Jeffrey C. Lagarias
Editor

THE KEPLER CONJECTURE

The Hales-Ferguson Proof

By
Thomas C. Hales
Samuel P. Ferguson

Springer

IOANNIS KE-PLERI S.C. MAIEST. MATHEMATICI STRENA

Seu

De Niue Sexangula.

Cum Priuilegio S. Cæſ. Maieſt. ad annos xv.

FRANCOFVRTI AD MOENVM, apud Godefridum Tampach.

Anno M. DC. XI.

Jeffrey C. Lagarias
Editor

The Kepler Conjecture

The Hales-Ferguson Proof

by

Thomas Hales
Samuel Ferguson

Including
A Special Issue of
Discrete & Computational Geometry

Springer

Editor
Jeffrey C. Lagarias
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109-1043
USA
lagarias@umich.edu

ISBN 978-1-4614-1128-4 e-ISBN 978-1-4614-1129-1
DOI 10.1007/978-1-4614-1129-1
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011941118

Mathematics Subject Classification (2010): 52C17, 11H31, 05B40, 03B35, 68T15

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Kepler conjecture asserts that the densest packing of three-dimensional Euclidean space by equal spheres is attained by the "cannonball" packing, or face-centered-cubic (FCC) packing, which fills space with density $\frac{\pi}{\sqrt{18}} \approx 0.74048$. This conjecture, formulated by Kepler in his booklet "*Strena, seu, de Niue Sexangula*," was published in 1611, exactly four hundred years ago. Notably, in 1900 Hilbert included the sphere packing problem in his famous problem list, as part of his 18th problem. More than a century later, in a landmark result, the Kepler conjecture was solved in work of Thomas C. Hales and Samuel P. Ferguson. An abridged version of their proof appeared in the *Annals of Mathematics* in 2005, followed a year later by the publication of a detailed proof.

This book presents the Hales-Ferguson proof of the Kepler conjecture, together with supporting material and commentary. It begins with an introductory overview chapter, followed by a chapter on the local density approach to sphere packing bounds. This is followed by the six papers of Hales and Ferguson giving their detailed proof, as published in 2006 in a special issue of *Discrete & Computational Geometry*. Next comes a 2010 paper by Hales (with five other authors) making a slight revision to the 2006 proof, and listing corrections. It concludes with two of Hales's initial papers on the problem, published in 1997. All chapters except for the first are papers reprinted from *Discrete & Computational Geometry*.

This book is divided into four parts, as follows.

Part I: Introduction and Survey

The editor has written the two chapters in this part. The first chapter is introductory and features a brief summary of the history of work on the problem, together with a description of Hales and Ferguson's 1998 preprints, details of the peer review process of the Hales and Ferguson papers (which took eight years), and subsequent developments. It also includes remarks on the reliability of the proof and the subsequent approach of Hales to obtain a formal proof of the Kepler Conjecture (in a formal logical system), entirely checkable by computer.

The second chapter describes the general program of obtaining sphere packing upper bounds using local density inequalities, an approach that can be applied to sphere packing in any dimension. It was published as a paper in 2003 in *Discrete & Computational Geometry*. After considering the general case, it specializes to the three-dimensional case, and desecribes the main features of the Hales-Ferguson local density inequality as presented in their 1998 preprints; the 2006 published proof established a slightly modified inequality.

Part II: Proof of the Kepler Conjecture

These six chapters comprise the heart of the volume. They reprint the six papers of Hales with Ferguson that together provide the detailed proof of the Kepler Conjecture. As noted earlier, these papers appeared as a special issue (2006) of *Discrete & Computational Geometry*. However, in this book the short index to definitions that appeared in this special issue is omitted; it is replaced by the two indexes at the end of the book.

In 2010 Hales published an addendum and a list of errata to the published proof; we have added asterisks in the margin of the reprinted papers locating these corrections, which are listed in Part III of the volume on pp. 361–374.

Part III: A Revision to the Proof of the Kepler Conjecture

This part presents an important follow-up paper of Hales (with five coauthors) that was published in *Discrete & Computational Geometry* in 2010. It explains Hales's program to obtain a formal proof of the Kepler Conjecture. The initial process of formalization uncovered one logical gap in the original proof, and this follow up paper provides a correction filling that gap. It also supplies a list of errata to the original papers.

Part IV: Initial Papers of the Hales Program

This part presents two early papers of Hales on Kepler's conjecture, which were published in *Discrete & Computational Geometry* in 1997. These papers give his original formulation of an approach to proving the Kepler Conjecture via a local density inequality, and carry out some initial steps of this approach. They explain and establish a basic framework followed in the subsequent proof. As it turned out, to obtain a proof, this approach required some modification. This modification included changes to the local density inequalities as described in Part III. The proofs given in Parts II and III are independent of these two papers.

In reading this volume, it should be helpful to start with the introductory Chapter 1. Next one might study Chapter 2, which describes the general framework for obtaining upper bounds on sphere packing density, in any dimension. At present it is not known in which dimensions optimal such inequalities (i.e., tight inequalities) may exist: they are known to exist in dimensions 1, 2, and now, by the Hales-Ferguson proof, in dimension 3. It seems likely that optimal inequalities might exist in dimensions 8 and 24 as well. One might then read the historical survey of Hales in Chapter 3, which also describes some features of the proof given in the next five chapters. (This chapter could alternatively be read before reading Chapter 2.) Next one could look at some of the details of the formulation of the proof (Chapter 4). It would also be useful to look at the Revision paper (Chapter 9) to see features of the ongoing work towards a formal proof of the Kepler conjecture. Finally the reader may consider the remaining chapters in the volume.

I thank many people who helped me with this editing project. Ricky Pollack gave useful general advice. Tom Hales made many useful comments regarding the introductory Chapter 1. JoAnn Sears (University of Michigan Library) helped obtain a phographic copy of Kepler's 1611 volume from Cornell University Library. Finally I thank Ann Kostant and John Spiegelman for work on the preparation of this volume. During the preparation of this work I received support from NSF grants DMS-0500555 and DMS-0801029.

