Introduction to the Flyspeck Project

Thomas C. Hales

Math Department, University of Pittsburgh

Abstract. This article gives an introduction to a long-term project called *Flyspeck*, whose purpose is to give a formal verification of the Kepler Conjecture. The Kepler Conjecture asserts that the density of a packing of equal radius balls in three dimensions cannot exceed $\pi/\sqrt{18}$. The original proof of the Kepler Conjecture, from 1998, relies extensively on computer calculations. Because the proof relies on relatively few external results, it is a natural choice for a formalization effort.

1 Introduction

In 1998, Sam Ferguson and I presented a proof of the Kepler Conjecture, the assertion that no packing of congruent balls in three dimensions has density greater than the face-centered cubic packing. That density is $\pi/\sqrt{18}$, or about 0.74.

The face-centered cubic packing is the pyramid arrangement frequently used to stack oranges at fruit stands and cannonballs at war memorials. The face-centered cubic packing is by no means the only packing that achieves the optimal density. There are many trivial modifications, such as shifting an entire rows of balls in the packing, that do not affect the density.

The 1998 proof of the Kepler Conjecture is about 300 pages of text and relies on computer calculations based on about 40,000 lines of computer code. The proof was under review by a team of referees for five years before finally being accepted for publication. Even then, the team of referees concluded that they could not completely certify the proof. As an editor wrote me, "They checked many local statements in the proof, and each time they found that what you claimed was in fact correct. Some of these local checks were highly non-obvious at first, and required weeks to see that they worked out. The fact that some of these worked out is the basis for the 99% statement of Fejes Tóth that you cite." (G. Fejes Tóth, the chief referee, had previously written me that he was 99% certain of the correctness of the proof.)

In response to the lingering doubt about the correctness of the proof, at the beginning of 2003, I launched the *Flyspeck* project, whose aim is a complete formal verification of the Kepler Conjecture. In truth, my motivations for the project are far more complex than a simple hope of removing residual doubt from the minds of few referees. Indeed, I see formal methods as fundamental to the long-term growth of mathematics.

Dagstuhl Seminar Proceedings 05021
Mathematics, Algorithms, Proofs
http://drops.dagstuhl.de/opus/volltexte/2006/432

The name *FLYSPECK* is derived from the acronym FPK, for the Formal Proof of the Kepler Conjecture. The word flyspeck can mean to examine in detail, which seems quite appropriate for a formalization project.

The response from the formal theorem proving community has been supportive and enthusiastic. T. Nipkow, R. Zumkeller, S. Obua, and I gave presentations on various aspects the Flyspeck project during an afternoon session at the 2005 Dagstuhl meetings. A Flyspeck group meeting at the same conference was attended by J. Avigad, T. Hales, T. Nipkow, S. Obua, R. Solovay, F. Wiedijk, R. Zumkeller. A forthcoming thesis by G. Bauer attacks the problem of verifying the correctness of the computer code to classify tame plane graphs.

A very rough estimate, based on F. Wiedijk's rough estimate that it takes about one week to formalize one page of textbook mathematics, is that the Flyspeck project will require about 20 work-years. A project web page appears at [13] and a discussion forum at [6].

2 The Kepler Conjecture

There are now a number of written accounts of the proof of the Kepler Conjecture from various points of view. History appears in [11]. An introduction for a broad mathematical audience appears in [12]. A summary of some of the main algorithms and ideas appears in [10]. J. Lagarias describes the general framework of how the problem can be reduced to an optimization problem in a finite number of variables [15]. The proof appears in abridged form in [8] and in unabridged form in [9].

Here we mention just a few of the main ideas in simplified form, with just enough detail to permit a discussion of the Flyspeck project.

2.1 Linearization

Density is not linear. The density within the union of two sets is a weighted average of the densities within the two individual sets. A simple observation is that when proving a bound on densities

$$\mathit{density} = \frac{A}{B} \leq \frac{\pi}{\sqrt{18}}$$

the denominator can be cleared

$$B\left(\frac{\pi}{\sqrt{18}}\right) - A \geq 0. \tag{1}$$

In this form, the inequality to be proved is linear: the term A (giving the volume of the balls in a region) the term B (giving the full volume of the region) are both additive over a disjoint collection of regions.

