THE PROBLEM OF THE THIRTEEN SPHERES

In the Euclidean plane one can surround a unit circle by six other unit circles, all touching the central one. It is easy to see that six is already the maximal number of touching neighbors and moreover, such a configuration is unique (up to a rotation) and can be replicated to the whole plane, giving an infinite packing of unit circles in which every circle touches six neighbors. In the physical space, one can also ask the same kind of questions: how many unit spheres can simultaneously touch a central unit sphere? and what are those infinite packings of unit spheres such that in which every sphere touches the maximal touching neighbors? We will discuss the issues of the second question in next chapter. For the first question, twelve is clearly possible, as one can arrange those twelve neighbors touching the central unit sphere at the vertices of an inscribing regular icosahedron:

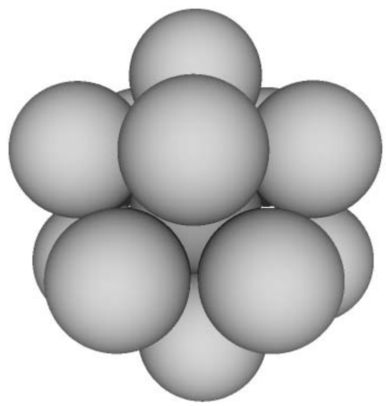

Figure 1.1

Unlike the case of circle packings, this symmetric arrangement cannot be replicated to the whole space. Moreover, it leaves a lot of space between the touching neighbors (as one can see that all the twelve touching neighbors are free to roll a little bit on the central unit sphere). It is natural to ask whether it is possible to rearrange them in a suitable way so that the space created can allow us to place a thirteenth touching neighbor on top of it. This is in fact a famous problem in classical solid geometry, namely, "the problem of the thirteen spheres":

> "*Is it possible to arrange thirteen unit spheres all touching a central unit sphere simultaneously*?"

Dated back to 1694 at the Christ Church in Oxford, there was a recorded discussion between David Gregory and Isaac Newton on the above issue. Newton believed that twelve should already be the maximal, but Gregory thought that thirteen might be possible. Their discussion ended without a proof or an example. When this recorded discussion was discovered later, it immediately became a famous problem in classical solid geometry, sometimes also referred as the "Newton's problem".

The problem of the thirteen spheres turned out to be very challenging, and even after almost two centuries of work, there were only some *incorrect* proofs such as [Ben], [Hop], [Gun]. The answer was finally settled to the *impossibility of thirteen touching spheres* in

1953 by Schütte and van der Waerden [SW–2], and later, in a two-page article by Leech in 1956 [Lee] where he sketched another proof of the *impossibility*. In 1993 Hsiang published a proof of the famous Kepler Conjecture [Hsi–1], and in developing the techniques on the estimations of total buckling heights, he gave another proof of "thirteen is impossible" as an application of his area estimation techniques [Hsi–2].

In this chapter we are going to study the above three proofs in chronological order. In retrospect, all the above three proofs rely on area estimates on certain kinds of graphs (or configurations) on the sphere, and the fundamental reason for the non-existence of thirteen touching neighbors is exactly that the total surface area of the unit sphere is not sufficient to accommodate such a graph or configuration with thirteen vertices. In other words, the (spherical) area of certain kinds of *hypothetical* configurations must exceed 4π, thus impossible to exist.

1 The proof by van der Waerden

In the paper [SW–2] there were, in fact, two proofs on the *impossibility of thirteen touching neighbors*. The first, simpler proof was due to van der Waerden and the second, similar but slightly more involved proof was due to Schütte. Schütte's proof was included there because firstly, it has the priority and secondly, they also believed that the method involved should be applicable to the study of other problems of "spherical codes", namely, the arrangements of certain number of points on the unit sphere with maximal separation. But in here we will only discuss the proof by van der Waerden.

Irreducible graphs on the minimal sphere

In their previous work [SW–1], Schütte and van der Waerden had already developed a system of tools to study the spherical code problems. They associated to each arrangement of points on the unit sphere a graph, by joining those pairs of points with minimal Euclidean separation (among pairwise of points) using straight line intervals, then normalized the length of such intervals to be 1 by a suitable scaling. What they wanted to do was to use this graph to decompose the scaled sphere into spherical equilateral polygons by radial projection, but the graph obtained might consist of just a single edge only, so they introduced the following concepts of "minimal sphere" and "irreducible graph":

Definition: The minimal sphere of N points is the sphere with minimal radius r, on which N points with the above Euclidean separation (≥ 1) can still be maintained.

Then they showed the existence of such a value r. In other words, one can assume that the points are already arranged with maximal separation (which is 1 in Euclidean distance, or $a = 2r\sin^{-1}\frac{1}{2r}$ in spherical distance) on the corresponding minimal sphere of radius r.

Definition: By a procedure "push-away" ("*Wegschieben*") we mean a movement of a point P on the sphere in such a way that if the (Euclidean) separation between P and Q is "> 1", it keeps the separation to be still "> 1" (but may change the value); if the separation between P and Q is "$= 1$", it increases the separation to be "> 1".

Definition: A connected graph is called irreducible if it does not allow any "push-away" movement.