Jeffrey C. Lagarias
Ann Arbor, MI
January 11, 2011

Contents

Preface v

Part I Introduction and Survey

1 The Kepler Conjecture and Its Proof, by J. C. Lagarias 3
1. The Kepler Problem for Sphere Packing 5
2. Why the Kepler Problem Is Difficult 9
3. Local Density Approach: History 14
4. Hales Program and Hales-Ferguson Papers 16
5. Peer-Reviewing of the Hales-Ferguson Papers 17
6. Reliability of the Hales-Ferguson Proof 18
7. Formal Proof of the Kepler Conjecture 20
8. Applications of the Hales-Ferguson Proof Methodology 21
9. Contents of This Volume 22

2 Bounds for Local Density of Sphere Packings and the Kepler Conjecture, by J. C. Lagarias 27
1. Introduction 29
2. Local Density Inequalities 31
3. History 36
4. Hales-Ferguson Partition Rule and Score Function 40
5. Kepler Conjecture 47
6. Concluding Remarks 52

Part II Proof of the Kepler Conjecture

Guest Editor's Foreword 62

3 Historical Overview of the Kepler Conjecture, by T. C. Hales 65
1. Introduction 67
1.1. The face-centered cubic packing 67
1.2. Early history, Hariot, and Kepler 69
1.3. History 70
1.4. The Literature 71
2. Overview of the Proof 74
2.1. Experiments with other decompositions 74
2.2. Contents of the papers 77
2.3. Complexity 78
2.4. Computers 78

4 A Formulation of the Kepler Conjecture, by T. C. Hales and S. P. Ferguson 83
3. The Top-Level Structure of the Proof 87
3.1. Statement of theorems 87
3.2. Basic concepts in the proof 91
3.3. Logical skeleton of the proof 92
3.4. Proofs of the central claims 94
4. Construction of the Q-System 94
4.1. Description of the Q-system 95
4.2. Geometric considerations 97
4.3. Incidence relations 99
4.4. Overlap of simplices 103
5. V-Cells 106
5.1. V-cells 107
5.2. Orientation 110
5.3. Interaction of V-cells with the Q-system 111
6. Decomposition Stars 116
6.1. Indexing sets 116
6.2. Cells attached to decomposition stars 118
6.3. Colored spaces 120
7. Scoring 121
7.1. Definitions 122
7.2. Negligibility 128
7.3. Fcc-compatibility 128
7.4. Scores of standard clusters 129
7.5. Scores of simplices and cones 131
7.6. The example of a dodecahedron 132

5 Sphere Packings III. Extremal Cases, by T. C. Hales 135
8. Local Optimality 138
8.1. Results 138
8.2. Rogers simplices 139
8.3. Bounds on simplices 141
8.4. Breaking clusters into pieces 143

8.5. Proofs 148
9. The $\mathcal{S}$-System 151
9.1. Overview 151
9.2. The set $\delta(v)$ 152
9.3. Overlap 159
9.4. The $\mathcal{S}$-system defined 160
9.5. Disjointness 161
9.6. Separation of simplices of type A 162
9.7. Separation of simplices of type B 162
9.8. Separation of simplices of type C 163
9.9. Simplices of type C' 163
9.10. Scoring 164
10. Bounds on the Score in Triangular and Quadrilateral Regions 165
10.1. The function τ 165
10.2. Types 166
10.3. Limitations on types 169
10.4. Bounds on the score in quadrilateral regions 170
10.5. A volume formula 173

6 Sphere Packings IV. Detailed Bounds, by T. C. Hales 177
11. Upright Quarters 180
11.1. Erasing upright quarters 180
11.2. Contexts 181
11.3. Three anchors 182
11.4. Six anchors 183
11.5. Anchored simplices 183
11.6. Anchored simplices do not overlap 184
11.7. Five anchors 186
11.8. Four anchors 188
11.9. Summary 190
11.10. Some flat quarters 192
12. Bounds in Exceptional Regions 193
12.1. The main theorem 193
12.2. Nonagons 195
12.3. Distinguished edge conditions 196
12.4. Scoring subclusters 196
12.5. Proof 197
12.6. Preparation of the standard cluster 198
12.7. Reduction to polygons 199
12.8. Some deformations 200
12.9. Truncated corner cells 202
12.10. Formulas for truncated corner cells 203
12.11. Containment of truncated corner cells 204
12.12. Convexity 207
12.13. Proof that distances remain at least 2 208

13. Convex Polygons 211
13.1. Deformations 211
13.2. Truncated corner cells 211
13.3. Analytic continuation 212
13.4. Penalties 213
13.5. Penalties and bounds 214
13.6. Penalties 216
13.7. Constants 217
13.8. Triangles 219
13.9. Quadrilaterals 220
13.10. Pentagons 221
13.11. Hexagons and heptagons 222
13.12. Loops 222
14. Further Bounds in Exceptional Regions 225
14.1. Small dihedral angles 225
14.2. A particular 4-circuit 226
14.3. A particular 5-circuit 228

7 Sphere Packings V. Pentahedral Prisms, by S. P. Ferguson 235
15. Pentahedral Prisms 237
15.1. The main theorem 238
15.2. Propositions 238
16. The Main Propositions 243
16.1. Scoring 243
16.2. Dimension reduction 244
16.3. Proof of Proposition 15.4 246
16.4. Proof of Proposition 15.5: Top level 247
16.5. Proof of Proposition 15.5: Flat quad clusters 247
16.6. Proof of Proposition 15.5: Octahedra 248
16.7. Proof of Propositions 15.5: Pure Voronoi quad clusters 251
16.8. Pure Voronoi quad clusters: Acute case 252
16.9. Pure Voronoi quad clusters: Obtuse case 253
17. Calculations 264
17.1. Interval arithmetic 264
17.2. The method of subdivision 265
17.3. Numerical considerations 265
17.4. Calculations 266

8 Sphere Packings VI. Tame Graphs and Linear Programs, by T. C. Hales 275
18. Tame Graphs 279
18.1. Basic definitions 279
18.2. Weight assignments 280
18.3. Plane graph properties 282
19. Classification of Tame Plane Graphs 283

19.1. Statement of the theorem 283
19.2. Basic definitions 283
19.3. A finite state machine 284
19.4. Pruning strategies 285
20. Contravening Graphs 288
20.1. A review of earlier results 288
20.2. Contravening plane graphs defined 292
21. Contravention is Tame 294
21.1. First properties 294
21.2. Computer calculations and their consequences 295
21.3. Linear programs 296
21.4. A noncontravening 4-circuit 299
21.5. Possible 4-circuits 300
22. Weight Assignments 300
22.1. Admissibility 301
22.2. Proof that $\mathrm{tri}(v) > 2$ 302
22.3. Bounds when $\mathrm{tri}(v) \in \{3, 4\}$ 304
22.4. Weight assignments for aggregates 306
23. Linear Program Estimates 308
23.1. Relaxation 308
23.2. The linear programs 309
23.3. Basic linear programs 310
23.4. Error analysis 312
24. Elimination of Aggregates 313
24.1. Triangle and quad branching 313
24.2. A pentagonal hull with $n = 8$ 314
24.3. $n = 8$, Hexagonal hull 314
24.4. $n = 7$, Pentagonal hull 314
24.5. Type $(p, q, r) = (5, 0, 1)$ 316
24.6. Summary 316
25. Branch and Bound Strategies 316
25.1. Review of internal structures 316
25.2. 3-crowded and 4-crowded upright diagonals 318
25.3. Five anchors 319
25.4. Penalties 319
25.5. Pent and Hex branching 321
25.6. Hept and Oct branching 322
25.7. Branching on upright diagonals 324
25.8. Branching on flat quarters 325
25.9. Branching on simplices that are not quarters 326
25.10. Branching on quadrilateral subregions 327
25.11. Implementation details for branching 327
25.12. Variables related to score 327
25.13. Appendix: Hexagonal inequalities 329
25.14. Conclusion 337

