Local Density Bounds for Sphere Packings and the Kepler Conjecture

Jeffrey C. Lagarias
AT&T Labs - Research
Florham Park, NJ 07932-0971
jcl@research.att.com

(July 27, 2000)

ABSTRACT

This paper describes and formalizes the local density inequality approach to getting upper bounds for sphere packing densities in $\mathbb{R}^n$. This approach was first suggested by L. Fejes-Tóth in 1954 as a method to prove the Kepler conjecture that the densest packing of unit spheres in $\mathbb{R}^3$ has density $\frac{\pi}{\sqrt{18}}$, which is attained by the "cannonball packing." Local density inequalities give upper bounds for the sphere packing density formulated as an optimization problem of a nonlinear function over a compact set in a finite dimensional Euclidean space. The approaches of L. Fejes-Tóth, of W.-Y. Hsiang, and of T. C. Hales to Kepler's conjecture are each based on (different) local density inequalities. Recently T. C. Hales, together with S. P. Ferguson, has presented extensive details carrying out his approach to prove Kepler's conjecture. We describe the particular local density inequality underlying the Hales and Ferguson approach to prove Kepler's conjecture and sketch some features of their proof.

1. Introduction

The Kepler conjecture was stated by Kepler in 1611. It asserts that the face-centered cubic lattice gives the tightest possible packing of unit spheres in $\mathbb{R}^3$.

Kepler Conjecture. *Any packing Ω of unit spheres in $\mathbb{R}^3$ has upper packing density*

$$\bar{\rho}(\Omega) \le \frac{\pi}{\sqrt{18}} \simeq 0.740480 \ . \tag{1.1}$$

The definition of upper packing density is given in §2. The problem of proving the Kepler conjecture appears as part of Hilbert's 18th problem, see [Hi].

T. C. Hales has described an approach for proving Kepler's conjecture, and has announced a proof, completed with the aid of S. Ferguson, which is currently presented in a set of six preprints. The proof is computer-intensive, and involves checking over 5000 subproblems. The Hales approach is similar to earlier approaches in that it aims to prove a local density inequality that gives a (sharp) upper bound on the density. It involves several new ideas which are indicated in §4 and §5.

Local density inequalities obtain upper bounds for the sphere packing constant via an auxiliary nonlinear optimization problem over a compact set of "local configurations". They measure a "local density" in the neighborhood of each sphere center separately. The general approach to the Kepler conjecture is first to find a local optimization problem that actually attains the optimal bound $\frac{\pi}{\sqrt{18}}$ (assuming that one exists), and then to prove it. This approach was first suggested in the early 1950's by L. Fejes-Tóth[FT, pp. 174-181], who presented some evidence that an optimal local density inequality might exist in three dimensions.

The objects of this paper are:

(i) To formulate local density inequalities for sphere packings in $\mathbb{R}^n$, in sufficient generality to include the known candidates for optimal local inequalities.

(ii) To review the history of local density inequalities for three dimensional sphere packing and the Kepler conjecture.

(iii) To give a precise statement of the local density inequality considered in the Hales-Ferguson approach.

(iv) To outline some of features of the Hales-Ferguson proof.

In §2 we present a general framework for local density inequalities, which is valid in $\mathbb{R}^n$, given as Theorem 2.1. This framework is sufficient to cover the approaches of L. Fejes-Tóth, W.-Y. Hsiang, and T. Hales and S. Ferguson to Kepler's conjecture. In §3 we review the history of work on local optimization inequalities for Kepler's conjecture. In §4 we describe the precise local optimization problem formulated by Ferguson and Hales in [FH], which putatively attains $\frac{\pi}{\sqrt{18}}$. In §5 we make remarks on some details of the proof strategy taken in the papers of Hales [I], [II], [III], [IV], Hales and Ferguson [FH], and Ferguson [V]. In §6 we make some concluding remarks.

The current status of the Hales-Ferguson proof is that it appears to be sound. The proof has reputedly been examined in fairly careful detail by a team of reviewers, but it is so long and complicated that it seems difficult for any one person to check it. This paper is intended as an aid in understanding the overall structure of the Hales-Ferguson proof approach. For another account of the Hales and Ferguson work, see Oesterlè [Oe]. For Hales' own perspective, see Hales [H00].

Two appendices are included which contain some information relevant to the Hales-Ferguson proof. Appendix A describes some of the Hales-Ferguson scoring functions. Appendix B lists references in the Hales and Ferguson preprints for proofs of lemmas and theorems stated without proof in §4 and §5.

Notation. $\mathbf{B}_n := B_n(\mathbf{0};1) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \le 1\}$ is the unit n-sphere. It has volume $\kappa_n := \pi^{\frac{n}{2}}/\Gamma(\frac{n}{2}+1)$, with $\kappa_2 = \pi$ and $\kappa_3 = \frac{4\pi}{3}$. We let $\mathbf{C}_n(\mathbf{x}, T) := \mathbf{x} + [0,T]^n$ denote an n-cube of sidelength T, with sides parallel to the coordinate axes, and lowest corner at $\mathbf{x} \in \mathbb{R}^n$.

2. Local Density Inequalities

In this section we present a general formulation of local density inequalities.

We recall the standard definition of sphere packing densities, following Rogers [Ro]. Let Ω denote a set of unit sphere centers, so that $\|\mathbf{v} - \mathbf{v}'\| \ge 2$ for distinct $\mathbf{v}, \mathbf{v}' \in \Omega$.

Definition 2.1 (i) For a bounded region S in $\mathbb{R}^n$, and a sphere packing $\Omega + \mathbf{B}_n$ specified by the sphere centers Ω, the *density* $\rho(S) = \rho(\Omega, S)$ of the packing in the region S is

$$\rho(S) := \frac{vol(S \cap (\Omega + \mathbf{B}_n))}{vol(S)} . \tag{2.1}$$

(ii) For $T > 0$ the *upper density* $\rho(\Omega, T)$ is the maximum density of the packing Ω over all cubes of size T, i.e.

$$\bar{\rho}(\Omega, T) := \sup_{\mathbf{x} \in \mathbb{R}^3} \rho([0,T]^n + \mathbf{x}) \ . \tag{2.2}$$

Then the *upper packing density* of Ω is

$$\bar{\rho}(\Omega) := \limsup_{T \to \infty} \bar{\rho}(\Omega, T) \ . \tag{2.3}$$

(iii) The *sphere packing density* $\delta(\mathbf{B}_n)$ of the ball $\mathbf{B}_n$ of unit radius is

$$\delta(\mathbf{B}_n) := \sup_{\Omega} \bar{\rho}(\Omega). \tag{2.4}$$

Definition 2.2 A sphere packing Ω is *saturated* if no new sphere centers can be added to it.

To obtain sphere packing bounds it obviously suffices to study saturated sphere packings, and in what follows we assume that all packings are saturated unless otherwise stated.

Definition 2.3 An *admissible partition rule* is a rule assigning to each saturated packing Ω in $\mathbb{R}^n$ a collection of closed sets $\mathcal{P}(\Omega) := \{R_\alpha = R_\alpha(\Omega)\}$ with the following properties.

(i) *Partition.* Each set R_α is a finite union of bounded convex polyhedra. The sets R_α cover $\mathbb{R}^3$ and have pairwise disjoint interiors.

(ii) *Locality.* There is a positive constant C (independent of Ω) such that each region R_α has

$$\mathit{diameter}(R_\alpha) \le C \ . \tag{2.5}$$

Each R_α is completely determined by the set of sphere centers $\mathbf{w} \in \Omega$ with

$$\mathit{distance}(\mathbf{w}, R_\alpha) \le C \ . \tag{2.6}$$

There are at most C regions intersecting any cube of side 1.

(iii) *Translation-Invariance.* The partition assigned to the translated packing $\Omega' = \Omega + \mathbf{x}$ consists of the sets $\{R_\alpha(\Omega) + \mathbf{x}\}$.

Definition 2.4 An *admissible weight function* (or *admissible score function*) σ for an admissible partition rule in $\mathbb{R}^n$ assigns to each region $R_\alpha \in \mathcal{P}(\Omega)$ and each $\mathbf{v} \in \Omega$ a real weight $\sigma(R_\alpha, \mathbf{v})$ which satisfies $|\sigma(R_\alpha, \mathbf{v})| < C^*$ for an absolute constant C^*, and which has the following properties.

(i) *Weighted Density Average.* There are positive constants A and B (independent of Ω) such that for each set R_α,

$$\sum_{\mathbf{v} \in \Omega} \sigma(R_\alpha, \mathbf{v}) = (A\ \rho(R_\alpha) - B) vol(R_\alpha) \ , \tag{2.7}$$

where

$$\rho(R_\alpha) vol(R_\alpha) = vol(R_\alpha \cap (\Omega + \mathbf{B}_n)) \tag{2.8}$$

measures the volume covered in R_α by the sphere packing Ω with unit spheres.

(ii) *Locality.* There is an absolute constant C (independent of Ω) such that each value $\sigma(R_\alpha, \mathbf{v})$ is completely determined by the set of sphere centers $\mathbf{w} \in \Omega$ with $\|\mathbf{w} - \mathbf{v}\| \le C$. Furthermore

$$\sigma(R_\alpha, \mathbf{v}) = 0 \quad \text{if} \quad \mathit{dist}(\mathbf{v}, R_\alpha) > C \ . \tag{2.9}$$

(iii) *Translation-Invariance.* The weight function σ' assigned to the translated packing $\Omega' = \Omega + \mathbf{x}$ satisfies

$$\sigma'(R_\alpha + \mathbf{x}, \mathbf{v} + \mathbf{x}) = \sigma(R_\alpha, \mathbf{v}) \ . \tag{2.10}$$

Note that this definition specifically allows negative weights.

The "local density" is measured by the sum of the weights associated to a given vertex $\mathbf{v}$ in a saturated packing.

Definition 2.5 (i) The *vertex D-star* (or *decomposition star*) $\mathcal{D}(\mathbf{v})$ at a vertex $\mathbf{v} \in \Omega$ consists of all sets $R_\alpha \in \mathcal{P}(\Omega)$ such that $\sigma(R_\alpha, \mathbf{v}) \neq 0$.

(ii). The *total score* assigned to a vertex D-star $\mathcal{D}(\mathbf{v})$ at $\mathbf{v} \in \Omega$ is

$$Score(\mathcal{D}(\mathbf{v})) := \sum_{R_\alpha \in \mathcal{D}(\mathbf{v})} \sigma(R_\alpha, \mathbf{v}) \ . \tag{2.11}$$

The total score at $\mathbf{v}$ depends only on regions entirely contained within distance C of $\mathbf{v}$. Any admissible partition and weight function $(\mathcal{P}, \sigma)$ together yield a local inequality for the density of sphere packings, as follows.

