A Journey To Low Spherical Discrepancy

By

Kejia Wang

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Mathematics)

at the

UNIVERSITY OF WISCONSIN – MADISON

2016

Date of final oral examination: December 8, 2016

The dissertation is approved by the following members of the Final Oral Committee:

Nigel Boston, Professor , Mathematics

Jordan Ellenberg, Professor , Mathematics

Shamgar Gurevitch, Associate Professor, Mathematics

Philip Wood, Assistant Professor , Mathematics

Stephen Wainger, Professor Emeriti, Mathematics

ABSTRACT

Discrepancy is a measurement of how uniform a point distribution is. The lower the discrepancy, the more uniform the distribution is. In the d-dimensional unit cube the notion of low discrepancy is well studied, and low discrepancy sequences are well understood. In recent years, this field has been enriched with sophisticated sequence construction techniques using arithmetic curves over finite fields, known as the Niederreiter-Xing method. However, the spherical discrepancy on the 2-dimensional unit sphere remains largely unexplored. In fact, the definition of low spherical discrepancy is not even officially established. Most "well-spaced" spherical sequences found in literature are obtained by lifting well-spaced sequences form the unit square to the sphere via certain maps (for example, the Lambert Transformation). In this thesis, we will investigate direct sequence construction algorithms on the sphere and the related spherical cap discrepancy. The point distribution is done by a greedy algorithm and triangulating the unit sphere. Counting the number of points inside an arbitrary spherical cap remains the challenge.

Contents

Chapter 1

Introduction

1.1 Why Discrepancy

Given an *integrable function* f over the d-dimensional *unit cube*, the integral $\int_{[0,1]^d} f dx$ can be estimated with a finite sum:

$$\int_{[0,1]^d} f(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} x_i f(x_i). \tag{1.1}$$

There are many ways to choose the N points $x_1, x_2, ..., x_N$. For example, one can simply choose them to be the points of some regular grid of $[0,1]^d$. In practice, Monte Carlo Integration and Quasi-Monte Carlo Integration are commonly used techniques for numerical integration. Monte Carlo Integration uses pseudorandom numbers while Quasi-Monte Carlo Integration uses quasirandom numbers. Consequently, the error bounds that Monte Carlo Integration yields are probabilistic while the error bounds

Quasai-Monte Carlo Integration yields are deterministic. Certain sequences of *quasirandom numbers* are also called *sub-random sequences* or **low-discrepancy sequences**. **Discrepancy** is a measure of distribution of a sequence of points. The lower the discrepancy is the more uniformly distributed the points are.

Definition 1.1.1. [1] The discrepancy of an infinite sequence X in $[0,1]^d$ is defined as the following:

$$D(X(N)) = \sup_{B\subset[0,1]^d} |\frac{\#(X(N)\bigcap B)}{N} - m(B)|, \tag{1.2}$$

where $B = \Pi_{i=1}^d [a_i, b_i]$ for $[a_i, b_i] \subset [0,1]$, $X(N)$ is the finite subsequence consisting of the first N elements of X and m is Lebesgue measure.

The notion of **Star Discrepancy** D^* is often used in place of discrepancy D. The definition of Star Discrepancy is very similar.

Definition 1.1.2. [1] The star discrepancy of an infinite sequence X in $[0,1]^d$,

$$D^*(X(N)) = \sup_{B\subset[0,1]^d} |\frac{\#X(N)\bigcap B}{N} - m(B)|, \tag{1.3}$$

where the rectangles B are of the form $\Pi_{i=1}^d [0, u_i]$ for $u_i \in [0,1]$.

The following relation between D and D^* is well-known and shows that they are equivalent for many purposes.

Theorem 1.1.3. [1] $D^*(X(N)) \leq D(X(N)) \leq 2^d D^*(X(N))$.

One of the reasons that discrepancy of a sequence is important is that it is directly involved in computing the error bounds of estimation of integrals by finite sums. To extend the integral estimation 1.1 to multi-dimension, we need the notion of the *variation in the sense of Hardy and Krause.*

Definition 1.1.4. ([1] pp. 19)

For a function f on $\bar{I}^d$ and a subinterval $\mathcal{J} \subset \bar{I}^d$, let $\Delta(f;\mathcal{J})$ be an alternation sum of the values of f at the vertices of $\mathcal{J}$ (i.e., the function values at adjacent vertices have opposite signs). *The variation of f on $\bar{I}^d$ in the sense of Vitali* is defined by

$$V^{(d)}((f)) = sup_{\mathcal{P}} \sum_{\mathcal{J} \subset \mathcal{P}} |\Delta(f;\mathcal{J})|, \tag{1.4}$$

where the supremum is extended over all partitions $\mathcal{P}$ of $\bar{I}^d$ into subintervals. The more convenient formula

$$V^{(s)}(f) = \int_0^1 ... \int_0^1 |\frac{\partial^d f}{\partial u_1 ... \partial u_d}| du_1 ... du_d \tag{1.5}$$

holds whenever the indicated partial derivative is continuous on $\bar{I}^d$. For $i \leq k \leq d$ and $1 \leq i_1 < i_2 < ... < i_k \leq d$ let $V^{(k)}(f; i_1, i_2, ..., i_k)$ be the variation in the sense of Vitali of the restriction of f to the k-dimensional face $\{(u_1, u_2, ..., u_s) \in \bar{I}^d | u_j = 1 \ for \ j \neq i_1, i_2, ..., i_k\}$. Then

$$V(f) = \sum_{k=1}^{d} \sum_{1 \leq i_1 < i_2 < ... < i_k \leq d} V^{(k)}(f; i_1, i_2, ..., i_k) \tag{1.6}$$

is called *the variation of f on $\bar{I}^d$ in the sense of Hardy and Krause*, and f is of bounded variation in this sense if $V(f)$ is finite.

Theorem 1.1.5. (The Koksma-Hlawka Inequality)([1] pp.20)

Let f be a function over $[0,1]^d$ with bounded Hardy-Krause Variation *$V(f)$, then*

$$|\frac{1}{N}\sum_{i=1}^{N} f(x_i) - \int_{[0,1]^d} f(x)dx| \leq V(f)D^*(x_1, x_2, ..., x_N). \tag{1.7}$$

The Koksma-Hlawka Inequality gives the best bound in the following sense:

Theorem 1.1.6. *Given set of points $\{x_1, x_2, ..., x_N\}$ and $\epsilon > 0$, there exists a function f with $V(f) = 1$ such that*

$$|\frac{1}{N}\sum_{i=1}^{N} f(x_i) - \int_{[0,1]^d} f(x)dx| > D^*(x_1, x_2, ..., x_N) - \epsilon. \tag{1.8}$$

Therefore, the accuracy of the estimation of the integral ultimately depends on the discrepancy of the sequence. For the obvious reason we would like to have the discrepancy of the sequence to be as low as possible. However, many believe that the discrepancy of a sequence cannot be arbitrarily low, as we indicate next.

Conjecture ([1] pp.32) *It is widely believed that in the unit cube $[0,1]^d$ any N-element point set $x_1, x_2, ..., x_N$ satisfies*

$$D^*(x_1, x_2, ..., x_N) \geq c_d \frac{(\log N)^{d-1}}{N} \tag{1.9}$$

for some constant c_d that depends on the dimension d.

When $d = 1$, the conjecture in this case, say $D^*(x_1, x_2, ..., x_N) > \frac{1}{2N}$, can be easily verified. For $d = 2$, equation 1.9 was proven by Schmidt in 1972 [14]. For dimension 3 or higher, the conjecture still remains open with the best general bounds given by

Roth [15]:

$$D^*(x_1, x_2, ..., x_N) \geq c_d \frac{(\log N)^{(d-1)/2}}{N}. \tag{1.10}$$

In 1935, the Dutch mathematician J. G. van der Corput created sequences over $[0, 1]$ by reversing *base n representation* of natural numbers. For example, the *base 10 Van der Corput sequence* is {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.01, 0.11, 0.21, 0.31, 0.41, 0.51, 0.61, 0.71, 0.81, 0.91, 0.02, 0.12, 0.22, 0.32, ...}, and the *base 2 Van der Corput sequence* is {0.1, 0.01, 0.11, 0.001, 0.101, 0.011, 0.111, 0.0001, 0.1001, 0.0101, 0.1101, 0.0011, 0.1011, 0.0111, 0.1111,...}. They are now known as *the Van der Corput sequences. Halton sequences*, created during 1960s, generalize *Van der Corput sequences.* They are produced using *prime bases* and are well distributed in lower dimensions. For example, the *Van der Corput sequence* in base 2 is $\{\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}, ...\}$ and the *Van der Corput sequence* in base 3 is $\{\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{4}{9}, \frac{5}{9}, \frac{7}{9}...\}$. Pairing them up, the first sequence being the first coordinates while the second sequence being the second coordinates, we get a *Halton Sequence* in $[0, 1]^2$, $\{(\frac{1}{2}, \frac{1}{3}), (\frac{1}{4}, \frac{2}{3}), (\frac{3}{4}, \frac{1}{9}), (\frac{1}{8}, \frac{4}{9}), (\frac{3}{8}, \frac{5}{9}) ...\}$. Both *Van der Corput sequences* and *Halton sequences* are well-known quasirandom sequences. They all satisfy a common inequality:

$$D^*(x_1, x_2, ..., x_N) \leq C \frac{(\log N)^d}{N}. \tag{1.11}$$

This common inequality evolved into the the definition of low discrepancy sequences.

Definition 1.1.7. *An infinite sequence* X is of *low discrepancy* if for all N

$$D^*(x_1, x_2, ..., x_N) \leq C \frac{(\log N)^d}{N}. \tag{1.12}$$

Remark 1.1.8. A defined above, *a low discrepancy sequence* X is *an infinite sequence.* The inequality 1.12 is satisfied by *any finite truncation* of X of length N.

1.2 Niederreiter-Xing Sequences

Other than the *Van der Corput sequences* and *Halton sequences*, *Sobol' Sequences* and *Niederreiter-Xing sequences* are more "contemporary" low discrepancy sequences in the unit cube $[0,1]^d$. The *Niederreiter-Xing method* is the best known technique in application to construct low discrepancy sequences. In order to discuss these methods, some new notions need to be introduced.

Definition 1.2.1. ((t,m,d) Net) ([1], page 48)

Let $b \geq 2$ and $0 \leq t \leq m$ be integers. A point set $S_{b^m} \subset [0,1]^d$ of size b^m is called a (t,m,d)-net in base b if for all non-negative integers $k_1, k_2, ..., k_d$ the elementary interval $\Pi_{i=1}^d [\frac{a_i}{b^{k_i}}, \frac{a_i+1}{b^{k_i}}]$ of hypervolume b^{t-m} contains exactly b^t points from S_{b^m}.

Definition 1.2.2. ((t,d) Sequence) ([1], page 48)

Let $b \geq 2$ be an integer. A sequence $\{x_1, x_2, x_3, ...\} \in [0,1]^d$ is called a (t,d)-sequence in base b if for all integers $0 \leq t \leq m$ and $k \geq 0$ the point set $\{x_{kb^m}, x_{kb^m+1}, x_{kb^m+2}, ..., x_{kb^{m+1}-1}\}$ is a (t, m, d)-net.

Sobol' introduced the concepts of net and (t,d) sequences for base $b = 2$ [2]. The formal definitions for general b were given by Niederreiter [3]. The current record holder of the sequences of in $[0,1]^d$ with "the lowest discrepancy" (with the constant C as small as possible) are constructed by Harald Niederreiter and Chaoping Xing using the arithmetic of curves over finite fields, and are known as the *Niederreiter-Xing sequences.*

1.2.1 Discrepancy Bounds of (t,d) Sequences

By the definitions of the (t,m,d) nets and (t,d) sequences, the smaller t is the finer the distribution. In fact, the star-discrepancy of (t,d) sequences are known.

Theorem 1.2.3. ([1], Theorem 4.17) *For any* (t,d) *sequence in base* b*,*

$$D^*(X(N)) \leq C(d,b)b^t\frac{(\log N)^d}{N} + \mathbf{O}(b^t\frac{(\log N)^{d-1}}{N}),$$

where the constant $C(d,b)$ *depends on the dimension* d *and base* b *only*

$$C(d,b) = \begin{cases} \frac{1}{d}(\frac{b-1}{2\log b})^d, & \textit{if } d=2, \textit{ or } b=2 \textit{ and } d=3,4 \\ \\ \frac{1}{d!}\frac{b-1}{2\lfloor b/2\rfloor}(\frac{\lfloor b/2\rfloor}{\log b})^d, & \textit{otherwise.} \end{cases} \tag{1.1}$$

The details on how the inequality is obtained can be found in [1] Chapter 4. Clearly, the magnitude of the star discrepancy of the first N elements of a (t,d) sequence is in direct connection with the magnitudes of t and $C(d,b)$. The smaller t and $C(d,b)$ are the lower the discrepancy is.

1.2.2 The Smallest Possible t

Van der Corput sequences are $(0,1)$ *sequences* for various base b. $t=0$ implies the strongest regularity. Before we proceed, the first natural question to ask is under what condition $(0, d)$ sequences exist for base b.

Definition 1.2.4. (Mutually Orthogonal Squares) ([4])

Given an integer $b \geq 2$, two b^2-tuples $e = (e(0), e(1), ..., e(b^2-1))$ and $f = (f(0), f(1), ..., f(b^2-1))$ with entries from the same set of cardinality b are called orthogonal if the b^2 ordered pairs $(f(i) \neq e(i))$ are all distinct $i = 0, 1, ..., b^2 - 1$. The b^2-tuples $e_1, e_2, ..., e_s$ with entries from the same set of cardinality b are called mutually orthogonal if e_i and e_j are orthogonal for all $1 \leq i < j \leq s$. The entries of a b^2-tuple can be arranged a prescribed manner in a square matrix with b rows and b columns. With such an identification of b^2-tuples with $b \times b$ matrices, we speak of orthogonal squares of order b and mutually orthogonal squares of order b.

Definition 1.2.5. (Latin Square) A $b \times b$ matrix is called a latin square of order b if each row and each column is a permutation of the same set of cardinality b.

Remark 1.2.6. $M(b)$ denotes the maximal cardinality of a set of mutually orthogonal latin squares of order b. For all $b \geq 2$ $M(b) \leq b-1$ ([5] pp.158, [6] pp.80). The existence of $b - 1$ mutually orthogonal latin squares of order b is equivalent to the existence of a finite projective plane of order b ([4], pp.209-210). When b is a prime power, $M(b) = b-1$ ([7], Thm.9.83).

Theorem 1.2.7. ([1], pp.62)

A $(0, d)$ sequence in base b only exists if $d \leq M(b) + 1$.

Corollary 1.2.8. ([1], pp.62)

A (0,d) sequence in base b can only exist when $d \leq b$.

1.2.3 Best Known $C(d,b)$ by Niederreiter-Xing Method

With the smallest possible value $t_b(d)$, the only way to lower the discrepancy of a $(t_b(d), d)$ sequence is to sharpen the constant coefficient $C(d, b)$. When b happens to be a prime power, there exists a a curve over the finite field $\mathbb{F}_b$ with genus t and $d + 1$ rational places. Niederreiter and Xing developed a method of constructing (t, d)-sequence over this finite field with the smallest known constant $C(d, b)$. This method is known as the Niederreiter-Xing method.

1.3 Discrepancies on S^n

Unlike the unit cube $[0, 1]^d$, the subject of discrepancies of the unit sphere S^n is not as well-studied and largely unknown. In fact, there are various notions of spherical discrepancies. We will only discuss the discrepancies on S^2 here. For the rest of the discussion σ will denote the normalized surface measure on S^2, i.e $\sigma(S^2) = 1$.

Recall from the previous section, in $[0, 1]^d$ low discrepancy sequences are used in *Quasi-Monte Carlo Integration*, and *the Koksma-Hlawka Inequality* gives the error bound in integral estimation. Analogously, what if we are to estimate an integral on the unit sphere with a finite sum using a similar technique, $\int_{S^2} f d\sigma \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i)$? Unfortunately, the answer is that there isn't even a satisfactory analogous notion of "Hardy-Krasue Variation" on S^2. The closet version of "Spherical Koksma-Hlawka Inequality" we have is the following:

Theorem 1.3.1. [18] *Using the unit operator* $\mathbf{D} = (-2\triangle^*)^{1/2}(-\triangle^* + \frac{1}{4})^{1/4}$ *of order* 3/2 *(*$\triangle^*$ *is the Beltrami Operator on* S^2*) we have*

$$|\frac{1}{N}\sum_{i=1}^{N} f(x_i) - \int_{S^2} f d\sigma| \leq \sqrt{6}D(X(N), \mathbf{D})\|f\|_{3/2}, \tag{1.1}$$

where f *is from the Sobolev Space* $\mathcal{H}^{3/2}(S^2)$*.* $D(X(N), \mathbf{D})$ *is called the generalized discrepancy associated with* $\mathbf{D}$ *and can be computed by*

$$4\pi D(\{x_1, x_2, ..., x_N\}, \mathbf{D}) = 1 - \frac{1}{N^2}\sum_{k\neq l} \log(1 + \|x_k - x_l\|/2)^2. \tag{1.2}$$

1.3.1 Various Spherical Discrepancies on S^2

Other than *the generalized discrepancy associated with* $\mathbf{D}$ introduced previously, the L_2 spherical cap discrepancies, which averages the local discrepancy for *spherical caps* (*defined in the next subsection*), is defined as:

Definition 1.3.2. (L_2 **Discrepancy)** [17] Let $P_N = \{x_1, ..., x_N\}$. The L_2 discrepancy D_2 of P_N is defined as

$$D_2(P_N) = (\int_{-1}^{1}\int_{S^2} |\frac{\#(P_N \cap C(t)}{N}) - \sigma(C(t))|^2 d\sigma dt)^{1/2}. \tag{1.3}$$

$C(t)$ is a cap on S^2 of height t.

The L_2 spherical cap discrepancy has the close connection to the distance sums of the N points.

Theorem 1.3.3. (Stolarsky Invariance) [16]

$$\frac{1}{N^2}\sum_{k\neq l}\|x_l - x_k\| + 4(D_2(x_1, x_2, ..., x_N))^2 = \int_{S^2}\int_{S^2}\|x - z\|d\sigma d\sigma = \frac{4}{3}, \tag{1.4}$$

where D_2 *is the* L_2 discrepancy *on* S^2*.*

The discrepancy we will focus on is introduced below, *the spherical cap discrepancy.*

1.3.2 Spherical Cap Discrepancy

Definition 1.3.4. (Spherical Cap Discrepancy) For $-1 \leq t \leq 1$, the ***spherical cap*** centered at ω of height t is defined as $C_{\omega,t} = \{x \in S^2 | \langle x, \omega\rangle \leq t\}$.

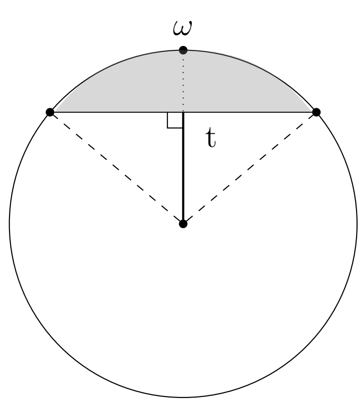


Figure 1.1: A spherical cap of height t centered at ω

Given a sequence X on S^2, the spherical discrepancy of the subsequence $X(N)$ is defined as:

$$D(X(N)) = \sup_{\omega\in S^2}\sup_{t\in[-1,1]}|\frac{\#(X(N)\cap C_{\omega,t})}{N} - \sigma(C_{\omega,t})|. \tag{1.5}$$

To date much about spherical cap discrepancy on S^2 is unknown. J. Beck's lower and upper bounds are the record holder so far.

Theorem 1.3.5. (Beck, lower bound) [13] *Given a set $\mathcal{P}$ of N points on S^2, there exists a spherical cap C with discrepancy*

$$|\#(\mathcal{P} \cap C)1 - N\sigma(C)| > c(2)N^{1/4}, \tag{1.6}$$

where $c(2)$ is a constant.

Theorem 1.3.6. (Beck, upper bound) [13] *For an arbitrary integer $N \geq 2$, there exists an N-element set $\{z_1, z_2, ..., z_N\} \subset S^2$ such that for any spherical cap $C \subset S^2$,*

$$|\#(\{z_1, ..., z_N\} \cap C) - N\sigma(C)| < c'(2)N^{1/4}(\log N)^{1/2}, \tag{1.7}$$

where $c'(2)$ is a constant.

In other words, for any sequence of length N, $P_N \subset S^2$, the lower bound for the spherical cap discrepancy

$$c(2)N^{-3/4} \leq D(P_N) \tag{1.8}$$

is always satisfied. And for a fixed length N there exists some sequence P_N such that the upper bound

$$D(P_N) \leq c'(2)N^{-3/4}\log N \tag{1.9}$$

is satisfied.

Beck's proofs of the upper and lower bounds are probabilistic and non-constructive.

Various ways of constructing uniformly distributed spherical sequences can be found in the literature. Many such sequences are obtained by mapping uniformly distributed sequences in $\mathbb{R}^2$ to S^2. However, it hasn't been verified that any of the existing sequences achieves the above bounds by Beck. In fact, a commonly agreed notion of ***low spherical discrepancy*** doesn't even exist. The main goal of this paper is to investigate efficient algorithms to distribute points on S^2 and what would be proper definition for "*low spherical discrepancy*". Before we do that, some related background of spherical geometry needs to be introduced.

Chapter 2

Spherical Geometry

By constructing spherical sequences and measuring discrepancies using spherical caps, we are entering a different geometric setting: the spherical geometry on S^2. While sharing lots analogous properties of the Euclidean geometry, spherical geometry has its own uniqueness, which might at first seem counterintuitive to those who are used to planar geometry. In Chapter 3, we will construct sequences directly on S^2, mainly using spherical triangulations. In Chapter 4, we will estimate the related distances of points on S^2 and, toward the end, the spherical cap discrepancy. All of these heavily rely on familiarities with knowledge on spherical geometry. The main reference of this Chapter is from the book " *Least Action Principle Of Crystal Formation Of Dense Packing Type And Kepler's Conjecture*" by W.Y Hsiang. We will introduce the definitions and theorems that are directly related to the discussion in upcoming chapters. Some facts and results are indirectly related to further discussion but fundamental in this subject.

We will introduce those too, for completeness.

2.1 Some Basic Background

The straight lines in spherical geometry are the great circles. The *spherical distance* between two points A and B on S^2 is the length in radius of *the shorter arc* of a great circle passing them. Unlike in plane geometry, there might be more than one *such shorter arc.* However, when A is not antipodal of B, the shorter arc segment passing A and B is unambiguous, which we denote by $\overset{\frown}{AB}$. The *spherical angle* at vertex A is denoted by $\measuredangle A$, in contrast to the Euclidean angle $\angle A$.

Notation 1. *To distinguish from plane geometry, we need to introduce some new notations. For three distinct points A, B, C on the unit sphere there exists a unique **spherical triangle** passing them with edges $\overset{\frown}{AB}$, $\overset{\frown}{BC}$ and $\overset{\frown}{AC}$. This **spherical triangle** is denoted by $\overset{\frown}{\triangle}\ ABC$, whereas the plane triangle is denoted by $\triangle ABC$. Every spherical triangle we consider in this paper has all of its edges less than π.*

Remark 2.1.1. The Triangle Inequality still holds true but only for the great circle segments that are less than π.

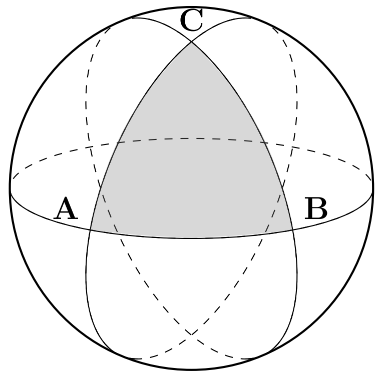


Figure 2.2: A good-looking spherical triangle $\overset{\frown}{\triangle}\, ABC$

Despite of many significant differences, spherical triangles and plane triangles do share some important similarities. For example, the *Side-Side-Side Theorem*, the *Side-Angle-Side Theorem*, the *Angle-Side-Angle Theorem*, the *Triangle Inequality* etc, still hold true in the spherical settings. According to the *Angle-Angle-Angle Theorem* $\measuredangle A$ $\measuredangle B$ and $\measuredangle C$ uniquely determine the spherical triangle $\overset{\frown}{\triangle}\, ABC$. However, one of the most striking results that spherical geometry doesn't share with planar geometry is that the area of a spherical triangle can be expressed in terms of the sum of the angles in a very short but exquisite way.

Theorem 2.1.2. (Area of Spherical Triangle)

$$\measuredangle A + \measuredangle B + \measuredangle C - \pi = \text{area}(\overset{\frown}{\triangle}\, ABC) \tag{2.1}$$

In contrast to the well-known fact that the sum of all angles of an Euclidean triangle is always equal to π, this theorem immediately gives us:

Corollary 2.1.3. *The sums of all the angles of a spherical triangle is always strictly* *greater than* π*. Each angle of an equilateral spherical triangle is strictly bigger than* $\pi/3$*.*

Notation 2. *When there's no ambiguity, we may use* $\overset{\frown}{\triangle} ABC$ *in place of* $\mathrm{area}(\overset{\frown}{\triangle} ABC)$*. O will denote the origin, unless specified otherwise.* $\boldsymbol{a}$*,* $\boldsymbol{b}$ *and* $\boldsymbol{c}$ *denote the arcs* $\overset{\frown}{BC}$*,* $\overset{\frown}{AC}$ *and* $\overset{\frown}{AB}$ *respectively.*

For the rest of this section, we will introduce more notions, quantities and fundamental computational rules in spherical geometry.