Part III A Revision to the Proof of the Kepler Conjecture

9 A Revision of the Proof of the Kepler Conjecture, by T. C. Hales, J. Harrison, S. McLaughlin, T. Nipkow, S. Obua, and R. Zumkeller . . 341
Part 1. Formal Proof Initiative 344
1. The Flyspeck project 344
2. Blueprint edition of the Kepler Conjecture 346
3. Formalizing the ordinary mathematics 348
4. Standard ML reimplementation of code 350
5. Proving nonlinear inequalities with Bernstein bases 352
6. Tame graph enumeration 356
7. Verifying linear programs 359
Part 2. Addendum to and Errata in the Original Proof 361
8. Biconnected graphs 361
9. Errata listing 372

Part IV Initial Papers of the Hales Program

10 Sphere Packings I, by T. C. Hales 379
1. Introduction 381
2. The Program 384
3. Quasi-Regular Tetrahedra 388
4. Quadrilaterals 393
5. Restrictions 397
6. Combinatorics 402
7. The Method of Subdivision 405
8. Explicit Formulas for Compression, Volume, and Angle 406
9. Floating-Point Calculations 417
Appendix. Proof of Theorem 6.1 428

11 Sphere Packings II, by T. C. Hales 433
1. Introduction 435
2. Some Polyhedra 438
3. The Score Attached to a Delaunay Star 441
4. The Main Theorem 444
Appendix 448

Errata E1

Index of Symbols 451

Index of Subjects 453

Part I

Introduction and Survey

Introduction to Part I

Part I contains two introductory chapters.

The first chapter in an introductory paper written for this volume. It gives a summary of the history of the problem, starting with the work of Kepler. It describes the difficulty of the problem, gives a description of Hales and Ferguson's 1998 preprints, including details of the peer review process of the Hales and Ferguson papers (which took 8 years), and subsequent developments. It also includes remarks on the reliability of the proof and the subsequent approach of Hales to obtain a formal proof of the Kepler Conjecture (in a formal logical system), entirely checkable by computer.

The second chapter provides a technical survey of local density approaches to obtain upper bounds on sphere packing density in any dimension. It then views the Hales-Ferguson approach and other approaches to the Kepler Conjecture within this context. It also provides an overview of the Hales-Ferguson proof, as given in the six 1998 preprints. This paper appeared in 2002, during the 1998–2006 peer reviewing process of the Kepler Conjecture proof, and affected the final revised form of the proof.

1

The Kepler Conjecture and Its Proof, by J. C. Lagarias

This introductory chapter gives a brief history of the Kepler problem and indicates sources of its difficulty. It describes various approaches to solving the problem, based on local density inequalities. It then discusses the proof of Hales and Ferguson and the peer-reviewing of their papers. It also discusses further developments, towards obtaining a formal proof of the Kepler conjecture, and other applications of the proof methodology.

Contents

1. The Kepler Problem for Sphere Packing
2. Why the Kepler Problem is Difficult
3. Local Density Approach: History
4. Hales Program and Hales-Ferguson Papers
5. Peer-Reviewing of the Hales-Ferguson Papers
6. Reliability of the Hales-Ferguson Proof
7. Formal Proof of the Kepler Conjecture
8. Applications of the Hales-Ferguson Proof Methodology
9. Contents of This Volume

Acknowledgments. I thank Gabor Fejes-Toth for editorial work on the DCG special issue in July 2006. I thank him, Tom Hales, Steven G. Krantz, and Richard Pollack for helpful comments on this paper. This work was supported in part by NSF grant DMS-0801029.

J.C. Lagarias (ed.), *The Kepler Conjecture: The Hales-Ferguson Proof*,
DOI 10.1007/978-1-4614-1129-1_1,

JOHANN KEPLER,

MATHEMATICIAN TO HIS IMPERIAL MAJESTY

A NEW YEAR'S GIFT

or

On the Six-Cornered Snowflake.

Published by GODFREY TAMPACH at
FRANKFORT ON MAIN,
in the year 1611.

The Six-Cornered Snowflake translated by Hardie (1966) Title page. By permission of Oxford University Press.

The Kepler Conjecture and Its Proof

Jeffrey C. Lagarias

Abstract This paper describes work on the Kepler conjecture starting from its statement in 1611 and culminating in the proof of Hales-Ferguson in 1998–2006. It discusses both the difficulty of the problem and of its solution.

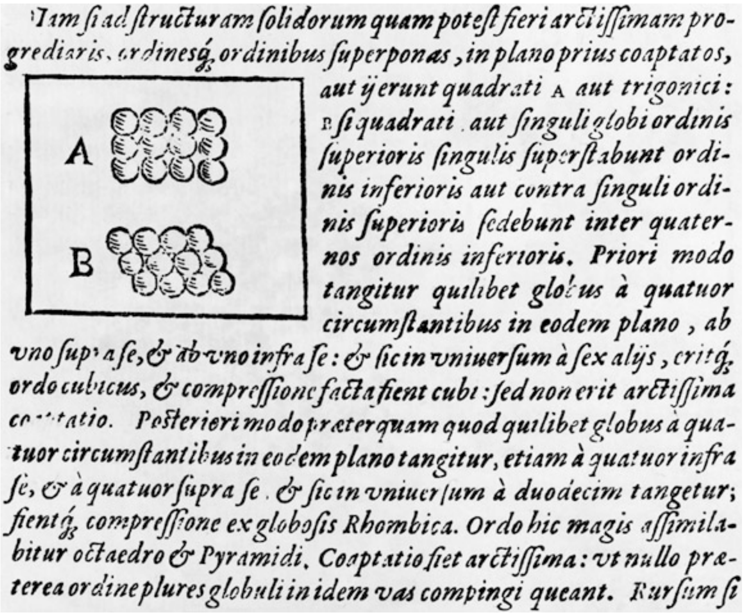
1 The Kepler Problem for Sphere Packing

The Kepler conjecture asserts that a densest packing of three-dimensional Euclidean space by equal spheres is given by the "cannonball" packing, or face-centered-cubic (FCC) packing, which fills space with density $\frac{\pi}{\sqrt{18}} \approx 0.74048$. This conjecture was formulated by Kepler in 1611 [43, pp. 14–17], as follows. After discussing cubical packing in two-dimensional layers versus hexagonal packing in two-dimensional layers, Kepler says (see Figure 1):