Theorem 2.1 *Given an admissible partition in $\mathbb{R}^n$ and weight function $(\mathcal{P}, \sigma)$, set*

$$\theta = \theta_{\mathcal{P},\sigma}(A, B) := \sup_{\Omega \ saturated} \left(\sup_{\mathbf{v} \in \Omega} \ Score(\mathcal{D}(\mathbf{v})) \right) \ . \tag{2.12}$$

and suppose that $\theta < \kappa_n A$, where κ_n is the volume of the unit n-sphere. Then the maximum sphere packing density satisfies

$$\delta(\mathbf{B}) \leq \frac{\kappa_n B}{\kappa_n A - \theta} \ . \tag{2.13}$$

Remark. (1) We let $f(A, B, \theta) := \frac{\kappa_n B}{\kappa_n A - \theta}$ denote the packing density bound as a function of the score constants A and B. The sphere packing density bound actually depends only on the *score constant ratio* $\frac{B}{A}$, rather than on B and A separately, since θ is a homongeneous linear function of A and B. This ratio detemines the relative weighting of covered and uncovered volume used in the inequality.

(2) A natural approach to sphere packing bounds, used in many previous upper bounds, is to partition space into pieces $R(\mathbf{v})$ corresponding to each sphere center $\mathbf{v}$, with each piece containing the unit sphere around $\mathbf{v}$, and aims to establish an upper bound

$$\rho(R(\mathbf{v})) = \frac{\kappa_n}{vol(R(\mathbf{v}))} \leq \theta \ , \quad all \quad \mathbf{v} \in \Omega. \tag{2.14}$$

Then one obtains $\bar{\rho}(\Omega) \leq \theta$. A bound of this sort is *volume-independent,* i.e. it does not take account of the fact that the pieces $R(\mathbf{v})$ may be of different volumes. This property potentially restricts the strength of the density bounds that can be obtained. In contrast the inequality of Theorem 2.1 does take into account the volumes of the individual pieces in the vertex D-star.

Proof of Theorem 2.1. We may assume that Ω is saturated. Given $T > 0$, and any $\epsilon > 0$ we choose a point $\mathbf{x} \in \mathbb{R}^n$ which attains the density bound $\bar{\rho}(\Omega, T)$ on the cube $\mathbf{C}_n(\mathbf{x}, T)$ to within ϵ. We evaluate the scores of all vertex D-stars of vertices $\mathbf{v} \in \Omega \cap \mathbf{C}_n(\mathbf{x}, T)$ in two ways. First, by definition of θ,

$$\sum_{\mathbf{v} \in \Omega \cap \mathbf{C}_n(\mathbf{x}, T)} Score(\mathcal{D}(\mathbf{v})) \leq \theta \# |\Omega \cap \mathbf{C}_n(\mathbf{x}, T)| \ . \tag{2.15}$$

However we also have

$$\begin{aligned}
\sum_{\mathbf{v}\in\Omega\cap\mathbf{C}_n(\mathbf{x},T)} Score(\mathcal{D}(\mathbf{v})) &= \sum_{\mathbf{v}\in\Omega\cap\mathbf{C}_n(\mathbf{x},T)} \left(\sum_{\alpha}\sigma(R_\alpha,\mathbf{v})\right) \\
&= \sum_{\substack{\alpha \\ R_\alpha\subseteq\mathbf{C}_n(\mathbf{x},T)}} \left(\sum_{\mathbf{v}\in\Omega\cap\mathbf{C}_n(\mathbf{x},T)}\sigma(R_\alpha,\mathbf{v})\right) + O(T^{n-1}) \\
&= \sum_{R_\alpha\subseteq\mathbf{C}_n(\mathbf{x};T)} (A\ \ \rho(R_\alpha) - B)\ vol(R_\alpha) + O(T^{n-1}) \\
&= \kappa_n A\#|\Omega\cap\mathbf{C}_n(\mathbf{x},T)| - B\ \ vol(\mathbf{C}_n(\mathbf{x},T)) + O(T^{n-1}) \\
&= \kappa_n A\#|\Omega\cap\mathbf{C}_n(\mathbf{x},T)| - BT^n + O(T^{n-1})\ . \qquad (2.16)
\end{aligned}$$

Here we use that fact that $\{R_\alpha\}$ partitions $\mathbb{R}^n$, so covers the cube, and the $O(T^{n-1})$ error terms above occur because the counting is not perfect within a constant distance C of the boundary of the cube. Combining these evaluations yields

$$(\kappa_n A\ -\ \theta)\#|\Omega\cap\mathbf{C}_n(\mathbf{x},T)| \leq\ B\ T^n + O(T^{n-1}.)$$

If $\theta < \kappa_n A$, then we can rewrite this as

$$\frac{\#|\Omega\cap\mathbf{C}_n(\mathbf{x},T)|}{T^n} \leq \frac{B}{\kappa_n A-\theta}\ + O(\frac{1}{T}). \qquad (2.17)$$

By assumption

$$\begin{aligned}
\bar{\rho}(\Omega,T)-\epsilon\ &\leq\ \frac{vol(\mathbf{C}(\mathbf{x},T)\cap(\Omega+\mathbf{B}))}{T^3} \\
&=\ \kappa_n\frac{\#|\Omega\cap\mathbf{C}(\mathbf{x},T)|}{T^n} + O\left(\frac{1}{T}\right)\ .
\end{aligned}$$

Together with (2.17), this yields

$$\bar{\rho}(\Omega,T)-\epsilon \leq \frac{\kappa_n B}{\kappa_n A-\theta} + O\left(\frac{1}{T}\right)\ , \qquad (2.18)$$

with an O-symbol constant independent of ϵ. Letting $\epsilon\to 0$ and then $T\to\infty$ gives the inequality for $\bar{\rho}(\Omega)$. Since this holds for all saturated packings the result follows. ∎

Determining the quantity $\theta_{\mathcal{P},\sigma}(A,B)$ for fixed A,B can be viewed as a nonlinear optimization problem over a compact set. The translation-invariance property of $(\mathcal{P},\sigma)$ allows the supremum (2.12) to be taken over the smaller set with $\mathbf{v}=\mathbf{0}$ and admissible Ω containing $\mathbf{0}$. The locality property shows that that the vertex D-star at $\mathbf{0}$ is completely determined by $\mathbf{w}\in\Omega$ with $\|\mathbf{w}\|\leq C$. The set of such configurations of nearby sphere centers forms a compact set in the Euclidean topology. Actually the partition and weight functions may be discontinuous functions of the locations of sphere centers, so the optimization problem above is not genuinely over a compact set. One must compactify the space of allowable vertex D-stars by allowing some sets of sphere centers to be assigned more than one possible vertex D-star. In practical cases there is a finite upper bound on the number of possibilities.

Definition 2.6 A local density inequality in $\mathbb{R}^n$ is *optimal* if

$$f(A,B,\theta_{\mathcal{P},\sigma}) = \delta(\mathbf{B}_n)\ . \qquad (2.19)$$

Optimal local density inequalities exist in one and two dimensions. In discussing the three-dimensional case, we shall presume that $\delta(\mathbf{B}_3) = \frac{\pi}{\sqrt{18}}$, so that an optimal density inequality in $\mathbb{R}^3$ will refer to one achieving this value. The evidence indicates that there are many different possible optimal local density inequalities in three dimensions, including that of the Hales and Ferguson proof.

There are currently four candidates for local density inequalities that may be optimal in three dimensions. The first is that of L. Fejes-Toth, described in §3, which uses averages over Voronoi domains, in which the score constant ratio $\frac{B}{A} = \frac{\pi}{\sqrt{18}}$. The second is that of Hsiang [Hs], which is a modification of the Fejes-Tóth averaging, and uses the same score constant ratio. The third is due to Hales [I], and is based on the Delaunay triangulation, using a modified scoring rule described in §3. The fourth is that given in Ferguson and Hales [FH], and uses a combination of Voronoi-type domains and Delaunay simplices, with a complicated scoring rule, described in §4. In the latter two cases the score constant ratio is

$$\frac{B}{A} = \delta_{oct} = \frac{-3\pi + 12\arccos(\frac{1}{\sqrt{3}})}{\sqrt{8}} \approx 0.720903.$$

In all four cases the compact set of local configurations to be searched has very high dimension. Each sphere center has three degrees of freedom, and the number of sphere centers involved in these methods to determine a vertex D-star seems to be around 50, so the search space consists of components of dimension up to roughly 150.

It is unknown whether *optimal* local density inequalities exist for the sphere packing problem in $\mathbb{R}^n$ in any dimension $n \geq 4$. In dimensions 4, 8 and 24 it seems plausible that the minimal volume Voronoi cell in any sphere packing actually occurs in the densest lattice packing. If so, the Voronoi cell decomposition would yield an optimal local inequality in these dimensions, and the densest packing would be a lattice packing in these dimensions. Another question asks: in which dimensions is the maximal sphere packing density attained by a sphere packing whose centers form a finite number of cosets of an n-dimensional lattice? Perhaps in such dimensions an optimal local density inequality exists. The state of the art in sphere packings in dimensions four and above is given in Conway and Sloane [CS].

3. History

We survey results on local density inequalities in three dimensions. The work on local density bounds was originally based on two partitions of $\mathbb{R}^3$ associated to a set Ω of sphere centers: the Voronoi tesselation and the Delaunay triangulation. Since they will play an important role, we recall their definitions.

Definition 3.1 The *Voronoi domain (*or *Voronoi cell)* of $\mathbf{v} \in \Omega$ is

$$V_{vor}(\mathbf{v}) = V_{vor}(\mathbf{v}, \Omega) := \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x} - \mathbf{v}\| \leq \|\mathbf{x} - \mathbf{w}\| \quad \textit{for all} \quad \mathbf{w} \in \Omega\}\ . \tag{3.1}$$

The *Voronoi tesselation* for Ω is the set of Voronoi domains $\{V_{vor}(\mathbf{v}) : \mathbf{v} \in \Omega\}$.

The Voronoi tesselation is a partition of space, up to boundaries of measure zero. If Ω is a saturated sphere packing, then all Voronoi domains are compact sets with diameter bounded by $4\sqrt{2}$.

Definition 3.2 The *Delaunay triangulation* associated to a set Ω is dual to the Voronoi tesselation. It contains an edge between every pair of vertices that have Voronoi domains that share a common face. Suppose now that the points of Ω are in general position, which means that each corner of a Voronoi domain has exactly four incident Voronoi domains. In this case these Voronoi domains have between them four faces that touch this corner, and these faces in turn determine (four edges of) a Delaunay simplex. The resulting Delaunay simplices partition $\mathbb{R}^3$ and make up the Delaunay triangulation. In the case of non-general position Ω the Delaunay triangulation is not unique. The possible Delaunay triangulations are determined locally as limiting cases of general position points. (There are only finitely many triangulations possible in any bounded region of space.)

All the simplices in a Delaunay triangulation have vertices $\mathbf{v}_i \in \Omega$ and contain no other point $\mathbf{v} \in \Omega$. We define, more generally:

Definition 3.3 A *D-simplex* (or *weak Delaunay-simplex*) for Ω is any tetrahedron T with vertices $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4 \in \Omega$ such that no other $\mathbf{v} \in \Omega$ is in the closure of T. We denote it $D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$.

In the literature a *Delaunay simplex* associated to a point set Ω is any simplex with vertices in Ω whose circumscribing sphere contains no other vertex of Ω in its interior. All simplices in a Delaunay triangulation of Ω are Delaunay simplices, so are necessarily D-simplices, but the converse need not hold.

