] *MACSYMA Reference Manual*, Version 9, Mathlab Group, Laboratory for Computer Science, MIT, Cambridge, MA, 1977.

[13] G. Voronoi, *Sur quelques proprie'te'sdes formes quadratiques positives parfaites*, (Part 1), J. reine angew. Math., **133** (1907), pp. 97-178.

[14] G. Voronoi, *Recherches sur les paralle'loe'dres primitifs* (Part 1), *ibid.*, **134** (1908), pp. 198-287.

[15] P. Zador, *Topics in the asymptotic quantization of continuous random variables*, IEEE Trans. Information Theory, **IT-28** (1982), pp. 139-149.