Remark 2.1.4. The three vertices A, B, C of a spherical triangle $\overset{\frown}{\triangle} ABC$ lie both on S^2 and in $\mathbb{R}^3$. The *circumcircle* of $\overset{\frown}{\triangle} ABC$ on S^2 is also the *circumcircle* of the Euclidean triangle $\triangle ABC$ in $\mathbb{R}^3$. $\measuredangle A$ *is precisely the* Euclidean angle *between the plane passing* O, A, B *and the plane passing* O, A, C*, and the three arcs* a*,* b*,* c *denoting* $\overset{\frown}{BC}$*,* $\overset{\frown}{AC}$*,* $\overset{\frown}{AB}$ *are the angles between* $\overrightarrow{OB}$ *and* $\overrightarrow{OC}$*,* $\overrightarrow{OA}$ *and* $\overrightarrow{OC}$*,* $\overrightarrow{OA}$ *and* $\overrightarrow{OB}$ *respectively.*

Lemma 2.1.5. *Let* $D = \det(\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC})$*, the determinant of the* 3×3 *matrix with columns* $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$*. Then*

$$D = (1 + 2\cos\boldsymbol{a}\cos\boldsymbol{b}\cos\boldsymbol{c} - \cos^2\boldsymbol{a} - \cos^2\boldsymbol{b} - \cos^2\boldsymbol{c})^{1/2} \tag{2.2}$$

One characterization of the relation between D, the volume of the parallelepiped generated by $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$, and the area of the spherical triangle $\overset{\frown}{\triangle} ABC$ is the following lemma.

Lemma 2.1.6. [9], *pp.*37*,* Lemma 2.1.1

Let D *be defined as above. Set* $u = 1 + \cos\boldsymbol{a} + \cos\boldsymbol{b} + \cos\boldsymbol{c}$*. Then*

$$\tan\frac{\triangle}{2} = \frac{D}{u}, \tag{2.3}$$

where the symbol $\triangle$ stands for the area of the spherical triangle.

Theorem 2.1.7. (Spherical Rules of Sines) ([9] *pp.* 30)

$$\frac{\sin \boldsymbol{a}}{\sin A} = \frac{\sin \boldsymbol{b}}{\sin B} = \frac{\sin \boldsymbol{c}}{\sin C} = \frac{\sin \boldsymbol{a} \sin \boldsymbol{b} \sin \boldsymbol{c}}{D}. \tag{2.4}$$

Theorem 2.1.8. (Spherical Rules of Cosines)([9] *pp.* 31)

$$\cos \boldsymbol{a} = \cos \boldsymbol{b} \cos \boldsymbol{c} + \sin \boldsymbol{b} \sin \boldsymbol{c} \cos A$$

$$\cos \boldsymbol{b} = \cos \boldsymbol{a} \cos \boldsymbol{c} + \sin \boldsymbol{a} \sin \boldsymbol{c} \cos B$$

$$\cos \boldsymbol{c} = \cos \boldsymbol{a} \cos \boldsymbol{b} + \sin \boldsymbol{a} \sin \boldsymbol{b} \cos C$$

When one of the angles is $\pi/2$, we have a special case of *the cosine rule*.

Corollary 2.1.9. ***(Traditional Spherical Pythagorean Theorem)***

$$\textit{If } \measuredangle C = \pi/2, \textit{ then } \cos \boldsymbol{c} = \cos \boldsymbol{a} \cos \boldsymbol{b}.$$

The cosine rule for $\measuredangle C = \pi/2$ leads us to a sequence of natural questions. Is this *Traditional Spherical Pythagorean Theorem* the analogue of the famous Euclidean *Pythagorean Theorem* in the spherical case? Is an spherical triangle with one its angle equal to $\pi/2$ the analogue of a planar right triangle? What would be a proper definition of **spherical right triangle**? And how about **spherical acute triangles** and **spherical obtuse triangles**? It turns out that being a *spherical right triangle* isn't as simple as possessing a 90 degree angle.

Definition 2.1.10. (Traditional Spherical Right Triangle) $\overset{\frown}{\triangle}\ ABC$ is called a *traditional spherical right triangle* when one of its angles is equal to $\pi/2$.

Before we can proceed, we need to introduce the *circumcircle* and *circumcenter* of $\overset{\frown}{\triangle}\, ABC$.

Definition 2.1.11. (Circumcircle and Circumdisk) The plane passing A, B, C intersects S^2 by a circle, the circumcircle of $\overset{\frown}{\triangle}\, ABC$. There exists a unique point on S^2 whose distances to A, B, C are equal. This unique point is the circumcenter of $\overset{\frown}{\triangle}\, ABC$. The closure of the interior of the circumcircle is called the circumdisk.

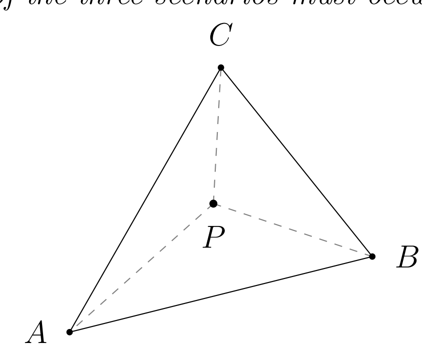
Remark 2.1.12. The circumcircle of the spherical triangle $\overset{\frown}{\triangle}\, ABC$ and the circumcircle of the planar triangle $\triangle ABC$ coincide.

The circumcenter of $\overset{\frown}{\triangle}\, ABC$ can be inside, or outside, or on an edge of $\overset{\frown}{\triangle}\, ABC$. The position of the circumcenters directly linked to the sizes of the angles, and, further, it divides all spherical triangles into three categories.

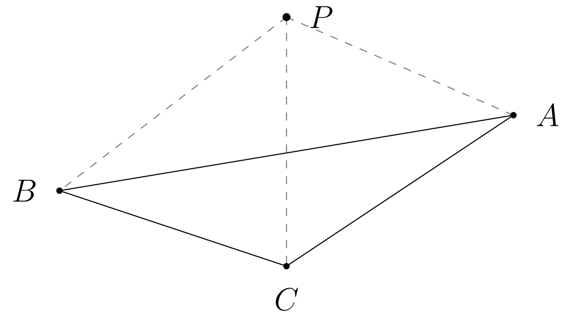
Theorem 2.1.13. [8], Theorem 2.

Let P be the circumcenter of $\overset{\frown}{\triangle}\, ABC$. One of the three scenarios must occur:

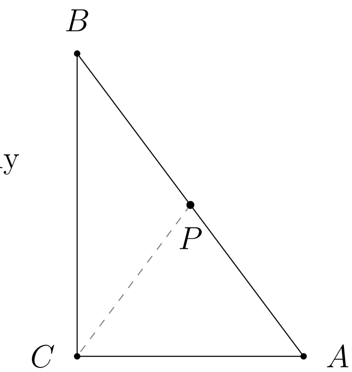
(1). C and P is on the same side of $\overset{\frown}{AB}$ if and only if $\measuredangle A + \measuredangle B > \measuredangle C$.



(2). C and P is on the opposite sides of $\overset{\frown}{AB}$ if and only if $\measuredangle A + \measuredangle B < \measuredangle C$.



(3). C and P is on the edge $\overset{\frown}{AB}$ if and only if $\measuredangle A + \measuredangle B = \measuredangle C$.



Definition 2.1.14. When $\measuredangle A + \measuredangle B < \measuredangle C$, we say $\overset{\frown}{\triangle}\, ABC$ is an **obtuse spherical triangle**. When $\measuredangle A+\measuredangle B > \measuredangle C$, $\measuredangle B+\measuredangle C > \measuredangle A$ and $\measuredangle A+\measuredangle C > \measuredangle B$, we say $\overset{\frown}{\triangle}\, ABC$ is a **acute spherical triangle**.

Corollary 2.1.15. *Let P be the circumcenter of the spherical triangle $\overset{\frown}{\triangle}\, ABC$.*

(i) If $\overset{\frown}{\triangle}\, ABC$ is acute then P lies in the interior of $\overset{\frown}{\triangle}\, ABC$.

(ii) If $\overset{\frown}{\triangle}\, ABC$ is obtuse and P and C lie on the opposite side of $\overset{\frown}{AB}$, then P and A lie the same side of $\overset{\frown}{BC}$ and P and B lie on the same side of $\overset{\frown}{AC}$.

Proof. The great circles passing $\overset{\frown}{AB}$, $\overset{\frown}{AC}$ and $\overset{\frown}{BC}$ divide the sphere into the following region: □

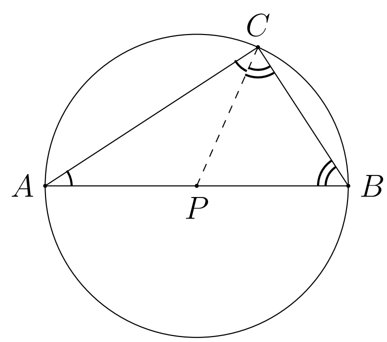
A spherical triangle with one of its angles equal to the sum of the other two is particularly of interest.

Definition 2.1.16. (Preferred Spherical Right Triangle)

If $\measuredangle A + \measuredangle B = \measuredangle C$, we say $\overset{\frown}{\triangle}\, ABC$ is a **preferred spherical right triangle**. Since in this case one angle is the sum of the other two, a preferred spherical right triangle is also called a **spherical half-sum triangle**.

Spherical half-sum triangles gain the name "preferred" for a reason. The definition immediately gives the following corollaries.

Corollary 2.1.17. *Let $\overset{\frown}{\triangle}\, ABC$ be a spherical half-sum triangles with circumcenter P and $\measuredangle A + \measuredangle B = \measuredangle C$, as the figure below suggested.*



The circumcenter of $\overset{\frown}{\triangle}\, ABC$ is the mid-point of the longest edge of the triangle, i.e the one facing $\measuredangle C$, which is defined to be the hypotenuse *of $\overset{\frown}{\triangle}\, ABC$. $\overset{\frown}{\triangle}\, APC$ and $\overset{\frown}{\triangle}\, CPB$ are* isosceles spherical triangles.

In fact, a spherical triangle can never be both a spherical right-angle triangle and spherical half-sum triangle.

Corollary 2.1.18. *A spherical triangle with an angle equal to $\pi/2$ **cannot** be a right spherical triangle.*

When $\measuredangle A + \measuredangle B = \measuredangle C$, clearly the area theorem gives $\measuredangle C = \frac{1}{2}\pi + \frac{1}{2} \overset{\frown}{\triangle} ABC$. The *Spherical half-sum triangles* lead us to answer some previously asked questions. One of such is the spherical analogue of the classic Pythagorean Theorem.

Theorem 2.1.19. *(Preferred Spherical Pythagorean Theorem)* [8], Theorem 3 *In a spherical half-angle triangle with hypotenuse* $\overset{\frown}{AB} = \boldsymbol{c}$

$$\sin^2(\frac{\boldsymbol{a}}{2}) + \sin^2(\frac{\boldsymbol{b}}{2}) = \sin^2(\frac{\boldsymbol{c}}{2}). \tag{2.5}$$

Corollary 2.1.20. *In a spherical half-angle triangle with hypotenuse* $\overset{\frown}{AB} = \boldsymbol{c}$*,*

$$\cos^2(\frac{\boldsymbol{a}}{2}) + \cos^2(\frac{\boldsymbol{b}}{2}) = \cos^2(\frac{\boldsymbol{c}}{2}). \tag{2.6}$$

Notation 3. *Now that we know the* Spherical half-sum triangles *are the very "right spherical right triangle". In the remaining discussion, by "spherical right triangle" or even "right triangle", we are only referring to* the spherical half-sum triangles.

2.2 More on Spherical Geometry (I)

One of our upcoming topics is to generate points directly over the unit sphere. More specifically, the points generated are going to be the *circumcenters* of certain spherical triangles. As the number of points grows, the distribution gets exponentially more complicated. When two spherical triangles, say $\overset{\frown}{\triangle} ABD$ and $\overset{\frown}{\triangle} BCD$, share a common edge, $\overset{\frown}{BD}$ the location of the circumcenters of $\overset{\frown}{\triangle} ABD$ and $\overset{\frown}{\triangle} BCD$ provides crucial information for our point generating algorithm. When A, B, C, D lie on the same circum circle, we say they are *co-circular*.

Notation 4. *When A and C are on different sides of edge $\overset{\frown}{BD}$, we say $\overset{\frown}{\triangle}\ ABD$ and $\overset{\frown}{\triangle}\ BCD$ are adjacent or they are "neighbours". The common edge they share, $\overset{\frown}{BD}$, is also the intersection of the circumdisks of $\overset{\frown}{\triangle}\ ABD$ and $\overset{\frown}{\triangle}\ BCD$ (or equivalently the intersection of the two planes that the circumcircles of $\overset{\frown}{\triangle}\ ABD$ and $\overset{\frown}{\triangle}\ BCD$ lie in).*

Theorem 2.2.1. [8], Theorem 5.

Assume A, B, C, D are co-circular. Then $\sin(\measuredangle A - \frac{1}{2}\overset{\frown}{\triangle}\ ABD) = \sin(\measuredangle C - \frac{1}{2}\overset{\frown}{\triangle}\ BCD)$. Further, one of the following two cases must occur.

(1) If A and C are on the same side of $\overset{\frown}{BD}$, $(\measuredangle A - \frac{1}{2}\overset{\frown}{\triangle}\ ABD) = (\measuredangle C - \frac{1}{2}\overset{\frown}{\triangle}\ BCD)$.

(2) Otherwise, $(\measuredangle A - \frac{1}{2}\overset{\frown}{\triangle}\ ABD) + (\measuredangle C - \frac{1}{2}\overset{\frown}{\triangle}\ BCD) = \pi$.

Corollary 2.2.2. [9] *pp.* 26

Let $\overset{\frown}{\triangle}\ ABC$ and $\overset{\frown}{\triangle}\ A'BC$ have the same orientation.

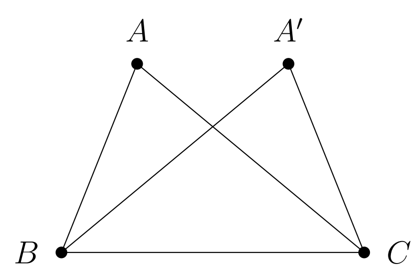


Figure 2.3: A, A' are on the same side of edge BC.

Then A, A', B, C are co-circular if and only if

$$\measuredangle ABC + \measuredangle ACB - \measuredangle BAC = \measuredangle A'BC + \measuredangle A'CB - \measuredangle BA'C. \qquad (2.1)$$

Remark 2.2.3. That $\overset{\frown}{\triangle}\ ABC$ and $\overset{\frown}{\triangle}\ A'BC$ have the same orientation is equivalent to that A and A' are on the same side of edge $\overset{\frown}{BC}$.

Corollary 2.2.4. *Let $\overset{\frown}{\triangle}\ ABC$ and $\overset{\frown}{\triangle}\ A'BC$ have the same orientation. Then,*

(i) A' is outside the circumcircle of $\overset{\frown}{\triangle}\ ABC$ if and only if $\measuredangle ABC + \measuredangle ACB - \measuredangle BAC < \measuredangle A'BC + \measuredangle A'CB - \measuredangle BA'C$.

(ii) A' is inside the circumcircle of $\overset{\frown}{\triangle}\ ABC$ if and only if $\measuredangle ABC + \measuredangle ACB - \measuredangle BAC > \measuredangle A'BC + \measuredangle A'CB - \measuredangle BA'C$.

Proof. The great circle passing A and the circumcenter of $\overset{\frown}{\triangle}\ ABC$ intersect with the circumcircle of $\overset{\frown}{\triangle}\ A'BC$. We may assume A' is the intersection.

When A' is outside the circumcircle of $\overset{\frown}{\triangle}\ ABC$, A is inside the $\overset{\frown}{\triangle}\ A'BC$. So,

$$\measuredangle ABC < \measuredangle A'BC, \text{ and } \measuredangle ACB < \measuredangle A'CB. \tag{2.2}$$

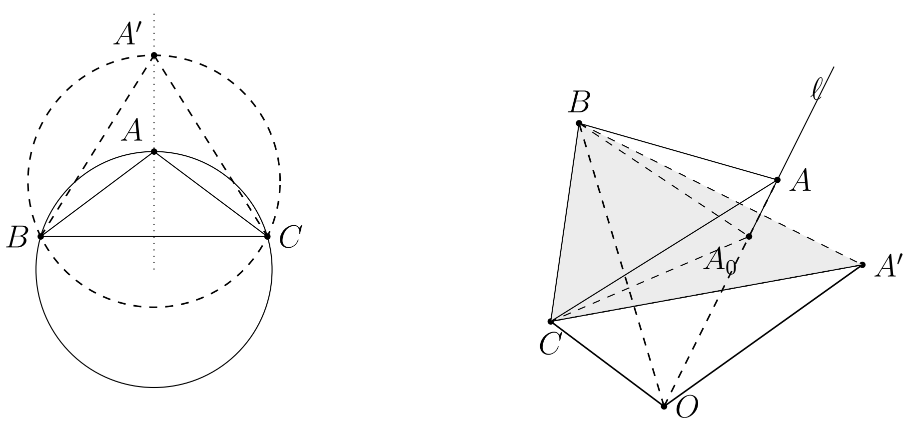


Figure 2.4

$\measuredangle BAC$ is the angle between the plane passing $\{O, A, B\}$ and the plane passing $\{O, A, C\}$. *The (Euclidean) straight line* passing A and O intersect the plane of *the plane triangle*

$\triangle$ at point A_0. Then

$$\angle BA_0C < \measuredangle BAC. \tag{2.3}$$

Since A_0 is in the interior of *the plane triangle* $\triangle BA'C$,

$$\angle BA'C < \angle BA_0C. \tag{2.4}$$

Meanwhile, since $\measuredangle BA'C$ is the angle between the plane passing $\{O, A', B\}$ and the plane passing $\{O, A', C\}$,

$$\angle BA'C > \measuredangle BA'C. \tag{2.5}$$

So we have

$$\measuredangle BA'C < \measuredangle BAC. \tag{2.6}$$

Therefore, combining inequality 2.2 and inequality 2.6, we get

$$\measuredangle ABC + \measuredangle ACB - \measuredangle BAC < \measuredangle A'BC + \measuredangle A'CB - \measuredangle BA'C. \tag{2.7}$$

This proves the *"if" direction* of statement (i). The *"if" direction* of statement (ii) follows by symmetry: interchanging the positions of A and A'.

The *only if direction* of (i) can be proven by contradiction. Assume $\measuredangle ABC + \measuredangle ACB - \measuredangle BAC < \measuredangle A'BC + \measuredangle A'CB - \measuredangle BA'C$. First, since this is a strict inequality, by Corollary 2.2.2, A' and A cannot be cocircular. If A' is inside the circumcircle of $\widehat{\triangle}\ ABC$, then the if direction of (ii) would give $\measuredangle ABC + \measuredangle ACB - \measuredangle BAC > \measuredangle A'BC + \measuredangle A'CB - \measuredangle BA'C$, a contradiction. The *"only if direction"* of (ii) following by symmetry of A and A'. $\square$

Definition 2.2.5. (Spherical Quadrilateral) [9] *pp.* 28

When the four edges $\overset{\frown}{AB}, \overset{\frown}{BC}, \overset{\frown}{CD}, \overset{\frown}{AD}$ are non-crossing, the four points A, B, C, D form a spherical quadrilateral, denoted by $\overset{\frown}{\square}\ ABCD$. The quadrilateral is called convex is each of its inner angle is at most π.

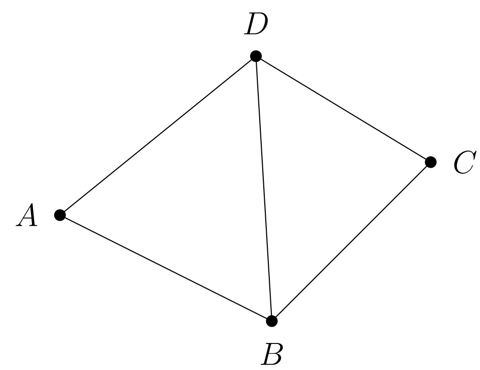


Figure 2.5: A Spherical Quadrilateral

Corollary 2.2.6. [9] *pp.* 28 *The four points of the spherical quadrilateral* $\overset{\frown}{\square}\ ABCD$ *are co-circular if and only if*

$$\measuredangle A + \measuredangle C = \measuredangle ABC + \measuredangle ADC. \tag{2.8}$$

Remark 2.2.7. For the purpose of our discussion, a spherical quadrilateral will always be assumed to have all four vertices on the same hemisphere.

Corollary 2.2.8. *For a spherical quadrilateral* $\overset{\frown}{\square}\ ABCD$*, if the four vertices are not co-circular, one of the following must occur.*

(i) C *is outside the circumcircle of* $\overset{\frown}{\triangle}\ ABD$ *if and only if* $\measuredangle A + \measuredangle C < \measuredangle B + \measuredangle D$.

(ii) C *is inside the circumcircle of* $\overset{\frown}{\triangle}\ ABD$ *if and only if* $\measuredangle A + \measuredangle C > \measuredangle B + \measuredangle D$.

Proof. (i) Assume C is outside the circumcircle of $\overset{\frown}{\triangle}\, ABD$. Denote the circumcenter of $\overset{\frown}{\triangle}\, ABD$ by I. The great circle passing I and C will intersect the circumcircle of $\overset{\frown}{\triangle}\, ABD$ at C'.

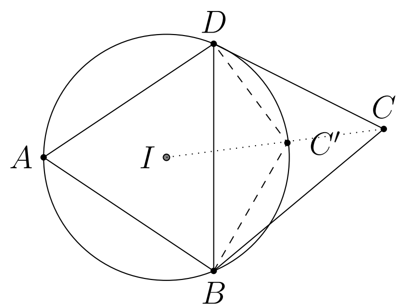


Figure 2.6

Since C' is cocircular with A, B, D, by Corollary 2.2.6, we know that

$$\measuredangle BAD + \measuredangle BC'D = \measuredangle ABC' + \measuredangle ADC'. \tag{2.9}$$

Now C and C' are on the same side of edge $\overset{\frown}{BD}$, Corollary 2.2.4 applies. We have

$$\measuredangle CDB + \measuredangle CBD - \measuredangle BCD > \measuredangle C'DB + \measuredangle C'BD - \measuredangle BC'D. \tag{2.10}$$

Subtract $\measuredangle BAD - \measuredangle ABD - \measuredangle ADB$ on both sides of the inequality above, we get

$$\measuredangle ABC+\measuredangle ADC-\measuredangle BAD-\measuredangle BCD > \measuredangle ADC'+\measuredangle ABC'-\measuredangle BAD-\measuredangle BC'D = 0. \tag{2.11}$$

Therefore,

$$\measuredangle ABC + \measuredangle ADC - \measuredangle BAD - \measuredangle BCD > 0, \tag{2.12}$$

and equivalently,

$$\measuredangle BAD + \measuredangle BCD < \measuredangle ABC + \measuredangle ADC. \tag{2.13}$$

This proves *the if direction* of (i). (ii) The proof is done similar way. If C is inside the circumcircle of $\overset{\frown}{\triangle}\ ABD$, then the great circle passing the circumcenter of $\overset{\frown}{\triangle}\ ABD$, I, and C will intersect the circumcircle of $\overset{\frown}{\triangle}\ ABD$.

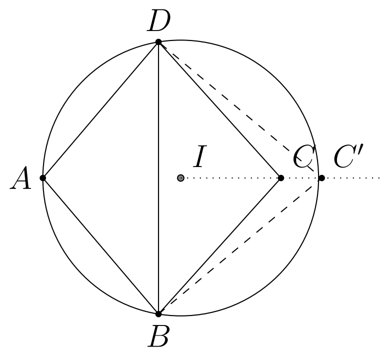


Figure 2.7

By Corollary 2.2.6,

$$\measuredangle BAD + \measuredangle BC'D = \measuredangle ABC' + \measuredangle ADC'. \tag{2.14}$$

Since C and C' are on the same side of edge $\overset{\frown}{BD}$ and C' is inside the circumcircle of $\overset{\frown}{\triangle}\ C'BD$, by Corollary 2.2.4,

$$\measuredangle CDB + \measuredangle CBD - \measuredangle BCD < \measuredangle C'DB + \measuredangle C'BD - \measuredangle BC'D. \tag{2.15}$$

Subtracting both sides of the above equation by $\measuredangle BAD - \measuredangle ABD - \measuredangle ADB$, we get

$$\measuredangle ABC + \measuredangle ADC - \measuredangle BAD - \measuredangle BCD < 0. \tag{2.16}$$

The *only if directions* of both (i) and (ii) follow easily. $\square$

Theorem 2.2.9. (Lexell's Theorem) [9] *pp.* 26-27

Let $\overset{\frown}{\triangle}\ ABC$ *and* $\overset{\frown}{\triangle}\ \tilde{A}BC$ *be two spherical triangles with the same orientation of vertices and the same area. Let* B' *and* C' *are the antipodal points of* B *and* C*. Then* $A, \tilde{A}, B'$ *and* C' *are co-circular.*

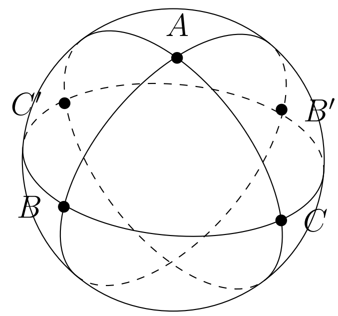


Figure 2.8: Lexel-Circle with base $\overset{\frown}{BC}$

Definition 2.2.10. (Lexell-circle) [9] *pp.* 27-28

Let B' and C' be the antipodal points of B and C. The locus of a point A such that $\overset{\frown}{\triangle}\ ABC$ has constant oriented area is an open circular arc with end points B' and C'. The circle passing A, B', C' is called the Lexell-circle of A (with base $\overset{\frown}{BC}$).

2.3 More on Spherical Geometry (II)

In this section, we will present and prove results that are specifically tailored to assist some later discussion and proofs.

Definition 2.3.1. An isosceles spherical triangle is a spherical triangle with two equal size, or equivalently a spherical triangle with two equal angles.

Remark 2.3.2. In spherical geometry it is possible to for an isosceles triangle to have the two equal angles bigger than $\pi/2$. Suppose in $\overset{\frown}{\triangle}\ ABC\ \measuredangle ACB = \measuredangle ABC$. The great circles passing A, B and A, C intersect at the antipodal point of A, denoted by $-A$. In $\overset{\frown}{\triangle}\ ABC\ \measuredangle ACB = \measuredangle ABC \leq \pi/2$ if and only if in $\overset{\frown}{\triangle}\ (-A)BC\ \measuredangle(-A)CB = \measuredangle(-A)BC \geq \pi/2$. So, with the extra assumption that every edge in an triangle is less than $\pi/2$ the two equal angles of an isosceles triangle will be less than $\pi/2$.