2.2 Reduction to a Finite Dimensional Problem

Space can be partitioned into countably many regions such that

$$B_i(\frac{\pi}{\sqrt{18}}) - A_i \geq f_i, \tag{2}$$

for some fudge factor f_i depending on the region i. The fudge factors are chosen so that in an appropriate asymptotic sense they sum to zero:

$$\sum_i f_i \sim 0. \tag{3}$$

By summing Inequality 2 over all the regions, and using the fact that the fudge factors become negligeable, we obtain Inequality 1.

The proof of the Kepler Conjecture then reduces to a discrete collection of Inequalities 2. We can do even better. The fudge factors are chosen so that there is a smooth extension to the form

$$B(x)(\frac{\pi}{\sqrt{18}}) - A(x) \geq f(x), \tag{4}$$

for x in some compact set $K \subset \mathbb{R}^n$. The functions $B(x)$, $A(x)$, and $f(x)$ are continuous on K, and this one inequality over K implies that for the individual regions i. In this way, the Kepler Conjecture reduces to proving a nonlinear inequality on a compact subset of Euclidean space. Let us call Inequality 4 the *main inequality.*

2.3 Planar Graphs

Each point in the compact set K can be described geometrically as a cluster of a finite number of spheres around one fixed sphere, with coordinates chosen so that the fixed sphere is centered at the origin. We fix the radius of the spheres to be 1, so that the distance between the centers of the spheres is at least 2.

We can describe the combinatorial structure of a cluster of spheres by a planar graph as follows. To discretize the problem, we fix a cutoff parameter T greater than 2 whose exact value does not concern us.[1] We define a vertex in the planar graph for every sphere in the cluster whose distance from the fixed central sphere is at most T. We connect two vertices by an edge if they correspond to two spheres whose centers are separated by distance at most T. If the constant T is small enough, it can be shown that the resulting graph is planar.

A major part of the proof of the Kepler Conjecture is an investigation of the properties that a planar graph must have, if it is associated with a cluster of spheres that gives a potential counterexample $x \in K$ to the the main inequality. Eventually, a large list of such properties are established. Any planar graph satisfying these properties is referred to as a *tame graph.* A major theorem states that a counterexample to the main inequality is associated with a tame

[1] The value is rather arbitrarily set at 2.51.

graph. (But the converse is not true: a tame graph does not necessarily give a counterexample to the main inequality.)

The question then arises: what is the classification of tame planar graphs? The original proof of the Kepler Conjecture produced a list of several thousand graphs and a computer proof that every tame planar graph is isomorphic to one on this list. T. Nipkow has reduced this list considerably. G. Bauer is tackling the formal verification of the classification of tame plane graphs.

2.4 Linear Programming

The compact set K can be partitioned into a union

$$K = \sqcup_G K_G$$

with G running over the set of planar graphs. On each subset K_G, there is considerable information about the combinatorial structure of the clusters of spheres. The verification of the main inequality breaks into a finite series of separate verifications, one for each K_G, as G runs over tame plane graphs, up to isomorphism.

This verification is nonlinear but it can be relaxed to a series of linear programs. The idea is to bound a nonlinear function

$$B(x)\frac{\pi}{\sqrt{18}} - A(x) - f(x)$$

from below by a collection of linear inequalities. A lower bound on the collection of linear inequalities is then a lower bound on the nonlinear function. To minimize a linear system of inequalities, linear programming methods can be used. There are about 100,000 such inequalities, each involving some 200 variables and about 2000 constraints. Eventually, the nonlinear inequality is proved as a consequence of these large systems of linear programs. S. Obua is undertaking the project of formally verifying the linear programs that arise in the proof of the Kepler Conjecture.

There is a separate difficult problem of verifying that the collection of hyperplanes give strict lower bounds to the nonlinear function. In the original proof, these verifications were done by computer using interval arithmetic to control round-off error. R. Zumkeller is developing the tools to prove such inequalities within a formal system.

3 Extracting Algorithms from the Proof of Kepler

G. Gonthier, in his lecture announcing the formal proof of the Four Color Theorem, described the trade-off between the generality of deduction and the efficiency of computation. The designers of formal proofs should try to organize the proof to benefit from the efficiencies of computation [5].