The admissible partitions that have been seriously studied all consist of a domain $V(\mathbf{v})$ associated to each vertex $\mathbf{v} \in V$, which we call a *V-cell*, together with a collection of certain D- simplices $D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ which we call the *D-system* of the partition. We use the term *D-set* to refer to a D-simplex included in the D-system. We note that a V-cell may consist of several polyhedral pieces, and may even be disconnected.

The original approach of L. Fejes-Tóth to getting local upper bounds for sphere packing in $\mathbb{R}^3$ used the Voronoi tesselation associated to Ω. If Ω is a saturated packing, then each Voronoi domain $V_{vor}(\mathbf{v})$ is a bounded polyhedron consisting of points within distance at most $4\sqrt{2}$ of $\mathbf{v}$. Examples are known of Voronoi domains in a saturated packing that have 44 faces; an upper bound for the number of faces of a Voronoi domain of a saturated packing is 49. The *Voronoi partition* takes the V-sets $V(\mathbf{v})$ to be the Voronoi domains of $\mathbf{v}$, with no D-sets, and the vertex D-star $\mathcal{D}_{vor}(\mathbf{v})$ is just $V_{vor}(\mathbf{v})$. A *Voronoi scoring rule* is:

$$Score(\mathcal{D}_{vor}(\mathbf{v})) := (\rho(V(\mathbf{v})) - B) \;\; vol(V(\mathbf{v})), \; . \tag{3.2}$$

with the score constants $A = 1$ and B is to be chosen optimally. Such scoring rules are admissible. However it has long been known that no Voronoi scoring rule gives an optimal inequality (2.19). The dodecahedral conjecture states that the maximum packing density of a Voronoi domain is attained for a local configuration of 12 spheres touching at the center of faces of a circumscribed regular dodecahedron.

Dodecahedral Conjecture. *For a Voronoi domain $V_{vor}(\mathbf{v})$ of a unit sphere packing,*

$$\rho(V_{vor}(\mathbf{v})) \leq \frac{\pi}{15(1 - \cos\frac{\pi}{5}\tan\frac{\pi}{3})} \simeq 0.754697 \tag{3.3}$$

and equality is attained for the dodecahedral configuration.

A proof of the dodecahedral conjecture has been announced by Hales and McLaughlin [Dod], based on similar ideas to the Hales' approach to the Kepler conjecture.

In 1953 L. Fejes-Tóth [FT] proposed that an optimal inequality might exist based on a weighted averaging over Voronoi domains near a given sphere center, and in 1964 he made a specific proposal for such an optimal inequality. In the notation of this paper he used the Voronoi partition and an admissible scoring function of the form

$$\sigma(V_{vor}(\mathbf{w}),\mathbf{v}) := \omega(\mathbf{w},\mathbf{v})\{(A\rho(V_{vor}(\mathbf{w})) - B)vol\ (V_{vor}(\mathbf{w}))\}, \tag{3.4}$$

in which $A = 1$ and $B = \frac{\pi}{\sqrt{18}}$ and the *weights* $\omega(\mathbf{w},\mathbf{v})$ are given by

$$\omega(\mathbf{w},\mathbf{v}) := \begin{cases} \frac{1}{12} & \text{if} \quad 2 \le \|\mathbf{w}-\mathbf{v}\| \le 2+t\ , \\ 0 & \text{if} \quad \|\mathbf{w}-\mathbf{v}\| > 2+t\ , \end{cases} \tag{3.5}$$

and

$$\omega(\mathbf{v},\mathbf{v}) := 1 - \sum_{\mathbf{w}\neq\mathbf{v}} \omega(\mathbf{w},\mathbf{v}). \tag{3.6}$$

Here $t \ge 0$ is a fixed constant. The resulting inequality in Theorem 2.1 is optimal if the associated constant $\theta_{\mathcal{P},\omega}(A,B) = 0$. The Fejes-Tóth scoring function corresponds to a weighted averaging over the spheres touching a central sphere at $\mathbf{v}$, where the weight assigned the central sphere depends on how many spheres touch it. L. Fejes-Tóth considered choosing t as large as possible consistent with requiring that $\omega(\mathbf{v},\mathbf{v}) \ge 0$, which is equivalent to requiring that it is impossible to pack 13 spheres around a given sphere, with all 13 sphere centers within distance $2+t$ of the center of the given sphere. In 1964 he suggested [FT1, p.299] that one could take the value $t = 0.0534$, and we consider this to be Fejes-Tóth's candidate for an optimal inequality. It has not been demonstrated that $\omega(\mathbf{v},\mathbf{v}) \ge 0$ holds for this value of t, but note that the argument of Theorem 2.1 is valid for any value of t, even if some $\omega(\mathbf{v},\mathbf{v})$ are negative. The only issue is whether the resulting sphere packing bound is optimal. Fejes-Tóth [FT1] explicitly noted that establishing an optimal inequality, if it is true, reduces the problem in principle to one in a finite number of variables, possibly amenable to solution by computer.

In 1993 Wu-Yi Hsiang [Hs] studied a variant of the Fejes-Tóth approach. He used the Voronoi partition and an admissible scoring function [1] of the form (3.4), with the same $A = 1$ and $B = \frac{\pi}{\sqrt{18}}$ and with the *weights* $\omega(\mathbf{w},\mathbf{v}) \ge 0$ given by

$$\omega(\mathbf{w},\mathbf{v}) := \begin{cases} \frac{1}{1+N(\mathbf{v})} & \text{if} \quad \|\mathbf{w}-\mathbf{v}\| \le \frac{218}{100}\ , \\ 0 & \text{if} \quad \|\mathbf{w}-\mathbf{v}\| > \frac{218}{100}\ , \end{cases} \tag{3.7}$$

where

$$N(\mathbf{v}) := \#\left\{\mathbf{w} \in \Omega :\ 0 < \|\mathbf{w}-\mathbf{v}\| \le \frac{218}{100}\right\} \tag{3.8}$$

counts the number of "near neighbors" of $\mathbf{v}$. Hsiang announced that his local inequality is optimal (with $\theta_{\mathcal{P},\omega}(A,B) = 0$), and that he had proved it, which would then constitute a proof of Kepler's conjecture. However his proof of optimality is regarded as incomplete by

[1] We have converted the locally averaged density in Hsiang [Hs, Section 3] to the form given in §2 by clearing denominators, and we have cancelled out Hsiang's factor of 13. Note also that each Voronoi domain contains exactly one sphere, so that $\rho(V(\mathbf{v}))vol(V(\mathbf{v})) = \frac{4\pi}{3}$.

the mathematical community, see G. Fejes-Tóth's review of Hsiang's paper in Mathematical Reviews, the critique in [Ha94], and Hsiang's rejoinder [Hs2].

In 1992 Hales ([H1], [H2]) studied the Delaunay triangulation, which partitions $\mathbb{R}^3$ into D-sets. There are a finite number of (local) choices for Delaunay triangulations of a neighborhood of a fixed $\mathbf{v} \in \Omega$. Hales used the following function in defining his associated weight function.

Definition 3.4 The *compression* of a finite region R in $\mathbb{R}^3$ with respect to a sphere packing Ω is

$$\Gamma(R) := (\rho(R) - \delta_{oct}) vol(R) \tag{3.9}$$

in which

$$\delta_{oct} := \frac{-3\pi + 12\arccos(\frac{1}{\sqrt{3}})}{\sqrt{8}} \approx 0.720903 \tag{3.10}$$

is the packing density of the regular octahedron of sidelength 2 with unit spheres centered at its vertices.

Hales initially considered the admissible weight function

$$\sigma(D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4), \mathbf{v}_i) := \Gamma(D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)). \tag{3.11}$$

The vertex D-star of $\mathbf{v}$ consists of all the simplices in the Delaunay triangulation that have $\mathbf{v}$ as a vertex; we call this set of simplices the *Delaunay D-star* $\mathcal{D}_{Del}(\mathbf{v})$ at $\mathbf{v}$. Hales used score constants $A = 4$ and $B = 4\delta_{oct}$, with $A = 4$ used since each simplex is counted four times. However he discovered that the *pentagonal prism* attained a score value exceeding what is needed to prove Kepler's conjecture. The pentagonal prism is conjectured to be extremal for this score function.

The fact that the (conjectured) extremal configurations for the Voronoi tesselation and Delaunay triangulation do not coincide suggested to Hales that a hybrid scoring rule be considered that combines the best features of the Voronoi and Delaunay scoring function. In 1997 Hales again considered a Delaunay triangulation, but modified the scoring rule to depend on the shape of the D-simplex $D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$. For some simplices he used the weight function above, while for others he cut the simplex into four pieces, one for each vertex, call the pieces $V(D, \mathbf{v}_i)$, and assigned the weights[2]

$$\sigma(D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4), \mathbf{v}_i) = 4\Gamma(V(D, \mathbf{v}_i)),$$

for $1 \le i \le 4$. He also partitioned a vertex D-star into pieces called "clusters" whose score functions could be evaluated separately and added up to get the total score. Each "cluster" is a finite union of Delaunay simplices filling up that part of the vertex D-star at $\mathbf{v}$ lying in a pointed cone with vertex $\mathbf{v}$. This vertex cone subdivision facilitates computer-aided proofs by decomposing the problem into smaller subproblems. Hales ([I], [II]) presented evidence that this modified scoring function satisfies an optimal local inequality. He showed that the two known local extremal configurations [3] gave local maxima of the score of the Delaunay D-star in the configuration space with

$$Score(\mathcal{D}_{Del}(\mathbf{v})) = 8pt\ , \tag{3.12}$$

[2]More precisely, he used the "analytic continuation" of this scoring function that is described in Appendix A.

[3]These correspond to Voronoi cells being the rhombic dodecahedron or trapezo-rhombic dodecahedron in Fejes-Tóth [FT1, p. 295].

where

$$pt := \frac{11\pi}{3} - 12\arccos\left(\frac{1}{\sqrt{3}}\right) \simeq 0.0553736 \ , \tag{3.13}$$

which is the optimal value. He also showed that this was a global upper bound over the subset of configurations described by a vertex map[4] $\mathcal{G}(\mathbf{v})$ that is triangulated. Hales [I, Conjecture 2.2] conjectured that the modified score function achieved the optimal inequality

$$Score(\mathcal{D}_{Del}(\mathbf{v})) \le 8pt \ , \tag{3.14}$$

for all Delaunay D-stars $\mathcal{D}_{Del}(\mathbf{v})$. However he and his student S. P. Ferguson [FH] discovered that a pentagonal prism configuration comes very close to violating the inequality (3.14). Furthermore there turned out to be many similar difficult configurations which might possibly violate the inequality. These and other difficulties indicated that it was not numerically feasible to prove (3.14) by a computer proof, assuming that (3.14) is actually true.

Hales and Ferguson together then further modified both the partition rule $\mathcal{P}$ and the scoring rule σ, to obtain a rule with the following properties.

(i) It makes the score inequality stronger on the known bad cases related to the pentagonal prism configuration.

(ii) It uses a more complicated notion of "cluster", which includes Voronoi pieces as well as D-sets, and which retains the "decoupling" property that it is completely determined by vertices of Ω in the cone above it.

(iii) It chooses a scoring function which when combined with "truncation' on clusters is still strong enough to rule out most configurations. The "truncation" operation greatly reduces the number of configurations to be checked, at the cost of weakening the inequality to be proved.

In §4 we give a precise description of the Hales-Ferguson rules $(\mathcal{P}, \sigma)$.