The symmetry of isosceles triangles gives the following proposition:

Proposition 2.3.3. *Given two distinct points A and B on S^2, the set of points on S^2 that are of equal distance to A and B is the great circle perpendicular to $\overset{\frown}{AB}$ at its mid-point.*

Proposition 2.3.4. (Larger Angle Face Larger Edge)

$$\textit{In } \overset{\frown}{\triangle}\ ABC,\ \measuredangle C \geq \measuredangle A \textit{ if and only if } \boldsymbol{c} \geq \boldsymbol{a}.$$

Proof. Since $\measuredangle C \geq \measuredangle A$, there exists a point D on edge $\overset{\frown}{AB}$ such that $\measuredangle A = \measuredangle ACD$.

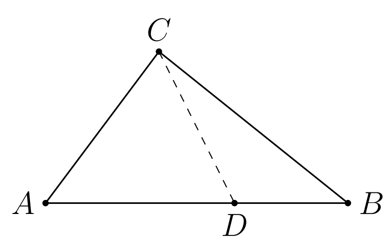


Figure 2.9

The triangle $\overset{\frown}{\triangle}$ is isosceles, $\overset{\frown}{AD}$=$\overset{\frown}{CD}$. By triangle inequality $\overset{\frown}{CD} + \overset{\frown}{BD} \geq \overset{\frown}{BC}$. That is $\overset{\frown}{AB}$=$\overset{\frown}{AD} + \overset{\frown}{BD} \geq \overset{\frown}{BC}$. $\mathbf{c} \geq \mathbf{a}$. ∎

If $\measuredangle C \geq \measuredangle A$, by Proposition 2.3.4 $\mathbf{c} \geq \mathbf{a}$. In other words, $\angle AOB \geq \angle BOC$. So, the Euclidean segments $AB \geq BC$. Consequently, in the planar triangle $\triangle ABC$, we have $\angle C \geq \angle A$. This can be characterized as:

Corollary 2.3.5. (Bigger Spherical Angle Bigger Planar Angle)
$\measuredangle C \geq \measuredangle A$ in $\overset{\frown}{\triangle}\ ABC$ if and only if in the corresponding planar triangle $\triangle ABC$ $\angle C \geq \angle A$.

Notation 5. *Let $\mathcal{C}$ be a circle on the unit sphere. $A, B \in \mathcal{C}$. The great circle segment $\overset{\frown}{AB}$ divides the circle into two pieces, each piece we call an "arc" of the circle $\mathcal{C}$.*

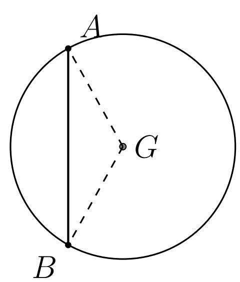


Figure 2.10

We call the longer arc between A, B "the heavier arc" and the shorter arc between A, B "the light arc". The area bounded by $\overset{\frown}{AB}$ and the heavier arc is called "the heavier side" of the circle, while the area bounded by $\overset{\frown}{AB}$ and the lighter arc is called "the light side" of the circle. Equivalently, the heavier side is the side contains the circumcenter while the lighter side doesn't.

Definition 2.3.6. (Central Angle of An Edge)

Using the same notation introduced above, the smaller $\measuredangle AGB$ (i.e the one $\leq \pi$) is called the central angle of edge $\widehat{AB}$.

Remark 2.3.7. If $\widehat{AB}$ happens to the intersection of two circles centered at G and I respectively.

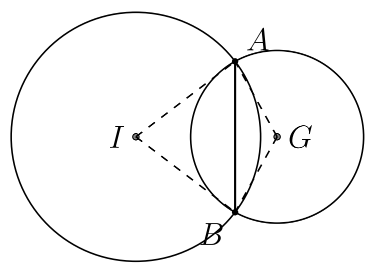


Figure 2.11

The circle centered at I is bigger than the circle centered at G if and only if the central angels satisfy $\measuredangle AIB < \measuredangle AGB$.

Chapter 3

Constructing Spherical Sequences

3.1 The First 50 Points

In this section we will present a recursive method of constructing "optimal" spherical sequences(the definition of optimal will be given later). In this method new points are added to the existing finite sequences using some "greedy algorithms" such that the newly added points have the "farthest distances" to the existing ones.

Definition 3.1.1. ***(The Spherical Distances)*** Given two points $x_1,\ x_2 \in S^2$ let $arc(x_1, x_2) = \langle x_1, x_2 \rangle$ and $dist(x_1, x_2) = \cos^{-1}(\langle x_1, x_2 \rangle)$, where $\langle \cdot\ ,\ \cdot \rangle$ is the standard inner product in $\mathbb{R}^3$. If $X \subset S^2$ is a set of points, define $arc(X, x_1) = \max\{arc(x, x_1)|\ x \in X\}$ and $dist(X, x_1) = \cos^{-1}(arc(X, x_1))$. For two sets of points $X, Y \subset S^2$ define $arc(X, Y) = \max\{arc(x, y)|\ x \in X, y \in Y\}$ and $dist(X, Y) = \cos^{-1}(arc(X, Y))$. $arc(X) = \max\{arc(x, y)|\ x, y \in X, x \neq y\}$ and $dist(X) = \cos^{-1}(arc(X))$.

Remark 3.1.2. $dist(x, y)$ is the length of the shorter great circle segments with ending points x and y on the unit sphere. Using the notation from Chapter 2, $dist(x, y) = |\ \widehat{xy}\ |$.

Notation 6. *Let $C(P, r)$ be the interior of the cap centered at P of spherical radius r. $\overline{C}$ is the closure of cap C. $rad(C)$ and $cen(C)$ denote the spherical radius and the center of the cap C respectively.*

Notation 7. *For $x \in S^2$ we will denote the antipodal point of x by $-x$.*

3.1.1 Method of The Largest Circumference

Given a sequence of N points $\mathbf{S}_N = \{P_1, P_2, ..., P_N\}$ on S^2 where $N \geq 3$ there is a unique circle passing through every 3 distinct points P_i, P_j, P_k. Every such circle divides the sphere into two disjoint caps. Let $C^1_{i,j,k}$ and $C^2_{i,j,k}$ denote the larger and smaller caps respectively. Among all these caps there exists a largest cap $C_{i',j',k'}$, by three points $P_{i'}, P_{j'}, P_{k'}$, whose interior doesn't contain any point of $\mathbf{S}_N$.

Definition 3.1.3. The existence of $C_{i',j',k'}$ in the previous paragraph is *rarely* unique. Let C_N denote the set of such caps and D_N denote the set of centers of caps in C_N. Elements in D_N are called "*deep holes*" by analogy to Lattice Theory, meaning the points farthest to the previous ones. Clearly C_N and D_N have the same number of elements.

We now start an explicit construction of one such sequence of 50 points using the recursion described above. The initial step of this recursive method requires 3 points. Start with the north pole $P_1 = (0, 0, 1)$; the furthest point from P_1 is the South Pole

$P_2 = (0, 0, -1)$. Any point on the equator has the furthest distance to P_1 and P_2. Choose the third point to be $(1, 0, 0)$.

The First 14 Points $\mathbf{S}_{14}$

C_3 consists of two caps: $C^1_{1,2,3}$ and $C^2_{1,2,3}$ are two disjoint hemispheres centered at $(0, 1, 0)$ and $(0, -1, 0)$. Choose the center of one of them to be P_4 say $(0, 1, 0)$, and $P_5 = (0, -1, 0)$. Among all the caps generated by P_1, P_2, P_3, P_4 and P_5, the hemisphere centered at $(-1, 0, 0)$ is the unique largest one containing no previous points. $C_5 = \{C((-1, 0, 0), \frac{\pi}{2})\}$. So, $P_6 = (-1, 0, 0)$.

Lemma 3.1.4. *The* deep holes *of* $\mathbf{S}_6$ D_6 *consists of 8 points:* $(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$*, i.e the circumcenters of the 8 octants.*

All the *deep holes* in D_6 are fairly far apart, as the following lemma states more precisely.

Lemma 3.1.5. *All the caps in* C_6 *has the same spherical radius* $r = \arccos(\frac{1}{\sqrt{3}})$*. For two distinct caps* $C(x, r), C(y, r)$*, where* x, y *are* ***deep holes*** D_6 $dist(x, y) > \arccos(\frac{1}{\sqrt{3}})$*. In other words, none of the caps in* C_6 *contains the centers of others.*

$\mathbf{S}_{14}$ is obtained by adding D_6 to $\mathbf{S}_6$.

Remark: The choice of the first 14 points is unique up to ordering. However, this is not the case as the recursion proceeds.

Introducing The Greedy Algorithm

By adding points only from the *deep holes* to the existing sequence, we get a new sequence. We say the existing sequence *extends* to the new one, and call the new sequence *the extension* of the existing one, e.g $\mathbf{S}_6$ extends to $\mathbf{S}_{14}$ and $\mathbf{S}_{14}$ is an extension of $\mathbf{S}_6$. This recursive *sequence extension* is a *greedy algorithm* and we will call it "*The Greedy Algorithm of The Largest Circumference*". Clearly, this algorithm produces finitely many *extended* sequences for each length n.

Definition 3.1.6. (*Optimal Sequences*)

Let $f(n) = \min\{arc(S)|S \subset S^2$ is a set of points of length n constructed using *The Greedy Algorithm of The Largest Circumference*$\}$ and $g(n) = \max\{dist(S)|S \subset S^2$ is a set of points of length n constructed using *The Greedy Algorithm of The Largest Circumference*$\}$. A sequence S_n of length n is *optimal* if $arc(S_n) = f(n)$ or, equivalently, $dist(S_n) = g(n)$.

We will give a table listing the approximated values of $f(n)$ for n up to 50 after introducing an *optimal* sequence of size 50.

The Next 12 Points

C_{14} and D_{14} are both of size 24. There exists a root $\alpha_{3,1}$ of $13x^4 - 10x^2 + 1$ and $\alpha_{3,3} = \sqrt{1 - (\alpha_{3,1})^2}$ is a root of the polynomial $13x^4 - 16x^2 + 4$ such that $\alpha_{3,1} \approx 0.806898221355073$ and $\alpha_{3,3} \approx 0.590690494568872$.

Notation 8. *Here we start using double subscript for the coordinates of the points. The first digit* 3 *in the double subscripts indicates this is the third set of* deep holes*, whereas the first and second sets of* deep holes *are* $\{P_i \mid i = 1, ..., 6\}$ *and* $\{P_j \mid j = 7, 8, 9, ..., 14\}$ *respectively.*

Computation shows $D_{14} = \{(\pm\alpha_{3,1}, \pm\alpha_{3,3}, 0), (0, \pm\alpha_{3,1}, \pm\alpha_{3,3}), (\pm\alpha_{3,3}, 0, \pm\alpha_{3,1}), (\pm\alpha_{3,1}, 0, \pm\alpha_{3,3}), (0, \pm\alpha_{3,3}, \pm\alpha_{3,1}), (\pm\alpha_{3,3}, \pm\alpha_{3,1}, 0)\}$. However, not all the 24 elements of D_{14} should be added to form the new sequence. The reason is the following: clearly all the 24 points from D_{14} lie on the "edges" between two adjacent octants, for example $x = (\alpha_{3,1}, \alpha_{3,3}, 0)$ and $y = (\alpha_{3,3}, \alpha_{3,1}, 0)$ and $dist(x, y) < dist(x, S_{14})$. A maximum of 12 elements can be chosen to form the new sequence $\mathbf{S}_{26}$: $\{P_{15}, ..., P_{26}\} = \{(\pm\alpha_{3,1}, 0, \pm\alpha_{3,3}), (0, \pm\alpha_{3,3}, \pm\alpha_{3,1}), (\pm\alpha_{3,3}, \pm\alpha_{3,1}, 0)\}$.

Definition 3.1.7. Denote the elements of the Abelian group $\mathbb{Z}_2\times\mathbb{Z}_2\times\mathbb{Z}_2$ by $\{(\pm 1, \pm 1, \pm 1)\}$. Let $\mathrm{N} \subset \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ be the subgroup $\{(\pm 1, 1, \pm 1)\}$ of order 4. $N \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. S_3 is the permutation group of degree 3 and $\langle(1\ 2\ 3)\rangle \subset \mathrm{S}_3$ is the subgroup of the 3 cycles. The group $(\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes \mathrm{S}_3$ is known as *the signed symmetric group* $\mathbb{Z}_2 \wr \mathrm{S}_3$, or *the octahedral group*, $\mathrm{S}_2 \wr \mathrm{S}_3$

We introduce the following group actions.

Notation 9. N *acts on elements of* D_{14} *by coordinatewise multiplication. For instance, if* $\tau = (-1, 1, -1)$ *then* $\tau\colon (\alpha_{3,3}, 0, \alpha_{3,1}) \to (-\alpha_{3,3}, 0, -\alpha_{3,1})$.

Notation 10. *Let* $\alpha_{3,2} = 0$. S_3 *acts on the elements of* D_{14} *by permuting the second digit* *of the subscript. For instance,* $\sigma = (1\ 2\ 3)$ *is the three cycle and* $\sigma\colon (-\alpha_{3,1}, 0, \alpha_{3,3}) \to (-\alpha_{3,3}, \alpha_{3,1}, 0)$.

Lemma 3.1.8. *The group* $\mathrm{G} = \mathrm{N} \rtimes \mathrm{S}_3$ *acts on* D_{14} *transitively.*

Lemma 3.1.9. *Let* $\mathrm{H} = \mathrm{N} \rtimes \mathrm{A}_3$. $\mathrm{H} \lhd \mathrm{G}$. H *acts on the 12 chosen elements from* D_{14} $\{(\pm\alpha_{3,1}, 0, \pm\alpha_{3,3}),\ (\pm\alpha_{3,3}, \pm\alpha_{3,1}, 0),\ (0, \pm\alpha_{3,3}, \pm\alpha_{3,1})\}$ *transitively.*

Proof. H is of index 2 and, hence, a normal subgroup of G. The transitivity of the two sets can be proven by computing the orbits of $(\pm\alpha_{3,1}, 0, \pm\alpha_{3,3})$ under G and H. $\square$

As we have commented before that the above choice of P_{15} to P_{26} *isn't* unique even up to re-ordering: one obvious substitution is take the other 12 elements in D_{14}. However, they are not too far away from being unique, as the following lemma states:

Lemma 3.1.10. *There are two choices for the 15th to 26th points: our choice* $\{P_{15}, ..., P_{26}\}$ *above, and its complement in* D_{14}. *The two choices are isometric.*

Proof. Let Ω denote $\{P_{15}, ..., P_{26}\}$. Let its complement in $D_{14} \setminus \Omega$ be denoted by Ω^C. Ω is the image of $(\alpha_{3,1}, 0, \alpha_{3,3})$ under the index 2 normal subgroup H of G. Ω is mapped to Ω^C by any two cycle of S_3. This proves the two options are isometric.

To show Ω and Ω^C are the only two possible options, let $\Gamma \subset D_{14}$ be a subset of 12 elements. $|\Omega \cap \Gamma| = d$ for some $d = 0, 1, ..., 12$, and $|\Omega^C| = 12 - d$. As a simple application of the *Pigeonhole Principle*, Γ contains at least a pair of points x, y such that $\eta(x) =$

$(\alpha_{3,1}, 0, \alpha_{3,3})$ and $\eta(y) = (\alpha_{3,3}, 0, \alpha_{3,1})$. Hence, $dist(x, y) = dist(\eta(x), \eta(y)) \le dist(\mathbf{S}_{26})$. where the equality holds *if and only if* $d = 0$ or 12. $\square$

Remark 3.1.11. The two choices above for $\{P_{15}, ..., P_{26}\}$ are the only ones leading to the first 50 optimal points. We will later give an explicit example of optimal 51 points where $\{P_{15}, ..., P_{26}\}$ is chosen differently.

The Next 12 Points P_{27} ***To*** P_{38}

C_{26} and D_{26} are both of size 24. Just like the case of D_{14}, the 24 elements follow certain patterns of symmetry and *not all* the 24 elements can be chosen to *extend* the sequence.

Notation 11. *We will continue using the same group action, the* 48 *symmetries of the group of* signed permutation $\mathbb{Z}_2 \wr \mathrm{S}_3$, *where* $\mathbb{Z}_2$ *acts on the sign of each coordinate and* S_3 *permutes the* 3 *coordinates.*

Let λ_4 be the root of the polynomial $12409t^8 - 12268t^6 + 2286t^4 - 124t^2 + 1$ with $\lambda_4 \approx 0.874423504330819$. Let $\alpha_{4,1} > \alpha_{4,2} > \alpha_{4,3} > 0$ be such that $arc((\alpha_{4,1}, \alpha_{4,2}, \alpha_{4,3}), S_{26}) = \lambda_4$. $\alpha_{4,1}$ happens to be equal to λ_4. $\alpha_{4,2}$ is a root of the polynomial $12409x^8 - 10284x^6 + 2798x^4 - 252x^2 + 1$ and $\alpha_{4,3}$ is a root of the polynomial $12409x^8 - 27084x^6 + 19814x^4 - 5100x^2 + 169$. Their approximated numerical values are $\alpha_{4,2} \approx 0.443562574092605$, and $\alpha_{4,3} \approx 0.196559858409978$.

Lemma 3.1.12. *The group* $\mathbb{Z}_2 \wr \mathrm{A}_3$ *acts on* D_{26} *transitively.*

Lemma 3.1.13. *Let* Ω *be the image of* $(\alpha_{4,1}, \alpha_{4,2}, \alpha_{4,3})$ *under the action of group* $(\mathbb{Z}_2 \times 1 \times \mathbb{Z}_2) \rtimes \mathrm{A}_3$*, which is an index 2 subset of* $\mathbb{Z}_2 \wr \mathrm{A}_3$*.* $|\Omega| = 12$*. Furthermore, if* $\Omega_k \subset \Omega$ *is of size* k*,* $arc(\Omega_k \cup \mathbf{S}_{26}) = f(26+k)$*.*

Adding all elements of Ω to $\mathbf{S}_{26}$, we get the extended sequence $\mathbf{S}_{38}$.

Remark 3.1.14. We choose Ω to be $\{P_i \mid i = 27, ..., 38\}$. A counting argument shows that there are 4^6 choices for $\{P_i \mid i = 27, 28, ..., 38\}$ that achieve $f(38)$. They are *not* all isometric to Ω. Each of the 4^6 configurations can lead to 50 optimal points, although their further extensions beyond 50 points will be suboptimal.

The first 50 points S_{50}

$|D_{38}| = 12$. Given a "deep hole" $x \in D_{38}$ denote $arc(x, S_{38})$ by λ_5, where λ_5 is a root of the polynomial $13637110513t^{16} - 16758767896t^{14} + 5707856588t^{12} - 821722248t^{10} + 59096342t^8 - 2239848t^6 + 43436t^4 - 376t^2 + 1$ and $\lambda_5 \approx 0.885967389267793$. There exists $(\alpha_{5,1}, \alpha_{5,2}, \alpha_{5,3}) \in D_{38}$ such that $\alpha_{5,1} = \lambda_5$, $\alpha_{5,2} \approx 0.372960097967509$ and $\alpha_{5,3} \approx 0.275613044825668$.

Lemma 3.1.15. *Let* Ω *be the orbit of* $(\alpha_{5,1}, \alpha_{5,2}, \alpha_{5,3})$ *under the action by* $\mathbb{Z}_2 \wr \mathrm{A}_3$*. Let* Γ *be the orbit of* $(\alpha_{5,1}, \alpha_{5,2}, \alpha_{5,3})$ *under the action of the index* 2 *subgroup* $(\mathbb{Z}_2 \times 1 \times \mathbb{Z}_2) \rtimes \mathrm{A}_3$*. Then* $D_{38} = \Omega \setminus \Gamma$*.*

Lemma 3.1.16. $arc(D_{38} \cup \mathbf{S}_{38}) = \lambda_5$*.*

By adding all 12 elements of D_{38} to $\mathbf{S}_{38}$ we extend $\mathbf{S}_{38}$ to $\mathbf{S}_{50}$. Let $\alpha > \beta > \gamma > 0$ be

Theorem 3.1.17. $\mathbf{S}_{50}$ *is optimal. (See Definition 3.1.6 for optimal sequences)*

Remark 3.1.18. The sequence $\mathbf{S}_{50}$ constructed above *cannot* be *extended* further to any S_{51} without making $dist(S_{51}) < dist(\mathbf{S}_{50})$. The following table lists some numerical data of the *spherical distance* of subsequences of $\mathbf{S}_{50}$:

$n \geq 3$	$\text{arc}(\mathbf{S}_n) = f(n)$
3, ..., 6	0
7, ..., 14	$\frac{1}{\sqrt{3}} \approx 0.577350269189625$
15, ..., 26	$\alpha_{3,1} \approx 0.806898221355073$
27, ..., 38	$\lambda_4 \approx 0.874423504330819$
39, ..., 50	$\lambda_5 \approx 0.885967389267793$

Table 3.1

3.1.2 Extension/Subsequence of Sequences

However, despite of the careful construction of choosing "the furthest points" each step, an optimal sequence of length N may not be able to *extend* to a longer optimal sequence.

Lemma 3.1.19. *There exists a sequence of 51 points* $\mathbf{K}_{51}$ *constructed using* The Greedy Algorithm of Largest Circumference *such that* $arc(\mathbf{K}_{51}) < arc(S_{51})$ *or equivalently* $dist(\mathbf{K}_{51}) > dist(S_{51})$ *where* S_{51} *is an* extension $\mathbf{S}_{50}$.

From the lemma immediately follows

Theorem 3.1.20. *There doesn't exist an infinite sequence whose finite truncations are all optimal.*

In the following proof of *Lemma 10* $\mathbf{K}_{51}$ will be constructed using the same greedy algorithm, i.e *The Greedy Algorithm of The Largest Circumference.* At each step we choose the maximum number of points from the set of *deep holes*. However, unlike the construction of $\mathbf{S}_{50}$ above, our choices for elements of $\mathbf{K}_{51}$ may not follow any group pattern. As we will see, this causes the first 50 elements of $\mathbf{K}_{51}$ to be suboptimal.

Proof. Start with the same first 14 points as $\mathbf{S}_{50}$, i.e $\mathbf{K}_{14}$ consists of the 6 poles and the circumcenters of the 8 octants.

The next 12 elements: Recall from the previous, the set of *deep holes* for $\mathbf{K}_{14}$ consists of 24 elements. Let α and β be equal to $\alpha_{3,1}$ and $\alpha_{3,3}$ respectively. $\mathbf{K}_{26}$ is obtained by adding the following 12 elements to $\mathbf{K}_{14}$:

$$\{(0,\beta,\alpha),(\beta,0,\alpha),(\beta,\alpha,0),\ (0,-\beta,-\alpha),(-\beta,0,-\alpha),(-\beta,-\alpha,0),$$
$$(0,-\beta,\alpha),(0,\alpha,-\beta),(\alpha,-\beta,0),\ (-\alpha,\beta,0),(\alpha,0,-\beta),(-\beta,0,\alpha)\}.$$

$arc(\mathbf{K}_{26}) = f(26)$***.*** $\mathbf{K}_{26}$ ***is optimal.***

The next 11 elements: $\mathbf{K}_{26}$ has 24 *deep holes*. Let $\alpha > \beta > \gamma > 0$ be equal to $\alpha_{4,1}$, $\alpha_{4,2}$ and $\alpha_{4,3}$ respectively. Unlike the case $\mathbf{S}_{26}$, only 11 deep holes of $\mathbf{K}_{26}$ can be selected. Adding the following 11 elements to $\mathbf{K}_{26}$ we get **an optimal sequence of**

length 37 $\mathbf{K}_{37}$:

$$\{(\beta, -\gamma, -\alpha), (-\alpha, -\gamma, \beta), (-\alpha, -\gamma, -\beta), (-\gamma, -\alpha, -\beta),$$
$$(\alpha, \gamma, \beta), (\beta, -\alpha, \gamma), (\gamma, \alpha, \beta), (\alpha, \beta, -\gamma),$$
$$(-\beta, \alpha, \gamma), (-\gamma, \beta, -\alpha), (-\gamma, -\alpha, \beta)\}.$$

Remark: $\mathbf{K}_{37}$ *cannot be extended to an optimal sequence of length 38.*

The extension to $\mathbf{K}_{43}$: $\mathbf{K}_{37}$ has 8 *deep holes*. Let $\alpha > \beta > \gamma > 0$ be equal to $\alpha_{5,1}$, $\alpha_{5,2}$ and $\alpha_{5,3}$ respectively. The 8 *deep holes* of $\mathbf{K}_{37}$ are

$$\{\ (\alpha, -\gamma, \beta),\ (-\beta, \alpha, -\gamma),\ (-\alpha, \gamma, \beta),\ (-\alpha, \gamma, -\beta),\ (\gamma, -\alpha, -\beta),\ (\beta, -\alpha, -\gamma),\ (\beta, \gamma, -\alpha),$$
$$(\gamma, \beta, -\alpha)\ \}.$$

Because $(\gamma, -\alpha, -\beta)$ is too close to $(\beta, -\alpha, -\gamma)$ and $(\beta, \gamma, -\alpha)$ is too close to $(\gamma, \beta, -\alpha)$, only one from each pair can be chosen. Adding the 6 *deep holes* (**other than** $(\beta, -\alpha, -\gamma)$ and $(\gamma, \beta, -\alpha)$) to $\mathbf{K}_{37}$ we get $\mathbf{K}_{43}$.

The 44th element: $\mathbf{K}_{43}$ has two deep holes, h_1 and h_2. Let $\lambda = arc(h_1, \mathbf{K}_{43}) = arc(h_2, \mathbf{K}_{43})$. $\lambda \approx 0.890617459428756$ happens to be a root of $833902585633x^{16} - 940416185384x^{14} + 249532655932x^{12} - 23661702488x^{10} + 1015878310x^{8} - 21031832x^{6} + 208252x^{4} - 872x^{2} + 1$. And $h_1 \approx$ (-0.890617459428756, -0.451777166189461, 0.051941631380693) and $h_2 \approx$ (-0.890617459428756, -0.451777166189461, -0.051941631380693). Clearly they are too close to be both chosen. Adding h_1 extends $\mathbf{K}_{43}$ to $\mathbf{K}_{44}$.