A different defense of algorithms in pure mathematics was recently made by I. Daubechies in her 2005 Gibbs lecture. She argued that the constraints imposed by the finiteness conditions of algorithms lead to the discovery of elegant structures in pure mathematics [1].

With these two sources of motivation, we can ask whether there are parts of the 300 page written proof of the Kepler Conjecture that could be better implemented by general computer algorithm? There are reasons to be optimistic. In the original 1998 proof, I tried to avoid computers, except when paper calculations would have been out of the question. As a result, many results are proved by hand that could have been done in a simpler manner by computer. Also, anyone familiar with the written proof is aware that the same arguments tend to be used again and again with minor variations. This gives hope that some general tactics can be designed that are capable of proving many of these results at once.

We give one concrete suggestion, based on tensegrities.

3.1 Tensegrities

For our purposes, a tensegrity is a collection of constraints on a finite set of points in Euclidean space. The constraints come in two sorts: struts and cables. A *strut* is a lower bound constraint between a pair of points and a *cable* is an upper bound constraint between a pair of points. The picture to imagine is a collection of tubes strung together by cables. The ends of a cable can come arbitrarily close together, but cannot be separated farther than the extended length of the cable. The ends of a strut can be arbitrarily far apart, but no closer than the compressed length of the strut.

As an aside, we mention that *tensegrity* is an architectural term coined by Buckminster Fuller [4]. Sculptures made from cables and tubes, often appearing to defy gravity, can often be described as tensegrities. (For examples, search for "tensegrity" in Google images.)

Bob Connelly first pointed out the link between tensegrities and the proof of the Kepler Conjecture. The centers of a finite cluster of spheres provide the points of the tensegrity. The constraint that no two spheres overlap gives a lower bound strut between each pair of points. The truncation parameter T, discussed above, gives cable constraints between various pairs of points.

Let us say that a tensegrity exists, if there is a collection of points in Euclidean space (in three dimensions) satisfying the system of constraints. A unifying feature of the 300 page proof of the Kepler Conjecture is that many of the geometrical statements can be expressed as existence problems for tensegrities and for slight generalization of tensegrities. A list of simple examples appears in [10].

By picking coordinates for the points of the tensegrity, each existence problem can be expressed as a statement in the elementary theory of the real numbers. Thus, by quantifier elimination methods, such problems are always decidable. Such decision procedures have recently been implemented within HOL Light [18]. Unfortunately, many of these problems involve up to about 20 variables. Problems of this size are generally regarded as outside the practical range of quantifier elimination methods.

To express the full range of existential problems in geometry that arise in the proof of the Kepler Conjecture, we can create a small decidable theory with more than just cables and struts. Let d be the Euclidean distance function on $\mathbb{R}^3$. Let η be the circumradius of a triangle as a function of the three vectors giving the vertices. To generalize the notion of tensegrity, we define predicates:

$\mathit{cable}\ u\ v\ r = (d\ u\ v \leq r)$
$\mathit{strut}\ u\ v\ r = (r \leq d\ u\ v)$
$\mathit{strict\_cable}\ u\ v\ r = (d\ u\ v < r)$
$\mathit{strict\_strut}\ u\ v\ r = (r < d\ u\ v)$
$\mathit{face\_cable}\ u\ v\ w\ r = (\eta\ u\ v\ w \leq r)$
$\mathit{face\_strut}\ u\ v\ w\ r = (r \leq \eta\ u\ v\ w)$

So far this just allows a variant on cables and struts with strict inequality constraints, and a variant that allows something analogous to cables and struts giving bounds on the circumradius formed by three points. We can add variants *strict_face_cable* and *strict_face_strut* that make the inequalities strict. In addition to these, we need predicates: *convex_meet, affine_meet, convex_affine_meet, cone, strict_cone, circular_cone,* and *strict_circular_cone.*

The predicate *convex_meet* is true if the convex hull of one list of points meets the convex hull of a second list of points. The predicate *affine_meet* works similarly for affine sets, and *convex_affine_meet* is the mixed version that tests for a nonempty intersection of a convex hull with an affine hull. The predicate *cone* tests whether a given point lies in the closed cone generated over a second point by a finite list of points. The predicate *strict_cone* is similar for the open cone. The predicate *circular_cone* can be defined in terms of the dot product as

$\mathit{circular\_cone}\ u\ v\ r\ w =$
$(d\ w\ u) * (d\ v\ u) * (d\ v\ u) \leq \&2 * r * ((w - u) \cdot (v - u))$

Geometrically, it states that w lies in the infinite circular cone with vertex u and that passes through the circle at distance r from u and v.