Then they claimed that on the minimal sphere (of $N > 6$ points) the corresponding induced graph could be decomposed into irreducible graphs, and possibly with some isolated points. Moreover, an irreducible graph possess the following nice properties:

(a) All the angles in an irreducible graph are less than π;

(b) If the radius of the minimal sphere is at most 1, then an irreducible graph will decompose the sphere into triangles, quadrilaterals and pentagons only;

(c) For the case $N > 6$, no isolated point can lie in the interior of a triangle or a quadrilateral. For the case $N > 12$, no isolated point can lie in the interior of a pentagon.

Van der Waerden wanted to prove the impossibility of thirteen touching neighbors by showing that corresponding minimal sphere must have radius r greater than 1. So he assumed the contrary that $r \leq 1$, and therefore the induced graph would decompose the sphere into *equilateral* spherical triangles, quadrilaterals and pentagons *without isolated point*. He tried to obtain a contradiction based on the estimates on what he called the "angle-excesses" of those polygons, which are in fact "area-excesses" by a scaling to unit sphere, and assuming the side length a to be at least $\frac{\pi}{3}$.

Total sum of angle-excesses of a configuration

Set α to be the inner angle of an equilateral spherical triangle, and β to be that of a square (i.e. equilateral quadrilateral with four equal angles). Van der Waerden introduced the following concept of "angle-excess" of a spherical equilateral polygon with n sides:

$$w = \sigma - (n-2) \cdot 3\alpha \tag{1}$$

where σ is the sum of interior angles of the polygon. Note that by definition triangles have no angle-excess. The total angle-excess is denoted by W, which is simply the sum of individual angle-excesses w over all the faces:

$$W = \sum\sigma - \sum(n-2) \cdot 3\alpha = N \cdot 2\pi - (2E - 2F) \cdot 3\alpha \tag{2}$$

where N, E and F are respectively the number of points, edges and faces of the polygonal decomposition of the sphere. By Euler formula, one can replace $E - F$ by $N - 2$, so we have the following total sum of angle-excesses:

$$\begin{aligned} W &= N \cdot 2\pi - 2(N-2) \cdot 3\alpha \\ &= 26\pi - 66\alpha \qquad\qquad \text{(for } N = 13\text{)} \end{aligned} \tag{3}$$

The main idea of his proof is to estimate the lower bound of angle-excesses for the individual spherical polygons (resp. a certain collection of polygons) and then obtain a lower bound estimate of the total sum of angle-excesses which turns out to be larger than the above value of W, thus obtaining a contradiction that proves the impossibility.

Outline of the lower bound estimation of angle-excesses

Van der Waerden first claimed that the minimal angle-excess of a pentagon was achieved by an isosceles "triangle" with side-lengths $2a$ and base-length a, where $a = 2r \sin^{-1} \frac{1}{2r}$.

Hence the angle-excess w_5 of a pentagon is bounded below by the angle-excess Q of a square, namely

$$w_5 \geq Q = 4\beta - 6\alpha \geq 4\pi - 10\alpha \qquad (\alpha + \beta \geq \pi \text{ for } r \leq 1) \tag{4}$$

Therefore in the hypothetical configuration there could exist at most one pentagon, or at most one square, since $2Q > W$ for $r \leq 1$.

Next van der Waerden introduced the terminologies of "large" and "small" angles in a quadrilateral. Each quadrilateral in the configuration contained a larger pair of equal angles $\{\gamma, \gamma\}$ and a smaller pair of equal angles $\{x, x\}$. Hence, by definition, γ and x should satisfy the following bounds

$$\alpha < x \leq \beta, \quad \beta \leq \gamma < 2\alpha \tag{5}$$

The second claim of van der Waerden was the lower bound estimate on the *partial sum* of angle-excesses w_P counting only those quadrilaterals *with a small angle at* P. For the cases without pentagons involved, he claimed that

$$w_P \geq w_0 \tag{6}$$

where w_0 denoted the angle-excess of the quadrilateral with small angle $2\pi - 4\alpha$. By an easy elimination of the case of having all angles being large at a point P (cf. Case 3 near the end of this section), he did the total accounting as follows:

Case 1: *The configuration contains only triangles and non-square quadrilaterals.*

Summing up all w_P for the thirteen vertices, and since the angle-excess of each quadrilateral will be counted twice in the sum of LHS of (6), we have

$$2W = \textstyle\sum_P w_P \geq 13 w_0 \tag{7}$$

which is impossible by a simple numerical checking. For example, when $r = 1$, one has $\alpha = \cos^{-1}\frac{1}{3}$, $2W = 0.876174904$ and $13w_0 = 1.0414911$.

Case 2: *The configuration contains one pentagon or one square.*

Since the angle-excess of a pentagon w_5 (resp. a square Q) is at least

$$w_5 \geq Q \geq 4\pi - 10\alpha \tag{4$'$}$$

This is already so big that even we count the angle-excesses of those quadrilaterals adjacent to the pentagon (resp. square) only once, we still can obtain the following lower bound estimate:

$$\begin{aligned} 2W &\geq 2w_5 + 8w_P \geq 2w_5 + 8w_0 \\ (\text{resp. } &\geq 2Q + 9w_P \geq 2Q + 9w_0) \end{aligned} \tag{8}$$

which is again impossible, and this completes the proof by van der Waerden on the impossibility of thirteen touching neighbors.

Remark: The lower bound estimate on w_5 can be justified by a lemma appeared in Schütte's proof (Satz 5 in [SW–2]). However, in arriving at the estimate (6) : $w_P \geq w_0$, van der Waerden relied on a sequence of deformations in which he only explained how to perform the deformations but did not provide justifications to guarantee the decrement of the *partial sum* of angle-excesses w_P, or the deformations could be performed in a consistent way globally.