The 45th to the 51st points: $\mathbf{K}_{44}$ has 14 *deep holes*. The spherical distance between a *deep hole* h and the existing 44 points $arc(h, \mathbf{K}_{44}) \approx 0.907689792617791$ happens to be a root (of multiplicity 1)of $241x^8 - 380x^6 + 182x^4 - 28x^2 + 1$. A maximum of 7 of them can be added to extend to $\mathbf{K}_{51}$, and $arc(\mathbf{K}_{51}) = arc(h, \mathbf{K}_{44})$. □

Proof. (of Theorem) Clearly the first 50 points of $\mathbf{K}_{51}$ $arc(\mathbf{K}_{50}) > f(50)$. However, any extension of $\mathbf{S}_{50}$ by an *deep hole* S_{51} would have $arc(S_{51}) > arc(\mathbf{K}_{51})$. □

Extensions of $\mathbf{K}_{51}$

$\mathbf{K}_{51}$ cannot be extended further without changing the spherical distance of the resulting sequence. It turns out $\mathbf{K}_{51}$ has 3 *deep holes*, all of which are far apart enough that all three of them can be added to entend the sequence to $\mathbf{K}_{54}$. Even further, $arc(\mathbf{K}_{54}) \approx$ 0.914193954804357 is a simple root of $25270910733829842817x^{16} - 66588078845325380648x^{14} + 70522196178243041020x^{12} - 39067864879587856088x^{10} + 12307100102256438694x^{8} - 2238864443006380$

$2280270552799111 00x^4 -$

$119105145789985 04x^2 + 245664276089761$.

Since $\mathbf{K}_{51}$ has better spherical distance than any extension of $\mathbf{S}_{50}$ but yet can't be extended while maintaining the same spherical distance, it is natural question to ask whether there exists a longer sequence with the same sphercal distance. As experiments show there exists a sequence of length 53 with spherical distance equal to $arc(\mathbf{K}_{51})$. This completes our first 53 values for f:

$n \geq 3$	$f(n)$
3, ..., 6	0
7, ..., 14	$\approx 0.577350269189625$
15, ..., 26	$\approx 0.806898221355073$
27, ..., 38	$\approx 0.874423504330819$
39, ..., 50	$\approx 0.885967389267793$
51, 52, 53	$\approx 0.907689792617791$

Table 3.2

3.1.3 Separation - The"Minimal" Distance

The Method of Largest Circumference doesn't produce a unique sequence (unless the sequence is of length 14). Also, a sequence of length N, S_N, generated by this method can be sub-optimal. However, such a sequence S_N cannot be "*improved*" in the sense that the distance between the Nth point with the previous S_{N-1} is *maximized* and is the shortest distance between ay pair of points in S_N.

Definition 3.1.21. (*Separation of a Point Set*) ([9] *pp.* 66-67)

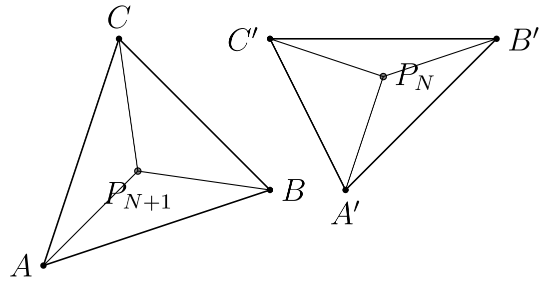
Let Σ be a finite point set on S^2. The minimal distance between any pair of points in Σ is called the separation of Σ, denoted by $\delta(\Sigma)$.

Definition 3.1.22. (*ϵ-saturated*) ([9] *pp.* 66-67)

A finite point set Σ on S^2 is called ϵ-saturated, if adding one more point to Σ will change *the separation* $\delta(\Sigma)$ *from at least* ϵ *to more than* ϵ*, i.e* $\delta(\Sigma) \geq \epsilon$ *but* $\delta(\Sigma \cup \{p\}) < \epsilon$*.*

Theorem 3.1.23. (Minimal Distance) S_N *is a sequence of length* N *created using the Method of The Largest Circumference. Let* P *be a deep hole of* S_N*. Then* $dist(P, S_N) = dist(S_N \cup \{P\})$*. In other words, the separation of the sequence* S_{N+1}*,* $\delta(S_{N+1}) = dist(P, S_N)$*.*

Proof. We will show the distance between S_N and its next deep hole is the shortest distance between any pair of points in S_{N+1} by induction. The inital step is when $N = 8$. The 9th point is the circumcenter of one of the 8 octants. The assertion clearly holds.
Let P_i denote the ith element of such a sequence. Assume $dist(P_N, S_{N-1}) = \delta(S_N)$ for $N \geq 9$. $S_{N+1} = S_N \cup \{P_{N+1}\}$, where P_{N+1} is a deep hole of S_N. We will show $dist(P_{N+1}, S_N) \leq dist(P_i, P_j)$ for any pair of points $P_i, P_j \in S_{N+1}$.
Suppose P_{N+1} is the circumcenter of $\overset{\frown}{\triangle}\, ABC$ and P_N is the circumcenter of $\overset{\frown}{\triangle}\, A'B'C'$, where $A, B, C, A', B', C' \in S_N$. $dist(P_N, S_{N-1})$ is equal to the circumradius of $\overset{\frown}{\triangle}\, A'B'C'$ and $dist(P_{N+1}, S_N)$ is equal to the circumradius of $\overset{\frown}{\triangle}\, ABC$.
Case 1: Suppose $A, B, C \in S_{N-1}$. The all $A, B, C, A', B', C' \in S_{N-1}$. In this case, $dist(P_{N+1}, S_N) = dist(P_{N+1}, S_{N-1})$. By the defintion of a deep hole, the newly added deep hole P_{N+1} must have circumraius no bigger than its previous deep hole P_N. $dist(P_{N+1}, S_{N-1}) \leq dist(P_N, S_{N-1})$. Hence, by the inductive hypothesis, $dist(P_{N+1}, S_N) \leq dist(P_i, P_j)$ for all $P_i, P_j \in S_N$.

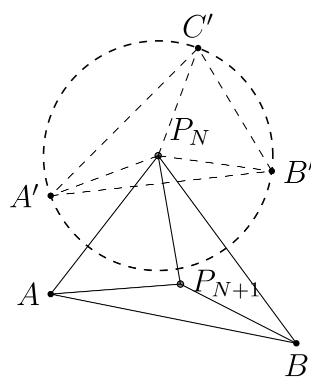


Suppose triangles exist in S_{N-1}. By the definition of deep holes, the circumradius of $\widehat{\triangle}\ A'B'C'$ is no smaller than the circumradius of $\widehat{\triangle}\ A'B'C'$.

Figure 3.12

Clearly, $dist(P_{N+1}, S_N) \leq dist(P_{N+1}, P_j$ for all $P_j \in S_N$. Therefore, $dist(P_{N+1}, S_N) = \delta(S_{N+1})$.

Case 2: Not all the three points A, B, C are in S_{N-1}. Or equivalently one of them is P_N. WLOG, assume $C = P_N$. In this case, $dist(P_{N+1}, S_N) = dist(P_{N+1}, P_N)$. It suffices to show $dist(P_N, P_{N+1}) \leq dist(P_N, S_{N-1})$, where $dist(P_N, S_{N-1})$ is the length of the circumradius of $\widehat{\triangle}\ A'B'C'$.



P_N is one of the three vertices of the spherical triangle whose circumcenter is P_{N+1}.

Figure 3.13

If $dist(P_N, P_{N+1}) > dist(P_N, S_{N-1})$, since P_{N+1} is a deep hole of S_N, $dist(P_{N+1}, P_i) > dist(P_N, S_{N-1})$ for all $P_i \in S_N$. This implies $dist(P_{N+1}, S_{N-1}) > dist(P_N, S_{N-1})$, which is a contradictiong to that P_N is a deep hole of S_{N-1}. Therefore, $dist(P_N, P_{N+1}) \leq dist(P_N, S_{N-1})$. $dist(P_N, P_{N+1}) = \delta(S_{N+1})$, completing the induction. □

Corollary 3.1.24. (Non-increasing Separation) *The separation of a sequence S_N construncted using the Method of Largest Circumference, $\delta(S_N)$, is non-increasing as N increses. And the separation decreases precisely when the sequence is $\delta(S_N)$-*saturated.

3.1.4 $\mathbf{S}_{50}$ And $\mathbf{K}_{51}$ From Field Extension Viewpoint

In the construction of the sequences $\mathbf{S}_{50}$ and $\mathbf{K}_{51}$, the coordinates of each points and different values of $f(n)$ turn out to be algebraic numbers. The explanation is simple. Suppose n poins are placed on S^2 already, the $(n+1)$th point is the circumcenter of some spherical triangle with vertices A, B, C. Let $A = (a_1, a_2, a_3)$, $B = (b_1, b_2, b_3)$ and $c = (c_1, c_2, c_3)$. Let the $(n+1)$th point be denoted by $P = (x, y, z)$. The x, y, x must satisfy:

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r \\ r \\ r \end{bmatrix}$$

$$x^2 + y^2 + z^2 = 1$$

Solving the system of four equations gives that x, y, z and r are all algebraic numbers and lie inside a field extension of degree 2 of the existing number field.

Notation 12. *The circumradius can be computed [9] pp. 35. Let* $u = 1 + \cos \boldsymbol{a} + \cos \boldsymbol{b} + \cos \boldsymbol{c}$ *and* $w = \cos \boldsymbol{a} \cos \boldsymbol{b} + \cos \boldsymbol{b} \cos \boldsymbol{c} + \cos \boldsymbol{c} \cos \boldsymbol{a}$. *$D$ is the same as introduced in Chapter 2. Let R denote the circumradius of the spherical triangle with edges* ***a*****,** ***b*****,** ***c***.

Lemma 3.1.25. *([9] pp.34)*

$$\sec^2 R = \frac{1}{D}(4(1 + w) - u^2) \tag{3.1}$$

Corollary 3.1.26. *Let* $\alpha = \cos \boldsymbol{a}$, $\beta = \cos \boldsymbol{b}$ *and* $\gamma = \cos \boldsymbol{c}$. *Then,*

$$\cos^2 R = \frac{1 - \alpha^2 - \beta^2 - \gamma^2 - 2\alpha\beta\gamma}{4(\alpha\beta + \beta\gamma + \gamma\alpha + 1) - (1 + \alpha + \beta + \gamma)^2}. \tag{3.2}$$

From the corollary it is also clear that the field extension concerned here is quadratic.

Corollary 3.1.27. *The coordinates of all points of this sequence are constructible numbers.*

3.2 The Orbit of The 8 Octants

Our idea of how to construct a well-spaced spherical sequence is clear: given a finite sequence on S^2 there exists a set of deep holes, of which we choose a maximal number of elements to add to the given sequence, until the separation of the sequence starts to decrease, and then we look for the next set of deep holes. However, this is much easier said than be done. Which elements of the set of deep holes are selected at any step may change the next set of deep holes. The choice may result in better separation of the sequence in the short term but not in the long run, or vice versa. One ambitious conjecture one may wish to be true is that *all sequences* Σ_N *constructed using the greedy algorithm of largest circumference have the same separation as* N *goes to* ∞*, and ultimately, the same spherical cap discrepancy.* Unfortunately very little is known in this field. Conjectures like this are merely believable-sounding. As one of the first steps on the road to more abstract and generalized conclusions on this subject, we would like to construct *one particular* well-spaced spherical sequence in a definite and deterministic fashion. Unlike the unit cube, the unit sphere S^2 has fascinating properties of rotational symmetry, which can be assisting as well as restricting.

3.2.1 An Orbit of The 8 Octants

The concept of orbits comes from group action. We will define a group G which acts on $\mathcal{O}$, the set of the 8 octants. G has a subgroup of index 2, which will act on two orbits of $\mathcal{O}$ transitively. Each orbit consists of 4 octants, and any two adjacent octants are in different orbits. $\mathcal{O}_+$ will denote the orbit containing the first octant while $\mathcal{O}_-$ will denote the orbit containing the octants adjacent to the first octant.

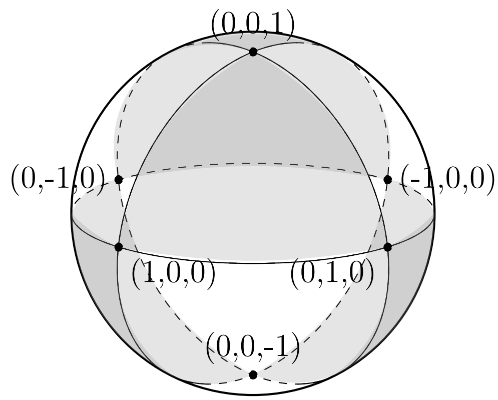


Figure 3.14: The shaded octants are of the same orbit $\mathcal{O}_+$.

Notation 13. *We will use 3-tuples of the "$+$" and "$-$" signs to denote each octant. For example, $(+++) = \{(x,y,z)|x,y,z > 0\}$. Notice, the octants are open sets. In other words, they do not contain their boundaries.*

3.2.2 The Group Action

Using the new notation, $\mathcal{O}_+ = \{(+++), (+--), (-+-), (--+)\}$ and $\mathcal{O}_- = \{(---), (-++), (+-+), (++-)\}$. It is not hard to observe that each octant in $\mathcal{O}_+$ has an odd number of "$+$" signs, and each octant in $\mathcal{O}_-$ has an odd number of "$-$" signs.

Definition 3.2.1. Let $\sigma_+ : \mathcal{O}_+ \longrightarrow \mathcal{O}_+$ such that

$$\sigma_+ : (+++) \longmapsto (+--),\ \sigma_+ : (+--) \longmapsto (-+-),$$

$$\sigma_+((-+-)) = (--+),\ \sigma_+ : (--+) \longmapsto (+++).$$

Corollary 3.2.2. *$\langle \sigma_+ \rangle \cong \mathbb{Z}_4$ acts on $\mathcal{O}_+$ by the definition of the map.*

We can define an analogous map and group action on $\mathcal{O}_-$:

Definition 3.2.3. Let $\sigma_- : \mathcal{O}_- \longrightarrow \mathcal{O}_-$ such that

$$\sigma_- : (- - -) \longmapsto (- + +), \sigma_- : (- + +) \longmapsto (+ - +),$$

$$\sigma_- : (+ - +) \longmapsto (+ + -), \sigma_- : (+ + -) \longmapsto (- - -).$$

Corollary 3.2.4. $\langle \sigma_- \rangle \cong \mathbb{Z}_4$ *acts on* $\mathcal{O}_-$ *by the definition of the map.*

The actions of σ_+ and σ_- are almost identical except for the sets they are acting on. It is natural to define the following:

Notation 14. *Define* $- : \mathcal{O} :\longrightarrow \mathcal{O}$ *such that* $-$ *changes the signs of each octant to the opposite, e.g* $-(+ + +) = (- - -)$*. And let* $+ : \mathcal{O} :\longrightarrow \mathcal{O}$ *be the identity map.* $\{+, -\} \cong \mathbb{Z}_2$.

Proposition 3.2.5. σ_+ *and* σ_- *are conjugates by* $- \circ \sigma_+ \circ - = \sigma_-$*.*

Let $\sigma = \sigma_+$ and define the group $G = \langle \sigma, - \rangle$. Clearly, G is nonabelian since σ doesn't commute with $-$.

Proposition 3.2.6. *G is isomorphic to the dihedral group of order 8.*

Proof. G is generated the same way as *the dihedral group of order 8.* $\square$

Proposition 3.2.7. *The subgroup of G $\langle \sigma \rangle$ act transitively on* $\mathcal{O}_+$ *and* $\mathcal{O}_-$*.*

In the next section, instead of considering the whole orbit of the 8 octants, we will consider the two orbits $\mathcal{O}_+$ and $\mathcal{O}_-$ separately.

3.2.3 Distributing Points Using The Orbits

The elements of G from the previous section induce maps on points on S^2. If $(x_0, y_0, z_0) \in S^2$ is a point in the first octant, define $\sigma : (x_0, y_0, z_0) \longmapsto (x_0, -y_0, -z_0)$. σ is now a bijective map sending elements of the octant $(+++)$ to points of the octant $(+--)$. Analogous maps by other elements of G can be defined similarly. In terms of distributing points on S^2, once a point (x_0, y_0, z_0) in the octant $(+++)$ is selected, we get 3 more points, one of each on another octant from the orbit $\mathcal{O}_+$. And once we select a point $(x_1, y_1, z_1) \in (---)$, the group action on $\mathcal{O}_-$ will give us 3 more points, one from each other octant of $\mathcal{O}_-$. $\langle\sigma\rangle$ is a normal subgroup of G, as the index of $\langle\sigma\rangle$ is 2, and the quotient group is isomorphic to $\mathbb{Z}_2$.

Theorem 3.2.8. *If we have a distribution of $\mathcal{L}$ points on the octant $(+++)$, and a distribution of $\mathcal{J}$ points on the octant $(-++)$, mapping these points to the other octants of each orbit of $\mathcal{O}_+$ and $\mathcal{O}_-$, we get a distribution of $12\mathcal{L} + 12\mathcal{J}$ points over the sphere S^2.*

Proof. In the following picture, the shaded octants are in the same orbits $\mathcal{O}_+$, and the rest is $\mathcal{O}_-$. On the boundaries of the octants, we can draw arrows as indicated in the picture, so that each octant in $\mathcal{O}_+$ has its boundary oriented counterclockwise, and each octant in $\mathcal{O}_-$ has its boundary oriented clockwise.

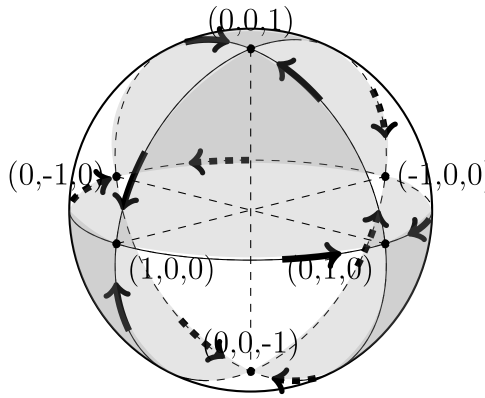


Figure 3.15: **The shaded octants are oriented counterclockwise; the unshaded ones are oriented clockwise**.

Suppose we are given a point in the octant $(+++)$ (x, y, z). Unless (x, y, z) is the center of the octant, we have (x, y, z), (z, x, y), and (y, z, x) are three distinct points inside $(+++)$. Then we get 9 more distinct points $\{(x, -y, -z), (z, -x, -y), (y, -z, -x)\}$, $\{(-x, y, -z), (-z, x, -y), (-y, z, -x)\}$ and $\{(-x, -y, z), (-z, -x, y), (-y, -z, x)\}$ lying inside $(+--)$, $(-+-)$, and $(--+)$, the octants of $\mathcal{O}_+$, respectively. Following similar fashion, if we have a point (u, v, w) from $(---)$, which is not the center of the octant, then $(u, v, w), (w, u, v)$ and (v, w, u) are distinct, and we have 9 more distinct points from the other 3 octants in $\mathcal{O}_-$. $\square$

The Theorem provides a shortcut to distribute points over S^2: instead of distributing points over the whole sphere, we can focus on a pair of adjacent octants, and the rest of the distribution will follow by symmetry of the orbits.

Proposition 3.2.9. *The orientation of the 8 octants given in the Theorem isn't the only* *way to orient the boundaries. Clearly, the reverse orientation would work too. These two orientations are bijective, with the bijection being defined as reverse the direction of each arrow in the picture, and there is no other way to orient the boundaries in such a way that the two orbits $\mathcal{O}_+$ and $\mathcal{O}_-$ have different orientations.*

If $(x, y, z) \in (+\,+\,+)$ is not the center of the octant, the set of 12 distinct points $\{(x, y, z), (z, x, y), (y, z, x)\} \cup \{(x, -y, -z), (z, -x, -y), (y, -z, -x)\} \cup \{(-x, y, -z), (-z, x, -y), (-y, z, -x)\} \cup \{(-x, -y, z), (-z, -x, y), (-y, -z, x)\}$ form an *equivalence class.* If (x, y, z) is the center of the octant $(+\,+\,+)$, i.e, $x = y = z = 1/\sqrt{3}$, then the 4 distinct points $\{(x, x, x), (x, -x, -x), (-x, x, -x), (-x, -x, x)\}$ is an *equivalence class.* A similar statement regarding equivalence relation for the point $(u, v, w) \in (-\,-\,-)$ can be said. We define the equivalence relation as follows:

Definition 3.2.10. Let $\mathcal{T}$ and $\mathcal{R}$ be two octants either both from $\mathcal{O}_+$ or both from $\mathcal{O}_-$. Let $(x_1, x_2, x_3) \in \mathcal{T}$ and $(u_1, u_2, u_3) \in \mathcal{R}$. $(x_1, x_2, x_3) \sim (u_1, u_2, u_3)$ if $(|x_1|, |x_2|, |x_3|) = (|u_{\sigma(1)}|, |u_{\sigma(2)}|, |u_{\sigma(3)}|)$.

Proposition 3.2.11. $\sim$ *is an equivalence relation on $\mathcal{O}$. An element in an equivalence class will be denoted by "(x, y, z) **(mod)** $\mathcal{O}_+$" or "(x, y, z) **(mod)** $\mathcal{O}_-$", depending on which orbit the octant that (x, y, z) is in belongs to.*

Remark 3.2.12. With the assist of the symmetry and the equivalence relation just defined, to distribute points on S^2, we need only consider the distribution over the two adjacent octants $(+\,+\,+)$ and $(+\,-\,+)$. Although using the equivalence relation of the octants doesn't seem to make much difference at the early stage of our sequence construction, as the number of points get large it does give some simplicity.

3.3 Memphis' Triangulation

Given a finite set of points on S^2 there exists a unique set of deep holes. However, as we have seen in constructing well-spaced sperhical sequences, when applying *the greedy algorithm of largest circumference*, it often happens that a pair of deep holes are so close that not both of them can be added to form the new sequence. The flexibility of the point choices poses great uncertainties to the algorithm and difficulties to computing the spherical discrepancies. The particular configuration of the optimal first 50 points $\mathbf{S}_{50}$ we presented demonstrates strong regularity in symmetry. We would like a definite algorithm that tells us exactly which points of the set of the deep holes to choose. The algorithm we describe in this section will agree with the greedy algorithm of the largest circumference(this is not trivial to see and will be proven in the next chapter). In particular, the first 50 points it generates agrees with $\mathbf{S}_{50}$. This particular algorithm we are about to define will be called *Memphis' Tiangulation*. Any sequence produced by this algorithm will be called a *Memphis' Sequence*.

3.3.1 The Initial Step

The recursive method starts with the 6 poles on S^2. Connecting the 6 poles, we get the very familiar triangulation of S^2: the 8 octants.

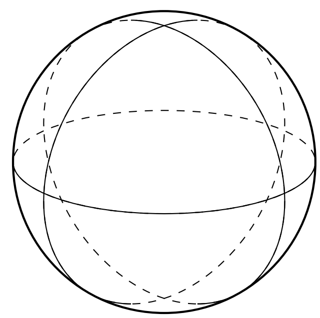


Figure 3.16: The 8 Octants Form a Triangulation of S^2

All the *edges* in this *triangulation* have the same length $\pi/2$. We have obtained the first 6 points $\mathscr{M}_6$.

Remark 3.3.1. In describing the point distribution algorithm, we need to use lots of pictures in demonstration. We will use the following " bird view" of S^2 from the " North Pole". The figure on the left is the north hemisphere. The figure on the right is the south hemisphere viewed from the north pole as if the north hemisphere were transparent. The two hemispheres share the same boundary, so the boundary circle of the south hemisphere is drawn with a dashed line.

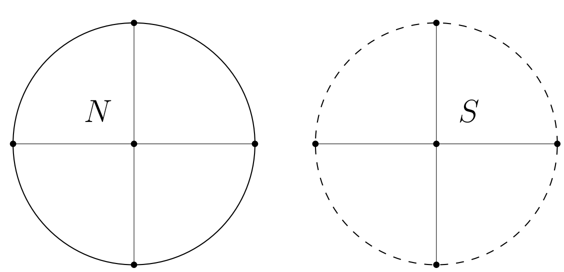


Figure 3.17: $\mathscr{M}_6$ on S^2

Notation 15. *In a triangulation of S^2, any spherical triangle has 3 adjacent triangles. We call adjacent spherical triangles* neighbors*. Two neighbor triangles form a* spherical quadrilateral*.*

3.3.2 The First Few Recursive Steps

Let $\mathscr{D}_6$ denote the set of deep holes of $\mathscr{M}_6$. Add an element from $\mathscr{D}_6$ to $\mathcal{M}_6$, say the center of the octant $(+\,+\,+)$, $\overset{\frown}{\triangle}\, ABC$, denoted by X. Connecting X with the vertices of $\overset{\frown}{\triangle}\, ABC$, and with the vertices of the neighbours of $\overset{\frown}{\triangle}\, ABC$, we have 6 more edges in the figure.

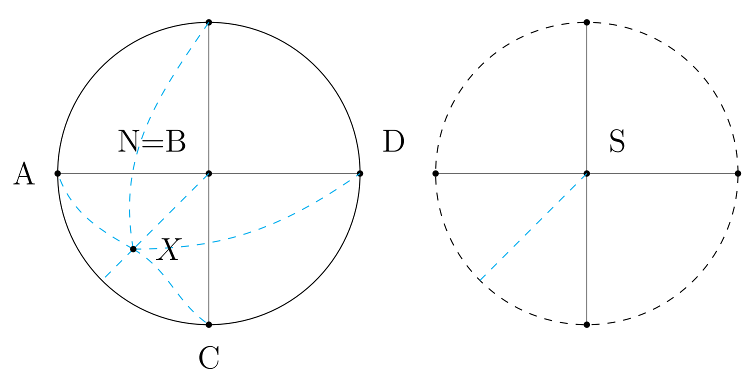


Figure 3.18: $\mathscr{M}_7$ with 3 pairs of intersecting edges, X is the circumcenter of $\overset{\frown}{\triangle}\, ABD$

However, some edges are intersecting with each other and, hence, must be removed, as there can be no intersecting edges in a triangulation. This raises the question: *which edges should be removed and which edges should stay?* The algorithm of removing and adding edges is often referred to "*the flip algorithm*".

3.3.3 Introducing *The Flip Algorithm*

We illustrate the algorithm with *spherical quadrilateral* $\overset{\frown}{\square}\ ADBX$, the two diagonals of which, $\overset{\frown}{BA}$ and $\overset{\frown}{XD}$, happen to intersect. One of the two cases must occur:

Definition 3.3.2. (The Flip Algorithm)

(I) If $\measuredangle DBX + \measuredangle XAD \geq \measuredangle BDA + \measuredangle BXA$, then remove $\overset{\frown}{DX}$. $\overset{\frown}{AB}$ stays.