The remarkable little secret about the proof of the Kepler Conjecture is that virtually all the low-dimensional geometry[2] can be expressed as decidable existential questions in terms of this short list of predicates, and other secondary predicates defined in terms of these. A generalized tensegrity can be defined as a finite collection of constraints expressed in terms of these predicates.

To give one concrete example of a generalized tensegrity, consider a tetrahedron in which each pair of vertices is connected by a cable of prescribed length. Does there exist a point in the interior of the tetrahedron that is connected by a strut of prescribed length to each of the four vertices of the tetrahedron? To express this question, we make use of cables and struts, but we also add the condition that the point lies in the interior of the tetrahedron, which can be expressed as a generalized tensegrity (using the *convex_meet* of two lists of points), but not in terms of cables and struts alone.

The proofs of most of these geometrical lemmas about generalized tensegrities are based on a small number of arguments used in a repetitive way. The

[2] We are excluding from discussion here the many properties of a simplex that are established by interval arithmetic.

generic proof of the nonexistence of a generalized tensegrities goes something like this. Assume for a contradiction, that the generalized tensegrity exists as a configuration P. The configuration P is stretched out, by a series of deformations, in such a way that all the constraints of the tensegrity continue to hold, but in such a way that as many cables as possible are stretched to the limit. Eventually, as P is stretched out, so many constraints hold that the configuration becomes rigid. At this stage, a coordinate system is introduced to measure the distances between pairs of points. It is found that one of the constraints is violated. Since the deformations did not violate any new constraints, the original configuration P must have violated a constraint. This is contrary to the initial hypotheses that P is realizes all the constraints of the tensegrity. Hence no such generalized tensegrity exists.

There is also a dual version that shrinks the generalized tensegrity until the struts become binding constraints that produce a rigid configuration.

An examination of the proof of the Kepler Conjecture shows that this short discussion of generalized tensegrities gives a generic description of a large part of the text. This provides real hope that generic tools can be designed to speed up the formalization process.

4 Prerequisites and Progress to Date

One nice feature of the Flyspeck project is that there are relatively few prerequisites: the proof of the Kepler Conjecture ultimately rests on (a long series) of relatively simple results in discrete geometry.

4.1 Jordan Curve Theorem

The deepest topological result that is used in the proof of the Kepler Conjecture is the Jordan curve theorem. The Jordan curve theorem is needed in the discussion about planar graphs.

While it is probably true that it would be possible to get by with a weak form for the Jordan curve theorem, say for polygonal closed curves form by geodesic arcs on a unit sphere, the Jordan curve theorem has been viewed for a hundred years "as a most important step in the direction of a perfectly rigorous mathematics" [24]. Thus, it seems worthwhile to give a formal proof of the theorem in its full generality.

In January 2005, I completed a formal proof of the Jordan curve theorem. The statement in HOL Light takes the following form.

```
let JORDAN_CURVE_THEOREM = prove_by_refinement(
  `∀C. simple_closed_curve top2 C ⇒
    (∃A B. top2 A ∧ top2 B ∧
      connected top2 A ∧ connected top2 B ∧
      (A ≠ ∅) ∧ (B ≠ ∅) ∧
      (A ∩ B = ∅) ∧ (A ∩ C = ∅) ∧ (B ∩ C = ∅) ∧
```

```
      (A ∪ B ∪ C =euclid 2))`,
   ···);;
```

A literal translation of this HOL Light code into English is as follows: [Let *top2* be the metric space topology on $\mathbb{R}^2$.] Let C be a simple closed curve, with respect to *top2*. Then there exist sets A and B with the following properties: A and B are open in the topology *top2*; A and B are connected with respect to the topology *top2*; A and B are nonempty; the sets A, B, and C are pairwise disjoint; and the union of A, B, and C is $\mathbb{R}^2$.