> Now if you proceed to pack the solid bodies as tightly as possible, and set the files that are first arranged on the level on top of others, layer on layer, the pellets will be either squared (A in diagram), or in triangles (B in diagram). If squared, either each single pellet of the upper range will rest on a single pellet of the lower, or, on the other hand, each single pellet of the upper range will settle between every four of the lower. In the former mode any pellet is touched by four neighbors in the same plane, and by one above and one below, and so on throughout, each touched by six others. The arrangement will be cubic, and the pellets, when subjected to pressure, will become cubes. But this will not be the tightest pack. In the second mode, not only is every pellet touched by its four neighbors in the same plane, but also by four in the plane above and four below, so throughout one will be touched by twelve, and under pressure spherical pellets will become rhomboid. This arrangement will be more comparable to the

This work was supported in part by NSF Grant DMS-0801029.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, USA
e-mail: lagarias@umich.edu



Iam si ad structuram solidorum quam potest fieri arctissimam pro-
grediaris, ordinesq; ordinibus superponas, in plano prius coaptatos,
aut ij erunt quadrati A aut trigonici:
B si quadrati, aut singuli globi ordinis
superioris singulis superstabunt ordi-
nis inferioris aut contra singuli ordi-
nis superioris sedebunt inter quater-
nos ordinis inferioris. Priori modo
tangitur quilibet globus à quatuor
circumstantibus in eodem plano, ab
vno supra se, & ab vno infra se: & sic in vniuersum à sex alijs, eritq;
ordo cubicus, & compressione facta fient cubi: sed non erit arctissima
coaptatio. Posteriori modo præterquam quod quilibet globus à qua-
tuor circumstantibus in eodem plano tangitur, etiam à quatuor infra
se, & à quatuor supra se, & sic in vniuersum à duodecim tangetur;
fientq; compressione ex globosis Rhombica. Ordo hic magis assimila-
bitur octaedro & Pyramidi. Coaptatio fiet arctissima: vt nullo præ-
terea ordine plures globuli in idem vas compingi queant. Rursum si

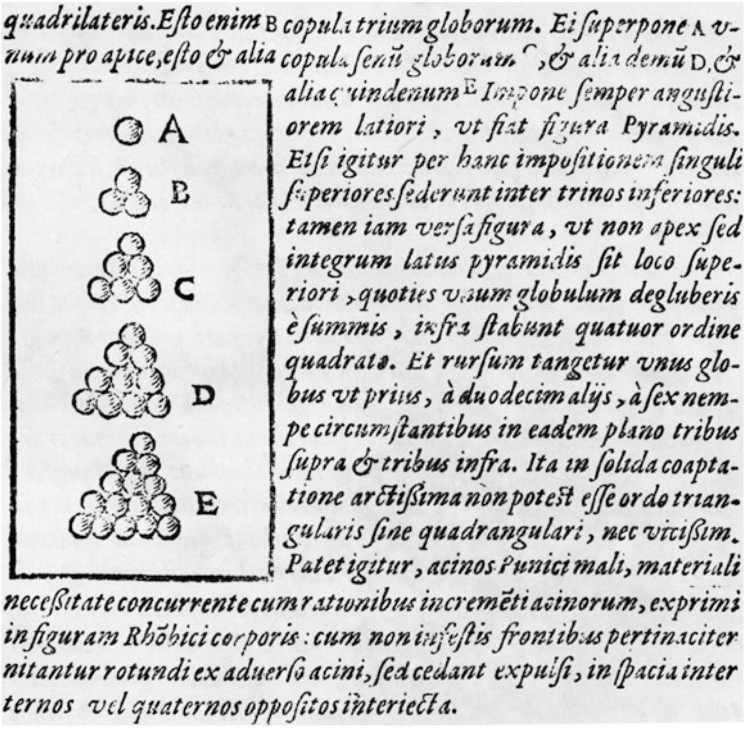
Fig. 1 Kepler's 1611 assertion. [Image courtesy of the Division of Rare and Manuscript Collections, Cornell University Libraries.]

octahedron and pyramid. The packing will be the tightest possible, so that in no other arrangement could more pellets be stuffed into the same container.[1]

Kepler goes on to describe the FCC packing in more detail, saying (see Figure 2):

Thus, let B be a group of three balls; set one A, on it as apex; let there be also another group, C, of six balls; another, D, of ten; and another, E, of fifteen. Regularly superimpose the narrower on the wider to produce the shape of a pyramid. Now, although in this construction each one in the upper layer is seated between three in the lower, yet if you turn the figure round so that not the apex but a whole side of the pyramid is uppermost, you will find, whenever you peel off one ball from the top, four lying below it in square pattern. Again as before, one ball will be touched by twelve others, to wit, by six neighbors in the same plane, and by three above and three below. Thus in the closest pack in three dimensions, the triangular pattern cannot exist without the square, and vice

[1] "Iam si ad structuram solidorum quam potest fieri arctissimam progredaris, ordinesque ordinibus superponas, in plano prius coaptatos aut ii erunt quadrati A aut trigonici B: si quadrati aut singuli globi ordinis superioris singulis superstabunt ordinis inferioris aut contra singuli ordinis superioris sedebunt inter quaternos ordinis inferioris. Priori modo tangitur quilibit globis a quattuour cirucmstantibus in eodem plano, ab uno supra se, et ab uno infra se: et sic in universum a six aliis, eritque ordo cubicus, et compressione facta fient cubi: sed non erit arctissima coaptatio. Posteriori modo praeterquam quod quilibet globus a quattuor circumstantibus in eodem plano tangitur etiam a quattuor infra se, et a quattuor supra se, et sic in universum a duodecim tangetur; fientque compressione ex globosis rhombica. Ordo hic magis assimilabitur octahedro et pyramidi. Coaptatio fiet arctissima, ut nullo praetera ordine plures globuli in idem vas compingi queant." [English translation: Colin Hardie [43, p. 15]]



quadrilateris. Esto enim B copula trium globorum. Ei superpone A v-
num pro apice, esto & alia copula senũ globorum C, & alia demũ D, &
alia quindenum E. Impone semper angusti-
orem latiori, vt fiat figura Pyramidis.
Etsi igitur per hanc impositionem singuli
superiores sederunt inter trinos inferiores:
tamen iam versa figura, vt non apex sed
integrum latus pyramidis sit loco supe-
riori, quoties vnum globulum degluberis
è summis, infra stabunt quatuor ordine
quadrato. Et rursum tangetur vnus glo-
bus vt prius, à duodecim alijs, à sex nem-
pe circumstantibus in eadem plano tribus
supra & tribus infra. Ita in solida coapta-
tione arctissima non potest esse ordo trian-
gularis sine quadrangulari, nec vicissim.
Patet igitur, acinos Punici mali, materiali
necessitate concurrente cum rationibus incremẽti acinorum, exprimi
in figuram Rhõbici corporis: cum non infestis frontibus pertinaciter
nitantur rotundi ex aduerso acini, sed cedant expulsi, in spacia inter
ternos vel quaternos oppositos interiecta.