(II) Otherwise, i.e $\measuredangle DBX + \measuredangle XAD < \measuredangle BDA + \measuredangle BXA$, remove $\overset{\frown}{AB}$. Edge $\overset{\frown}{XD}$ stays.

Remark 3.3.3. In the spherical quadrilateral $\overset{\frown}{\square}\ ADBX$, either $\measuredangle BDA + \measuredangle BXA \leq \measuredangle DBX + \measuredangle DAX$ or $\measuredangle BDA + \measuredangle BXA > \measuredangle DBX + \measuredangle DAX$. In the former case, according to corollary 2.2.4, we have both that point B is inside the circumcircle of $\overset{\frown}{\triangle}\ ADX$ and that point A is inside the circumcircle of $\overset{\frown}{\triangle}\ BXD$. However, in this case, it follows neither that X is inside the circumcirle of $\overset{\frown}{\triangle}\ ADB$ nor that D is inside the circumcircle of $\overset{\frown}{\triangle}\ ABX$, according to corollary 2.2.4.

Remark 3.3.4. $\measuredangle DBX + \measuredangle XAD = \measuredangle BDA + \measuredangle BXA$ precisely when the four points A, X, B, D are co-circular. In this case, removing either edge would be acceptable in the sense that none of points is inside the circumcircle of the spherical triangle consisting of the other three. However, in our algorithm, we make the " pre-existing" edge $\overset{\frown}{AB}$ stay.

From $\mathscr{M}_8$ To $\mathscr{M}_{14}$

Once this procedure is done, we can add another element of $\mathscr{D}_6$ to $\mathscr{M}_7$, and connecting the newly added point with the vertices of neighbours of the triangle it lies in, and removing crossing edges using *the flip algorithm*. Repeat the same process to obtain $\mathscr{M}_N$, until the separation of $\mathscr{M}_{N+1}$ is decreasing. And we obtain $\mathscr{M}_{14}$.

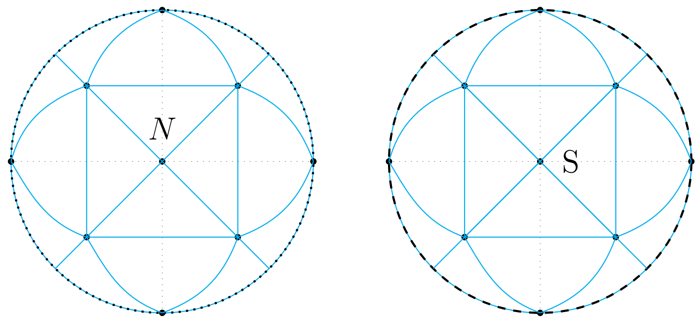


Figure 3.19: $\mathscr{M}_{14}$, the dotted lines no longer exist.

Remark 3.3.5. The points of $\mathscr{M}_{14}$ are the same as $\mathbf{S}_{14}$ obtained in section 3.1. However, we'd like to point out significance of the notation of $\mathscr{M}_N$ is that, much more important than being a point set or a sequence, it underlines a triangulation algorithm.

Remark 3.3.6. $\mathscr{M}_{14}$ is unique, but the way to get there from $\mathscr{M}_6$ isn't.

3.3.4 $\mathscr{M}_{26}$

The set of 24 deep holes of all lie on the boundaries of the octants. From our memories of $\mathbf{S}_{50}$ we know that not all these 24 points can be added. Although we will eventually choose the same next 12 points, our choice will be made in a different perception. Recall of the two ways to orient the 8 octant (refer to Figure 3.15):

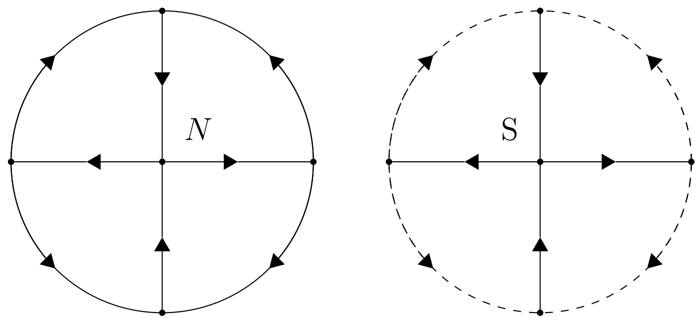


Figure 3.20: Points chosen from $\mathscr{D}_{14}$ are alone the arrows of *the orientation*.

The 24 deep holes $\mathscr{D}_{14}$ all lie on the edges, along which the arrow are marked. Two deep holes on each edge. We choose the next 12 points in the following way: on each edge, the element in $\mathscr{D}_{14}$ closer to the end point, to which the arrow points, will be chosen.

Remark 3.3.7. *Again, the 12 chosen elements of $\mathscr{D}_{14}$ are unique, but the order to add them one by one is not.*

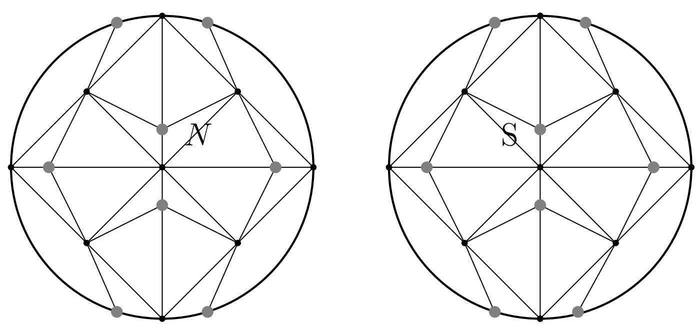


Figure 3.21: $\mathscr{M}_{26}$.

3.3.5 $\mathscr{M}_{27}$ to $\mathscr{M}_{38}$ on Octants of $\mathcal{O}_+$

$\mathscr{M}_{26}$ has 24 deep holes. We will choose the 12 blues points as marked in the figure below, in the (shaded) octants of $\mathcal{O}_+$. After adding in the 12 points one by one, we eventually get the triangulation for the 38 points, $\mathscr{M}_{38}$. To avoid the complications of all the edges, in the following figure, we are only presenting the 38 points in $\mathscr{M}_{38}$. Interested readers may try recovering the triangulations.

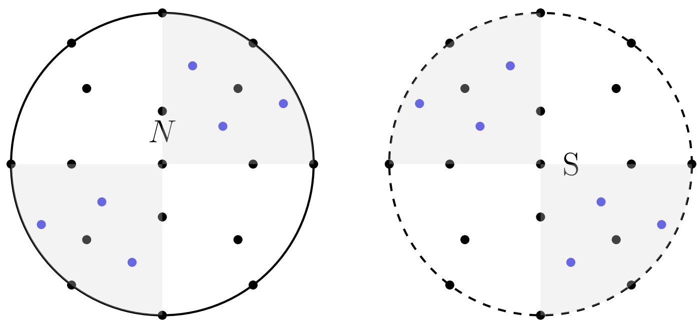


Figure 3.22: the points of $\mathscr{M}_{38}$

3.3.6 $\mathscr{M}_{39}$ to $\mathscr{M}_{50}$ on Octants of $\mathcal{O}_-$

$\mathscr{M}_{38}$ has 12 deep holes. Again, to avoid complication of all the edges, on the 50 points in $\mathscr{M}_{50}$ are presented below. The newly added deep holes are the blue points in the unshaded octants of $\mathcal{O}_-$.

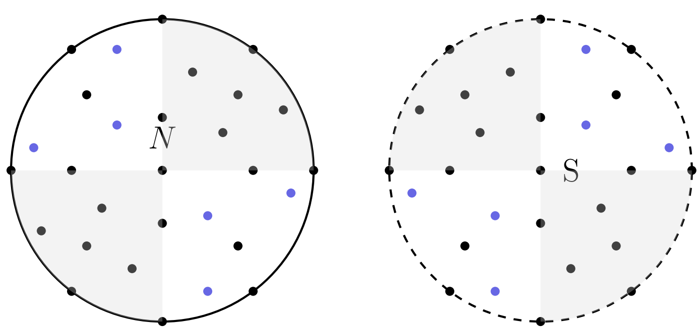


Figure 3.23: The 50 points of $\mathscr{M}_{50}$

The 50 points of $\mathscr{M}_{50}$ agree with $\mathbf{S}_{50}$.

3.3.7 From $\mathscr{M}_N$ to $\mathscr{M}_{N+1}$, and The Infinite Sequence $\mathscr{M}$

Notation 16. *(Size of A Spherical Triangle)*

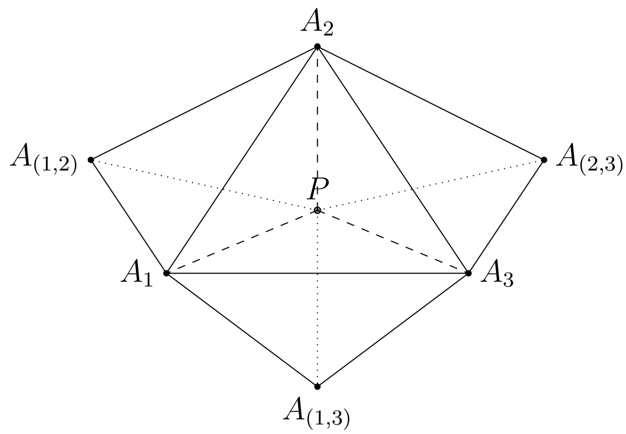
We define the size of a spherical triangle by its circumradius. For the rest of our discussion, a larger triangle means a triangle with larger circumradius. When a point is contained inside the circumcircle of a triangle, we say that the triangle covers the point or the point is covered by the triangle.

During the construction of $\mathscr{M}_{50}$, every step we know exactly where are deep holes are from the previous section. However, as N becomes large, the locations of the deep holes becomes a big unknown. For $\mathscr{M}_N$, we will choose the next point, the $(N+1)$th point to be the circumcenter of a largest triangle, and repeat the following procedure.

Definition 3.3.8. (Memphis' Triangulation)

Suppose $\overset{\frown}{\triangle}\, A_1A_2A_3$ is a largest triangle in $\mathscr{M}_N$, whose circumcenter P we choose to be the $(N+1)$th point.

(i) If $\overset{\frown}{\triangle}\, A_1A_2A_3$ is acute, P is inside the interior of the triangle.



(ii) If $\widehat{\triangle}\ A_1A_2A_3$ is right with hypotenuse $\widehat{A_1A_2}$, then P is the mid-pint of $\widehat{A_1A_2}$.

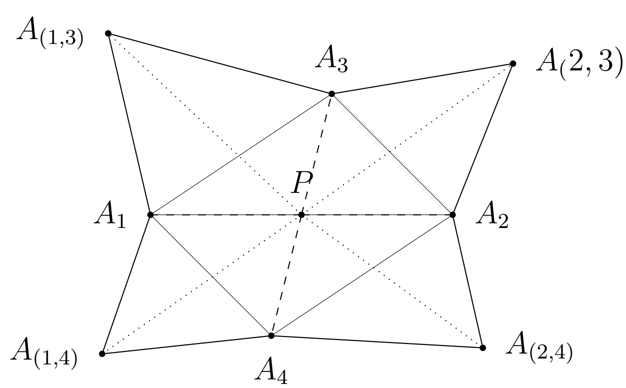


Figure 3.24: $\widehat{\triangle}\ A_1A_2A_3$ right.

(iii) If $\widehat{\triangle}\ A_1A_2A_3$ is obtuse with largest angle $\measuredangle A_3$, then P and A_3 lie on different sides of edge $\widehat{A_1A_2}$.

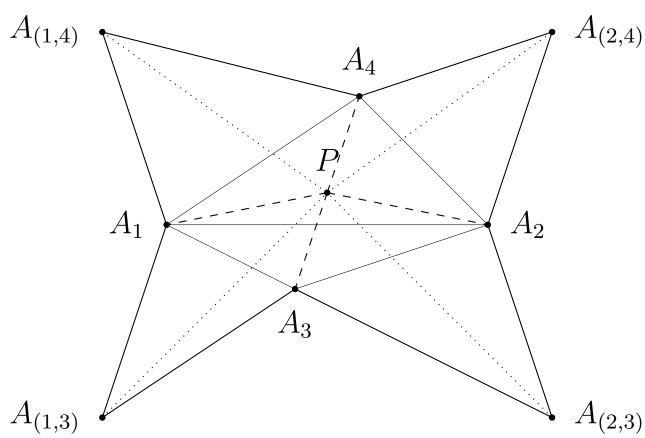
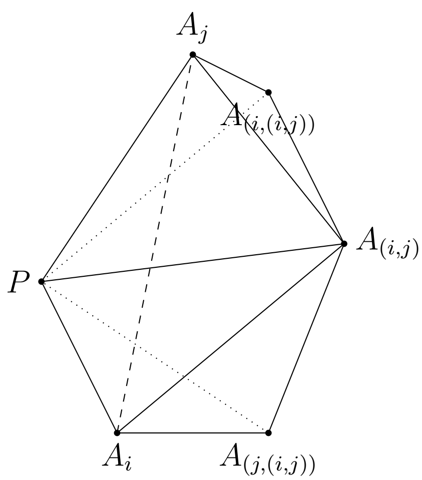


Figure 3.25: P is the circumcenter of the obtuse $\overset{\frown}{\triangle}\ A_1A_2A_3$

For each edge $\overset{\frown}{A_iA_j}$, $i, j = 1, 2, 3, 4$ and $i \neq j$, $\overset{\frown}{\triangle}\ PA_iA_j$ has a neighbor $\overset{\frown}{\triangle}\ A_iA_jA_{(i,j)}$.

Remark 3.3.9. The subindex (i, j) is not an ordered pair. In particular, $A_{(i,j)} = A_{(j,i)}$.

For each pair of neighbors $\overset{\frown}{\triangle}\ PA_iA_j$ and $\overset{\frown}{\triangle}\ A_iA_jA_{(i,j)}$, perform the flip algorithm (Definition 3.3.2). After performing the flip algorithm, if a new triangle (in fact a pair of new triangles) with P as a vertex is (are) generated,



If P is covered inside $\overset{\frown}{\triangle}\ PA_iA_j$, after performing the flip algorithm, edge $\overset{\frown}{PA_{(i,j)}}$ is added and $\overset{\frown}{A_iA_j}$ is removed. Two pairs of neighbors are created: $\overset{\frown}{\triangle}\ PA_iA_{(i,j)}$ is adjacent to $\overset{\frown}{\triangle}\ A_iA_{(i,j)}A_{i,(i,j)}$ and $\overset{\frown}{\triangle}\ PA_jA_{(i,j)}$ is adjacent to $\overset{\frown}{\triangle}\ A_jA_{(i,j)}A_{j,(i,j)}$. Repeat the flip algorithm to the new neighbors.

then repeat the flip algorithm to the new triangle and its neighbor. Every time a new triangle $\widehat{\triangle}\ PA_{.}A_{..}$ is generated, we will perform the flip algorithm to $\widehat{\triangle}\ PA_{.}A_{..}$ and its neighbor sharing the common edge $\widehat{A_{.}A_{..}}$ until no edge needs to be flipped. Then we obtain a triangulation of the $N+1$, denoted by $\mathscr{M}_{N+1}$. This recursive triangulation algorithm is called "Memphis' Triangulation".

Remark 3.3.10. *As we will prove in the next chapter, the deep holes can always be chosen to be the circumcenter of an non-obtuse triangle, i.e either acute or right. However, we will still describe Memphis' Triangulation for the case of obtuse triangles for the purpose of completeness.*

Definition 3.3.11. As N goes to ∞, we obtain an infinite extension of $\mathscr{M}_{50}$, denoted by $\mathscr{M}$. A sequence constructed using *Memphis' Triangulation* will be called *a Memphis' Sequence.*

3.3.8 Regularities of $\mathscr{M}_N$

Let $\mathscr{M}_N$ be the finite truncation of $\mathscr{M}$ is length N. $\mathscr{M}_N$ has very strong regularities. These regularities will be proven in the next Chapter. We will state them below.

(I: Maximal Separation) *None of the point in $\mathscr{M}_N$ is contained inside the circumcircle of any triangle in this triangulation.*

(II: Non-obtuse Triangles) *Every deep hole of a Memphis' Sequence can be chosen to be the circumcenter of an non-obtuse triangle.*

There are two types of edges in the algorithm of *Memphis' Triangulation*: the first type are those created by connecting *the circumcenter* of the chosen triangle with its vertices; the second type are those created by connecting *the circumcenter* of the chosen triangle with the vertex of its neighbour.

Notation 17. *We will call edges of the first type* ***radius edges****, and the second type* ***cross edges****.*

(III: Shortest Edges) Suppose P is the circumcenter of a largest triangle of $\mathscr{M}_N$. The distance between P and $\mathscr{M}_N$ as a point set is shortest distance between any pair of points in $\mathscr{M}_N \cup \{P\}$. In particular, the last added in radius edges are the shortest edges in the triangulation.

Chapter 4

Delaunay-Memphis' Triangulation

The main result we present in this chapter is that the triangulation constructed using *Memphis' Algorithm* is a *Delaunay Triangulation*.

4.1 Introduce Delaunay Triangulation

There are various versions of the definition of Delaunay triangulations. Many of them use lengthy and complicated-looking notations. Some authors define Delaunay triangulations are *the duals of Voronoi diagrams*. To assist our discussion of Memphis' Algorithm and spherical cap discrepancies, we will define a Delaunay triangulation as follows.

Definition 4.1.1. (Delaunay Triangulation) In the plane, a triangulation among N points is called a Delaunay Triangulation if none of the points is contained inside the circumcircle of any triangle. Analogously a spherical Delaunay triangulation among

N points on S^2 is a triangulation on the unit sphere such that none of these points is contained inside the circumcircle of all the spherical triangles. When a circumcircle doesn't contain any point in its interior, we say *the circumcircle is empty.*

Remark 4.1.2. In our definition of a Delaunay triangulation, no point is contained inside the circumcircle of any triangle. By "inside" we strictly mean the interior of the circumdisk.

Notation 18. *In a Delaunay triangulation, more than 3 points may be co-circular. Some sources consider such co-circular configurations a "degeneracy", as they make the Delaunay triangulations non-unique.*

Notation 19. *In the upcoming discussion, we will often use the word "Delaunay" as an adjective for simplicity. For example, when two triangles share a common edge, if neither of them contains the other inside the circumcircle, we say these two triangles are Delaunay. When a triangle doesn't contain a point P inside its circumcircle, we say the triangle doesn't cover P.*

Theorem 4.1.3. [12] (Existence and Uniqueness of Delaunay Triangulation in $\mathbb{R}^2$) *Given N points in the plane, if no four points are cocircular with empty circumcircle, there exists a unique Delaunay Triangulation.*

Remark 4.1.4. In the Theorem above, we do need the non-degeneracy mentioned in Notation 18 in order to get both existence and uniqueness.

4.2 $\mathscr{M}_N$ is a Delaunay Triangulation

In this section we will be proving there exists a Delaunay triangulation among the N points of $\mathscr{M}_N$. In particular, $\mathscr{M}_N$ is a Delaunay triangulation. But, before going through all the efforts of proving this statement, the first question is why we care about whether $\mathscr{M}_N$ is Delauany at all. In Chapter 3, in order to generate a well-spaced sequence, every new point, *a deep hole*, is in maximal distance to the previous ones. We named this method *the Method of Largest Circumference*. To search for the next *deep holes*, there are $\binom{N}{3}$ triangles among the N points to compare. The complexity of this algorithm grows fast as N gets big. Then we introduced *Memphis' Algorithm*, which recursively generates new points by triangulation. One obvious advantage of *Memphis' Algorithm* is that, instead of $\binom{N}{3}$ eligible candidates, there are only $2N - 4$ competing triangles. However, at each step, the new point selected by *Memphis'* is merely the the circumcenter of *a largest triangle* (i.e *a triangle existing in* $\mathscr{M}_N$ *with the largest circumradius*). What if this new point is not *a deep hole*? In that case, these two algorithms would disagree. Fortunately, such disagreement never happens, which is guaranteed by *Delaunayness* of the triangulation: as long as $\mathscr{M}_N$ is Delaunay, the circumcenter of the largest triangle is a deep hole. Further, the deep holes must be the circumcenter of a non-obtuse triangles in $\mathscr{M}_N$.

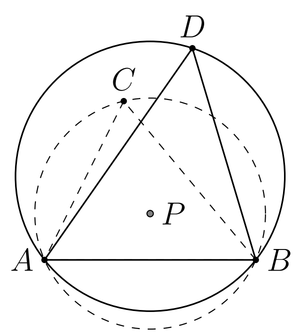
Theorem 4.2.1. (Memphis' Non-obtuse Triangles) *If* $\mathscr{M}_N$ *is Delauany, a deep hole of the* N *points can always be chosen to be the circumcenter of some non-obtuse triangle* *in this triangulation.*

Proof. (Theorem 4.2.1) Before proving the non-obtuse shape of the triangle in which the deep holes lies, we must first prove that the deep holes are circumcenters of triangles in $\mathscr{M}_N$.

Suppose the circumcenter P of $\overset{\frown}{\triangle}\, ABC$ is *a deep hole* of the N points. No triangle in $\mathscr{M}_N$ can have larger circumradius than $\overset{\frown}{\triangle}\, ABC$.

Remark 4.2.2. *A, B, C are merely points $\mathscr{M}_N$. $\overset{\frown}{\triangle}\, ABC$ is not assumed to be an existing triangle in $\mathscr{M}_N$.*

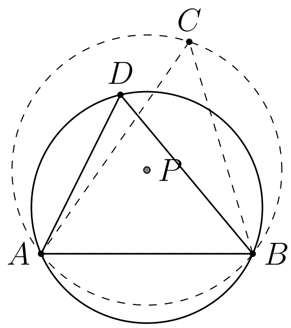
If A and B are connected in $\mathscr{M}_N$, then there exists a point D on the same side of $\overset{\frown}{AB}$ as C. C cannot be inside the circumcircle of $\overset{\frown}{\triangle}\, ABD$ because $\mathscr{M}_N$ is Delaunay.



The dashed lines may not exist $\mathscr{M}_N$. $\overset{\frown}{\triangle}\, ABD$ exists in the triangulation, which is by assumption Delaunay. So, C cannot be inside the circumcircle of $\overset{\frown}{\triangle}\, ABD$.

Figure 4.26

D cannot be inside the circumcircle of $\overset{\frown}{\triangle}\, ABC$ because P is a deep hole.

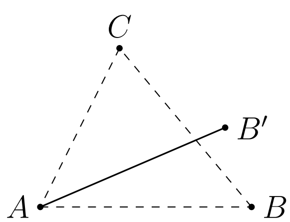


The dashed lines may not exist $\mathscr{M}_N$. However, $|\ \widehat{CP}\ |$ would be shorter than $|\ \widehat{PD}\ |$. This is a contradiction to the definition of a deep hole.

Figure 4.27

Hence, A, B, C, D must be co-circular. When any two points of A, B, C are connected in $\mathscr{M}_N$, the deep hole P is also the circumcenter of $\widehat{\triangle}\ ABD$ which exists in $\mathscr{M}_N$.

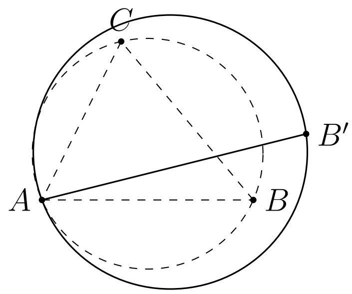
If no two points of A, B, C are connected in $\mathscr{M}_N$. Let $\widehat{\triangle}\ AB'C'$ be an existing triangle in $\mathscr{M}_N$ such that B and C lie on opposite sides of $\widehat{AB'}$.



The great circle passing A and B' cuts through edge $\widehat{BC}$. B' is not inside the circumcircle of $\widehat{\triangle}\ ABC$.

Figure 4.28

Since the circumcenter of $\widehat{\triangle}\ ABC$ is a deep hole, B' is not inside the circumcircle of $\widehat{\triangle}\ ABC$. Therefore, the circumcircle of $\widehat{\triangle}\ ABC$ and the circumcircle of $\widehat{\triangle}\ AB'C'$ must have at least one intersection, A. B' and A divide the circumcircle of $\widehat{\triangle}\ AB'C'$ into two arcs. *Not both of these arcs can be in the interior of the circumcircle of $\widehat{\triangle}\ ABC$.*



The heavier arc of the circumcircle of $\overset{\frown}{\triangle}\ AB'C'$ will contain either B or C inside, or maybe both B and C.

Figure 4.29

If one of the arcs of the circumcircle of $\overset{\frown}{\triangle}\ AB'C'$ between A and B' is outside the circumcircle of $\overset{\frown}{\triangle}\ ABC$, then that arc contains either B or C inside, contradicting to $\mathscr{M}_N$ being Delaunay. So the only possibility left is that both arcs between A and B' lie precisely on the circumcircle of $\overset{\frown}{\triangle}\ ABC$. That is A, B, C, B', C' are co-circular. The deep hole P is also the circumcenter of $\overset{\frown}{\triangle}\ AB'C'$, an existing triangle of $\mathscr{M}_N$. We have proven that the newly selected point by Memphis Algorithm is indeed a deep hole.

Now we are going to show the deep hole is the circumcenter of a non-obtuse triangle, *under the assumption $\mathscr{M}_N$ is Delaunay.*

Assume one of its *deep holes* is the circumcenter of some obtuse triangle $\overset{\frown}{\triangle}\ A\tilde{C}B$ with $\measuredangle\tilde{C} > \measuredangle A + \measuredangle B$. Then $\overset{\frown}{AB}$ is the longest edge of $\overset{\frown}{\triangle}\ A\tilde{C}B$. Let $\overset{\frown}{\triangle}\ ABC$ be its the neighbour sharing $\overset{\frown}{AB}$.

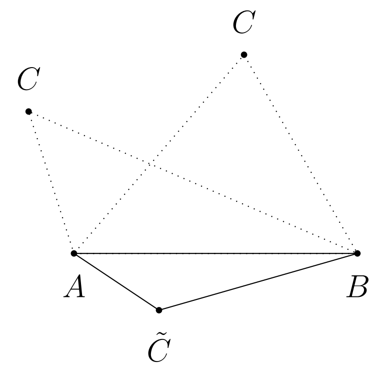


Figure 4.30: $\overset{\frown}{\triangle}\ A\tilde{C}B$ with neighbour $\overset{\frown}{\triangle}\ ABC$.