Or more idiomatically, a Jordan curve C partitions the plane into the three sets A, B, and C itself, where A and B are nonempty connected open sets.

The starting point for the formal proof is an elegant proof by C. Thomassen [22]. Thomassen's proof is based on the fact that the Jordan curve theorem is essentially equivalent to the nonplanarity of the complete bipartite graph K_{33} (the graph obtained by joining each of one set of three vertices to each of a second set of three vertices). The proof opens with a construction that shows how to produce a planar embedding of K_{33} from any counterexample to the Jordan curve theorem. A graph approximation theorem is then proved that takes any planar embedding of K_{33} and constructs from it a planar embedding with piecewise linear edges. Finally, a simple case of the Jordan curve theorem for polygons is proved, and this simple case is used to prove that no piecewise linear planar embedding of the K_{33} graph exists. In short, a counterexample to the Jordan curve theorem leads to a contradiction; and the theorem is established.

For added simplicity in the formal proof, the theorems about polygons and piecewise linear edges are modified so that all linear segments are parallel to the coordinate axes and each linear segment has integer length.

4.2 Multivariable Calculus

From analysis, the primary need (beyond what is already present in the HOL Light system) is multivariable calculus. Chapter 8 of Loomis and Sternberg's *Advanced Calculus* contains all of the foundational material that is needed from integral calculus [17]. This is an honors-level undergraduate textbook that includes the construction of a measure, basic results about null sets, a simple form of Fubini's theorem, and a change of variables formula. The full strength of Lebesgue integration is not needed because the volumes that arise in the proof of the Kepler Conjecture tend to be relatively simple shapes (such as polytopes or the intersection of balls with half-spaces).

It seems possible to avoid integration on general manifolds. Most integrals on surfaces are planar or subsets of the unit sphere. Integration on the unit sphere can be replaced by integration on the unit ball, using the fact that the volume of the unit ball is 1/3 the area of a sphere.

One very encouraging development in the most recent distribution of HOL Light is the amount of multivariable calculus that J. Harrison has added [14]. This gives hope that significant tools from multivariable calculus may soon be available.

In addition to the abstract theorems such as Fubini and a change-of-variables formula, there is also a need for practical tactics to compute explicit areas and volumes. For example, give a tactic that takes a polytope (presented say as the convex hull of a given list of vertices) and returns a theorem asserting its volume. Or prove that that the volume of a unit ball is $4\pi/3$. A long list of explicit integrals such as these is needed in the proof of the Kepler Conjecture.

4.3 Euclidean Geometry

In the 1998 proof of the Kepler Conjecture, there was no reason to avoid well-known facts in geometry. For example, the original proof makes use of the isoperimetric inequality for spherical polygons. (Used in [7], and proved at [2].) But it seems that the use of the isoperimetric inequality is inessential, and that the formal proof should avoid it altogether. Similar remarks apply to Lexell's theorem (which is used casually in various places and proved in [3, p.22]): it seems better to avoid it than to give a formal proof.

The proof of the Kepler Conjecture uses many of the standard identities of spherical trigonometry, such as the spherical law of cosines. Fortunately, the standard distribution of HOL Light contains many of the most common identities of planar trigonometry. It will be a relatively small additional effort to go from a formalization of planar trigonometry to spherical trigonometry. An undergraduate-level treatment can be found in [20].

There are also a large number of basic results about planar and solid geometry that are used and must be formally proved. Here I include basic results about convexity and affine spaces, as well as basic formulas in geometry. For example, prove Heron's formula for the circumradius of a triangle, the Cayley-Menger formula for the volume of a tetrahedron as a function of its edge-lengths, the Harriot-Euler formula for the area of a spherical triangle as a function of its angles, and Lagrange's formula for the area of a spherical triangle as a function of its vertices [16, pp.331–359]. (Or perhaps Lagrange's formula can be avoided altogether by making greater use of the Harriot-Euler formula.)

I do not consider these formulas in planar and solid geometry to be major prerequisites for the formal proof in the way that the Jordan curve theorem and multivariable calculus are, simply because such formulas can be generally be obtained by conventional means by brief calculations in a computer algebra system. On the other hand, there are many such formulas, so that they may require a significant cumulative effort to establish them formally.