4. Hales-Ferguson Partition and Score Function

Ferguson and Hales [FH] use the following partition and scoring rule. The partition uses two types of D-simplices, with a complicated rule for picking which ones to include as D-sets in the partition. Modified Voronoi domains $V(\mathbf{v})$ are used as V-sets. These differ from the usual Voronoi domain (with the D-sets removed) by mutually exchanging some regions called "tips". The scoring rule is also complicated: the weight function used on a D-simplex no longer depends on just its shape, but depends on the structure of nearby D-sets.

We begin by defining the two types of D-simplices.

Definition 4.1 *A QR-tetrahedron* (or *quasi-regular tetrahedron*) is any tetrahedron with all vertices in Ω and all edges of length $\le \frac{251}{100}$.

Definition 4.2 A *QL-tetrahedron* (or *quarter*) is any tetrahedron with all vertices in Ω and five edges of length $\le \frac{251}{100}$ and one edge with length $\frac{251}{100} < l \le 2\sqrt{2}$. The long edge is called the *spine* (or *diagonal*) of the QL-tetrahedron.

[4] See §5 for a definition of vertex map $\mathcal{G}(\mathbf{v})$.

For some purposes[5] the case of a spine of length exactly $\frac{251}{100}$ should be considered as either a QL-tetrahedron and QR-tetrahedron. Here we treat it exclusively as a QR-tetrahedron.

Neither kind of tetrahedron is guaranteed to be included in the Delaunay triangulation of Ω, but we do have:

Lemma 4.1 *All QR-tetrahedra and QL-tetrahedra are D-simplices.*

The Hales-Ferguson partition rule starts by selecting which D-simplices to include in the D-system. These consist of:

(i) All QR-tetrahedra.

(ii) Some QL-tetrahedra. The QL-tetrahedra included in the partition satisfy the *common spine condition* which states that for a given spine, either all QL-tetrahedra having that spine are included, or none are.

This collection of tetrahedra must (by definition of admissible partition) form a nonoverlapping set, where we say that two sets S_1 and S_2 *overlap* if $\bar{S_1} \cap \bar{S_2}$ has positive Lebesgue measure in $\mathbb{R}^3$. To justify (i) we have:

Lemma 4.2 *No two QR-tetrahedra overlap.*

QL-tetrahedra may overlap QR-tetrahedra or other QL-tetrahedra, hence one needs a rule for deciding which QL-tetrahedra to include. To begin with, QL-tetrahedra can overlap QR-tetrahedra in essentially one way.

Lemma 4.3 *If a QL-tetrahedron and QR-tetrahedron overlap, then the QR-tetrahedron has a common face with an adjacent QR-tetrahedron, and the two unshared vertices of these QR-tetrahedra are the endpoints of the spine of the QL-tetrahedron. The union of these two QR-tetrahedra can be partitioned into three QL-tetrahedra having the given spine, which includes the given QL-tetrahedron. Aside from these QL-tetrahedra, no other QL-tetrahedron overlaps either of these two QR-tetrahedra.*

This lemma shows that the QL-tetrahedra having a given spine have the property that either all of them or none of them overlap the set of QR-tetrahedra. We next consider how QL-tetrahedra can overlap other QL-tetrahedra. The following configuration plays an important role.

Definition 4.3 A *Q-octahedron* is an octahedron whose 6 vertices $\mathbf{v}_i \in \Omega$ and whose 12 edges each have lengths $2 \le l \le \frac{251}{100}$.

A Q-octahedron has three interior diagonals. If a diagonal has length $2 < l \le \frac{251}{100}$ then it partitions the Q-octahedron into four QR-tetrahedra. If a diagonal has length $\frac{251}{100} < l \le 2\sqrt{2}$ then it partitions the Q-octahedron into four QL-tetrahedra of which is the common spine. If a diagonal has length $l > 2\sqrt{2}$ it yields no partition. A Q-octahedron thus gives between zero and three different partitions into four QR-tetrahedra or QL-octahedra. We call it a *live Q-octahedron* if it has at least one such partition. Lemma 4.3 implies that if it has a partition into QR-octahedra then it has no other partition into QR-octahedra or QL-octahedra.

[5] In proving inequalities one wants to work on a compact set. In compactifying the space of configurations, this requires allowing the lower inequality in the definition of QL-tetrahedron to be an equality.

Lemma 4.4 *A QL-tetrahedron having spine a diagonal of a Q-octahedron does not overlap any QL-tetrahedron whose spine is not a diagonal of the same Q-octahedron.*

The rules for choosing which QL-tetrahedra to include either take all QL-tetrahedra having the same spine or take none of them. Thus the selection rule really specifies which spines to include.

Definition 4.4 Consider an edge $[\mathbf{v}_1, \mathbf{v}_2]$ with $\mathbf{v}_1, \mathbf{v}_2 \in \Omega$ and $\frac{251}{100} < \|\mathbf{v}_1 - \mathbf{v}_2\| \leq 2\sqrt{2}$. A vertex $\mathbf{w} \in \Omega$ is called an *anchor* of the edge $[\mathbf{v}_1, \mathbf{v}_2]$ if

$$\|\mathbf{w} - \mathbf{v}_i\| \leq \frac{251}{100} \quad \text{for} \quad i = 1, 2 .$$

Ferguson and Hales use the number of QL-tetrahedra having a spine $[\mathbf{v}_1, \mathbf{v}_2]$ and the number of anchors of that spine in deciding which QL-tetrahedra to include in the D-system. Call a QL-tetrahedron *isolated* if it is the only QL-tetrahedron on its spine $[\mathbf{v}_1, \mathbf{v}_2]$. The inclusion rule for an isolated QL-tetrahedron is:

(QL0) An isolated QL-tetrahedron is included in the D-system if and only if it overlaps[6] no other QL-tetrahedron or QR-tetrahedron.

Next consider spines $[\mathbf{v}_1, \mathbf{v}_2]$ which have two or more associated QL-tetrahedra. Such spines have at least three anchors, and the inclusion rules are:

(QL1) Each non-isolated QL-tetrahedron on a spine with 5 or more anchors is included in the D-system.

(QL2) Each non-isolated QL-tetrahedron on a spine with 4 anchors is included in the D-system, if the spine is not a diagonal of some Q-octahedron. In the case of a live Q-octahedron, we include all QL-tetrahedra having one particular diagonal. and exclude all QL-tetrahedra on other diagonals. For definiteness, we choose the spine to be the shortest diagonal. In case of a tie for shortest diagonal, a suitable tie-breaking rule is used.

(QL3) Each non-isolated QL-tetrahedron on a spine with 3 anchors is included in the D-system if each QL-tetrahedron on the spine does not overlap any other QL-tetrahedron or QR-tetrahedron, or overlaps only isolated QL-tetrahedra. It is excluded from the D-system if some tetrahedron on the spine overlaps either a QR-tetrahedron or a non-isolated QL-tetrahedron having four or more anchors. Finally, if some tetrahedron on the spine overlaps a nonisolated QL-tetrahedron having exactly three anchors, then the spine of the overlapped set is unique, and exactly one of these two sets of non-isolated QL-tetrahedra with 3 anchors is to be included in the D-system, according to a tiebreaking rule[7].

The set of QL-tetrahedra selected above are pairwise disjoint and are disjoint from all QR-tetrahedra. This is justified by the following lemma.

Lemma 4.5 *(i) If two QL-tetrahedra overlap, then at most one of them has* 5 *or more anchors.*
(ii) If two overlapping QL-tetrahedra each have 4 *anchors, then their spines are (distinct) diagonals of some Q-octahedron.*

[6]It cannot overlap a QR-tetrahedron by Lemma 4.3.

[7]The tiebreaking rule could be to include the spine with lowest endpoint using a lexicographic ordering of points in $\mathbb{R}^3$. It appears to me that Hales would permit an arbitrary choice of which one to include, see Lemma 4.5 (iii).

(iii) If a nonisolated QL-tetrahedron with 3 *anchors overlaps another QL-tetrahedron having three anchors, then each of their spines contains exactly two nonisolated QL-tetrahedra, and these four QL-tetrahedra overlap no other QL-tetrahedron or QR-tetrahedron.*

We call the set of QR-tetrahedra and QL-tetrahedra selected as above the Hales-Ferguson *D-system.* (Hales and Ferguson call this a Q-system.)

We now define the V-cells of the Hales-Ferguson partition. To begin with, we take the Voronoi domain $V_{vor}(\mathbf{v})$ at vertex $\mathbf{v}$ and remove from it all D-simplices in the D-system to obtain a reduced Voronoi region $V_{red}(\mathbf{v})$. Next we move certain regions of $V_{red}(\mathbf{v})$ called "tips" to neighboring reduced Voronoi regions to obtain modified regions $V_{mod}(\mathbf{v})$ and finally we define the V-cell $V(\mathbf{v})$ at $\mathbf{v}$ to be the closure of $V_{mod}(\mathbf{v})$.

Definition 4.5 Let T be any tetrahedron such that the center $\mathbf{x} = \mathbf{x}(T)$ of its circumscribing sphere lies outside T. A vertex $\mathbf{v}$ of T is *negative* if the plane H determined by the face F of T opposite $\mathbf{v}$ separates $\mathbf{v}$ from $\mathbf{x}$. The "*tip*" $\Delta(T, \mathbf{v})$ of T associated to a negative vertex $\mathbf{v}$ is that part of the Voronoi region of $\mathbf{v}$ with respect to the points $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ that lies in the closed half plane H^+ determined by H that contains $\mathbf{x}$. The "tip" region $\Delta(T, \mathbf{v})$ does not overlap T, and is a tetrahedron having $\mathbf{x}$ as a vertex, and has three other vertices lying on H.

Lemma 4.6 *(i) A QR-tetrahedron or QL-tetrahedron T has at most one negative vertex.*

(ii) If a negative vertex is present, then the three vertices of the associated "tip" that lie on H actually lie in the face of T opposite to the negative vertex.

(iii) The "tip" of any tetrahedron in the Hales-Ferguson D-system either does not overlap any D-simplex in the Hales-Ferguson D-system, or else is entirely contained in the union of the D-simplices in the D-system.

We say that a "tip" that does not overlap any D-set is *uncovered.* The lemma shows that uncovered "tips" lie in the union of the Voronoi regions $\{V_{red}(\mathbf{w}) : \mathbf{w} \in \Omega\}$, so that rearrangement of uncovered "tips" is legal. There is an a priori possibility that two "tips" may overlap[8] each other.

Uncovered Tip Rearrangement Rule. Each $\mathbf{y} \in \mathbb{R}^3$ that belongs to an uncovered "tip" is reassigned to the nearest *vertex* $\mathbf{w} \in \Omega$ such that $\mathbf{y}$ is not in an uncovered "tip" of any pair $(T, \mathbf{w})$ where T is in the Hales-Ferguson D-system and $\mathbf{w}$ is a negative vertex of $\mathbf{v}$. (A tiebreaking rule is used if two nearest vertices $\mathbf{w}$ are equidistant.)

This rule cuts an uncovered "tip" into a finite number of polyhedral pieces and reassigns the pieces to different reduced Voronoi regions. This prescribes how $V_{mod}(\mathbf{v})$ is constructed, and thus defines the Hales-Ferguson V-cells $V(\mathbf{v})$.