Fig. 2 Kepler's description of the FCC "cannonball" packing. [Image courtesy of the Division of Rare and Manuscript Collections, Cornell University Libraries.]

> versa. It is therefore obvious that the loculi of the pomegranate are squeezed into the shape of a solid rhomboid.[2]

The FCC packing had been noted earlier by the English mathematician Thomas Hariot [Harriot] (1560–1621). Hariot was mathematics tutor to Sir Walter Raleigh, designed some of his ships, wrote a treatise on navigation, and went on an expedition to Virginia in 1585–1587. He computed a chart in 1591 on how to most efficiently stack cannonballs using the FCC packing, and computed a table of the number of cannonballs in such stacks (cf. Shirley [55, pp. 242–243]). Hariot supported the atomic theory of matter, in which case macroscopic objects may be packed arrangements of very tiny spherical objects, i.e., atoms [42, Chap. III]. He corresponded with Kepler in 1606–1608 on optics, and mentioned the atomic theory in a Dec. 1606 letter as a possible way of explaining why some light is reflected, and some refracted, when hitting a liquid. Kepler replied in 1607,

2 "Esto enim B copula trium globorum. Ei superpone A unum pro apice; esto et alia copula senum globorum C, et alia denum D et alia quindenum E. Impone semper angustiorem latiori, ut fiat figura pyramidis. Etsi igitur per hanc impositionem singuli superiores sederunt into trinos inferiores: tamen iam versa figura, ut non apex sed integrum latus pyramidis sit loc superiori, quoties unum globulum deglberis e summis, infra stabunt quattuor ordine quadrato. Et rursum tangetur unus globus ut prius, et duodecim aliis, a sex nempe circumstantibus in eodem plano tribus supra et tribus infra. Ita in solida coaptatione arctissima non potest ess ordo triangularis sine quadrangulari, nec vicissim. Patet igitur, acinos punici mali, materiali necessitate concurrente cum rationaibus incrementi acinorum, exprimi in figuri rhombici corporis" [English translation by Colin Hardie [43, p. 17]]

not supporting the atomic theory. The known correspondence of Hariot with Kepler does not deal directly with sphere packing.

Questions on sphere packing attracted the attention of Isaac Newton. A discussion with the mathematician David Gregory on 4 May 1694 concerning the brightest stars was summarized in a memorandum of Gregory [50, Vol III, p. 317] as:

> To discover how many stars there are of a given magnitude, he [Newton] considers how many spheres, nearest, second from them, third etc. surround a sphere in a space of three dimensions, there will be 13 of first magnitude, 4×13 of second, $9 \times 4 \times 13$ of third.[3]

Newton's own star table "A Table of ye fixed Starrs for ye yeare 1671" records 13 first magnitude stars, 43 of the second magnitude, 174 of third magnitude; cf. [50, Vol. II, p. 394]. In an (unpublished) notebook Gregory considered the packing problem in 2-dimensions and 3-dimensions and recorded that 13 spheres might touch a given equal sphere [50, Vol. III, p. 321]. It has been asserted that Newton believed only 12 spheres of fixed radius could touch a single sphere of the same radius, but I do not know of a primary reference for this assertion. Certainly every sphere in the FCC packing touches exactly 12 neighbors. It is now known that the maximum number of disjoint equal spheres that can touch a given equal spheres (the "kissing number") is 12, as observed by Hoppe [36] in 1874 and shown rigorously in 1954 by Schütte and van der Waerden [54]. See Conway and Sloane [10, Sec 1.2] and Casselman [7] for further discussion and references.

In the 19th century it was discovered that besides the FCC packing there exists another equally dense sphere packing possessing a full translation group of symmetries, the hexagonal close packing (HCP). This packing was described by William Barlow [3] in 1883, in connection with the possible internal symmetries of crystals.

In 1900 Hilbert [34] included the sphere packing problem in his famous problem list, as part of his 18th problem. He wrote:

> I point out the following question, related to the preceding one, and important to number theory and perhaps sometimes useful in physics and chemistry: How can one arrange most densely in space an infinite number of equal solids of given form, e.g., spheres with given radii or regular tetrahedra with given edges (or in prescribed positions), that is, how can one so fit them together that the ratio of the filled to the unfilled space may be as great as possible?[4]

See Milnor [48] for a further discussion and history of work on the various parts of Hilbert's 18th problem.

The topic of sphere packing is of interest in much wider situations in mathematics, physics and materials science; in the higher dimensional case it is also important in

[3] "Ut noscatur quot sunt stellae magnitudinis 1 ae, 2 dae, 3 ae & tc. considerando quot spherae proximae, seundae ab his 3 ae &tc spheram in spatio trium dimensionis circumstent: erunt 13 primae, 4×13 2-dae, $9 \times 4 \times 13$ 3 ae."

[4] "Ich weise auf die heirmit in Zusammenhang stehende, für die Zahlentheorie wichtige und vielleicht auch der Physik und Chemie einmal Nutzen bringende Frage hin, wie man undendlich viele Körper von das gleichen vorgeschriebenen Gestalt, etwa Kugeln mit gegebenem Radius oder reguläre Tetraeter mit gegebener Kante (bez. in vorgeschriebener Stellung) im Raume am dictesten einbetten, d. h. so lagern kann, daß des Verhältnis des erfülten Raumes zum nichterfüllten Raume möglichst groß ausfällt." [English translation: Dr. Mary Winston Newson]

communications and coding theory. For a physics viewpoint see Aste and Weaire [2] and for both mathematical and communications aspects see Conway and Sloane [10]. For purely mathematical aspects, see Böröczky [5], Fejes-Toth [15], Rogers [53], and Zong [61].

This volume includes the complete proof of the Kepler conjecture by Thomas C. Hales with Samuel P. Ferguson, first presented in six preprints posted on the math arXiv in 1998, and then after extensive revisions, published in six papers in 2006 in *Discrete & Computational Geometry*. (An abridged version of this proof was published earlier by Hales in *Annals of Mathematics* in 2005.) The proof consists of a large body of mathematical arguments and a massive computer verification of many inequalities. There is also a follow-up paper written in 2009 ("A revision of the proof of the Kepler Conjecture") and published in 2010, supplying more details about one point in the 2006 proof, and describing progress towards constructing a formal proof of the Kepler conjecture, in a formal language, completely checkable and certifiable by computer. When completed, this will be a second-generation proof.

We view these papers not just as mathematics but also as historical documents. Below we describe some aspects of the history of work on the problem, and how the work of Hales and Ferguson came about. Then we describe the peer-review process, the reliability of this proof, and the ongoing work of Hales to obtain a formal proof of the Kepler conjecture.