A reformulation of van der Waerden's estimation on angle-excesses in a simpler, more clean-cut way

First of all, by a constant scaling, we can simply replace the lower bound estimate on angle-excesses by the lower bound estimate of area-excesses. Then the "Satz 5" of Schütte's proof is in fact a special case of the following lemma:

Lemma 1.1 (cf. Lemma 2.1.2 of [Hsi–2]): *A quadrilateral with four given side-lengths attains its maximal area when it is cocircular. Shearing deformation (i.e. deformation by varying the length of a chosen diagonal with the four side-lengths fixed) further away from cocircularity is monotonic area-decreasing.*

Verification on the lower bound estimate of w_5:

Given a pentagon $ABCDE$, we choose four of the vertices to form a quadrilateral, say $ABCD$. Then we can select a suitable shearing direction so that the deformation is further away from cocircularity:

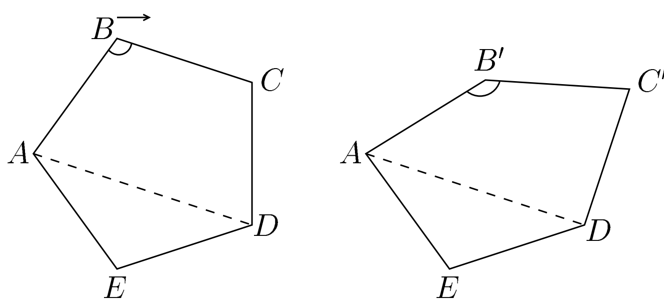

Figure 1.2

Since for $r \leq 1$, $\alpha + \beta \geq \pi$ and therefore this area-decreasing shearing deformation can be performed as long as the interior angle at B' is still less than π (see condition (a) for irreducible graphs). The same deformation can also be applied to quadrilateral $AC'DE$. As a result, the area of a pentagon will be bounded below by the isosceles triangle as specified (by pushing to the limiting case). □

Next, to prove the estimate (6) : $w_P \geq w_0$, we can apply a simple corollary of the following classical result in spherical geometry:

Lemma 1.2 (Lexell's Theorem): *Let $\triangle ABC$ and $\triangle ABC'$ be two triangles with the same base $\overline{AB}$ and same oriented area. Then C, C', A^* and B^* are cocircular, where A^* and B^* are the antipodal points of A and B.*

Corollary: *For a cluster of isosceles triangles with the same side lengths sharing a common top vertex P with a fixed sum of central angles, the more lopsided is the distribution of the central angles, the smaller is the total area.*

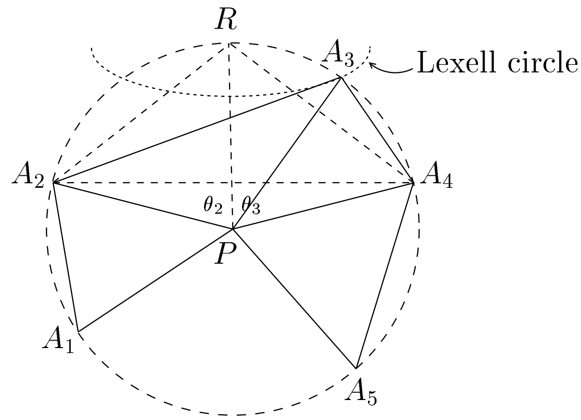

Figure 1.3

Proof: As indicated in Figure 1.3, the Lexell circle which collects points A_3' so that area$(\triangle A_2A_4A_3')$ = area$(\triangle A_2A_4A_3)$ is shown. Obviously, the more balanced is the distribution of θ_2, θ_3, the bigger is the area of $\triangle A_2A_4R$ and hence the total area of the isosceles triangles. □

Verification on the lower bound estimate of (6) : $w_P \geq w_0$:

Recall the definition that w_P counts the area-excesses of only those quadrilaterals *with small angles at P*. As $6\alpha > 2\pi$, and irreducible graph has no 2-fork vertex, we need only to consider the following three cases:

Case 1: *P is a 5-fork vertex.*

In this case, we cannot have a large angle at P since $4\alpha + \beta > 2\pi$. As long as the total area-excess is concerned, we may rearrange the positions of the polygons so that all the quadrilaterals are placed together. By cutting each quadrilateral into two congruent isosceles triangles using the diagonal joining its two large angles, we obtain a cluster of isosceles triangles at point P.

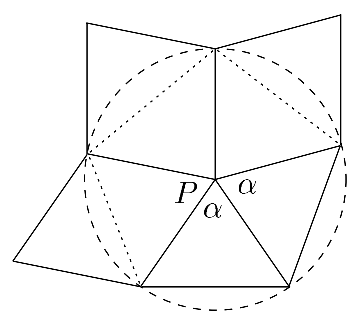

Figure 1.4

Now, by the above Corollary, we obtain a lower bound of the total area of the cluster, hence a lower bound of the total area-excess w_P, which is the one with central angle distribution being the most lopsided possible, namely $(\alpha, \alpha, \alpha, \alpha, 2\pi - 4\alpha)$. The corresponding area-excess is exactly w_0.