4.4 Related Projects

There are a few problems in discrete geometry whose methods are closely related to the those that appear in the Kepler Conjecture. It might be worthwhile to give a formal proof of some of these other problems as a way of developing tools that will be useful to the Flyspeck Project.

The problem of 13 spheres In a famous discussion between Isaac Newton and David Gregory, they asked whether 13 spheres could be arranged tangent to a $14th$ sphere. Newton suspected that this was impossible; only a proof was lacking.

The problem of 13 spheres has a long history of dubious proofs. The first proper solution was given by K. Schütte and B. L. van der Waerden in 1953 [21]. A recent article by O. Musin presents the proof in a way that would make it an attractive intermediate step in the Flyspeck project [19].

Thue's bound The sphere packing problem in two dimensions was solved in by A. Thue [23]. Complaints have been made about Thue's original proof, but there are now several elementary proofs of this result. My favorite proof is described in [12]. As soon as multivariable calculus has been developed within a formal system, a short but valuable application of the theory of areas would be a formal proof of Thue's theorem.

References

1. I. Daubechies, Gibbs Lecture, "The interplay between analysis and algorithms", January 5, 2005.
2. L. Fejes Tóth, Isoperimetric problems concerning tesselations, Acta Math. Acad. Sci. Hungar 14 1963 343–351.
3. L. Fejes Tóth, Lagerungen in der Ebene auf der Kugel und im Raum, Springer-Verlag, 1972.
4. B. Fuller, Synergetics, MacMillan, 1975.
   Online at http://www.rwgrayprojects.com/synergetics/toc/status.html
5. G. Gonthier, lecture "The four color theorem in COQ" TYPES conference, Campus Thalès, Dec 17, 2004.
6. Google Groups Flyspeck, http://groups-beta.google.com/group/flyspeck.
7. T. Hales, Sphere Packings, I, Discrete Comput Geometry 17:1–51 (1997).
8. T. Hales, A Proof of the Kepler Conjecture (Annals Version),
   http://www.math.pitt.edu/˜ thales/kepler03/fullkepler.pdf
9. T. Hales, A Proof of the Kepler Conjecture (DCG Version),
   http://www.math.pitt.edu/˜ thales/kepler04/fullkepler.pdf
10. T. Hales, Some algorithms arising in the proof of the Kepler conjecture, Discrete and computational geometry, 489–507, Algorithms Combin., 25, Springer, Berlin, 2003.
11. T. Hales, An Overview of the Kepler conjecture, math.MG/9811071.
12. T. Hales, Cannonballs and honeycombs. Notices Amer. Math. Soc. 47 (2000), no. 4, 440–449.
13. T. Hales, The Flyspeck Project Fact Sheet,
    http://www.math.pitt.edu/˜ thales/flyspeck/index.html
14. J. Harrison, HOL Light, Jan 2005 version, preliminary release, available at
    http://www.math.pitt.edu/ thales/flyspeck/hol_light_050124.tar.gz.
15. J. Lagarias, Bounds for Local Density of Sphere Packings and the Kepler Conjecture, Discrete Comput Geom 27:165–193 (2002).
16. J.-L. Lagrange, *Oeuvres*, Vol 7, Gauthier-Villars, Paris, 1857–1892.

17. L. H. Loomis and S. Sternberg, Advanced Calculus, Addison-Wesley, 1968.
18. S. McLaughlin and J. Harrison, A Proof Producing Decision Procedure for Real Arithmetic, preprint 2005.
19. O. R. Musin, The kissing problem in three dimensions, preprint 2004.
20. P. Ryan, Euclidean and Non-Euclidean Geometry, an analytic approach, Cambridge University Press, 1986.
21. K. Schütte and B. L. van der Waerden, Das Problem der dreizehn Kugeln, Math. Annalen 125 (1953), 325–334.
22. C. Thomassen, The Jordan-Schönflies Theorem and the Classification of Surfaces, Amer. Math Monthly, Feb 1992, 116–130.
23. A. Thue, Om nogle geometrisk taltheoretiske Theoremer, *Forandlingerneved de Skandinaviske Naturforskeres* 14 (1892), 352–353.
24. O. Veblen, Theory on Plane Curves in Non-Metrical Analysis Situs, Trans. of the AMS, Vol 6, No 1 (Jan 1905), 83–98.