By the definition of *deep holes*, the circumcircle of $\overset{\frown}{\triangle}\ A\tilde{C}B$ doesn't cover any point of $\mathscr{M}_N$ in its interior. Hence, in the quadrilateral $\overset{\frown}{\square}\ A\tilde{C}BC$, we must have

$$\measuredangle CA\tilde{C} + \measuredangle CB\tilde{C} \geq \measuredangle ACB + \measuredangle A\tilde{C}B. \tag{4.1}$$

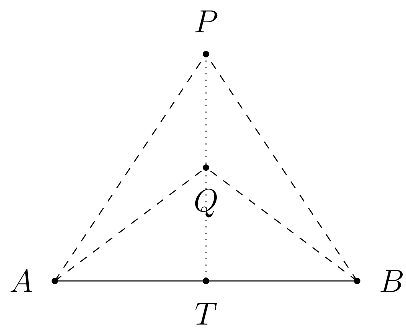
Let P be the circumcenter of $\overset{\frown}{\triangle}\ ABC$. Since $\overset{\frown}{\triangle}\ A\tilde{C}B$ is obtuse with largest angle $\measuredangle\tilde{C}$, P and C are on the same side of $\overset{\frown}{AB}$. Let Q be the circumcenter of $\overset{\frown}{\triangle}\ A\tilde{C}B$. P and Q must be on the same side of $\overset{\frown}{AB}$. Otherwise, P and $\tilde{C}$ would be in the same side, which would imply $\overset{\frown}{\triangle}\ ABC$ is obtuse with largest angle $\measuredangle C$ and C would be contained in the interior of the circumcircle of $\overset{\frown}{\triangle}\ A\tilde{C}B$, a contradiction to equation (4.1).

Remark 4.2.3. *We don't know the shape of $\overset{\frown}{\triangle}\ ABC$ yet, whether it is acute, right, or obtuse. However, by equation (4.1), if $\overset{\frown}{\triangle}\ ABC$ happens to be obtuse or right, then $\measuredangle ACB$ cannot be its largest angle.*

Connect P, Q with A and B. The locus of points (on S^2) whose distances to A and B are equal is the great circle passing the mid-point T of $\overset{\frown}{AB}$ and perpendicular to $\overset{\frown}{AB}$.

Since $|AP| = |BP|$ and $|AQ| = |BQ|$, P and Q must lie on this great circle. P, Q and T are on the same great circle.

Case 1: If P is outside $\overset{\frown}{\triangle}\ ABQ$.



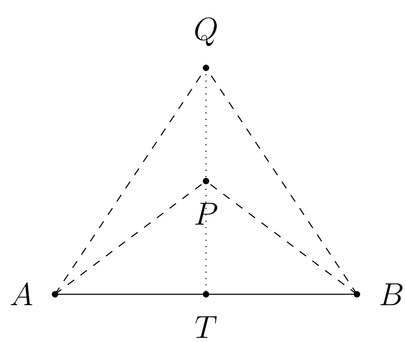
Applying the Cosine Law to the *right-angle triangles* $\overset{\frown}{\triangle}\ ATP$ and $\overset{\frown}{\triangle}\ ATQ$, we have

$$0 = \sin \overset{\frown}{AT} \sin \overset{\frown}{PT} \cos \pi/2 = \cos \overset{\frown}{AP} - \cos \overset{\frown}{AT} \cos \overset{\frown}{PT}$$

$$0 = \sin \overset{\frown}{AT} \sin \overset{\frown}{QT} \cos \pi/2 = \cos \overset{\frown}{AQ} - \cos \overset{\frown}{AT} \cos \overset{\frown}{QT}$$

Since Q is inside $\overset{\frown}{\triangle}\ ABP$, $\overset{\frown}{PT}$ is an extension of $\overset{\frown}{QT}$. In other words, $\overset{\frown}{PT}>\overset{\frown}{QT}$. Hence, $\cos \overset{\frown}{AP} - \cos \overset{\frown}{AQ} = \cos \overset{\frown}{AT} \cos \overset{\frown}{PT} - \cos \overset{\frown}{AT} \cos \overset{\frown}{QT} < 0$. $\mid \overset{\frown}{AP} \mid > \mid \overset{\frown}{AQ} \mid$. This is a contradiction to the assumption that the circumcenter of $\overset{\frown}{\triangle}\ A\tilde{C}B\ Q$ is a *deep hole.*

Case 2: P is inside $\overset{\frown}{\triangle}\ ABQ$. In this case we do get $\overset{\frown}{AQ}$ is the longer edge than $\overset{\frown}{AP}$.



There are 3 possibilities in this case: $\widehat{\triangle}\, ABC$ is acute, right, or obtuse. In each of these possibilities, we will obtain contradiction to the known relation 4.1.

Since P is inside $\widehat{\triangle}\, ABQ$,

$$\measuredangle QAP > 0 \tag{4.2}$$

and , hence,

$$\measuredangle QA\tilde{C} = \measuredangle PA\tilde{C} + \measuredangle QAP > \measuredangle PA\tilde{C} \tag{4.3}$$

Similarly, we have

$$\measuredangle QBP > 0 \tag{4.4}$$

and , hence,

$$\measuredangle QB\tilde{C} = \measuredangle PB\tilde{C} + \measuredangle QBP > \measuredangle PB\tilde{C}. \tag{4.5}$$

Subcase 1: If $\widehat{\triangle}\, ABC$ is acute, then P is inside the triangle $\widehat{\triangle}\, ABC$.

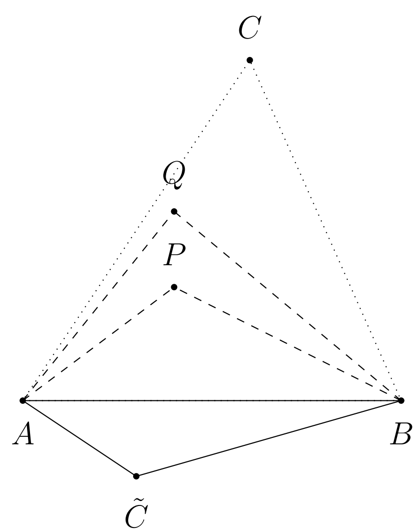


Figure 4.31: The position of C to Q is unknown

By inequalities 4.3, $\measuredangle PA\tilde{C} < \measuredangle QA\tilde{C}$. Therefore,

$$\measuredangle CA\tilde{C} = \measuredangle CAP + \measuredangle PA\tilde{C} < \measuredangle CAP + \measuredangle QA\tilde{C}. \tag{4.6}$$

Similarly, by inequality 4.5,

$$\measuredangle CB\tilde{C} = \measuredangle CBP + \measuredangle PB\tilde{C} < \measuredangle CBP + \measuredangle QB\tilde{C}. \tag{4.7}$$

On the other hand, the isosceles triangle $\overset{\frown}{\triangle}\, CPB$ and $\overset{\frown}{\triangle}\, CAP$ give

$$\measuredangle ACB = \measuredangle CAP + \measuredangle CBP, \tag{4.8}$$

and the isosceles triangle $\overset{\frown}{\triangle}\, QA\tilde{C}$ and $\overset{\frown}{\triangle}\, Q\tilde{C}B$ give

$$\measuredangle A\tilde{C}B = \measuredangle QA\tilde{C} + \measuredangle QB\tilde{C}. \tag{4.9}$$

So, combining all these inequality relations we have:

$$\measuredangle ACB + \measuredangle A\tilde{C}B = \measuredangle CAP + \measuredangle CBP + \measuredangle A\tilde{C}B, \text{ by equality (4.8)} \tag{4.10}$$

$$= \measuredangle CAP + \measuredangle CBP + \measuredangle QA\tilde{C} + \measuredangle QB\tilde{C}, \text{ by equality (4.9)} \tag{4.11}$$

$$> \measuredangle CA\tilde{C} + \measuredangle CBP + \measuredangle QB\tilde{C}, \quad \text{by inequality (4.6)} \tag{4.12}$$

$$> \measuredangle CA\tilde{C} + \measuredangle CB\tilde{C}, \quad \text{by inequality (4.7).} \tag{4.13}$$

Clearly, (4.13) is the opposite of what inequality (4.1) states. This is a contradiction.

Subcase 2: $\overset{\frown}{\triangle}\, ABC$ is a right triangle. (4.1) determines that $\measuredangle C$ cannot be the largest angle. So, WLOG, we assume

$$\measuredangle CAB + \measuredangle ACB = \measuredangle ABC. \tag{4.14}$$

Then the circumcenter P is the mid-point of the edge $\overset{\frown}{AC}$.

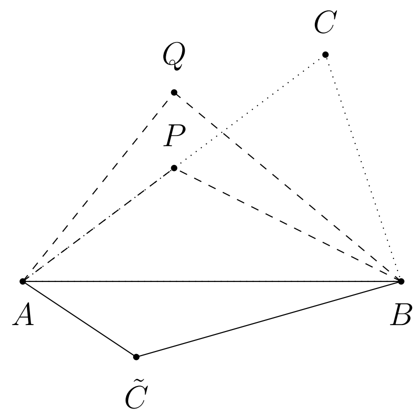


Figure 4.32: $\measuredangle PAC = 0$ and $\measuredangle PA\tilde{C} = \measuredangle CA\tilde{C}$

In the isosceles triangles $\overset{\frown}{\triangle}\, BCP$,

$$\measuredangle ACB = \measuredangle PBC. \tag{4.15}$$

This gives us:

$$\measuredangle ACB + \measuredangle A\tilde{C}B = \measuredangle PBC + \measuredangle A\tilde{C}B \tag{4.16}$$

$$= \measuredangle PBC + \measuredangle QA\tilde{C} + \measuredangle QB\tilde{C} \tag{4.17}$$

$$> \measuredangle PBC + \measuredangle PA\tilde{C} + \measuredangle PB\tilde{C}, \text{ by inequalities (4.3) and (4.5)} \tag{4.18}$$

$$= \measuredangle CB\tilde{C} + \measuredangle CA\tilde{C}. \tag{4.19}$$

This strict inequality contradicts the inequality (4.1).

Subcase 3: $\overset{\frown}{\triangle}\, ABC$ is obtuse. And WLOG, we assume $\measuredangle ABC > \measuredangle ACB + \measuredangle CAB$.

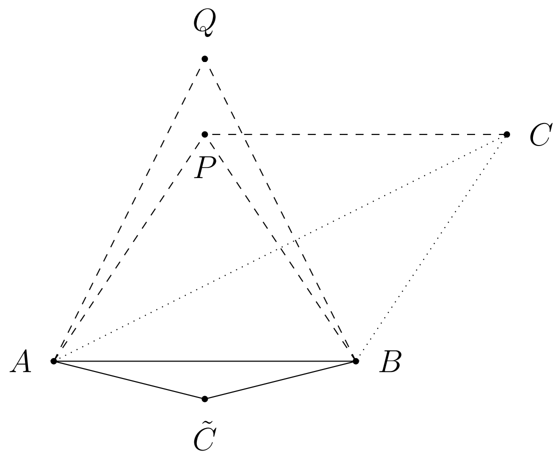


Figure 4.33: P and B are on the opposite sides of edge $\overset{\frown}{AC}$.

$$\measuredangle ACB = \measuredangle PBC - \measuredangle PAC. \tag{4.20}$$

Hence,

$$\measuredangle ACB + \measuredangle A\tilde{C}B = \measuredangle PBC - \measuredangle PAC + \measuredangle QA\tilde{C} + \measuredangle QB\tilde{C}. \tag{4.21}$$

However,

$$\measuredangle QA\tilde{C} - \measuredangle PAC > \measuredangle PA\tilde{C} - \measuredangle PAC, \text{ by (4.3)} \tag{4.22}$$

$$= \measuredangle CA\tilde{C}. \tag{4.23}$$

Applying this to (4.21), we get

$$\measuredangle ACB + \measuredangle A\tilde{C}B > \measuredangle PBC + \measuredangle QB\tilde{C} + \measuredangle CA\tilde{C} \tag{4.24}$$

$$> \measuredangle PBC + \measuredangle PB\tilde{C} + \measuredangle CA\tilde{C}, \text{ by inequality (4.5)} \tag{4.25}$$

$$> \measuredangle CB\tilde{C} + \measuredangle CA\tilde{C}, \tag{4.26}$$

which is a contradiction to (4.1). Now we have completed checking *all the three subcase* in **case 2**.

So far we have ruled out the case that P is outside $\overset{\frown}{\triangle}\ ABQ$ and the case that P is inside $\overset{\frown}{\triangle}\ ABQ$. The only left possibility is that P and Q coincide.

Case 3: If $P = Q$, $A, B, C, \tilde{C}$ are co-circular. If $\overset{\frown}{\triangle}\ ABC$ happens to be non-obtuse, then we are done. $\overset{\frown}{\triangle}\ ABC$ is the the non-obtuse triangle whose circumcenter is *a deep hole* we choose. If $\overset{\frown}{\triangle}\ ABC$ is obtuse, then, again by equation 4.1 $\measuredangle C$ cannot be its largest angle. WLOG, we assume $\measuredangle ABC$ is its largest angle. Then $\overset{\frown}{AC}$ will be its longest edge. The circumcenter P and B lie on different sides of $\overset{\frown}{AC}$.

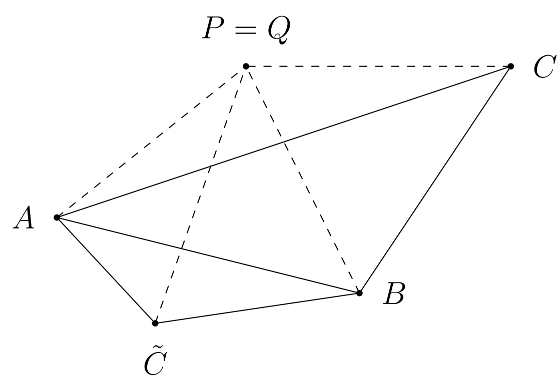


Figure 4.34: *All dashed lines are of equal length, since $A, B, C, \tilde{C}$ are co-circular.*

Then $\mid \overset{\frown}{AC} \mid > \mid \overset{\frown}{AB} \mid$. So $\overset{\frown}{\triangle}\ ABC$ is another obtuse triangle and its circimcenter is *a deep hole*. Then the same analysis we have been doing on the obtuse triangle $\overset{\frown}{\triangle}\ A\tilde{C}B$ can be applied to the new obtuse triangle $\overset{\frown}{\triangle}\ ABC$ and its neighbour sharing its edge $\overset{\frown}{AC}$. Each time we apply the same argument, the process *either* terminates when *the deep hole* happens to be the circumcenter of an non-obtuse triangle *or* we end up with another obtuse triangle. This process cannot repeat forever, as we only have finitely many triangles.

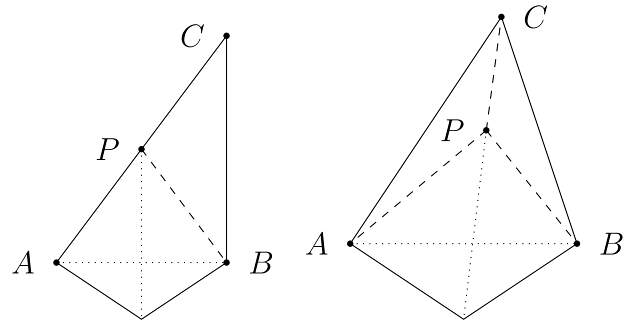


Figure 4.35: *The chosen deep hole P is the circumcenter of $\overset{\frown}{\triangle}\ ABC$*

Since $\mathscr{M}_N$ has only finitely many *deep holes*, there are finitely many obtuse triangles to start the search with, and each time the chosen *deep hole* will end up being the circumcenter of some non-obtuse triangle. The proof is completed. □

Remark 4.2.4. In the statement of Theorem 4.2.1, the description is that the deep hole *"...can be chosen as ..."*, not *"...can only be chosen as..."*. According to the theorem, when a deep hole happens to be the circumcenter of an obtuse triangle, this obtuse triangle must be co-circular with some non-obtuse triangle.

By "*a non-obtuse triangle*" we mean either an acute triangle or a right triangle. When a deep hole of $\mathscr{M}_N$ happens to be the circumcenter of a right triangle, Theorem 4.2.1 gives us a neat corollary.

Corollary 4.2.5. *If P happens to be the circumcenter of the right triangle $\overset{\frown}{\triangle}\ P_iP_jP_k$, say P is the mid-point of the hypotenuse $\overset{\frown}{P_iP_j}$, then the adjacent triangle $\overset{\frown}{\triangle}\ P_iP_jP_l$ must be a right triangle too, with hypotenuse $\overset{\frown}{P_iP_j}$. The four points P_i, P_j, P_k, P_l are co-circular with circumcenter P.*

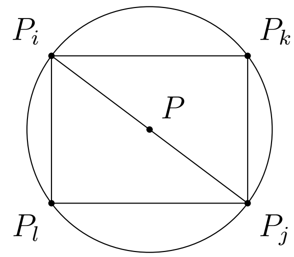
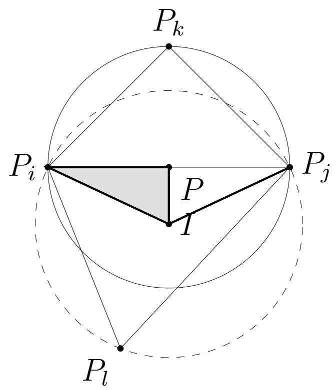


Figure 4.36

Remark 4.2.6. *The proof of Corollary 4.2.5 uses very similar argument in the proof of Theorem 4.2.1. The proof provided below is divided into three similar cases: when* $\overset{\frown}{\triangle}\ P_iP_jP_l$ *is acute, obtuse, or right.*

Proof. (Corollary 4.2.5) Since the *deep hole* P is the circumcenter of the right triangle $\overset{\frown}{\triangle}\ P_iP_jP_k$ with hypotenuse $\overset{\frown}{P_iP_j}$, we know that $\overset{\frown}{\triangle}\ P_iP_jP_k$ is one of the largest triangles (*recall that the size of a triangle is measured by its circumradius*). We will prove the assertion by contradiction, assuming the triangle $\overset{\frown}{\triangle}\ P_iP_jP_l$ is not a triangle with hypotenuse $\overset{\frown}{P_iP_j}$.

Case 1: $\overset{\frown}{\triangle}\ P_iP_jP_l$ is acute.



I is the circumcenter of the acute triangle $\widehat{\triangle}\ P_iP_jP_l$. In the grey triangle $\measuredangle IPP_i = \pi/2$. $\widehat{P_iI}$ is longer than $\widehat{PP_i}$.

Figure 4.37: Case 1: $\widehat{\triangle}\ P_iP_jP_l$ *is acute*

Then following the same argument as in the proof of statement (I), we get the acute triangle as bigger circumradius than its neighbour right triangle. This is a contradiction as we know $\widehat{\triangle}\ P_iP_kP_j$ is the largest.

Case 2: $\widehat{\triangle}\ P_iP_jP_l$ is obtuse.

First, the longest edge of the obtuse triangle $\widehat{\triangle}\ P_iP_jP_l$ cannot be $\widehat{P_iP_j}$. Otherwise, $\measuredangle P_iP_lP_j$ would be the largest angle in $\widehat{\triangle}\ P_iP_jP_l$ with relation:

$$\measuredangle P_iP_lP_j > \measuredangle P_lP_jP_i + \measuredangle P_lP_iP_j. \tag{4.27}$$

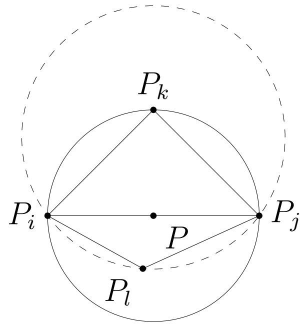
In the right triangle $\widehat{\triangle}\ P_iP_jP_k$, we have

$$\measuredangle P_iP_kP_j = \measuredangle P_kP_iP_j + \measuredangle P_kP_jP_i. \tag{4.28}$$

Consequently,

$$\measuredangle P_iP_lP_j + \measuredangle P_iP_kP_j > \measuredangle P_lP_jP_i + \measuredangle P_lP_iP_j + \measuredangle P_kP_iP_j + \measuredangle P_kP_jP_i, \tag{4.29}$$

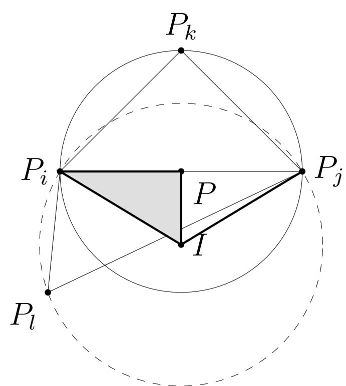
which indicates that P_l is inside the circumcircle of $\widehat{\triangle}\ P_iP_jP_k$, a contradiction.



As a result of $\overset{\frown}{P_iP_j}$ being the longest edge in $\overset{\frown}{\triangle}$ $P_iP_jP_l$ we would reach the contradiction that P_l is contained inside the circumcircle of $\overset{\frown}{\triangle}$ $P_iP_jP_k$.

Figure 4.38: Case 2: *$\overset{\frown}{P_iP_j}$ cannot be the longest edge of $\overset{\frown}{\triangle}$ $P_iP_jP_l$*

However, if $\overset{\frown}{P_iP_j}$ is not the longest edge, WLOG assume $\overset{\frown}{P_lP_j}$ is the longest edge. We run into a similar situation as in case 1. The circumcenter of $\overset{\frown}{\triangle}$ $P_iP_lP_j$ will be on the same side as P_l of the great circle passing P_i, P_j. Denote this circumcircle by I



I is the circumcenter of the obtuse triangle $\overset{\frown}{\triangle}$ $P_iP_jP_l$. In the grey triangle $\measuredangle IPP_i = \pi/2$. $\overset{\frown}{P_iI}$ is longer than $\overset{\frown}{PP_i}$.

Figure 4.39: Case 2: *$\overset{\frown}{\triangle}$ $P_iP_jP_l$ is obtuse with longest edge $\overset{\frown}{P_lP_j}$*

As explained in the figure above, the circumradius of the $\overset{\frown}{\triangle}$ $P_iP_jP_l$ will turn out to be longer than the circumradius of $\overset{\frown}{\triangle}$ $P_iP_jP_k$, a contradiction.

Case 3: $\overset{\frown}{\triangle}$ $P_iP_jP_l$ is right but its hypotenuse is not $\overset{\frown}{P_iP_j}$. Assume WLOG that $\overset{\frown}{P_iP_l}$ is the hypotenuse of the right triangle $\overset{\frown}{\triangle}\ P_iP_jP_l$. Then

$$|\ \overset{\frown}{P_iP_l}\ | > |\ \overset{\frown}{P_iP_k}\ |. \tag{4.30}$$

We know that the circumradius of $\overset{\frown}{\triangle}\ P_iP_jP_l$ is $\frac{1}{2}|\ \overset{\frown}{P_iP_l}\ |$ and the circumradius of $\overset{\frown}{\triangle}\ P_iP_jP_k$ is $\frac{1}{2}|\ \overset{\frown}{P_iP_j}\ |$. This leads to the conclusion that $\overset{\frown}{\triangle}\ P_iP_jP_l$ is bigger than $\overset{\frown}{\triangle}\ P_iP_jP_k$, a contradiction. □

We need one more lemma before proving $\mathscr{M}_N$ is Delaunay.

Lemma 4.2.7. *$\overset{\frown}{\triangle}\ ABC$, $\overset{\frown}{\triangle}\ ABB'$ and $\overset{\frown}{\triangle}\ ACC'$ are adjacent as indicated in the following figure.*

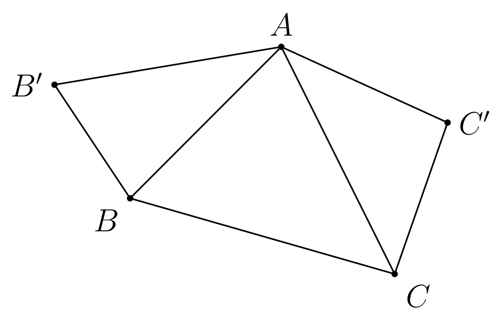


Figure 4.40

If $\overset{\frown}{\triangle}\ ABC$ and $\overset{\frown}{\triangle}\ ABB'$ are Delaunay, and $\overset{\frown}{\triangle}\ ABC$ and $\overset{\frown}{\triangle}\ ACC'$ are Delaunay, then C' is not contained inside the circumdisc of $\overset{\frown}{\triangle}\ ABB'$ and B' is not contained inside the circumdisc of $\overset{\frown}{\triangle}\ ACC'$.

Proof. To show that C' is not contained inside the circumdisc of $\overset{\frown}{\triangle}\ ABB'$, draw an auxiliary line (i.e great arc segment) between C' and B, it suffices to show that $\overset{\frown}{\triangle}\, ABB'$ and $\overset{\frown}{\triangle}\, ABC'$ are Delaunay.

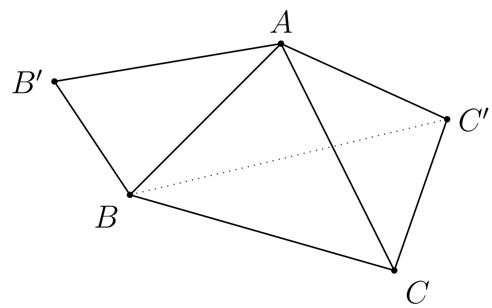


Figure 4.41

Since $\overset{\frown}{\triangle}\, ABC$ and $\overset{\frown}{\triangle}\, ABB'$ are Delaunay,

$$\measuredangle B' + \measuredangle ACB \leq \measuredangle B'BC + \measuredangle B'AC, \tag{4.31}$$

which is equivalent to

$$\measuredangle B' \leq \measuredangle B'BC + \measuredangle B'AC - \measuredangle ACB. \tag{4.32}$$

Since $\overset{\frown}{\triangle}\, ABC$ and $\overset{\frown}{\triangle}\, ACC'$ are Delaunay, C' is not inside the circumdisc of $\overset{\frown}{\triangle}\, ABC$. So,

$$\measuredangle BAC + \measuredangle ABC - \measuredangle ACB \leq \measuredangle C'AB + \measuredangle C'BA - \measuredangle AC'B, \tag{4.33}$$

which is equivalent to

$$\measuredangle AC'B \leq \measuredangle BAC' + \measuredangle ABC' - \measuredangle BAC - \measuredangle ABC + \measuredangle ACB. \tag{4.34}$$

Adding inequality 4.32 and inequality 4.34, we get

$$\measuredangle B' + \measuredangle AC'B \leq \measuredangle BAC' + \measuredangle ABC' - \measuredangle BAC - \measuredangle ABC + \measuredangle B'BC + \measuredangle B'AC. \tag{4.35}$$

Regrouping the right-hand side inequality 4.35, we have

$$\measuredangle BAC' + \measuredangle B'AC - \measuredangle BAC = \measuredangle B'AC', \tag{4.36}$$

and

$$\measuredangle B'BC + \measuredangle ABC' - \measuredangle ABC = \measuredangle B'BC'. \tag{4.37}$$

Therefore, equality 4.35 becomes

$$\measuredangle B' + \measuredangle AC'B \leq \measuredangle B'AC' + \measuredangle B'BC'. \tag{4.38}$$

That is $\overset{\frown}{\triangle}\ AB'B$ and $\overset{\frown}{\triangle}\ AC'B$ are Delaunay, and C' is not contained inside the circumdisc of $\overset{\frown}{\triangle}\ AB'B$. It follows by symmetry that B' is not contained inside the circumdisc. $\square$

The rest of this section is devoted to the proof of the main theorem.