Case 2: *P is a 4-fork vertex.*

It is impossible to have three large angles at P. If there are two large angles at P, then since $\alpha + \beta \geq \pi$, we must have $r = 1$ and central angle distribution $(\alpha, \alpha, \beta, \beta)$, which is impossible because of the existence of two squares. If there is only one large angle γ at P, then since $\gamma < 2\alpha$, so the sum of remaining three angles must be larger than $2\pi - 2\alpha$. Hence the total area-excess of the remaining three triangles/quadrilaterals (i.e. w_P) must be larger than the one with central angle distribution $(\alpha, \alpha, 2\pi - 4\alpha)$, which is again w_0.

If there is no large angle at P, the lower bound estimate of w_P becomes trivial because the most lopsided distribution will include one β and $Q > w_0$.

Case 3: *P is a 3-fork vertex.*

If there is a small angle x, then $x > 2\pi - 4\alpha$ (because the sum of the other two angles is less than 4α), hence $w_P \geq w_0$. If all the three angles are large angles (i.e. $w_P = 0$), then the total area-excess of these three quadrilaterals is at least larger than that of the triple quadrilaterals with central angle distribution $(2\alpha, \beta, \beta)$, which is impossible.

Hence the estimate (6) : $w_P \geq w_0$ is also justified. □

Remark: In fact, the above case division also appeared in Schütte's proof. But in there he was discussing other issues.

In retrospect, the key idea of van der Waerden's proof is the existence of an *irreducible graph*. It provides a decomposition of the sphere into polygons which are all *equilateral*. This facilitates the lower bound estimation of the area-excesses of those polygons, from which a contradiction of the assumption of thirteen touching neighbors arises. However, the construction, or the proof of the existence, of an irreducible graph is rather sophisicated and it is highly non-trivial to pin down in a clean-cut way.

2 The proof by Leech

Next let us discuss the proof by Leech on the impossibility of thirteen touching neighbors. In his short article of 1956 [Lee], Leech outlined his idea of the proof which did not require a sophisicated construction of irreducible graph with equal lengths as van der Waerden did. Instead, by a carefully chosen upper bound of $\cos^{-1}\frac{1}{7}$ on edge lengths, he simply joined those pairs of vertices with separation less than the chosen bound and applied the area estimation techniques and combinatorial (or topological) analysis to show the non-existence of such hypothetical configurations. However, the part on area estimation and combinatorial analysis became more involved than in van der Waerden's case. In his own words, Leech wrote, "*certain details which are tedious rather than difficult being omitted*".

An outline of Leech's proof

Leech started with a given number of points on the unit sphere with $\frac{\pi}{3}$ (spherical) separation. As mentioned before, he joined those pair of vertices with separation strictly less than $\cos^{-1}\frac{1}{7}$ by a great circle arc. He pointed out that, with brief explanations, the graph obtained will satisfy the following two properties:

(a) Edges will not cross over each other, and by a deformation if necessary, one can assume that the graph will subdivide the unit sphere into polygons;

(b) The graph contains no vertex with degree more than five. It is because the inner angles of a polygon are always bigger than $\frac{\pi}{3}$, as all of them should be bounded below by the base angle of a $\frac{\pi}{3}$-isosceles triangle with base length $\cos^{-1}\frac{1}{7}$, which is exactly $\frac{\pi}{3}$.

The next key step in Leech's approach was the following claims on the area estimates of the polygons obtained:

(i) the area of a triangle is at least equal to the area of a $\frac{\pi}{3}$-equilateral triangle, whose area is equal to $\triangle_{\frac{\pi}{3}} = 3\cos^{-1}\frac{1}{3} - \pi = 0.55128\ldots$;

(ii) the area of a quadrilateral is at least equal to the area of a $\frac{\pi}{3}$-equilateral quadrilateral with one diagonal length being $\cos^{-1}\frac{1}{7}$, whose area is equal to $2(\cos^{-1}(-\frac{1}{7}) + \frac{2}{3}\pi - \pi) = 1.334\ldots = 2\triangle_{\frac{\pi}{3}} + 0.231\ldots$;

(iii) the area of a pentagon is at least equal to the area of a $\frac{\pi}{3}$-equilateral pentagon with two non-crossing diagonals of lengths both equal to $\cos^{-1}\frac{1}{7}$, whose area is about $2.226 = 3\triangle_{\frac{\pi}{3}} + 0.572$.

and so on. In his short article, Leech simply stated the above area estimates without proofs. Then he applied the Euler formula, firstly to obtain an upper bound on the number of vertices V:

$$(2V-4)\cdot\triangle_{\frac{\pi}{3}} \leq 4\pi \quad \Rightarrow \quad V \leq 13 \tag{9}$$

and secondly, similar to the argument of van der Waerden's proof, to obtain the following area-excess estimate in the case of $V = 13$:

$$4\pi - 22\triangle_{\frac{\pi}{3}} = 0.438\ldots \geq 0.231F_4 + 0.572(F_5 + \ldots) \tag{10}$$

where F_n denotes the number of n-gons in the hypothetical configuration with thirteen vertices. From the above estimate on area-excesses he concluded that if thirteen touching neighbors were possible, then $F_5 = F_6 = \ldots = 0$ and $F_4 = 0$ or 1. He could easily eliminate the possibility of $F_4 = 0$ because those 33 edges sharing among 13 vertices must create a vertex with degree at least six. For the case of $F_4 = 1$, Leech claimed that it is again not possible to have such a combinatorial type, but he wrote (in his own words), "*I know of no better proof of this than sheer trial*". And then he only outlined the idea of his "*sheer trial*" in a few lines following this claim.