As background, we first explain why the problem is hard.

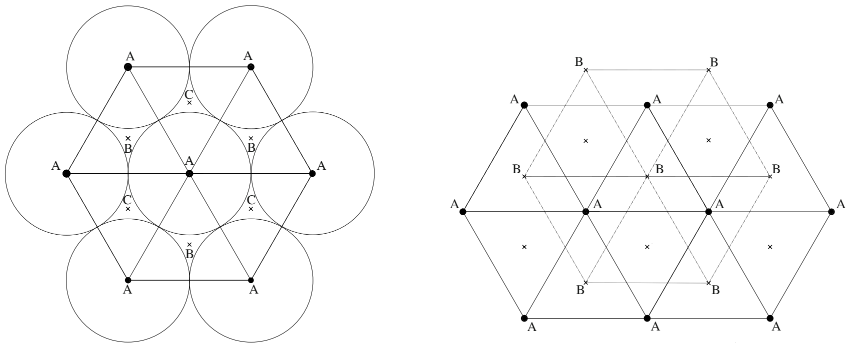
2 Why the Kepler Problem Is Difficult

It is not immediately apparent that Kepler's Conjecture is a problem of extraordinary difficulty.

A first difficulty is to define a rigorous notion of density of packing of spheres. This is resolved using limiting notions of packing a large region and letting its diameter increase to infinity. This notion was not clarified until the 20th century. One can then prove a result asserting that a packing exists that attains a maximal limiting density; cf. [53, Chap. 1].

A second difficulty is that notions of limiting density are very crude in the sense that one can remove spheres from an arbitrarily large finite region without affecting the limiting density. One may therefore wish to impose further local restrictions on the notion of a "densest" packing, for example; that it contain no large "holes." We define a *saturated packing* of spheres to be one into which no new sphere can be inserted, i.e., there is no "hole" of diameter 2 or larger in the packing. Besides this, one might be able to increase density locally by removing a finite collection of spheres in a region and repacking that region to squeeze in one more sphere. This sort of condition seems difficult to analyze, but it already shows that one may wish to take account of "local" conditions specifying density of a packing, compare Bezdek et al [4].

A third difficulty, peculiar to three dimensions, is that there exist uncountably many essentially different "optimally dense" packings. Here we consider packings as essentially different if they are not congruent under a Euclidean motion of space. Consider the packing that starts with a planar layer of hexagonally close packed spheres. That is, there is a planar slice through this layer that intersects all the sphere centers, giving a circle packing of the plane, and this packing is the optimal two-dimensional hexagonal



(a) Packing layer A

(b) First layer A, second layer B

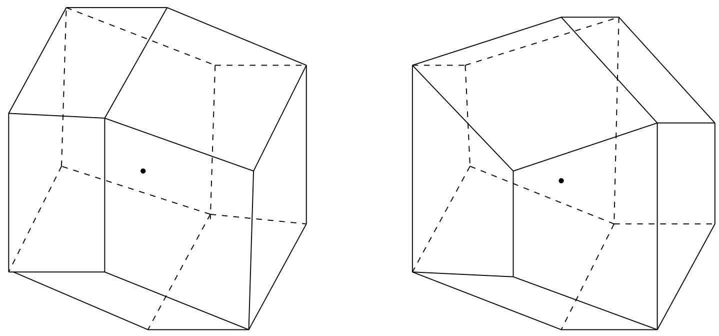
Fig. 3 Hexagonal packing layers viewed from above

circle packing. Now one can fit a second identical layer of spheres on top of this layer, so that the spheres fit as deep as possible in indentations between the spheres in the first layer. If we mark one sphere as the center of the first layer, there are two possible ways to do this: each such packing occupies 3 of the 6 holes formed by a hexagon and there are two choices. One can repeat this in the layer below the first layer, and continue adding layers in this fashion, making a choice at each layer. All such packings have the same limiting density. If the stacked layers are viewed vertically from above, then spheres in all the layers can be seen to line up in three possible positions of the hexagonal lattice, which we can label A, B, C, with the packing layer with the marked sphere labelled A. These are pictured in Figure 3. Note that the centers in B and C layers are “deep holes” in layer A.

In effect the choices at each layer of the packing can be labelled using letters $\{A, B, C\}$, with no two consecutive letters the same, to uniquely label a packing (with a single marked sphere serving as origin) with a doubly infinite string of such letters. It can be shown that two packings are essentially the same (up to a Euclidean motion) if their doubly infinite strings of letters can be lined up to agree. Two of these packings are especially nice, the *face-centered-cubic lattice packing (FCC packing)* and the *hexagonal close packing (HCP packing)*, described further below. The FCC packing corresponds to a repeating pattern ABCABCABC... while the HCP packing corresponds to a repeating pattern ABABAB... (or ACACAC... or BCBCBC...). The FCC packing was found by Kepler, while the HCP packing was first described by Barlow [3] in 1883; cf. Coxeter [11, Sec. 22.4]. These packings are described at the beginning of Hales [24] (in this volume).

The collection of “optimally dense” packings just described are locally optimal in a very strong sense. A *Voronoi domain* or *Voronoi cell* around a given sphere center is the set[5] of all points in space closer to that sphere center than to any other sphere center. In saturated packings all Voronoi cells are polyhedrons. For the packings just described, all

[5] This defines the interior of the Voronoi domain. The Voronoi domain itself is the closure of this set, adding boundary points which are certain points equidistant to two sphere centers.



(a) Rhombic Dodecahedron

(b) Trapezoidal Dodecahedron

Fig. 4 Voronoi Domains of (a) Type 1 (FCC) (b) Type 2 (HCP)

Voronoi domains consist of one of two shapes, each having 12 faces and 14 vertices [15, p. 173]. (See Figure 4.)

These are:

Type 1. Rhombic Dodecahedron (all 12 faces are rhombi).

Type 2. Trapezoidal Dodecahedron (6 faces trapezoids, 6 faces rhombi)

Here Type 1 occurs for all Voronoi cells in the middle layer of consecutive layers labelled ABC, ACB, BAC, BCA, CAB, CBA, while type 2 occurs for all Voronoi cells in consecutive layers labelled ABA, ACA, BAB, BCB, CAC, CBC. Voronoi cells of types 1 and 2 are known to have the same volume and surface area. For each of these cells, the ratio of the volume of each sphere to the volume of the Voronoi cell containing it is exactly $\frac{\pi}{\sqrt{18}}$. Thus in these packings optimality holds locally; that is, it is attained simultaneously in each Voronoi cell. In the FCC (Face-Centered Cubic) lattice packing, all Voronoi cells are of type 1. In the HCP (Hexagonal Close Packing) packing, all Voronoi cells are of type 2. (See [58].) In all remaining packings above, Voronoi cells of both types occur in the packing. The packings can be told apart by their different arrangements of types 1 and 2 Voronoi cells. All these packings have at least a two-dimensional lattice of translational symmetries; the FCC and HCP packings, and countably many others, have a full three-dimensional lattice of translational symmetries. From the mathematical viewpoint, as an optimization problem, this means that there are (at least) two different local optima, represented by the two types of Voronoi cells above.