A two-dimensional analogue of a "tip" is pictured[9] in Figure 1.

In this figure the triangle $T = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ plays the role of a D-simplex, with $\mathbf{v}_2$ as a negative vertex and the "tip" is the shaded region. The points $\mathbf{c}_{012}, \mathbf{c}_{013}, \mathbf{c}_{123}$ are centroids of the triangles determined by the corresponding $\mathbf{v}_i$'s. The shaded triangle $[\mathbf{c}_{012}, \mathbf{c}_{013}, \mathbf{c}_{123}]$ is in the Voronoi cell $V_{vor}(\mathbf{v}_0)$ while the remainder of the "tip" is in the Voronoi cell $V_{vor}(\mathbf{v}_2)$. The uncovered tip rearrangement rule partitions the part in $V_{vor}(\mathbf{v}_2)$ into three triangles which are reassigned to the V-cells $V(\mathbf{v}_0), V(\mathbf{v}_1)$ and $V(\mathbf{v}_3)$, e.g. $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{c}_{012}]$ is reassigned to $V(\mathbf{v}_1)$. The reassignment of the "tip" ensures that the pointed cone over $\mathbf{v}_2$ generated by the D-simplex

[8] I don't know if this possibility can occur.

[9] See Figure 2.1 of Hales [II] for another example.

$[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ does not contain any part of the V-cell at $\mathbf{v}_2$. In this example the V-cell at $\mathbf{v}_0$ does not feel the effect of the vertex $\mathbf{v}_2$, due to the rearrangement.

We now turn to the Hales-Ferguson scoring rules. These use the compression function $\Gamma(S)$ given in (3.9). The compression function is additive: If $S = S_1 \cup S_2$ is a partition, then

$$\Gamma(S) = \Gamma(S_1) + \Gamma(S_2) \ . \tag{4.1}$$

For a D-simplex T,

$$vol(T)\rho(T) = \sum_{i=1}^{4} \frac{(\text{solid angle})_i}{3} \ , \tag{4.2}$$

where a full solid angle is 4π.

The Hales-Ferguson weight function for a V-cell is as follows.

(S1) For a V-cell $V(\mathbf{v})$,

$$\sigma_{HF}(V(\mathbf{v}), \mathbf{w}) = \begin{cases} 4\Gamma(V(\mathbf{v})) & \text{if} \quad \mathbf{v} = \mathbf{w} \ , \\ \\ 0 & \text{if} \quad \mathbf{v} \neq \mathbf{w} \ . \end{cases} \tag{4.3}$$

We next consider the weight function for D-sets. Let $(T, \mathbf{v})$ denote a D-simplex together with a vertex $\mathbf{v}$ of it.

Definition 4.6 The *Voronoi measure* $vor(T, \mathbf{v})$ is defined as follows. If the center of the circumscribing sphere of T lies inside T, then T is partitioned into four pieces

$$V_{vor}^{+}(T, \mathbf{v}_i) := \{\mathbf{x} \in T : \|\mathbf{x} - \mathbf{v}_i\| \leq \|\mathbf{x} - \mathbf{v}_j\| \quad \text{for} \quad 1 \leq j \leq 4\}$$

and then

$$vor(T, \mathbf{v}) := \Gamma(V_{vor}^{+}(T, \mathbf{v})) \ . \tag{4.4}$$

There is an analytic formula for the right side of (4.4) given in Appendix A, and this formula is used to define $vor(T, \mathbf{v})$ in cases where the circumcenter falls outside T.

In cases where the circumcenter is outside $\mathbf{v}$, and $\mathbf{v}$ is a negative vertex, then

$$vor(T, \mathbf{v}) = \Gamma(V_{vor}(T, \mathbf{v}) \cup \text{“tip”}) \tag{4.5}$$

while for the other three vertices parts of the “tip” are counted with a negative weight, in such a way that

$$\sum_{i=1}^{4} vor(T, \mathbf{v}_i) = 4\Gamma(T) \tag{4.6}$$

holds in all cases. The weight function for a D-set is given as follows:

(S2) For a QR-tetrahedron $T = D(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$ in the D-system,

$$\sigma_{HF}(T, \mathbf{v}) = \begin{cases} \Gamma(T) & \text{if the circumradius of } T \text{ is at most } \frac{141}{100}. \\ \\ vor(T, \mathbf{v}) & \text{if the circumradius of } T \text{ exceeds } \frac{141}{100}. \end{cases} \tag{4.7}$$

The QL-tetrahedron scoring function is complicated. For a QL-tetrahedron T. let $\eta^+(T)$ be the maximum of the circumradii of the two triangular faces of T adjacent to the spine of T, and define the function

$$\mu(T, \mathbf{v}) := \begin{cases} \Gamma(T) & \text{if } \eta^+(T) \le \sqrt{2} \\ vor(T, \mathbf{v}) & \text{if } \eta^+(T) > \sqrt{2} \end{cases} \tag{4.8}$$

Then the QL-tetrahedron scoring function is defined by:

(S3) ("Flat quarter" case) For a QL-tetrahedron T and a vertex $\mathbf{v}$ not on its spine,

$$\sigma_{HF}(T, \mathbf{v}) := \mu(T, \mathbf{v}) \ . \tag{4.9}$$

(S4) ("Upright quarter case") For a QL-tetrahedron T with vertex $\mathbf{v}$ on its spine, let $\hat{\mathbf{v}}$ denote the opposite vertex on the spine. If T is an isolated QL-tetrahedron, set

$$\sigma_{HF}(T, \mathbf{v}) := \mu(T, \mathbf{v}) \ . \tag{4.10}$$

If T is part of a Q-octahedron, set

$$\sigma_{HF}(T, \mathbf{v}) := \frac{1}{2}(\mu(T, \mathbf{v}) + \mu(T, \hat{\mathbf{v}})) \ . \tag{4.11}$$

In all other cases, set

$$\sigma_{HF}(T, \mathbf{v}) := \frac{1}{2}(\mu(T, \mathbf{v}) + \mu(T, \hat{\mathbf{v}})) + \frac{1}{2}(vor_0(T, \mathbf{v}) - vor_0(T, \hat{\mathbf{v}})) \ , \tag{4.12}$$

in which $vor_0(T, \mathbf{v})$ is a "truncated Voronoi measure" that only counts volume within radius $\frac{1}{2}(\frac{251}{100})$ of vertex $\mathbf{v}$, which is defined in Appendix A, and in [FH, pp. 9-11].

The scoring rule (S4) is the most complicated one. In it the definition (4.11) plays an important role in obtaining good bounds for the pentagonal prism case treated in Ferguson [V], while the definition (4.12) is important in analyzing general configurations using truncation in Hales [IV].

Theorem 4.1 *The Hales-Ferguson partition and scoring function* $(\mathcal{P}_{HF}, \sigma_{HF})$ *are admissible, with score constants* $A = 4$ *and* $B = 4\delta_{oct}$.

Proof. It is easy to verify that the definitions for scoring QR-tetrahedra and QL-tetrahedra satisfy the weighted density average property

$$\sum_{i=1}^{4} \sigma_{HF}(T, \mathbf{v}_i) = 4\Gamma(T) \ , \tag{4.13}$$

which correspond to $A = 4$ and $B = 4\delta_{oct}$, using (4.6). Most of the remaining admissibility conditions are verified by Lemmas 4.1–4.6 except for locality. For locality, a conservative estimate indicates that the rules for removing and adding "tips" to determine the V-cell $V(\mathbf{v})$ are determined by sphere centers $\mathbf{w} \in \Omega$ with $\|\mathbf{w} - \mathbf{v}\| \le 12\sqrt{2}$. Finally the score function on the D-simplices is determined by vertices within distance $6\sqrt{2}$ of $\mathbf{v}$. ∎

Theorem 2.1 associates to $(\mathcal{P}_{HF}, \sigma_{HF})$ a sphere-packing bound that the Hales program asserts is optimal. To establish the Kepler bound

$$\bar{\rho}(\Omega) \leq \frac{\pi}{\sqrt{18}} , \tag{4.14}$$

via (2.15), one must prove that

$$\theta := \theta_{\mathcal{P}_{HF}, \sigma_{HF}}(4, 4\delta_{oct}) = 8pt , \tag{4.15}$$

where

$$pt := \frac{11\pi}{3} - 12 \arccos\left(\frac{1}{\sqrt{3}}\right) \simeq 0.0553736 . \tag{4.16}$$

The score function $Score(\mathcal{D}_{HF}(\mathbf{v}))$ is discontinuous as a function of the sphere centers in Ω near $\mathbf{v}$, because it is a sum of contributions of pieces which may appear and disappear as sphere centers move, and discontinuities occur when QL-tetrahedra convert to QR-tetrahedra. To deal with this, one compactifies the configuration space by allowing some sphere center configurations to have more than one legal decomposition into pieces (but at most finitely many). The optimization problem can then be split into a finite number of subproblems on each of which σ_{HF} is continuous.

The complexity of the definition of $(\mathcal{P}_{HF}, \sigma_{HF})$ is designed to yield a computationally tractable nonlinear optimization problem. The introduction of QL-tetrahedra and the complicated score function on them is designed to help get good bounds for the pentagonal prism case and similar cases. The rule for moving "tips" is intended to facilitate decomposition of the nonlinear optimization problem into more tractable pieces via Theorem 5.5 below, and the use of "truncation."

5. Kepler Conjecture

The main result to be established by the Hales program is the following.

Theorem 5.1 (Main Theorem) *For the Hales-Ferguson partition and scoring rule $(\mathcal{P}_{HF}, \sigma_{HF})$, and any $\mathbf{v} \in \Omega$ in a saturated sphere-packing, the vertex D-star $\mathcal{D}_{HF}(\mathbf{v})$ at $\mathbf{v}$ satisfies*

$$Score(\mathcal{D}_{HF}(\mathbf{v})) \leq 8pt , \tag{5.1}$$

where $pt := \frac{11\pi}{3} - 12 \arccos\left(\frac{1}{\sqrt{3}}\right) \simeq 0.0553736.$

The Kepler conjecture follows by Theorem 2.1.

To prove the inequality (5.1), by translation-invariance we can reduce to the case $\mathbf{v} = \mathbf{0}$ and search the set of all possible vertex stars, which by §4 are determined by those points $\mathbf{w} \in \Omega$ with $\|\mathbf{w}\| \leq 12\sqrt{2}$. From now on we assume $\mathbf{0} \in \Omega$ and $\mathbf{v} = \mathbf{0}$.

The space of possible sphere centers $\{\mathbf{w} \in \Omega : \|\mathbf{w}\| \leq 12\sqrt{2}\}$ is compact. It can be decomposed into a large number of pieces, on each of which the score function is continuous. To obtain *compact* pieces, we must compactify the configuration space by assigning more than one possible local D-star $\mathcal{D}(\mathbf{v})$ to certain arrangements of sphere centers. The compactification assigns at most finitely many to each arrangement with an absolute upper bound on the number of possibilities.

The definition of the score $Score(\mathcal{D}(\mathbf{v}))$ involves a sum over the V-sets and D-sets. The usefulness of the compression measure $\Gamma(S)$ is justified by the following lemma.