Remark: In 1998, Aigner and Ziegler in their book [AZ] tried to provide a detailed proof of the impossiblity of thirteen touching neighbors using Leech's approach, especially they supplied a systematic argument to replace Leech's "*sheer trial*". However, detailed justifications on the area estimates were still missing from their proof.

Justifications of Leech's claims on area estimates

In fact, by direct applications of Lemma 1.1 (cf. Lemma 2.1.2 of [Hsi–2]), namely, shearing deformation further away from cocircularity is always monotonic area-decreasing, one can justify Leech's claims on area estimates in a straightforward, clean-cut way. Now we proceed to verify the three claims on area estimates as follows:

<u>Area estimate of a triangle</u>: Let $\{\ell_1, \ell_2, \ell_3\}$ be the side lengths of a given triangle. We can piece another copy of the triangle along an edge, say the one with length ℓ_3, to produce a quadrilateral. Note that a cocircular quadrilateral with side lengths $\{\ell_1, \ell_2, \ell_1, \ell_2\}$ will have its diagonal lengths both equal to $\cos^{-1}(\cos\ell_1 + \cos\ell_2 - 1)$ (cf. Eq (75) on p.45 of [Hsi–2]), which is at least $\frac{\pi}{2}$ as $\frac{1}{7} < \cos\ell_1, \cos\ell_2 \leq \frac{1}{2}$. Hence, shearing deformation by shortening ℓ_3 is area-decreasing. The shearing deformation can be performed until it is blocked by the $\frac{\pi}{3}$

lower bound on edge lengths. Repeat the same procedure for the other two edges, one can easily see that the minimal area should be attained by a $\frac{\pi}{3}$-equilateral triangle.

Area estimate of a quadrilateral: Again we perform area-decreasing shearing deformation on a given quadrilateral until the length of the shortening diagonal hitting its lower bound of $\cos^{-1}\frac{1}{7}$ (beyond that the quadrilateral will be decomposed into two triangles by definition). Similarly, a triangle with base length $\cos^{-1}\frac{1}{7}$ and side length bounds of $\frac{\pi}{3} \leq \ell < \cos^{-1}\frac{1}{7}$ will attain its minimal area when it is $\frac{\pi}{3}$-isosceles, whose area is equal to $(\cos^{-1}(-\frac{1}{7}) + \frac{2\pi}{3} - \pi) = 2\tan^{-1}\frac{\sqrt{3}}{5}$ (cf. area formula of triangle, Lemma 2.1.1 of [Hsi–2]) and hence the area of a quadrilateral is bounded below by the *double* of it.

Area estimate of a pentagon: We choose four out of the five vertices to form a quadrilateral and perform area-decreasing shearing deformation on it until the length of the shortening diagonal hitting the lower bound of $\cos^{-1}\frac{1}{7}$. The area of the triangle with base length $\cos^{-1}\frac{1}{7}$ is bounded below by $2\tan^{-1}\frac{\sqrt{3}}{5}$. For the adjacent quadrilateral, we again perform the area-decreasing shearing deformation until the length of another diagonal hitting the lower bound of $\cos^{-1}\frac{1}{7}$. Again the triangle with base length $\cos^{-1}\frac{1}{7}$ is bounded below by $2\tan^{-1}\frac{\sqrt{3}}{5}$, and the area of the $\cos^{-1}\frac{1}{7}$-isosceles triangle is bounded below by the one with base length $\frac{\pi}{3}$, whose area is $2\tan^{-1}\frac{\sqrt{143}}{25}$. Therefore the area of a pentagon will attain its minimal value when it is $\frac{\pi}{3}$-equilateral with two non-crossing diagonals of lengths both equal to $\cos^{-1}\frac{1}{7}$. The area of this pentagon is exactly $(2\tan^{-1}\frac{\sqrt{143}}{25} + 4\tan^{-1}\frac{\sqrt{3}}{5}) \approx 2.226$.

A simple argument on the non-existence of combinatorial type with $F_3 = 20$ and $F_4 = 1$

In the following we provide a simple argument to show the non-existence of a combinatorial type with thirteen vertices, in which there are only triangles and *exactly one quadrilateral*, and all the vertices are of degrees at most five. By Euler formula, there should be 32 edges and hence except one of them is a 4-fork vertex, all the remaining twelve are 5-fork vertices. We divide our discussion into two main cases according to whether the only 4-fork vertex belongs to the quadrilateral or not.

Case 1: *The* 4*-fork vertex belongs to the quadrilateral.*

In this case, the partial configuration around the quadrilateral must be the one as shown in Figure 1.5(a), and we need to add two points P, Q in the complement of the partial configuration as shown in Figure 1.5(b):

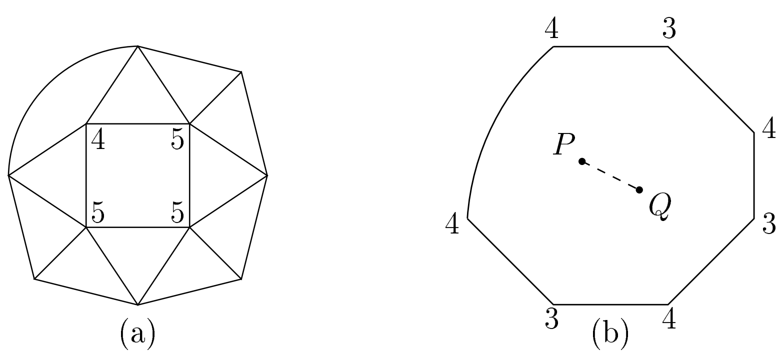

Figure 1.5

The numbers shown in Figure 1.5(b) are the degrees from the partial configuration in Figure 1.5(a). We need to add enough *non-crossing* edges to complete the configuration. Since there is no more quadrilateral in the configuration, we must join P and Q. As all the nine vertices in Figure 1.5(b) should be of degree five, we need to fill in edges so as to add 10 degrees to the boundary vertices and 8 degrees to the vertices P and Q. Therefore, we must join two boundary vertices by an edge, and with the restriction that this edge cannot block any boundary vertex from joining to P or Q. Clearly it is impossible to do so.