A fourth difficulty is that the optimization problem to maximize density is a priori an infinite-dimensional problem: one has infinitely many spheres to pack. Describing a packing requires infinitely many variables: three variables each for the coordinates of each sphere center. Approaches to make headway with this problem seek to prove stronger results which only involve finite-dimensional optimizations that encode local conditions. This is the "local density inequality" approach described below and in the next section. These approaches in effect assign, by some recipe, to each sphere in a sphere packing a (weighted) sum of the covered and uncovered volume near that sphere center. This recipe is "local" in that the weighted sum for a given sphere center is completely determined by the locations of all spheres in the sphere packing nearby, within a fixed distance C of the given sphere center. When the recipe quantities are added up over all spheres, it should count all volume with weight 1. If so, then an upper bound on the weighted area will give an upper bound on global sphere packing density. We say such a local density inequality is "optimal" if it will produce the upper bound $\frac{\pi}{\sqrt{18}}$ for the sphere packing density.

A fifth difficulty is that it is not clear that "optimal" local density inequalities exist, and if they do exist they (apparently) cannot be very simple. In support of this view, the two most natural ways to locally partition space attached to a sphere packing yield local density inequalities that are not "optimal," as we now explain.

The first natural way to partition space corresponding to sphere centers is to divide it into *Voronoi cells* around these points, as described above. The region assigned to each sphere center is its Voronoi cell. However, it is known that an arrangement of 12 unit spheres touching a given unit sphere with their sphere centers being the vertices of a regular dodecahedron yields a Voronoi cell that is a regular dodecaheron of inradius 1, having a ratio of covered to uncovered volume approximately 0.754697, which exceeds $\frac{\pi}{\sqrt{18}} \approx 0.74048$.

A second natural partition of space is the *Delaunay tessellation*, which is a partition of space $\mathbb{R}^3$ into simplices (tetrahedra) having vertices among the sphere center points. We describe it further below. Such Delaunay simplices have their vertices at four sphere centers, so that their edge lengths are necessarily 2 or greater. Now we can cut up these tetrahedra, for example by barycentric subdivision, and assign parts of the uncovered area to each of the four spheres at the corners of the tetrahedron. Each sphere is now assigned certain regions associated to each of the Delaunay simplices for which its center is one vertex. Hales notes ([24, p. 13] in this volume) that there is a Delaunay tessellation having an individual tetrahedron with volume of covered to uncovered volume of 0.78469. Even if one sums over all Delaunay tetrahedra associated to a given sphere center, there are examples of local configurations with density exceeding the Kepler bound. An original approach of Hales in the early 1990s, described in [17], [18], based on the Delaunay tessellation, ground to a halt due to combinatorial difficulty (cf. Hales [24, Sec. 2.1]).

The failure of these two natural decompositions to be "optimal" indicate the necessity to consider more complicated inequalities in which the region assigned to a sphere center will need in certain cases to "borrow" an excess of uncovered volume from some spheres nearby to this sphere. This raises the spectre that perhaps the distance over which volume has to be "borrowed" is in fact unbounded. If so, no "optimal" local density inequality exists. One would have instead an infinite sequence of such local inequalities, taken over larger regions, each giving a better upper bound, tending in the limit to $\frac{\pi}{\sqrt{18}}$.

A sixth difficulty, then, is that of designing candidate local density inequalities that may be "optimal." If an "optimal" local inequality does exist, then it can in principle be verified by a finite computation, i.e., it comprises a finite dimensional non-linear optimization problem over a large number of variables, specifying the possible locations of sphere centers in a ball of radius C around a given sphere center. The size of the problem rapidly goes up as the distance from the initial sphere increases. There can be 12 spheres touching a given sphere, and the next layer of spheres can contain more than 30 spheres. There now arises a psychological difficulty, which is that the "optimality" of the local density inequality is *only certified after the fact*, when a proof is found. This means that one must first do a very large amount of work, with the downside risk of eventually determining that the inequality is not optimal. Therefore, one would like to increase one's confidence in advance that the inequality considered is likely to be "optimal," by making various preliminary experimentation and computer checks, before attempting a full scale proof.

There are now several reasonable candidates that have been put forward to be "optimal" local density inequalities. These include one proposed by László Fejes-Tóth in 1953 [15] and again in 1964 [16], one proposed by W.-Y. Hsiang [37] in 1993, and several different such inequalites proposed by T. C. Hales, both alone and with S. P. Ferguson. The original local density inequality of Fejes-Tóth considers weighted averages of Voronoi cells of nearby sphere centers, and that of Hsiang is similar in spirit. In 1991 Hales [17], [18] gave an approach proposing "optimal" local inequalities based on the Delaunay tessellation, with some Voronoi cell correction terms. In 1994 and later, Hales [20], [21], and Hales and Ferguson [25] formulated various candidate "optimal" inequalities using a hybrid decomposition of space employing elements of both the Delaunay tessellation and the Voronoi domain tessellation. The analysis of Hales and Ferguson suggests that there is a large class of such inequalities that will be "optimal," and there is some flexibility in formulating them. These inequalities will however have quite complicated "scoring" functions.

A seventh, and most crucial difficulty for "optimal" local density inequalities is the computational size of the problem. Each separate local density inequality asserts an upper bound to an enormously complicated finite-dimensional nonlinear optimization problem. This problem may simply be too large to be computationally feasible. It involves finding the global maximum of a highly non-linear function, over a high dimensional space of possible local configurations, enumerated by the location of the centers of all the nearby spheres. Further, this seems to involve on the order of 40 spheres, so that the problem has approximately 120 dimensions. General nonlinear problems of this dimensionality, with no additional structure, are too large to solve at present. In addition, the partitioning of space typically results in many combinatorially distinct configurations, so that the space of all configurations has a huge number of connected components. The complexity of the problem will also depend on the landscape of local optima, i.e., how many there are and how close they are to the global optimum. Therefore the local density inequality to prove needs to be carefully engineered to have extra properties that will simplify the computations. One of the key features of the Hales and Ferguson approach is to design such a "local density" inequality, allowing it to be complicated and inelegant, in order to make it have properties that reduce the needed computations. Their "Formulation" paper fills 49 pages to describe the partition of space and "scoring" function comprising the local density measure (see [25] in this volume). The final proof is an intricate blend of theory, in part needed to simplify the problem, together with large-scale computations.