Theorem 4.2.8. (Delaunay Memphis Triangulation)

$\mathscr{M}_N$ is a Delaunay triangulation.

Proof. The proof is by induction. As the initial step, $\mathscr{M}_8$, the 8 octant, is clearly Delaunay. In fact, we have more than sufficient for our initial steps, $\mathscr{M}_N$, $N = 8, 9, ..., 14$, are all Delaunay triangulations.

Assume $\mathscr{M}_N$ is a Delauany triangulation. Denote the N points of $\mathscr{M}_N$ by $P_1, P_2, ..., P_N$, and for simplicity we will denote the $(N+1)$th point P_{N+1} by P to avoid over-typing subindices. We will show that there exists a Delaunay triangulation for $\mathscr{M}_N \cup \{P\}$, and this triangulation agrees with $\mathscr{M}_{N+1}$.

By Theorem 4.2.1, we know P is either the circumcenter of an acute triangle or a right triangle $\overset{\frown}{\triangle}\ P_iP_jP_k$. By Corollary 4.2.5, $\overset{\frown}{\triangle}\ P_iP_jP_k$ happens to be right, its adjacent triangle $\overset{\frown}{\triangle}\ P_iP_jP_l$ is right too and co-circular with $\overset{\frown}{\triangle}\ P_iP_jP_k$. Connecting the circumcenter P with P_i, P_j, P_k (*and* P_l *resp.*) with 3 (*or* 4 *resp.*) auxiliary lines (*by lines we mean great circle segments*). Extend the auxiliary lines to the $-P$ (*the antipodal point of* P), these auxiliary lines divide the sphere in to 3 (*or* 4 *resp.*) regions with disjoint interior.

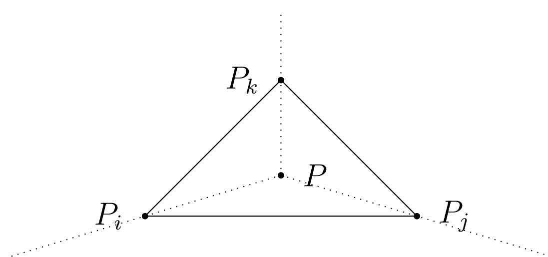


Figure 4.42: P is the circumcenter of the acute triangle $\overset{\frown}{\triangle}\ P_iP_jP_k$

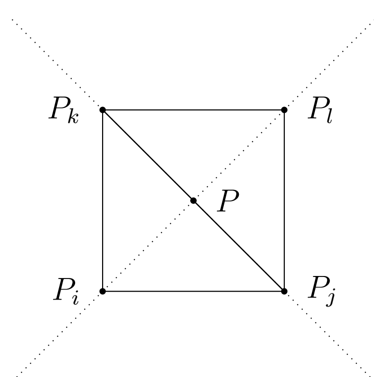


Figure 4.43: P is the circumcenter of the right triangle $\overset{\frown}{\triangle}\ P_iP_jP_k$

The number of the regions, *either* 3 *or* 4, is not important, as we will consider each region separately. WLOG we will only consider the region bounded by the auxiliary lines passing P, P_i and P, P_j between P and $-P$.

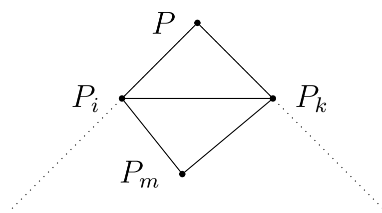


Figure 4.44

Let $\overset{\frown}{\triangle}\ P_iP_mP_j$ be the adjacent triangle of $\overset{\frown}{\triangle}\ PP_iP_j$.

Lemma 4.2.9. *If P_m is not in the interior of the region between P and $-P$ bounded by the two auxiliary lines, then $\overset{\frown}{\triangle}\ P_iP_mP_j$ and $\overset{\frown}{\triangle}\ PP_iP_j$ are Delaunay.*

Proof. (Lemma 4.2.9)

Case 1: P_m lies on one of the auxiliary line $\overset{\frown}{PP_i}$ (*or resp.* $\overset{\frown}{PP_j}$ *followed by symmetry*).

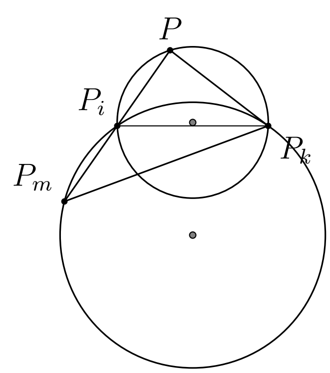


Figure 4.45

In this case, the point P_m lies on the extension of $\overset{\frown}{PP_m}$. Only the points on the arc $\overset{\frown}{PP_m}$ can be inside the circumcircle of $\overset{\frown}{\triangle}\ PP_iP_j$. Thus, P_m must be outside the circumcircle of $\overset{\frown}{\triangle}\ PP_iP_j$.

Case 2: P_m lies on the other side of line $\overset{\frown}{PP_i}$ (*or resp.* $\overset{\frown}{PP_j}$ *followed by symmetry*).

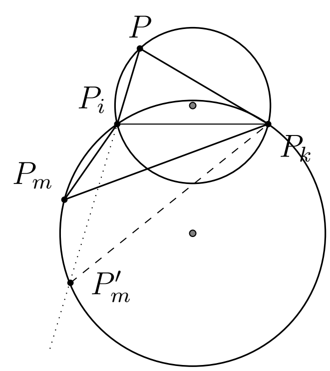


Figure 4.46

Since P_m and P_k are on different sides of the auxiliary line $\overset{\frown}{PP_i}$, the dotted auxiliary line must intersect the circumcircle of $\overset{\frown}{\triangle}\ P_mP_kP_i$ at some point denoted by P'_m. By the previous case, P'_m must lie outside the circumcircle of $\overset{\frown}{\triangle}\ PP_iP_k$. In other words,

$$\measuredangle P'_m + \measuredangle P < \measuredangle PP_iP'_m + \measuredangle PP_kP'_m. \tag{4.39}$$

Since P_m and P'_m are co-circular on the circumcircle of $\overset{\frown}{\triangle}\ P_iP_mP_j$, we know

$$\measuredangle P_m - \measuredangle P'_m = \measuredangle P_mP_iP_k + \measuredangle P_iP_kP_m - \measuredangle P'_mP_iP_k - \measuredangle P'_mP_kP_i. \tag{4.40}$$

Adding inequality 4.39 and equation 4.40, we get

$$\measuredangle P + \measuredangle P_m < \measuredangle P_mP_iP + \measuredangle P_mP_kP. \tag{4.41}$$

Therefore, $\overset{\frown}{\triangle}\ P_iP_mP_j$ and $\overset{\frown}{\triangle}\ PP_iP_j$ are Delaunay. $\square$

If the adjacent triangle $\overset{\frown}{\triangle}\ P_iP_mP_j$ is Delaunay with $\overset{\frown}{\triangle}\ P_iP_jP$ then we are done by Lemma (4.2.7). If not, then we draw another auxiliary line between P and $-P$ passing P_m.

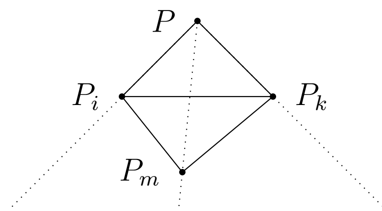
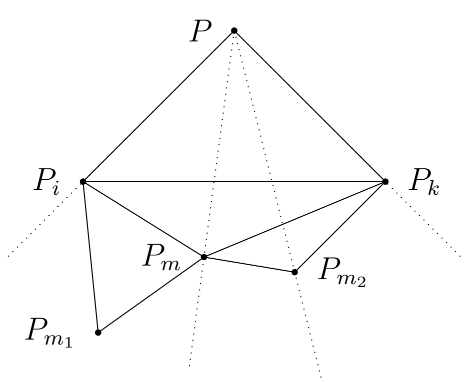


Figure 4.47

This splits the current region into two subregions. $\overset{\frown}{\triangle}\ P_iP_mP_k$ have two neighbours, one sharing the common edge $\overset{\frown}{P_kP_m}$ whiling the other sharing the common edge $\overset{\frown}{P_iP_m}$. Denote these two triangles by $\overset{\frown}{\triangle}\ P_iP_{m_1}P_m$ and $\overset{\frown}{\triangle}\ P_kP_{m_2}P_m$, as indicated in the figure below. *For exactly same reasoning as in Lemma 4.2.9* either P_{m_1} and P_{m_2} are in the interior of the subregions (*one point inside each subregion*)or the corresponding triangle (*or triangles*) $\overset{\frown}{\triangle}\ P_iP_{m_1}P_m$ or/and $\overset{\frown}{\triangle}\ P_iP_{m_2}P_m$ will not cover P inside their circumcircle(s). *Then we repeat the same process to the neighbours adjacent to* $\overset{\frown}{\triangle}\ P_iP_{m_1}P_m$ *and* $\overset{\frown}{\triangle}\ P_kP_{m_2}P_m$. *We will call this process "**stacking**".*



If $\overset{\frown}{\triangle}\ P_iP_{m_1}P_m$ *doesn't cover* P*, the process terminates and no auxiliary line passing* P_{m_1}*. If* $\overset{\frown}{\triangle}\ P_iP_{m_2}P_m$ *covers* P*, then we draw an auxiliary line passing* P_{m_2} *between* P *and* $-P$*.*

Figure 4.48

The *stacking* cannot be repeated forever because, first, there are finitely many triangles in $\mathscr{M}_N$, and if a point is "far away" enough then its triangle cannot cover P.

Lemma 4.2.10. *Let r denote the circumradius of $\widehat{\triangle}\ P_iP_jP_k$. Given a point Q in $\mathscr{M}_N$, if the distance between Q and P is no less than $2r$, then Q is not covered inside the circumdisc of $\widehat{\triangle}\ P_iP_jP_k$.*

Proof. Any two points inside the circumcircle of $\widehat{\triangle}\ P_iP_jP_k$ have distance less than $2r$. The lemma follows trivially. $\square$

So we will eventually end up with the diagram like the following shape, which we will refer to as *a* ***leaf*** *diagram. For the purpose of clarity, the points in the* leaf *are re-indexed $P, P_i, P_{n_1}, ..., P_x, P_j$.*

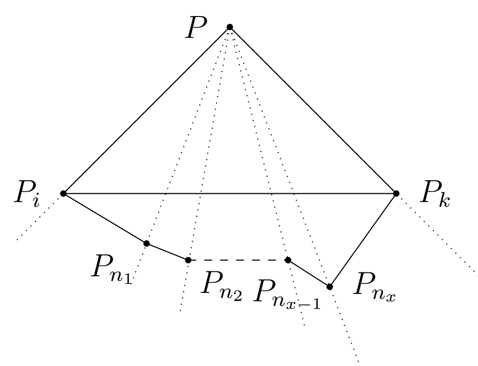


Figure 4.49: *A leaf diagram*

Following immediately from Lemma 4.2.9 we have the following:

Corollary 4.2.11. *If P_{n_1} is the (first) point on the boundary of the leaf connected to P_i, then the angle $\measuredangle PP_iP_{n_1}$ is strictly less than π. Similarly we have $\measuredangle PP_kP_{n_x}$ is strictly less than π, if P_{n_x} is the (first) point on the boundary of the leaf connected to P_k.*

Remark 4.2.12. *By the way the leaf diagram is constructed, there exists a one-one correspondence between each point of $P_i, P_{n_1}, ..., P_{n_x}, P_k$ and each auxiliary line.*

Notation 20. *We will denote the leaf diagram by $\mathcal{L}$ including the structure of the triangulation. Every triangle in $\mathcal{L}$ contains P inside its circumcircle.*

Lemma 4.2.13. (Boundary of $\mathscr{L}$) *All $P, P_i, P_{n_1}, P_{n_2}, ..., P_{n_x}, P_k$ lie on the boundary of $\mathscr{L}$.*

Proof. (Lemma 4.2.13) Clearly P, P_i, P_k are on the boundary of $\mathscr{L}$. Assume a point say $P_w \in \{P_{n_1}, P_{n_2}, ..., P_{n_x}\}$ is in the interior of $\mathscr{L}$. There are last least 3 triangles in $\mathscr{L}$ sharing P_w as the common vertex.

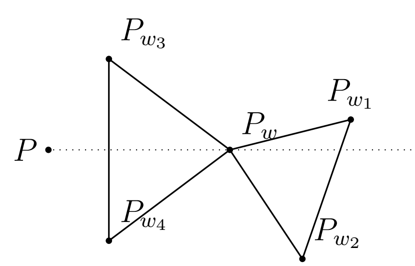
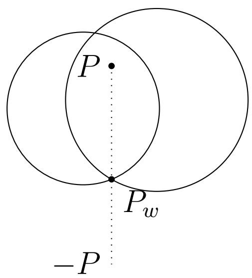


Figure 4.50: Auxiliary line passing P, P_w passes two triangles sharing P

The auxiliary line, initiated from P, passing P_w, ending at $-P$, passes two of triangles sharing P_w as a common vertex (by "passing" the triangles, we mean that the auxiliary line either passes the interior of the triangle or overlap with an edge). This auxiliary line is divided into two arcs by P_w, $\overset{\frown}{PP_w}$ and $\overset{\frown}{P_w(-P)}$. We denote the two triangles the auxiliary line passes by $\overset{\frown}{\triangle}\ P_wP_{w_1}P_{w_2}$ and $\overset{\frown}{\triangle}\ P_wP_{w_3}P_{w_4}$, and further assume $\overset{\frown}{PP_w}$ passes $\overset{\frown}{\triangle}\ P_wP_{w_3}P_{w_4}$ while $\overset{\frown}{P_w(-P)}$ passes $\overset{\frown}{\triangle}\ P_wP_{w_1}P_{w_2}$.

Meanwhile, we know that P lies inside both the circumcircle of $\overset{\frown}{\triangle}\ P_wP_{w_1}P_{w_2}$ and the circumcircle of $\overset{\frown}{\triangle}\ P_wP_{w_3}P_{w_4}$.



P lies in the intersection of the interiors of the circumcircles of $\overset{\frown}{\triangle}\ P_wP_{w_1}P_{w_2}$ and $\overset{\frown}{\triangle}\ P_wP_{w_3}P_{w_4}$.

Figure 4.51

So,the great circle segment $\overset{\frown}{PP_w}$ lies inside their intersection. Therefore, the great circle segment $\overset{\frown}{P_w(-P)}$ must be outside both circumcircles, and hence doesn't pass $\overset{\frown}{\triangle}\ P_wP_{w_1}P_{w_2}$. We have reached a contradiction. Therefore, P_w must be a boundary point of the leaf diagram $\mathscr{L}$. $\square$

Now we show the most important property of the leaf diagram:

Lemma 4.2.14. (Main Property of $\mathscr{L}$) *No four points of $P, P_i, P_{n_1}, P_{n_2}..., P_{n_x}, P_j$ can be cocircular with an empty circumcircle.*

Proof. (Lemma 4.2.14) By contradiction, assume there exist 4 points cocircular.

Case 1: P is one of the 4 points, say the 4 points are $P, P_{k_1}, P_{k_2}, P_{k_3}$. By the construction of the *leaf diagram*, there exist 3 auxiliary lines initiated from P passing through

$P_{k_1}, P_{k_2}, P_{k_3}$ individually. P_{k_2} will be denoting the point in between the two auxiliary lines passing P_{k_1} and P_{k_3}. In other words, P_{k_2} lies on one of the two arcs between P_{k_1} and P_{k_3}.

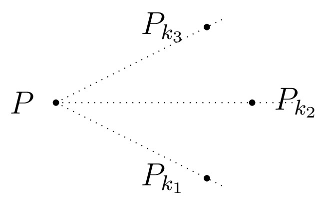


Figure 4.52

Since there is an auxiliary line passing P_{k_2} there exists a triangle $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ in $\mathcal{L}$ such that $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ contains P inside its circumcircle.

Remark 4.2.15. Although $P_{k_1}, P_{k_2}, P_{k_3}, P_y, P_z$ are all points in $\mathscr{L}$ with auxiliary lines passing them, $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ exists in $\mathscr{M}_N$ while $\overset{\frown}{\triangle}\ P_{k_1}P_{k_2}P_{k_3}$ may not. They are distinct triangles since one of them contains P inside its circumcircle and the other is co-circular with P. Hence, $\{P_y, P_z\} \neq \{P_{k_1}, P_{k_3}\}$. We may assume $P_{k_1} \neq P_y$ and $P_{k_1} \neq P_z$.

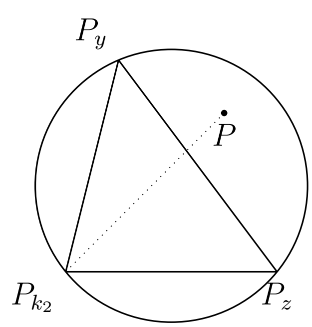
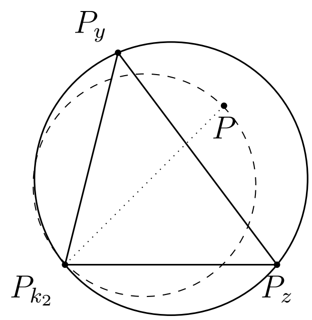


Figure 4.53: The circumcircle of $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ contains P inside

Then the circumcircle passing $P, P_{k_1}, P_{k_2}, P_{k_3}$ will intersect with the circumcircle of $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ at *either* 1 *or* 2 points.

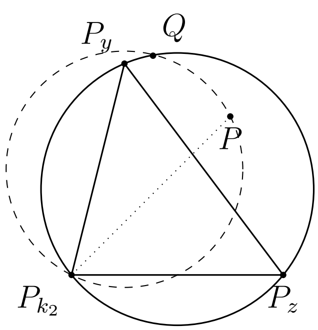


When P_{k_2} is the only intersection of the two circles, the circumcircle of $\overset{\frown}{\triangle}$ $P_{k_1}P_{k_2}P_{k3}$ is completely contained inside the circumcircle of $\overset{\frown}{\triangle}$ $P_{k_2}P_yP_z$.

Figure 4.54

Clearly, as indicated in figure 4.54, when the two circumcircle have only one intersection, i.e P_{k_2} P_{k_1}, P_{k_3} would be contained inside the circumcircle of $\overset{\frown}{\triangle}$ $P_{k_2}P_yP_z$. This is a contradiction to $\mathscr{M}_N$ being Delaunay.

When the circumcircle of $\overset{\frown}{\triangle}$ $P_{k_2}P_yP_z$ and the circumcircle of $\overset{\frown}{\triangle}$ $P_{k_1}P_{k_2}P_{k3}$ intersect at two points, one of the intersections is P_{k_2}. Denote the other intersection by Q (Q *could be one of* P_y *and* P_z).



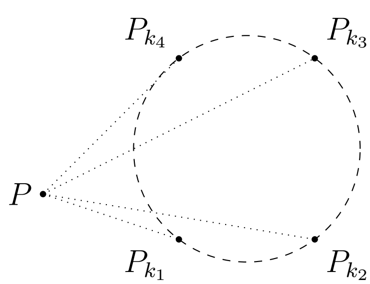
The dashed circle is the circumcircle of $\overset{\frown}{\triangle}$ $P_{k_1}P_{k_2}P_{k_3}$. Q may coincide with P_z or P_y. One arc between P and Q lies inside the circumcircle of $\overset{\frown}{\triangle}$ $P_yP_zP_{k_2}$.

Figure 4.55

Q and P_{k_2} divide the circumcircle into two arcs. One of those arcs must lie inside the circumcircle of $\overset{\frown}{\triangle}$ $P_yP_zP_{k_2}$. The arc inside the circumcircle of $\overset{\frown}{\triangle}$ $P_yP_zP_{k_2}$ must have either P_{k_1} or P_{k_3} on it. This is a contradiction to $\mathscr{M}_N$ being Delaunay.

Case 2: P is not one of the 4 points. We will label the 4 points with the following positioning: P_{k_2}, P_{k_3} lie in the interior of the area bounded between by the auxiliary lines passing P_{k_1} and P_{k_4}; P_{k_2} lies in the interior of the area bounded by the auxiliary lines passing P_{k_1} and P_{k_3}.

Remark 4.2.16. Later we will obtain a contradiction with the way the 4 points are positioned on the circle. An equivalent way of describe the labeling of the 4 points is: P_{k_1} and P_{k_4} divide the circle into two arcs, P_{k_2} and P_{k_3} lie on the same one (of the two); P_{k_3} further divides *this* arc into two *sub-arcs*, one with end points $P_{k_1}P_{k_3}$ while the other with end points $P_{k_3}P_{k_4}$; P_{k_2} lies on the former *sub-arc* end points $P_{k_1}P_{k_3}$.



The auxiliary line passing P, P_{k_2} as P_{k_3} cuts through the circumcircle of the 4 points. P_{k_3} and P_{k_4} are on the opposite side of P_{k_1}.

Figure 4.56

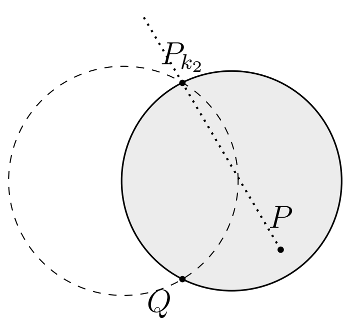
Remark 4.2.17. *P is outside the circumcircle and P_{k_i} for $i = 1, 2, 3, 4$ are co-circular, the auxiliary line passing P, P_{k_2} has no empty intersection with the interior of the circumcircle (of $P_{k_1}P_{k_2}P_{k_3}P_{k_4}$). So, this auxiliary divides this circle into two arcs with P_{k_2}*

being an end point. By the way the 4 points are positioned, P_{k_1} lies on the different arc as P_{k_3} and P_{k_4}.

Let $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ be the triangle in $\mathcal{L}$ that contains P inside. The circumcircle of $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$ cannot coincide with the circle passing the 4 points $P_{k_i}i = 1, 2, 3, 4$. Then $\{P_y, P_z\} \not\subset \{P_{k_1}, P_{k_3}, P_{k_4}\}$. P_{k_2} is one intersection of the circumcircle of $\{P_{k_1}, P_{k_2}, P_{k_3}, P_{k_4}\}$ and the circumcircle of $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$.

If P_{k_2} were the only intersection of these two circles, then the circle passing $P_{k_1}, P_{k_2}, P_{k_3}, P_{k_4}$ would be contained entirely inside the circumcircle of $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$, contradicting to $\mathscr{M}_N$ is a Delaunay Triangulation.

So we are left with the possibility that these two circumcircles have two intersections, one of which is P_{k_2}. Using the same notation as in case 1, let's denote the other intersection by Q. Q and P_{k_2} divide the circumcircle of $P_{k_1}P_{k_2}P_{k_3}P_{p_4}$ into two arcs, one of them lies entirely inside the circumcircle of $\overset{\frown}{\triangle}\ P_{k_2}P_yP_z$. Denote this arc by $\overset{\frown}{\mho}$.



Interior of the circumcircle of $P_{k_2}\overset{\frown}{P_y}P_z$ is colored grey. The circle passing $P_{k_i}, i = 1, 2, 3, 4$ is dashed. P lies in the in the interior of the grey circle and the exterior of the dashed circle.

Figure 4.57

The great circle segment $\overset{\frown}{P}\, P_{k_2}$ lies inside the circumcircle of $\overset{\frown}{\triangle}\, P_{k_2}P_yP_z$. $P\overset{\frown}{P}_{k_2}$ intersect with $\overset{\frown}{\mho}$. By remark 4.2.17, *either* P_{k_1} *or* both P_{k_3} and P_{k_4} lie on arc $\overset{\frown}{\mho}$. Hence, *either* P_{k_1} *or* both P_{k_3} and P_{k_4} lie in the interior of the circumcircle of $\overset{\frown}{\triangle}\, P_{k_2}P_yP_z$. This is a contradiction to that $\mathscr{M}_N$ is Delaunay. $\square$

Notation 21. (The Stereographic Projection)

The particular Stereographic Projection settings we use have P as the North Pole, with the plane tangent to the sphere at the South Pole $-P$.

Let $Q, Q_i, Q_{n_1}, ..., Q_k$ denote the image of $P, P_i, P_{n_1}, ...P_k$ on the plane under the map *Stereographic Projection.* Because no four points of $P, P_i, P_{n_1}, ..., P_{n_x}, P_k$ can be cocircular with empty circumcircle, no four points of $Q, Q_i, Q_{n_1}, ..., Q_k$ can be co-circular with empty interior. By Theorem 4.1.3 there exists a unique Delaunay Triangulation among $Q, Q_i, Q_{n_1}, ..., Q_k$. Project the triangulation back to the sphere, we get a Delaunay Triangulation of $P, P_i, P_{n_1}, P_{n_2}...P_k$. Let's denote this Delaunay Triangulation $\mathscr{T}$.