Case 2: *The 4-fork vertex does not belong to the quadrilateral.*

In this case, the partial configuration around the quadrilateral must be the one as shown in Figure 1.6(a), and we need to add one point P in the complement of the partial configuration as shown in Figure 1.6(b):

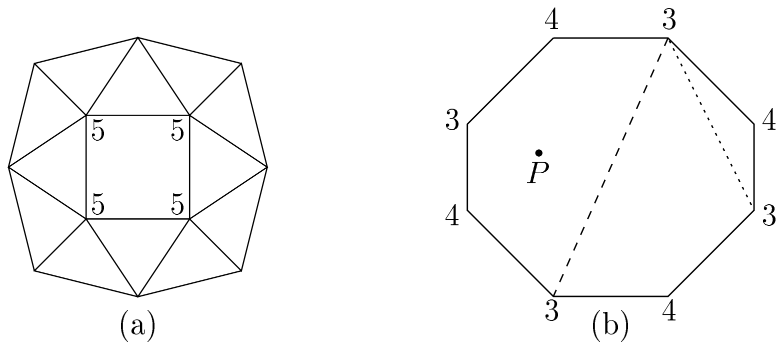

Figure 1.6

The partial configuration in Figure 1.6(a) contains 24 edges, so we need to add 8 edges in the complement of the configuration in Figure 1.6(b). As the degree of P is at most five, at least three of the edges are joining boundary vertices. It is straightforward to check that one can draw at most two edges in total starting at vertices with labels "4", so we must have an edge joining two vertices with labels "3".

If two nearby vertices with labels "3" are joined by an edge (as shown by the dotted line in Figure 1.6(b)), then there is a 4-fork boundary vertex. Hence $\deg P = 5$ and we need to add two more edges joining boundary vertices. Now it is straightforward to check that joining any two boundary vertices will either create an extra 3-fork or 4-fork vertex, or blocking too many vertices from joining to P (we need to join five vertices to P). Hence this case is impossible.

If two opposite vertices with labels "3" are joined by an edge (as shown by the dashed line in Figure 1.6(b)), then we consider the vertex with label "3" at the half of the complementary configuration without containing P. It is easy to see that adding one edge to it (make it become the only 4-fork vertex) will create another 4-fork vertex, and adding two edges to it will create two other 4-fork vertices. Thus this case is again impossible.

This completes the proof of the non-existence of combinatorial type with $F_3 = 20$ and $F_4 = 1$, in which every vertex is of degree at most five. $\square$

In retrospect, Leech's approach is quite simple. Starting with a much less restrictive graph than those irreducible graphs used in [SW–2] and using some simple lower bound area estimates on triangles (resp. quadrilaterals, pentagons), namely, 0.5513 (resp. 1.334, 2.226), the proof of impossibility of thirteen touching neighbors can be directly reduced to the *non-existence* of such a configuration with $F_3 = 20$ and $F_4 = 1$, which in fact, can simply be proved as above.

3 The proof by Hsiang

The proof by Hsiang was in fact given as a demonstration example (cf. p.72 in [Hsi–2]) of his area estimation techniques developed specifically for the estimations of total buckling heights in his proof of Kepler Conjecture [Hsi–1]. For a given set of finite number of points Σ on the unit sphere, he simply used a natural way to create a configuration by taking the Euclidean convex hull $I(\Sigma)$ of the points, and then radially projecting the edges and faces of $I(\Sigma)$ onto the unit sphere. The configuration obtained is called the *spherical configuration associated to* Σ, denoted by $\mathcal{S}(\Sigma)$.

Remark: Sometimes $\mathcal{S}(\Sigma)$ is also referred as the *standard configuration.* It is also the same as the Delaunay decomposition.

The spherical configuration $\mathcal{S}(\Sigma)$, by definition, satisfies the following two nice properties, which are both referred as the *convexity property* of $\mathcal{S}(\Sigma)$:

(a) All the spherical polygons obtained are cocircular;

(b) The interior of the circumcircle of any spherical polygon contains no points of Σ.

In application, we may also assume the following extra condition of a-saturacy:

Definition: A finite set of points Σ on the unit sphere is said to be a-saturated if the adding of any other point to Σ will change the separation from at least equal to a to less than a.

Then, Hsiang proved the following lemmas on area estimates:

Lemma 1.3 (cf. Lemma 2.2.1 in [Hsi–2]): *If Σ is an a-saturated set with at least four points, then all faces in $\mathcal{S}(\Sigma)$ are either triangles, quadrilaterals or pentagons, and the area of any face is at least equal to the area of the cocircular spherical polygon with the same number of sides of lengths equal to a.*

Let $\mathcal{S}'(\Sigma)$ denote a configuration obtained by adding non-crossing diagonals to subdivide any non-triangular face of $\mathcal{S}(\Sigma)$, if any, into triangles. Then

Lemma 1.4 (cf. Lemma 2.2.2 in [Hsi–2]): *The area of a $6\triangle$-star in a $\mathcal{S}'(\Sigma)$ of a $\frac{\pi}{3}$-saturated set of points is at least equal to* $2\cos^{-1}(-\frac{1}{3}) = 6\triangle_{\frac{\pi}{3}} + 0.51355\ldots$.