Lemma 5.1 *(i) Every QR-tetrahedron T satisfies*

$$\Gamma(T) \le pt \ , \tag{5.2}$$

with equality occurring only when T is a regular tetrahedron of edge length 2.

(ii) A QL-tetrahedron T has

$$\Gamma(T) \le 0 \ , \tag{5.3}$$

with equality occurring for those T having five edges of length 2 and a spine of length $2\sqrt{2}$.

Result (ii) illustrates a somewhat counterintuitive behavior of the local density function: when holding five edges of a tetrahedron fixed of length 2, and allowing the sixth edge to vary over $\frac{251}{100} \le l \le 2\sqrt{2}$, the local density measure is largest for a spine of *maximal* length.

The vertices $\mathbf{w} \in \Omega$ with $\|\mathbf{w}\| \le \frac{251}{100}$ play a particularly important role, for they determine all QR-simplices of Ω containing $\mathbf{0}$ as a vertex.

Definition 5.1 The *planar map* (or *graph*) $\mathcal{G}(\mathbf{v})$ associated to a vertex $\mathbf{v} \in \Omega$ consists of the radial projection onto the unit sphere $\partial B(\mathbf{v}; 1) = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x} - \mathbf{v}\| = 1\}$ centered at $\mathbf{v}$ of all vertices $\mathbf{w} \in \Omega$ with $\|\mathbf{w} - \mathbf{v}\| \le \frac{251}{100}$ plus all those edges $[\mathbf{w}, \mathbf{w}']$ between two such vertices which have length $\|\mathbf{w} - \mathbf{w}'\| \le \frac{251}{100}$.

Here we regard the planar map $\mathcal{G}(\mathbf{v})$ as being given with its embedding as a set of arcs on the sphere. The following lemma asserts that no new vertices are introduced other than those coming from points of Ω with $\|\mathbf{w}\| \le \frac{251}{100}$.

Lemma 5.2 *The radial projection of two edges $[\mathbf{w}_1, \mathbf{w}_2]$, $[\mathbf{w}'_1, \mathbf{w}'_2]$ as above onto the unit sphere $\partial B(\mathbf{v}; 1)$ give two arcs in $\mathcal{G}(\mathbf{v})$ which either are disjoint or which intersect at an endpoint of both arcs.*

We study local configurations classified by the planar map $\mathcal{G}(\mathbf{0})$. The planar map $\mathcal{G}(\mathbf{0})$, which is determined by the vertices $\|\mathbf{w}\| \le \frac{251}{100}$, does not in general uniquely determine the vertex D-star $\mathcal{D}_{HF}(\mathbf{0})$, but does determine all points $\mathbf{x}$ in it with $\|\mathbf{x}\| \le \frac{251}{200}$.

Definition 5.2 The part of $\mathcal{D}_{HF}(\mathbf{v})$ that lies in the pointed cone with base point $\mathbf{v}$ determined by a face of the map $\mathcal{G}(\mathbf{v})$ is called the *cluster* over that face. Note that the face need not be convex, or even simply connected- it could be topologically an annulus, for example.

The following lemma shows that the vertex D-star $\mathcal{D}_{HF}(\mathbf{0})$ can be cut up into clusters in a way compatible with the scoring function.

Lemma 5.3 *Each QR-tetrahedron or QL-tetrahedron in the D-star $\mathcal{D}_{HF}(\mathbf{0})$ is contained in a single cluster. Furthermore all such tetrahedra having a common spine are contained in a single cluster.*

In effect the partition of the vertex D-star into clusters partitions the V-cell into smaller pieces, while leaving the D-sets unaffected. The scoring function is additive over any partition of a V-cell into smaller pieces, according to (4.1) and (4.3). The *score* $\sigma_{HF}(F)$ of the cluster determined by a face F of $\mathcal{G}(\mathbf{v})$ is the sum of the scores of the QR-tetrahedra and QL-tetrahedra in the cluster, plus the Voronoi score $4\Gamma(R)$ of the remaining part R of the cluster. We then have

$$Score(\mathcal{D}_{HF}(\mathbf{v})) = \sum_{F \in \mathcal{G}(\mathbf{v})} \sigma_{HF}(F). \tag{5.4}$$

We now consider clusters associated to the simplest faces F in the graph $\mathcal{G}(\mathbf{v})$. Each triangular face corresonds to a QR-tetrahedron in $\mathcal{D}_{HF}(\mathbf{v})$, and, conversely, each QR-tetrahedron in $\mathcal{D}_{HF}(\mathbf{v})$ produces a triangular face. A *quad cluster* is a cluster over a quadrateral face. A Q-octahedron with spine ending at $\mathbf{0}$ results in a quadrilateral face, but there are many other kinds of quad clusters. In the case of faces F with ≥ 5 edges, the cluster may consist of a V-cell plus some QL-tetrahedra, in many possible ways. All the possible decompositions into such pieces have to be considered as separate configurations.

Lemma 5.4 *(i) A cluster over a triangular face F consists of a single QR-tetrahedron, and conversely. The score of such a cluster is at most 1 pt, and equality holds if and only if it is a regular tetrahedron of edge length 2.*

(ii) The sum of the score functions over any quad cluster is at most zero. Equality can occur only if the four sphere centers $\mathbf{v}_i$ corresponding to the vertices of the quad cluster each lie at distance 2 from $\mathbf{v}$ and also from each other, if they share an edge of the quad cluster.

(iii) The score of a cluster over any face with five or more sides is strictly negative.

The extremal graphs where equality is known to occur in (5.1) have eight triangular faces and six quadrilateral faces. The upper bound of $8pt$ for these cases is implied by this lemma. (It appeared first in Hales [II, Theorem 4.1].)

The following result rules out graphs $\mathcal{G}(\mathbf{v})$ with faces of high degree.

Theorem 5.2 *All decomposition stars $\mathcal{D}_{HF}(\mathbf{0})$ with planar maps $\mathcal{G}(\mathbf{0})$ satisfy*

$$Score(\mathcal{D}_{HF}(\mathbf{0})) \leq 8pt \tag{5.5}$$

unless the planar map $\mathcal{G}(0)$ consists entirely of (not necessarily convex) faces of the following kinds: polygons having at most 8 sides, in which pentagons and hexagons may contain an isolated interior vertex or a single edge from an interior vertex to an outside vertex, and a pentagon may exclude from its interior a triangle with two interior vertices.

There remain a finite set of possible map structures that satisfy the conditions of Theorem 5.2. Here we use the fact that there can be at most 50 vertices $\mathbf{v}$ with $\|\mathbf{v}\| \leq \frac{251}{100}$. The list is further pruned by various methods, and reduced to about 5000 cases. Since the (putative) extremal cases are already covered by Lemma 5.4, in the remaining cases one wishes to prove a strict inequality in (5.1), and such bounds can be obtained in principle by computer.

Most of the remaining cases are eliminated by linear programming bounds. The linear programs involve obtain upper bounds for the score function $Score(\mathcal{D}_{HF}(\mathbf{0}))$ for a planar map $\mathcal{G}$ of a particular configuration type, using as objective function the score function, in the form:

$$Maximize \;\; Score(\mathcal{D}_{HF}(0)) := \sum_{\substack{faces \\ F}} \sigma(F) \;, \tag{5.6}$$

where the variable $\sigma(F)$ is the sum of weights associated to the cluster over the face F. The use of linear programming relaxations of the nonlinear program seems to be a necessity in bounding the score function. For example the compression function $\Gamma(R)$ for different regions R is badly behaved: it is neither convex nor concave in general. The linear constraints include hyperplanes bounding the convex hull of the score function over the variable space.

One can decouple the contributions of the separate faces F of $\mathcal{G}(\mathbf{v})$ using the following result.

Lemma 5.5 (Decoupling Lemma) *Let $\mathbf{v} \in \Omega$ be a vertex of a saturated packing and let F be a face of the associated planar map $\mathcal{G}(\mathbf{v})$, and let $\mathcal{C}_F$ denote the (closed) pointed cone over F with vertex $\mathbf{v}$, and let $\mathcal{C}_{F,red}$ denote the closure of the cone over F obtained by removing from $\mathcal{C}_F$ all cones over D-sets with a corner at $\mathbf{v}$. Then the portion of the V-cell $V(\mathbf{v})$ that lies in $\mathcal{C}_F$ is completely determined by the vertices of Ω that fall in the smallest closed convex cone $\bar{\mathcal{C}}_F$ containing $\mathcal{C}_F$. In particular,*

$$V_F := V(\mathbf{v}) \cap \mathcal{C}_F = V(\Omega \cap \bar{\mathcal{C}}_F, \mathbf{v}) \cap \mathcal{C}_{F,red} \ . \tag{5.7}$$

To obtain such a decoupling lemma requires the exchange of "tips" between Voronoi domains, as described in §4.

The decoupling lemma permits the score function $\sigma(V(\mathbf{v}), \mathbf{v})$ to be decomposed into polyhedral pieces that depend on only a few of the nearby vertices. This decomposes the problem into a sum of smaller problems, to bound the scores of the pieces $\sigma(V(\mathbf{v}) \cap \mathcal{C}_F, \mathbf{v})$ in terms of these vertices. It will often be applied when the face F is convex, in which case $\bar{\mathcal{C}}_F = \mathcal{C}_F$.

A futher very important relaxation of the linear programs involves "truncation." The *truncated V-cell* is

$$V_{trunc}(\mathbf{v}) := V(\mathbf{v}) \cap \mathbf{B}(\mathbf{v} : \frac{251}{200}). \tag{5.8}$$

We may consider truncation of that part of the $V-cell$ over each face of $\mathcal{G}$ separately.

Lemma 5.6 *Let F be a face of $\mathcal{G}(\mathbf{v})$ and $\mathcal{C}_F$ the cone over that face. The region $V_{trunc}(\mathbf{v}) \cap \mathcal{C}_F$ is entirely determined by the vertices of $\mathcal{G}(\mathbf{v})$ in $\mathcal{C}_F$. If $\mathcal{V}_F$ denotes this set of vertices, together with $\mathbf{v}$ then this region is the closure of $(V_{vor}(\mathcal{V}, \mathbf{v}) \cap \mathcal{C}_F) - \{D-sets\}$. The compression function satisfies the bound*

$$\Gamma(V_{trunc}(\mathbf{v}) \cap \mathcal{C}_F) \geq \Gamma(V(\mathbf{v}) \cap \mathcal{C}_F). \tag{5.9}$$

The inequality (5.9) implies that replacing a Voronoi-type region by a truncated region can only increase the score, hence one can relax the linear program by using the score of truncated regions. If one is lucky the linear programming bounds using truncated regions will still be strong enough to give the desired inequality. The use of truncation greatly reduces the number of configurations that must be examined. Truncation bounds were also used in proving Theorem 5.2 above.

We add the following remarks about the construction of the linear programming problems.

(1) For each face F of a given graph type $\mathcal{G}$ Hales and Ferguson construct a large number of linear programming constraints in terms of the edge lengths, dihedral angles and solid angles of the polyhedral pieces making up the cluster of $\mathcal{D}_{HF}(\mathbf{v})$ over face F of the graph $\mathcal{G}$. The edge lengths, dihedral angles and solid angles are variables in the linear program. Some of the constraints embody geometric restrictions that a polyhedron of the given type must satisfy. Others of them are inequalities relating the weight function of the polyhedron, which is also a variable in the linear program, to the geometric quantities. The inequalities bound the score function on the cluster (either as a V-cell or as D-sets) in terms of these variables. There are also some global constraints in the linear program, for example that the solid angles of the faces around $\mathbf{v}$ add up to 4π.