In the Hales-Ferguson proof, in this volume, the local density inequality is designed so that the regions to be analyzed split up into (non-interacting) cones whose contributions to the "score" can be evaluated separately and then added. It is also necessary for the inequality to be proved to have a finite number of (sets of) global maxima. Then the proof will logically split into two parts: an exact analytic treatment in the neighborhood of each global maximum, proving this property, and then a cruder analysis of the remaining (very large) part of the space of possible configurations, establishing inequalities showing a bound strictly below the optimum. The second part is the vast bulk of the proof. In the actual proof there are two types of global maxima, corresponding to the two types of Voronoi domains above. (In the proof these comprise a finite number of different local configurations, called decomposition stars, that attain the maximum.) One might think that the difficulty in the proof is in checking near the global maxima, but in fact the greatest part of the difficulty is in dealing with the inequalities on the huge remaining portion of the configuration space. In fact, it is extremely hard to describe this configuration space in a way that is suitable for analysis and computation. The problem is reduced to many smaller nonlinear optimization problems, which are relaxed to convex optimization problems having a single maximum, that can be bounded above. The devil is in the details.

3 Local Density Approach: History

The first person to formulate a local density approach was László Fejes-Tóth [15], who made many deep insights into the sphere packing problem. In 1953 [15, pp. 174–181] he indicated the possibility of proving the Kepler conjecture by relating it to a nonlinear optimization problem over a compact set. More precisely, he proposed a specific inequality that might hold. This inequality involved averages of Voronoi cell volumes of close neighbors, and would imply Kepler's conjecture. In 1964 he restated this inequality [16, pp. 295–299] and added:

> Thus it seems that the problem can be reduced to the determination of a minimum of a function of a finite number of variables, providing a programme realizable in principle. In view of the intricacy of this function we are far from attempting to determine the exact minimum. But, mindful of the rapid development of our computers, it is imaginable that the minimum may be approximated with great exactitude.

Such optimization problems can generically be termed "local density inequalities" as they involve some weighted measure of density associated to a neighborhood of each sphere center in a packing. Establishing a local density inequality yields a result which is stronger than the Kepler conjecture, since for optimality it requires the existence of a packing maximizing the local inequality simultaneously at every sphere center.

In particular, each local density inequality asserts a different mathematical result. This is so, although each may imply Kepler's conjecture as a corollary. Assuming that the associated optimization problem has a finite set of isolated global maxima, and that the local minimality can be verified analytically in an open neighborhood of each of these maxima, then in principle the inequality can be proved by computer elsewhere using interval arithmetic. However the size of the resulting optimization problems seemed far beyond the range of what could be solved by a computer at that time, or even now, without new ideas to reduce the complexity of the problem.

A precise framework for local density inequalities is given in the next paper in this volume [45]. Local density is measured at a sphere center by a "score function" which assigns to it the volume nearby counted with certain penalties and credits. These penalties and credits must have the property that when summed over all sphere centers they cancel out, so that volume is then counted with weight at most 1 everywhere.

The two major starting points in designing such local density "score functions" are the Voronoi tessellation (also called the Dirichlet tessellation, and, in physics, Wigner-Seitz cells) and Delaunay tessellation. Given centers of spheres in a saturated sphere packing, the *Voronoi tessellation* partitions space into Voronoi cells, which for each sphere center is the (closure of) the set of points in space closer to that sphere center than any other sphere center; this region contains the unit sphere. For saturated sphere packings with unit radius spheres, a Voronoi cell is a finite polyhedron having no point at distance exceeding 2 from the sphere center.

A *Delaunay tessellation* for a set of points in space $\mathbb{R}^3$ is a subdivision of space into simplices (tetrahedra) such that the sphere circumscribing each tetrahedron has no points of the set in its interior. Delaunay [12] showed such tessellations exist. They are efficiently constructible; cf. Watson [59]. The resulting tessellation is topologically dual to the Voronoi tessellation, i.e., vertices of the Delaunay tessellation correspond to regions in the Voronoi tessellation, edges in the Delaunay tessellation correspond to faces of Voronoi cells, faces in the Delaunay tessellation correspond to edges in Voronoi cells, and the interior of a Delaunay cell corresponds to a vertex of a Voronoi cell lying in its interior. The Delaunay tessellation is unique for points in general position (no 5 points at equal distance from one point) but may be non-unique otherwise. In a finite region there are only a finite number of choices for the Delaunay simplices.

As noted above, the Voronoi and Delaunay partitions of space do not directly yield "optimal" local density inequalities. Further averagings of space are required. The approaches to local density inequalities proposed by Fejes-Tóth and by Hsiang [38], [41] are based on averaging of Voronoi cell densities of regions near a given sphere. Hales proposed a local density inequality based on Delaunay tessellation in 1991 [17, Sec. 4], which involves a functional defined on a compact space of (abstract) Delaunay stars. His original study indicated that this inequality is near optimal, and that to get an optimal inequality some correction terms had to be added; he proposed terms that add some information associated to Voronoi cells. He then formulated a candidate "optimal" inequality based on a superposition of Delaunay and Voronoi tessellations [17]. In work after 1994 Hales [20], [21] (papers in part IV of this volume) formulated "hybrid" local inequalities which decompose space starting from a Delaunay-like decomposition but use a Voronoi-type decomposition on that part of space on which the Delaunay simplices do not have prescribed shapes. The local density inequality of Hales and Ferguson [25] uses a decomposition of space starting from a set of simplices with corners at four sphere centers, having allowed shapes, without imposing the "empty sphere" condition of the Delaunay tessellation. (They did this to make their local conditions not interact across long distances.)

In 1990 Wu-Yi Hsiang proposed a local density inequality using averages of Voronoi cell volumes and announced a proof of it; this would prove the Kepler Conjecture. This announcement received much publicity, see Szpiro [56, pp. 144–152]. One idea of Hsiang was to use spherical geometry rather than Euclidean geometry to obtain the sphere packing bounds: some of the uncovered regions could be treated as spherical triangles, with volumes computed using spherical trigonometry. A paper was subsequently submitted to *Annals of Mathematics*. The reviewers there uncovered some difficulties in the proof. In response to the reviewer's criticisms, Hsiang did not resubmit his paper to the *Annals of Mathematics*, and instead submitted it elsewhere. He eventually was able to publish a revised 93 page paper in 1993 in the *International Journal of Mathematics* (Hsiang [38]). This journal records that the paper was received 17 November 1992, revised 9 March 1993, and was in print by the end of 1993, a rapid review for a paper of its length and complexity.

Objections were voiced to the published Hsiang proof by Conway, Hales, Muder and Sloane [9], and were detailed by Hales [19] in 1994 in the *Mathematical Intelligencer*. These objections applied not only to specific results claimed in the proof, but also asserted that some methods of argument used were invalid. To this critique Hsiang [40] gave a