Lemma 1.5 (cf. Remark on p.81 of [Hsi–2]): *The area of a $7\triangle$-star in a $\mathcal{S}'(\Sigma)$ of a $\frac{\pi}{3}$-saturated set of points is at least equal to* $7\triangle_{\frac{\pi}{3}} + 0.770329\ldots$.

The area estimates above are very strong because simply the area-excess of a $6\triangle$-star itself is already larger than the total area-excess $4\pi - 22\triangle_{\frac{\pi}{3}}$. Therefore, the Hsiang's proof of the impossibility of thirteen touching neighbors was just the following few lines:

Suppose for contrary that it is possible to place thirteen points on the unit sphere with $\frac{\pi}{3}$-separation. We may assume that the set of points is already $\frac{\pi}{3}$-saturated. By Euler formula, one has a triangulated spherical configuration with 22 triangles and 33 edges. Therefore, at least one of its thirteen vertices has more than five edges and hence the total area is at least equal to

$$16\triangle_{\frac{\pi}{3}} + 2\cos^{-1}(-\frac{1}{3}) > 4\pi \tag{11}$$

which is obviously a contradiction.

The key result used in Hsiang's approach is Lemma 1.4, and the proof of Lemma 1.4 (cf. §2.2.4 in [Hsi–2]) relies on a collection of one-parameter families of area-decreasing deformations (like the one in Lemma 1.1) which can be effectively applied to any $6\triangle$-star and reduces the lower bound area estimate to straightforward comparison of several critical cases where all geometric data can be explicitly computed. The area estimation techniques applied in the previous two sections follow exactly this line of thinking.

4 A comparative analysis on the three proofs of the impossibility of thirteen touching neighbors

Qualitatively speaking, the above three proofs all involve following three steps, namely, construction of a configuration, lower bound area (excess) estimations and combinatorial analysis.

(a) In van der Waerden's proof, the construction of an irreducible graph is rather complicated which requires some non-trivial deformations. But since all the edges are of equal length, the area estimates become quite simple and the combinatorial analysis is also simple.

(b) In Leech's proof, the construction of a configuration is quite simple, namely, just joins those within the chosen bound of $\cos^{-1}\frac{1}{7}$. The upper bound of $\cos^{-1}\frac{1}{7}$ is specifically chosen to eliminate the possibility of a 6-fork vertex in a configuration, and in the same time keeping a good lower bound area estimation. But, since the edge-length distribution becomes non-uniform, the lower bound area estimation of the polygons becomes slightly more involved and the combinatorial analysis of configuration is also slightly more complicated.

(c) In Hsiang's proof, the configuration is simply taken to be the natural one obtained from the convex hull of the points. There hardly exists a good control on the edge-length distribution, and so the lower bound area estimations are rather involved. However, the built-in convexity property of the spherical configuration exerts quite restrictive requirements on the partial configurations and hence a better area estimate can be obtained, which reduces the needed combinatorial analysis to triviality.

Therefore, each of the above three proofs has its own technical advantages and disadvantages. In order to obtain a more objective comparison on these three proofs, we try to employ their methods of proofs to study a directly related problem, which is more challenging and still remains open, namely, the spherical code problem of thirteen points.

Upper bound estimation on the maximal separation of thirteen points

The spherical code problem askes for the maximal separation one can obtained in arranging N points on the unit sphere. The impossibility of having thirteen touching neighbors simply says that the maximal separation δ_{13} is strictly less than $\frac{\pi}{3}$. By how much is δ_{13} less than $\frac{\pi}{3}$? We will try to employ the methods used in the above three proofs to obtain some upper bound estimates of δ_{13}. [Note: $\frac{\pi}{3} = 1.04719\ldots$]

Using van der Waerden's method:

In fact, van der Waerden's proof was published as a follow-up article after his joint work with Schütte on spherical codes [SW–1]. Therefore, his method is already in the setting of the problem. The key estimate in his proof is the *invalidity* of the following area-excesses estimate:

$$(7') \qquad 2W = (26\pi - 66\alpha) \geq 13w_0 \quad \text{is invalid when } a = \tfrac{\pi}{3}$$

By seeking the smallest edge-length a so that the above estimate still remains to be invalid, we can obtain an upper bound on δ_{13}. Note that

$$(12) \qquad \begin{aligned} \alpha &= \cos^{-1}\frac{\cos a}{1+\cos a}, \\ w_0 &= 2\cdot(2\pi - 7\alpha + 2\cos^{-1}\frac{\tan\sin^{-1}(\sin 2\alpha \sin a)}{\tan a}) \end{aligned}$$

and by direct numerical computations, we obtain $a \geq 1.04318$. Hence

$$(13) \qquad \delta_{13} \leq 1.04318$$

<u>Using Leech's method</u>:

The key estimate in Leech's proof is the deduction of $F_4 \leq 1$ from the following area-excess estimate

$$(10') \qquad 4\pi - 22\triangle_a = 26\pi - 66\alpha \geq (A_4 - 2\triangle_a)\cdot F_4 + \ldots$$

where A_4 is twice the area of an a-isosceles triangle with base angle equal to $\frac{\pi}{3}$.
It is straightforward to compute that

$$(14) \qquad A_4 = 4\cos^{-1}\frac{\tan\sin^{-1}(\sin\frac{\pi}{3}\sin a)}{\tan a} - \frac{2\pi}{3}$$

and in order to obtain $F_4 \leq 1$, we need $(A_4 - 2\triangle_a)\cdot 2 > 26\pi - 66\alpha$. Direct numerical computations show that $a \geq 1.04635$ and hence

$$(15) \qquad \delta_{13} \leq 1.04635$$

<u>Using Hsiang's method</u>:

Using the idea of Hsiang's proof, the minimal area of a $6\triangle$-star will be the one with two a-equilateral triangles, one a-square and two a-isosceles triangles with base angles both equal to $\pi - \alpha - \frac{\beta}{2}$, where $\beta = \cos^{-1}\dfrac{\cos a - 1}{\cos a + 1}$. Therefore, in order to obtain a contradiction, we need to ensure that the area-excess of such a minimal $6\triangle$-star, namely,

$$(16) \qquad 2\beta - 16\alpha + 4\pi + 4\cos^{-1}\frac{\tan\sin^{-1}(\sin(\alpha+\frac{\beta}{2})\sin a)}{\tan a}$$

is still larger than the total area-excess $4\pi - 22\triangle_a$. Direct numerical computations show that $a \geq 1.04455$ and hence

$$(17) \qquad \delta_{13} \leq 1.04455$$

Improvement on using Hsiang's method: In fact, from Leech's proof, the *non-existence* of a combinatorial type without 6-fork vertex and with only a single quadrilateral of Figure 1.5(a) type (resp. Figure 1.6(a) type) correspond to the non-existence of a triangulation *only* with a single 6△-star (resp. a pair of adjacent 6△-stars), namely, just by adding a diagonal as follows:

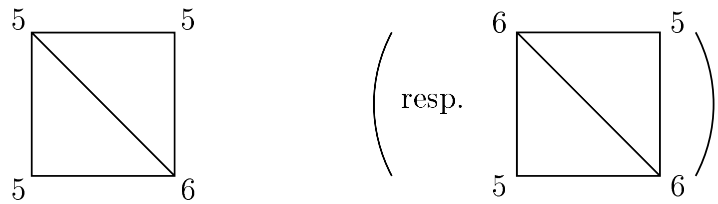

Figure 1.7

Therefore, in a triangulated spherical configuration of thirteen vertices, there must be at least two 6△-stars, and in the case of only two 6△-stars, there are no overlapping triangles between them. Therefore, the area-excess estimate in Hsiang's method can easily be improved to be the double of (16). The corresponding improvement on upper bound estimate of δ_{13} is:

$$\delta_{13} \leq 1.02746 \qquad (17')$$

In concluding the above comparison, we find that although the three methods all gave similar upper bound estimates of δ_{13}, the area estimation method employed by Hsiang however allows a further improvement when jointly applied with some simple combinatorial analysis (in fact, the original setup by Hsiang in the lower bound estimates of total buckling heights also involved some amount of combinatorial analysis). From the above brief discussion, Hsiang's approach appears to be able to provide a suitable set of tools for the study of the spherical code problems when combined with some more elaborate combinatorial analysis on the configurations and some refinements on area estimates.

References

[**AZ**] M. Aigner and G. M. Ziegler, *Proofs from The Book*, Springer, Berlin; New York, First ed. 1998.

[**Bar**] Barlow, *Nature*, 1883.

[**Ben**] C. Bender, *Bestimmung der grössten Anzahl gleicher Kugeln, welche sich auf eine Kugel van demselben Radius, wie die übrigen, auflegen lassen*, Archiv Math. Physik (Grunert) **56** (1874), 302–306.

[**Fej**] L. Fejes Tóth, *Lagerungen in der Ebene, auf der Kugel und im Raum*, Springer-Verlag, Berlin, 1971.

[**Gun**] S. Günther, *Ein stereometrisches Problem*, Archiv Math. Physik (Grunert) **57** (1875), 209–215.

[**Hop**] R. Hoppe, *Bemerkung der Redaction*, Archiv Math. Physik (Grunert) **56** (1874), 307–312.

[**Hsi–1**] W. Y. Hsiang, *On the sphere packing problem and the proof of Kepler's conjecture*, International J. of Math. **4** (1993), 739–831.

[**Hsi–2**] W. Y. Hsiang, *Least action principle of crystal formation of dense packing type and Kepler's Conjecture*, Nankai Tracts in Mathematics Vol. 3, World Scientific, 2001.

[Kep] J. Kepler, *The six-cornered snowflake*, Oxford Univ. Press, 1966. (English translation of latin booklet of 1611)

[Lee] J. Leech, *The problem of the thirteen spheres*, Math. Gazette **40** (1956), 22–23.

[SW–1] K. Schütte and B. L. van der Waerden, *Auf welcher Kugel haben* 5, 6, 7, 8 *oder* 9 *Punkte mit Mindesabstand Eins Platz?*, Math. Ann. **123** (1951), 96–124.

[SW–2] K. Schütte and B. L. van der Waerden, *Das Problem der dreizehn Kugeln*, Math. Ann. **125** (1953), 325–334.