(2) The weight function for D-sets does not permit subdivision of the simplex, but the weight function on the V-cell is additive under subdivision, so one can cut such regions up into smaller pieces if necessary, to get improved linear programming bounds, by including more stringent constraints.

(3) In the linear programming relaxation, a feasible solution to the constraints need not correspond to any geometrically constructible vertex D-star. All that is required is that every vertex D-star of the particular configuration type correspond to some feasible point of the linear program.

In this fashion one obtains a long list of linear programs, one for each configuration type, and to rule out a map type $\mathcal{G}$ one needs an upper bound for the linear program's objective function strictly below 8 pt. To rigorously obtain such an upper bound, it suffices to find a feasible solution to the dual linear program, and to obtain a good upper bound one wants the dual feasible solution close to a dual optimal solution. The value of the dual LP's objective function is then a certified upper bound to the primal LP. To obtain such a certification, it is useful to formulate the linear programs so that the dual linear program has only inequality constraints, with no equality constraints, so that the feasible region for it is full-dimensional. This way, one can guarantee that the dual feasible solution is *strictly* inside the dual feasible region which facilitates checking feasibility. This is necessary because the linear program put on the computer is only an approximation to the true linear program. For example, certain constraints of the true LP involve transcendental numbers like π, and one considers an approximation. The effect of these errors is to perturb the *objective function* of the dual linear program. Thus a rigorous bound on the effect of these perturbations on the upper bound can be obtained in terms of the dual feasible solution. In this way one can (in principle) get a certified upper bound [10] on the score for a map type $\mathcal{G}$, using a computer.

The linear programming bounds in the Ferguson-Hales approach above suffice to eliminate all map types $\mathcal{G}$ not ruled out by Theorem 5.2 except for about 100 "bad" cases. These are then handled by ad hoc methods. [I am not sure of the details about how these remaining "bad" cases are handled. Presumably they are split into smaller pieces, extra inequalities are generated somehow, and perhaps specific information on the location of vertices more than $\frac{251}{100}$ is incorporated into the linear programs.]

6. Concluding Remarks

The Kepler conjecture appears to be an extraordinarily difficult nonlinear optimization problem. The "configuration space" to be optimized over has an extremely complicated structure, of high dimensionality, and the function being optimized is highly nonlinear and non-convex, and lacks good monotonicity properties. The crux of the Hales approach is to select a formulation of an optimization problem that can be carried out (mostly by computer) in a reasonable length of time. This led to the Hales-Ferguson choice of an very complicated partition and score function, giving an inelegant local inequality, which however has good decomposition properties in terms of the nonlinear program. Much of the work in the proof lies in the reductions to reasonable sized cases, and the use of linear programming relaxations. The elimination of the most complicated cases, in Theorem 5.2 was a major accomplishment of this approach. The use of Delaunay simplices to cover most of the volume where density is high seems important to the proof and to the choice of score functions, since simple analytic formulae are available for tetrahedra. The Hales - Ferguson proof, assumed correct, is a tour de force of nonlinear optimization.

[10] The Hales proof in the preprints used a linear programming package CPLEX that does not supply such certificates. Therefore the linear programming part of the Ferguson-Hales proof needs to be re-done to obtain *guaranteed* certificates.

In contrast, the Hsiang approach formulates a relatively elegant local inequality, involving only Voronoi domains and a fairly simple weight function: only nearest neighbor regions are counted. It is conceivable that a rigorous proof of the Hsiang inequality can be established, but it very likely will require an enormous computer-aided proof of a sort very similar to the Hales approach. Voronoi domains do not seem well suited to computer proof: they may have 40 or more faces each, and the Hsiang approach requires considering up to twenty of them at a time. A computer-aided proof would likely have to dissect the Voronoi domains into pieces, further increasing the size of the problem.

Acknowledgments. I am indebted to T. Hales for critical readings of a preliminary version, with many suggestions and corrections. G. Ziegler provided comments and corrections.

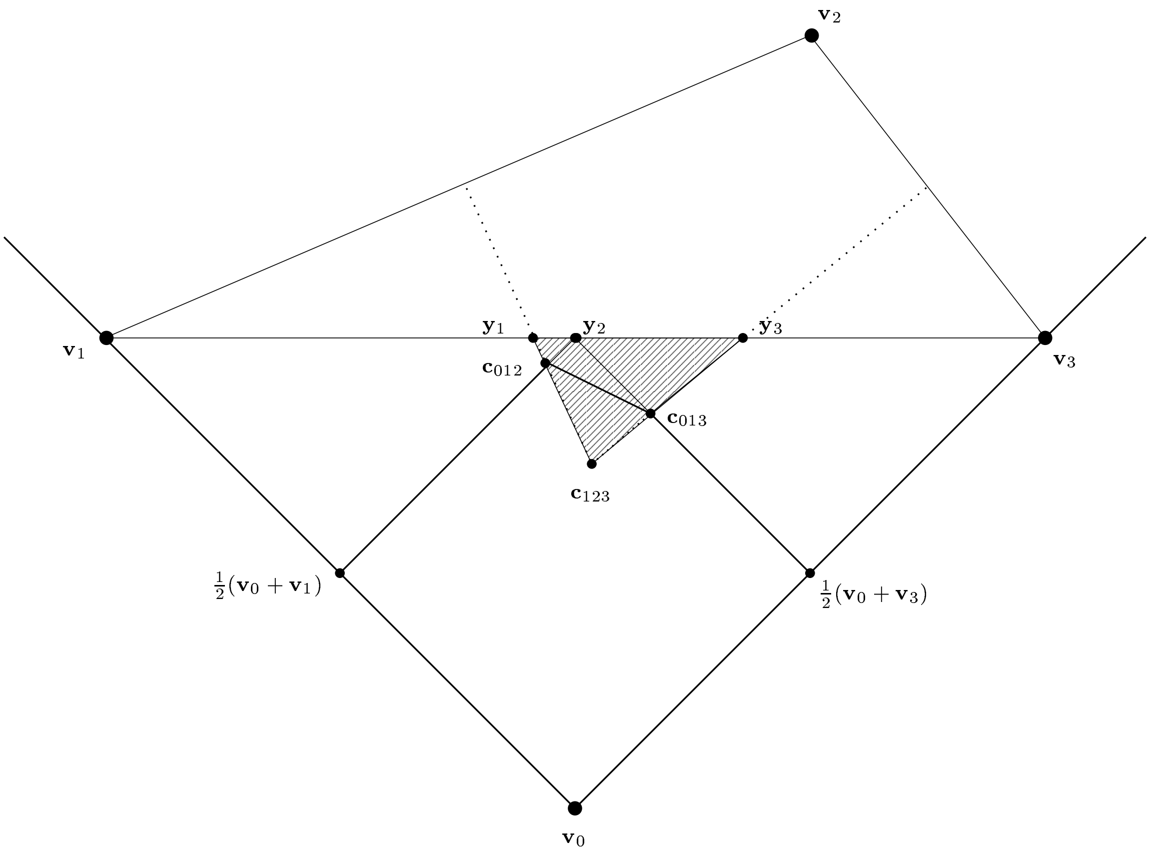

Figure 1: "Tip" of D-simplex $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$

Appendix A. Hales Score Function Formulas

These definitions are taken from in Hales [I, Section 8] and Ferguson and Hales [FH, p. 8–11]. A tetrahedron $T(l_1, l_2, l_3, \ldots, l_6)$ is uniquely determined by its six edge lengths l_i. Let the vertices of T be $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and number the edges as

$$l_i = \|\mathbf{v} - \mathbf{v}_i\| \quad \text{for} \quad 1 \le i \le 3,\ l_4 = \|\mathbf{v}_2 - \mathbf{v}_3\|, l_5 = \|\mathbf{v}_1 - \mathbf{v}_3\| \quad \text{and} \quad l_6 = \|\mathbf{v}_1 - \mathbf{v}_2\| \ . \tag{6.1}$$

We take $\mathbf{v}_0 = \mathbf{0}$ for convenience.

Suppose that the circumcenter $\mathbf{w}_c = \mathbf{w}_c(T)$ of T is contained in the pointed cone over vertex $\mathbf{v}_0$, determined by T. Let $\hat{T}_0$ denote the part of the Voronoi cell of $\mathbf{v}_0$ with respect to the set $\Omega = \{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\mathbf{v}_3\}$ of vertices of T that lies in T. Suppose in addition that the three faces of T containing $\mathbf{v}_0$ are each *non-obtuse* triangles. Then the set $\hat{T}_0$ subdivides into six pieces, called Rogers simplices by Hales [I, p. 31]. A *Rogers simplex* in T is the convex hull of $\mathbf{v}_0$, the midpoint of an edge emanating from $\mathbf{v}_0$, the circumcenter of one face of T containing that edge, and the circumcenter $\mathbf{w}_c = \mathbf{w}_c(T)$. If a denotes the half-length of an edge, b the circumradius of a face and $c = \|\mathbf{y}_c\|$ is the circumradius of T then the associated Rogers simplex has shape

$$R(a,b,c) := T(a,b,c,(c^2 - b^1)^{1/2}, (c^2 - a^2)^{1/2}, (b^2 - a^2)^{1/2}) \ , \tag{6.2}$$

with the positive square root taken. The intersection of a unit sphere centered at $\mathbf{0}$ with $R(a,b,c)$ has volume $\frac{1}{3}Sol(\mathbf{v}_0; R(a,b,c))$, where $Sol(\mathbf{v}_0; R(a,b,c))$ denotes the solid angle of $R(a,b,c)$ at $\mathbf{v}_0$, normalized so that a total solid angle is 4π. Set

$$x_i = l_i^2 \tag{6.3}$$

are the squares of the edge lengths.

Lemma 6.1 *The solid angle $Sol(\mathbf{v}_0, T)$ of a tetrahedron $T(l_1, l_2, l_3, l_4, l_5, l_6)$ is given by*

$$Sol(\mathbf{v}_0, T) := 2 \ arccot \left(\frac{2A}{\Delta^{1/2}} \right) \tag{6.4}$$

in which the positive square root of Δ is taken, the value of arccot lies in $[0, \pi]$, and

$$A(l_1, l_2, l_3, l_4, l_5, l_6) := l_1 l_2 l_3 + \frac{1}{2} l_1 (l_2^2 + l_3^2 - l_4^2) + \frac{1}{2} l_2 (l_1^2 + l_3^2 - l_5^2) + \frac{1}{2} l_3 (l_1^2 + l_3^2 - l_6^2) \tag{6.5}$$

and

$$\begin{aligned} \Delta(l_1, l_2, l_3, l_4, l_5, l_6) \ := \ & l_1^2 l_4^2 (-l_1^2 + l_2^2 + l_3^2 - l_4^2 + l_5^2 - l_6^2) \\ & + l_2^2 l_5^2 (l_1^2 - l_2^2 + l_3^2 + l_4^2 - l_5^2 + l_6^2) \\ & + l_3^2 l_6^2 (l_2^2 + l_2^2 - l_3^2 + l_4^2 + l_5^2 - l_6^2) \\ & - l_2^2 l_3^2 l_4^2 - l_1^2 l_3^2 l_5^2 - l_1^2 l_2^2 l_6^2 - l_4^2 l_5^2 l_6^2 \ . \end{aligned} \tag{6.6}$$

Definition A.1. (i) for a tetrahedron $T(l_1, l_2, \ldots, l_6)$ with vertex $\mathbf{v}_0$, if the circumcenter $\mathbf{w}_c$ of T falls inside the cone determined by T at $\mathbf{v}_0$, then we set

$$vor(T, \mathbf{v}_0) := 4 \sum_{i=1}^{6} \left\{ vol(R_i(a,b,c))(-\delta_{oct}) + \frac{1}{3} Sol(R_i, \mathbf{v}_0) \right\} \tag{6.7}$$

with

$$vol(R(a,b,c)) := \frac{a(b^2-a^2)^{1/2}(c^2-b^2)^{1/2}}{6}, \quad \text{for} \quad 1 \le a \le b \le c\ . \tag{6.8}$$

This formula satisfies $vor(T, \mathbf{v}_0) = 4\Gamma(\hat{T}_0)$.

(ii) The six tetrahedra $R_i(a,b,c)$ are still defined even when the circumcenter $\mathbf{w}_c$ falls outside the cone of T at vertex $\mathbf{v}_0$, and we still take the formula (6.7) to define $vor(T, \mathbf{v}_0)$, except that both $vol(R_i(a,b,c))$ and $Sol(R_i, \mathbf{v}_0)$ are counted with a negative sign: each tetrahedron $R_i(a,b,c)$ falls outside T, and has no interior in common with it.

Hales calls the definition (ii) the "analytic continuation" of case (i). It has a geometric interpretation.

The truncated Voronoi function $vor(T, \mathbf{v}_0; t)$ of a tetrahedron T at vertex $\mathbf{v}_0$ is intended to measure the compression $\Gamma(\hat{T}_0 \cap B(\mathbf{v}_0; t))$. Here we have truncated the region $\hat{T}_0$ by removing from it all points at distance greater than t from $\mathbf{v}_0$. We set

$$vor_0(T, \mathbf{v}_0) := vor\left(T, \mathbf{v}_0, \frac{251}{200}\right)\ . \tag{6.9}$$

The definition

$$vor(T, \mathbf{v}_0; t) := \Gamma(\hat{T}_0 \cap B(\mathbf{v}_0; t)) \tag{6.10}$$

is valid only when the circumcenter $\mathbf{w}_c$ of T lies in the cone generated from T at vertex $\mathbf{v}_0$. In the remaining case one must construct an analytic representation analogous to (6.7) for $vor(T, \mathbf{v}_0; t)$. This is done in [FH, pages 9–10].

Appendix B. References to the Hales Program Results

This paper was written in order to state the Hales-Ferguson local inequality in as simple a way as I could find, and does not match the order in which things are done in the preprints of Hales and Ferguson. Also, the lemmas and theorems stated here are not all stated in the Hales and Ferguson preprints; some of them are based on the talks that Hales gave at IAS. The pointers below indicate where to look in the preprints for the results I formulate as lemmas and theorems. *Warning*: The Hales-Ferguson partition and scoring function given in [FH], which are the ones actually used for the proof of the Kepler conjecture, differ from those used earlier by Hales in [I] and [II].

(0) The idea of considering local inequalities that weight total area and covered area by spheres in a ratio $\frac{B}{A}$ that is not equal to the optimal density occurs in Hales' original approach based on Delaunay triangulations, see [H1] [H2]. It also appears in Hales [I, Lemma 2.1] and in Ferguson and Hales [FH, Proposition 3.14]. I have inserted the parameters A and B in order to include the density inequality of Hsiang[Hs] in the same framework.

(1) *Definitions 4.1 and 4.2* appear in [FH, page 2].

(2) *Lemma 4.1* is Lemma 1.2 of [FH], proved in Lemma 3.5 of [I]. (The fact that no vertex of Ω occurs inside a face of a QL-tetrahedron or a QR-tetrahedron requires additional argument.)

(3) *Lemma 4.2* follows from Lemma 1.3 of [FH].

(4) *Lemma 4.3* is proved in [FH, p.3 bottom].

(5) *Lemma 4.4* is covered in the discussion on [FH, pages 5–6].

(6) *Lemma 4.5* (i) -(ii1) are covered in the discussion on [FH, page 5], including Lemma 1.8 of that paper.

(7) The notion of "tip" is discussed at length in section 2 of Hales [II]. In part II "tips" are not actually reassigned- although this is mentioned - their existence affects the scoring rule used for the associated Delaunay simplex which the "tip" is associated to. The rules for moving "tips" around to make V-cells in the Hales-Ferguson approach are discussed on [FH, page 8]. Warning: the way that "tips" are handled in part II and in [FH] may not be the same: [FH] takes priority.

(8) *Lemma 4.6* (i) is [II, Lemma 2.2] and [FH, Lemma 4.17]. Facts related to (ii) are discussed in [I, Sect. 8.6.7]. (For the second part I do not have a reference.) (iii) Hales mentioned this in IAS lectures, and sent me a proof sketch, which I expanded into the following: Let S be a simplex in the D-system that overlaps a "tip" protuding from $\mathbf{v}$. Say that the "tip" overlaps by pointing in to S along a face F of S. Thus F is a negatively oriented face of $S' = (F, \mathbf{v})$, which means that the simplex S' is a QR-tetrahedron or else a QL-tetrahedron with spine on F. Suppose first that S' is a QL-tetrahedron. It now follows that S must be QL-tetrahedron with its spine on F by [FH, Lemma 2.2]. So S' and S are adjacent QL-tetrahedra with spines on their common face F. Now S is in the D-system since S' is in the D-system. Thus the distance of $\mathbf{v}$ to the vertices in F is at most $\frac{251}{100}$, since $\mathbf{v}$ is not on the spine. We now suppose that the "tip" is not entirely contained in

S, and derive a contradiction. If it isn't contained in S, then it crosses out through a face F' of S. By the same argument, the distance from $\mathbf{v}$ to the vertices of F' is at most $\frac{251}{100}$. Thus $\mathbf{v}$ has distances at most $\frac{251}{100}$ from all vertices of S, which is impossible by [FH, Lemma 1.2] and [FH, Lemma 1.3]. Suppose secondly that S' is a QR-tetrahedron. Then one shows that S is also a QR-tetrahedron, hence is in the D-system. The rest of the argument goes as before, to the same contradiction.

(9) *Theorem 5.1.* The main theorem is first stated as Conjecture 3.15 in [FH, p. 13]. It is the theorem asserted to be proved in [KC].

(10) *Lemma 5.1* (i) appears as [FH, Lemma 3.13]. (ii) is a special case of [FH, Lemma 3.13] for a quad cluster, which can consist of four congruent QL-tetrahedra.

(11) The *standard regions* corresponding to the graph $\mathcal{G}(\mathbf{v})$ are defined on [FH, p. 4]. ("Planar map that breaks unit sphere into regions.")

(12) *Lemma 5.2* follows from Lemma 1.6 of [FH], which implies that crossing lines come from QL-tetrahedra only.

(13) *Lemma 5.3* is an immediate consequence of Lemma 5.2 and [FH, Lemma 1.3].

(14) *Lemma 5.4* appears as [FH, Lemma 3.13].

(15) *Theorem 5.2* follows from the Corollary to Theorem 4.4 of [IV]. See also Proposition 7.1 of [III].

(16) *Lemma 5.5* and *Lemma 5.6*. These results are briefly stated at the bottom of p. 8 of [FH]. There are also some relevant details in Hales [II, Sect. 2.2]. (I do not know an exact reference for detailed proof.)

	Hales terminology	Terminology in this paper
(1)	decomposition star	vertex D-star
(2)	quasiregular tetrahedron	QR-tetrahedron
(3)	quarter	QL-tetrahedron
(4)	diagonal (of quarter)	spine (of QL-tetrahedron)
(5)	Q-system	D-system
(6)	score $\sigma(R, \mathbf{v})$	weight function $\sigma(R, \mathbf{v})$
(7)	standard cluster	cluster

References

[CS] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Codes, Third Edition* Springer-Verlag: New York 1999.

[FT] L. Fejes Tóth, *Lagerungen in der Ebene auf der Kugel und im Raum*, Springer-Verlag: Berlin 1953. (Second Edition 1972.)

[FT1] L. Fejes Tóth, *Regular Figures*, MacMillan: New York 1964.

[V] S. P. Ferguson, Sphere Packings V, Thesis, U. of Michigan 1997. eprint: math.MG/9811077.

[FH] S. P. Ferguson and T. C. Hales, A Formulation of the Kepler Conjecture, eprint: math.MG/9811072.

[H1] T. C. Hales, The Sphere Packing Problem, J. Comp. App. Math. **44** (1992) 41–76.

[H2] T. C. Hales, Remarks on the density of sphere packings in three dimensions, Combinatorica, **13** (2), (1993), 181–197.

[Ha94] T. C. Hales, The status of the Kepler conjecture, Math. Intelligencer **16** (1994), no. 3, 47–58.

[I] T. C. Hales, Sphere Packings I, Disc. Comp. Geom. **17** (1997), 1–51, eprint: math.MG/9811073.

[II] T. C. Hales, Sphere Packings II, Disc. Comp. Geom. **18** (1997), 135–149, eprint: math.MG/9811074.

[III] T. C. Hales, Sphere Packings III, eprint: math.MG/9811075.

[IV] T. C. Hales, Sphere Packings IV, eprint: math.MG/9811076.

[KC0] T. C. Hales, An overview of the Kepler Conjecture, eprint: math.MG/9811071.

[KC] T. C. Hales, The Kepler Conjecture, eprint: math.MG/9811078.

[H00] T. C. Hales, Cannonballs and Honeycombs, Notices Amer. Math. Soc. **47**, No. 4 (2000) 440–449.

[Dod] T. C. Hales and S. McLaughlin, A proof of the dodecahedral conjecture, eprint: math.MG/9811079.

[Hi] D. Hilbert, Mathematical Problems, Bull. Amer. Math. Soc. **8** (1902), 437–479. Reprinted in: *Mathematical Developments Arising from Hilbert Problems*, Proc. Symp. Pure Math XXVIII, American Math. Soc.: Providence 1976.

[Hs] W.-Y. Hsiang, On the sphere problem and Kepler's conjecture, Intl. J. Math. **4** (1993), no. 5, 739–831. (MR 95g: 52032).

[Hs2] W.-Y. Hsiang, A rejoinder to Hales' article, Math. Intelligencer **17**, no. 1 (1995), 35–42.

[Oe] J. Oesterlé, Densité maximale des empilements de sphères en dimension 3 [d'après Thomas C. Hales et Samuel P. Ferguson], Séminaire Bourbaki, Exp. No. 863, Juin 1999.

[Ro1] C. A. Rogers, The packing of equal spheres, Proc. London Math. Soc. **8** (1958), 609–620.

[Ro] C. A. Rogers, *Packing and Covering*, Cambridge University Press: Cambridge 1964.