Remark 4.2.18. *The ultimate goal here is to create a Delaunay Triangulation for $\mathscr{M}_N \cup \{P\}$. The boundary of $\mathscr{T}$ might be different than the boundary of its leaf diagram $\mathscr{L}$. For example, it is not hard to imagine that the convexity of the boundaries could be very different, like the following figures imply.*

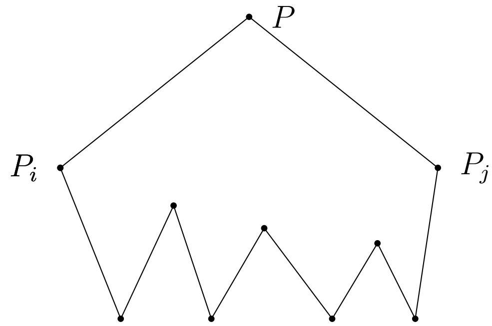


Figure 4.58: *The boundaries of $\mathscr{L}$ before the Stereographic Projection*

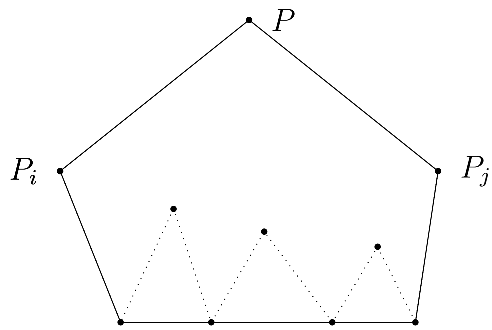


Figure 4.59: *The boundaries of $\mathscr{T}$ after the Stereographic Projection*

Now we know that after projecting the planar Delaunay Triangulation back to the sphere, the boundary of the leaf might change.

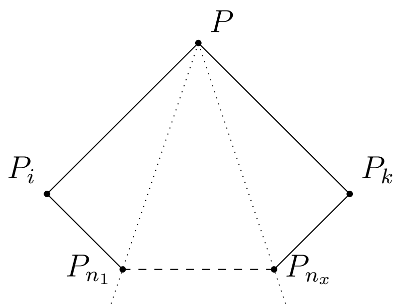
Notation 22. *If some edge* $\boldsymbol{e}$ *lies on the boundary of* $\mathscr{T}$ *but not on the boundary of* $\mathscr{L}$*, we say* $\boldsymbol{e}$ *is "unwanted".*

Notation 23. *If an unwanted edge* $\boldsymbol{e}$ *lies in a triangle of* $\mathscr{T}$ *whose other two edges (other than* $\boldsymbol{e}$*) are both on the boundary of* $\mathscr{L}$*, we say the unwanted edge* $\boldsymbol{e}$ *is "removable".*

To create an Delaunay Triangulation among the points with $\mathscr{L}$, we need to show all the *unwanted* (boundary) edges are *removable.* Denote the boundary of the leaf as by the ordered tuple $(P, P_i, P_{n_1}, ...P_{n_x}, P_k)$.

Remark 4.2.19. The boundary tuple $(P, P_i, P_{n_1}, ...P_{n_x}, P_k)$ can be viewed as connected *"loop" or "path"*. An equivalent way of saying all the unwanted (boundary) edges are removable is that *the same "path"* $(P, P_i, P_{n_1}, ...P_{n_x}, P_k)$ *still exists in* $\mathscr{T}$.

Remark 4.2.20. As we mentioned before, $\measuredangle P_iPP_k$, $\measuredangle PP_iP_{n_1}$, $\measuredangle PP_kP_{n_x}$ are all less than π. Part of the leaf looks like the following:



Angles in between the two dotted auxiliary lines are of unknown shapes, i.e they may or may not be convex. All that we know for sure is that $\measuredangle P_iPP_k$, $\measuredangle PP_iP_{n_1}$, $\measuredangle PP_kP_{n_x}$ *are all less than* π

Figure 4.60

In the process of mapping $\mathscr{L}$ to the plane by Stereographic Projection and then obtaining the planar unique Delaunay Triangulation with the image points, we have the following key observation:

Remark 4.2.21. **(A Key Observation)** The planar image of $\measuredangle P_iPP_k$, $\measuredangle PP_iP_{n_1}$, $\measuredangle PP_kP_{n_x}$ under our Stereographic Projection (with P being the North Pole, defined in Notation 21), $\measuredangle Q_iQQ_k$, $\measuredangle QQ_iQ_{n_1}$, $\measuredangle QQ_kQ_{n_x}$, are still less than π. The auxiliary lines passing $\widehat{PP_i}$ and $\widehat{PP_k}$, under the projection, become the infinite planar rays QQ_i and QQ_k initiated from Q. The planar Delaunay Triangulation are occurring in between the two infinite rays QQ_i and QQ_k.

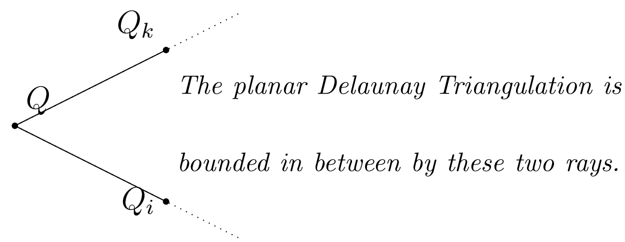


Figure 4.61

The point we are trying to make is that the planar path (Q_i, Q, Q_k) *lies on the boundary of the planar Delaunay Triangulation. After projecting the planar Delaunay Triangulation back to the sphere and obtaining* $\mathscr{T}$, *the path* (P_i, P, P_k) *remains the same.*

Now we need to focus on the path on the boundary $(P_i, P_{n_1}, ..., P_{n_x}, P_k)$ and proving it exists as a path in $\mathscr{T}$.

Notation 24. *We will call each of the following paths* (P_i, P_{n_1}), (P_{n_1}, P_{n_2}), ..., (P_{n_x}, P_k) *a segment of the path* $(P_i, P_{n_1}, ..., P_{n_x}, P_k)$.

The only way that $(P_i, P_{n_1}, ..., P_{n_x}, P_k)$ fails to appear the same in $\mathscr{T}$ as in $\mathscr{L}$ is that one of its segment (P_a, P_b) for $a, b \in \{i, n_1, n_2, .., n_x, k\}$, is removed. Assume $\widehat{\triangle}\ P_aP_{a_1}P_{a_2}$ is a new triangle formed (i.e didn't exist in $\mathscr{L}$) in $\mathscr{T}$ and $\widehat{P_{a_1}P_{a_2}}$ crosses (or, in other words, replaces) $\widehat{P_aP_b}$. Further assume $\widehat{\triangle}\ P_aP_bP_c$ was a triangle in $\mathscr{L}$.

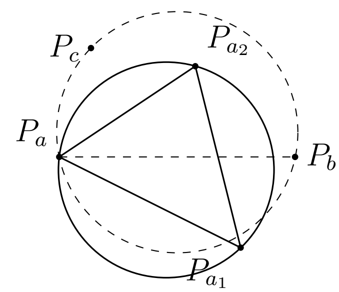


Figure 4.62: Formation of new edges in $\mathscr{T}$

Recall no 4 four points can be co-circular and $\mathscr{T}$ is Delaunay. P_b must lie outside the circumcircle of $\widehat{\triangle}\ P_aP_{a_1}P_{a_2}$. The circumcircle of $\widehat{\triangle}\ P_aP_bP_c$ intersect with the circumcircle of $\widehat{\triangle}\ P_aP_{a_1}P_{a_2}$ at one or two intersections (one of the intersections is P_a). If there is only one intersection, then the circumcircle of $\widehat{\triangle}\ P_aP_{a_1}P_{a_2}$ is contained inside the circumcircle of $\widehat{\triangle}\ P_aP_bP_c$. If there are two intersections, then one of the arcs between P_a and P_b lies outside the circumcircle of $\widehat{\triangle}\ P_aP_{a_1}P_{a_2}$. One of the points P_{a_1}, P_{a_2}, say P_{a_2}, must lie on the same side of $\widehat{P_aP_b}$ as that arc. Then P_{a_2} is contained inside the circumcircle of $\widehat{\triangle}\ P_aP_bP_c$, which is a contradiction. We have proven that none of the segment of path $(P_i, P_{n_1}, ..., P_{n_x}, P_k)$ changes in $\mathscr{T}$. As a consequence, all the unwanted edges are removable.

Remark 4.2.22. One important consequence of the analysis above is that if two points in $\mathscr{L}$ are not connected but become connected in $\mathscr{T}$, then *either* they form an unwanted edge, which can be removed, *or* one of these two points must be P. These proves the following corollary.

Corollary 4.2.23. *(Memphis' Triangulation) All* $P_i, P_{n_1}, ..., P_{n_x}, P_k$ *are connected to* P *in* $\mathscr{T}$*.*

Notation 25. *Removing all the unwanted edges from* $\mathscr{T}$*, let us denote the resulting triangulation* $\mathscr{T}_1$*.* $\mathscr{T}_1$ *has the same boundary as* $\mathscr{L}$*. They are different triangulation of the same set of points. Recall from the very beginning of this proof, we started with 3 (*or *4) regions, depending on whether the deep hole* P *is the circumcenter of an acute triangle (*or *right triangle* resp.*). Each region will give us a leaf diagram, and each region will have a Delaunay Triangulation (having the same boundary as its leaf diagram). We denote them by* $\mathscr{T}_1, \mathscr{T}_2, \mathscr{T}_3$ *(*and $\mathscr{T}_4$ resp.*).*

Remark 4.2.24. Each of $\mathscr{T}_1, \mathscr{T}_2, \mathscr{T}_3$ (*and* $\mathscr{T}_4$ *resp.*) is a Delaunay Triangulation. To finish the proof, we need to "paste" them together.

Given two adjacent triangulations of $\mathscr{T}_1, \mathscr{T}_2, \mathscr{T}_3$ (*and* $\mathscr{T}_4$ *resp.*), say $\mathscr{T}_1$ and $\mathscr{T}_4$, they share a common edge $\overparen{PP_i}$.

Lemma 4.2.25. *Using the same notation as above, if* $\overparen{\triangle}\ PP_iP_{n_1'}$ *is the triangle in* $\mathscr{T}_4$ *that is adjacent to* $\overparen{\triangle}\ PP_iP_{n_1}$*. They are Delaunay.*

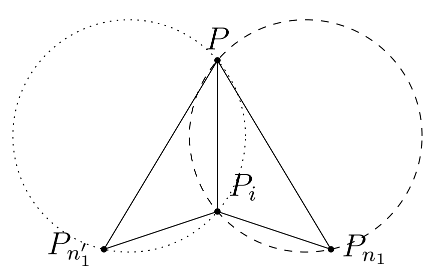


Figure 4.63

Proof. (Lemma 4.2.25) It suffices to show that $P_{n_1'}$ is outside the circumcircle of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$. Since P_{n_1} and $P_{n_1'}$ are not inside the circumcircle of $\overset{\frown}{\triangle}$ $P_iP_jP_i$, $\mid \overset{\frown}{PP_{n_1'}} \mid \geq \mid \overset{\frown}{PP_i} \mid$ and $\mid \overset{\frown}{PP_{n_1}} \mid \geq \mid \overset{\frown}{PP_i} \mid$. So, in each of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$ and $\overset{\frown}{\triangle}$ $PP_iP_{n_1'}$, *either* $\overset{\frown}{PP_i}$ is not the longest edge *or* the triangle is isosceles where the two larger edges have length $\mid \overset{\frown}{PP_i} \mid$. Therefore, in either case, the circumcenters of the two triangles $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$ and $\overset{\frown}{\triangle}$ $PP_iP_{n_1'}$ are on different sides of edge $\overset{\frown}{PP_i}$: P_{n_1} and the circumcenter of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$ are on the same side, and $P_{n_1'}$ and the circumcenter of $\overset{\frown}{\triangle}$ $PP_iP_{n_1'}$ are on the other side.

Remark 4.2.26. In particular, neither the circumcenter of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$ nor the circumcenter of $\overset{\frown}{\triangle}$ $PP_iP_{n_1'}$ can be (the midpoint) on the edge of $\overset{\frown}{PP_i}$.

P_{n_1} lies on the heavier arc between P and P_i of the circumcircle of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$ while $P_{n_1'}$ lies on the heavier arc between P and P_i of the circumcircle of $\overset{\frown}{\triangle}$ $PP_iP_{n_1'}$. The (interior of the) lighter arc of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$ between P and P_i is contained inside the area bounded by $\overset{\frown}{PP_i}$ and the heavier arc of $\overset{\frown}{\triangle}$ $PP_iP_{n_1'}$ between P and P_i. Therefore, $P_{n_1'}$ is outside the circumcircle of $\overset{\frown}{\triangle}$ $PP_iP_{n_1}$. □

Remark 4.2.27. By Lemma 4.2.25, in combination with Lemma 4.2.7, no point in $\mathscr{T}_4$ (other than P) is contained inside the circumcircle of any triangle of $\mathscr{T}_1$, and vice versa. Therefore, $\cup_i \mathscr{T}_i$ is Delaunay.

Remark 4.2.28. By Corollary 4.2.23 all the points in this union are connected to P. This configuration agrees with the one constructed with Memphis' Algorithm.

The last step of the proof relies on the following result, which states that no triangle outside the *leaves* can contain P inside its circumcircle.

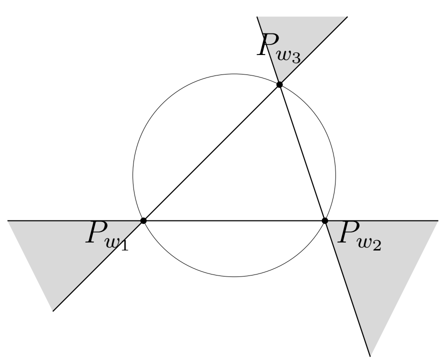
Lemma 4.2.29. *If $\overset{\frown}{\triangle}\ P_{w_1}P_{w_2}P_{w_3}$ is a triangle in $\mathscr{M}_N$ and at least one of $\{P_{w_1}, P_{w_2}, P_{w_3}\}$ is not in $\cup_i \mathscr{L}_i$, then P is not inside the circumcircle of $\overset{\frown}{\triangle}\ P_{w_1}P_{w_2}P_{w_3}$.*

(*The proof of Lemma 4.2.29 is provided after the proof of the theorem.*)

So, all the triangles of $\mathscr{M}_N$ outside of the leaves $\cup_i \mathscr{L}_i$ together with $\cup_i \mathscr{T}_i$ form a Delaunay triangulation of the N+1 points $\mathscr{M}_N \cup P$, denoted by $\mathscr{M}_{N+1}$. This completes the the proof. □

We provide the following proof to Lemma 4.2.29.

Proof. (Lemma 4.2.29)
Since $\overset{\frown}{\triangle}\ P_{w_1}P_{w_2}P_{w_3}$, none of the edges $\overset{\frown}{P_{w_1}P_{w_2}}, \overset{\frown}{P_{w_2}P_{w_3}}, \overset{\frown}{P_{w_1}P_{w_3}}$ lies in $\cup \mathscr{L}_i$ or $\cup \mathscr{T}_i$. In particular, none of the three edges lies on the boundary of $\cup \mathscr{L}_i$ or $\cup \mathscr{T}_i$. Locally the three great circles overlapping with edges $\overset{\frown}{P_{w_1}P_{w_2}}, \overset{\frown}{P_{w_2}P_{w_3}}, \overset{\frown}{P_{w_1}P_{w_3}}$ divide the sphere into disjoint regions:



The interior of circumcircle of $\overset{\frown}{\triangle}$ $P_{w_1}P_{w_2}P_{w_3}$ has empty intersection with the grey areas.

Figure 4.64

If P lies in one of the 3 shaded regions, then we are done. Suppose P is in one of the 3 unshaded regions, say the region sharing the side $\overset{\frown}{P_{w_2}P_{w_3}}$. Let $\overset{\frown}{\triangle}\ P_0P_{w_2}P_{w_3}$ be the neighbour of $\overset{\frown}{P_{w_2}P_{w_3}}$ sharing the common edge $\overset{\frown}{P_{w_2}P_{w_3}}$. Then P_0 and P is on the same side of the great circle passing P_{w_2}, P_{w_3}.

P and P_0 are on the same side of the great circle passing P_{w_2}, P_{w_3}.

P is outside the circumcircle of $\overset{\frown}{\triangle}\ P_0P_{w_2}P_{w_3}$.

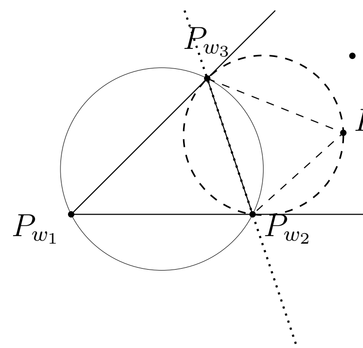


Figure 4.65

Since $\overset{\frown}{P_{w_2}P_{w_3}}$ is not an edge in $\cup_i \mathscr{L}_i$ or $\cup_i \mathscr{T}_i$, we know that $\overset{\frown}{\triangle}\ P_0P_{w_2}P_{w_3}$ is not a triangle in $\cup_i \mathscr{L}_i$ or $\cup_i \mathscr{T}_i$. Therefore, P is not inside the circumcircle of $\overset{\frown}{\triangle}\ P_0P_{w_2}P_{w_3}$. By Lemma 4.2.7, P is not inside the circumcircle of $\overset{\frown}{\triangle}\ P_{w_1}P_{w_2}P_{w_3}$ either, completing the proof. $\square$

4.3 Some Remarks

4.3.1 Non-uniqueness of $\mathscr{M}_N$

Now we know $\mathscr{M}_N$ is Delaunay. An immediate question to ask is whether this Delaunay triangulation is unique. The answer is also immediate: the degeneracy mentioned in Notation 18 can cause the triangulation to be non-unique. But can we get a conditional uniqueness if we "modulo the degeneracies"? Our goal is to investigate spherical cap discrepancies. How unique the Delaunay triangulation $\mathscr{M}_N$ is not of interest: it is not providing useful information for either of the sequence construction or the computation of the spherical discrepancy. However, we will discuss the degeneracy of $\mathscr{M}_N$ in more detail at the beginning of the next chapter.

4.3.2 Shortest Edges and Longest Edges of $\mathscr{M}_N$

By Theorem 4.2.1 and Theorem 4.2.8, every point of $\mathscr{M}_N$ is a deep hole of the previous points. Following from Theorem 3.1.23 stated and proved in the previous chapter, we get the *third regularity* property stated at the last section of Chapter 3.

Corollary 4.3.1. *(Shortest Edges of $\mathscr{M}_N$)*

The radius edges of the last added point are shortest edges in $\mathscr{M}_N$. The length of these radius edges is, $\delta(\mathscr{M}_N)$, the separation of $\mathscr{M}_N$ as a point set.

By Corollary 3.1.24, we have:

Corollary 4.3.2. *The length of the shortest edges of $\mathscr{M}_N$ is non-increasing as N grows. It decreases precisely when the N points are $\delta(\mathscr{M}_N)$-saturated.*

It is natural to ask is whether we can make similar statements about *the longest edge(s)* of $\mathscr{M}_N$. However, we don't have nearly as much information about *the longest edge(s)*: we don't know how many of them there are; we don't know how there are generated, i.e whether they are cross edges or radius edges of some previous deep holes; it seems natural to guess that the length of the longest edge(s) ought be be non-increasing as N increases, but we don't have enough evidence suggesting this either. The most we can say is the following remark on an upper bound of the longest edges.

Remark 4.3.3. Let γ_N denote the length of the shortest edges while Γ_N denote the length of the longest edge(s). The circumcircle of the triangle, who has an edge achieving length Γ_N, has radius at most $2\gamma_N$. Then we obtain a trivial bound that $\Gamma_N \leq 2\gamma_N$. Hence, as γ_N is non-increasing as N grows, we know this upper bound is a non-increasing function of N also.

Chapter 5

Spherical Cap Discrepancy

The previous two chapters introduced a recursive algorithm to produce well-separated points on S^2. Every newly generated point is a *deep hole* of the previous points. The sequence comes with a triangulation with strong regularity: the triangulation is Delaunay.

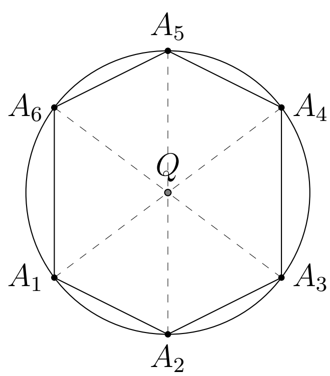
5.1 Degeneracy of $\mathscr{M}_N$

Suppose $\overset{\frown}{\triangle}\, A_1A_2A_3$ is a triangle in $\mathscr{M}_N$. If more than 3 points are co-circular on the circumcircle of this triangle, we have a degeneracy by Notation 18, in which case there are more than one triangulation among the co-circular points. However, $\mathscr{M}_N$ is Delaunay. So, intuitively, there shouldn't be too many points sharing the circumcircle.

Lemma 5.1.1. *(Maximal Cocirculation)*

Suppose $\overset{\frown}{\triangle}\ A_1A_2A_3$ *is a triangle in* $\mathscr{M}_N$. *There are at most two other points on the circumcircle of* $\overset{\frown}{\triangle}\ A_1A_2A_3$.

Proof. By contradiction. Assume there are 6 points co-circular. Denote these 6 points by $A_i, i = 1, 2, ..., 6$. WLOG we may assume that the relative positions of the 6 points on the circle are $A_1, A_2, ..., A_6$ counterclockwise. Let Q be the common circumcenter. Q is not a point in $\mathscr{M}_N$. And it may not be a deep hole of $\mathscr{M}_N$. $\mid \overset{\frown}{A_iQ} \mid$ is not longer than the shortest length between any pair of points in $\mathscr{M}_N$. Each edge of *the spherical hexagon* $\overset{\frown}{\hexagon}\ A_1A_2A_3A_4A_5A_6$ is no shorter than the shortest edge(s) of $\mathscr{M}_N$. Hence, none of the edges of $\overset{\frown}{\hexagon}\ A_1A_2A_3A_4A_5A_6$ is shorter than the circumradius of this hexagon.



Q is merely a circumcenter of some triangle in $\mathscr{M}_N$. The circumradius of $\overset{\frown}{\triangle}\ A_1A_2A_3$ is less than or equal to the length of the shortest edge(s) of $\mathscr{M}_N$. One of the 6 central angles must be less than or equal to $\pi/3$.

Figure 5.66

By the Pigeonhole Principle, at least one of the *central angles* of the 6 edges $\overset{\frown}{A_iA_{i+1}}, i = 1, 2, .., 5$ and $\overset{\frown}{A_1A_6}$ must be no bigger than $\pi/3$. Again, WLOG, we may assume the central angle $\measuredangle A_1QA_2 \leq \pi/3$. In the isosceles triangle $\overset{\frown}{\triangle}\ A_1QA_2$, $\overset{\frown}{A_1A_2}$ is the longest edge. Hence,

$$\measuredangle QA_1A_2 = \measuredangle QA_2A_1 \leq \measuredangle A_1QA_2 \leq \pi/3. \tag{5.1}$$

Consequently, $area(\overset{\frown}{\triangle}\ A_1QA_2) \leq 3 \cdot \pi/3 - \pi = 0$. This is impossible. $\square$

5.2 Estimating $\delta(\mathscr{M}_N)$, The Separation of $\mathscr{M}_N$

5.2.1 An Upper Bound

By inductively choosing every point as far from the previous as possible, our goal is to make the points as well separated as we can. However, by the following theorem, no matter how well separated the points are placed, there always exists some point that is "close" to the rest. More precisely, we have the following upper bound on the separation of $\mathscr{M}_N$.

Theorem 5.2.1. *[11] From $n > 2$ points on S^2 there can always be found two with spherical distance no bigger than*

$$\arccos \frac{\cot^2 \omega - 1}{2}, \omega = \frac{n}{n-2}\frac{\pi}{6}. \tag{5.1}$$

Rewriting $\omega = \frac{1}{2}\frac{n\pi}{3n-6}$ and applying the double angle formula, the above theorem can also be stated as:

Corollary 5.2.2. (An Upper Bound for $\delta(\mathscr{M}_N)$)

$$\delta(\mathscr{M}_N) \leq \arccos \frac{\cos \frac{\pi N}{3N-6}}{1 - \cos \frac{\pi N}{3N-6}}. \tag{5.2}$$

5.2.2 Memphis' Lower Bound

Memphis' Triangulation is an example of the *Euler Triangulation.* The *Euler Characteristic* formula applies.

$$\mathbf{V - E + F = 2,}$$

where we follow the convention that **V, E** and **F** stand for the numbers of *vertices, edges* and *faces* respectively. Recalling how Memphis' Algorithm runs: every time a new point is added to the existing sequence, the triangulation gains *3 more* faces and *2 more* edges, which is consistent with the *Euler Characteristic.*

Remark 5.2.3. It is not hard to see that the number of triangles in $\mathscr{M}_N$ is $2N-4$. So the average area of a triangle in $\mathscr{M}_N$ is $\frac{1}{2N-4}$.

Notation 26. *Denote a triangle in $\mathscr{M}_N$ with the largest area by $\overset{\frown}{\triangle}\, T_{max}$.* Recall in Chapter 3 we define the size of a triangle by its circumradius, i.e a larger triangle means a triangle with longer circumradius. *So, $\overset{\frown}{\triangle}\, T_{max}$ may or may not be a largest triangle. The circumcenter of $\overset{\frown}{\triangle}\, T_{max}$ may or may not be a* deep hole*. Let $\overset{\frown}{\triangle}\, \Omega$ denote a largest triangle in $\mathscr{M}_N$, whose circumcenter is a* deep hole*.*

Using the notations introduced in Chapter 2, let **a, b, c** be the edges of a spherical triangle, with circumradius R,

$$w = \cos\mathbf{a}\cos\mathbf{b} + \cos\mathbf{a}\cos\mathbf{c} + \cos\mathbf{b}\cos\mathbf{c};$$

$$u = 1 + \cos\mathbf{a} + \cos\mathbf{b} + \cos\mathbf{c};$$

$$D^2 = 1 + 2\cos\mathbf{a}\cos\mathbf{b}\cos\mathbf{c} - \cos^2\mathbf{a} - \cos^2\mathbf{b} - \cos^2\mathbf{c}.$$

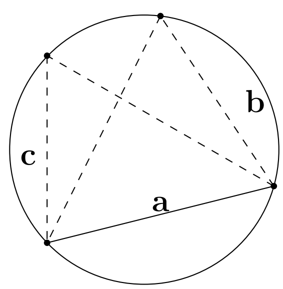
The formula to compute B in Lemma 3.1.25,

$$\sec^2 R = \frac{1}{D}(4(1+w) - u^2), \tag{5.3}$$

which can be rewritten as

$$\tan^2 R = \frac{2(1-\cos\mathbf{a})(1-\cos\mathbf{b})(1-\cos\mathbf{c})}{D^2}. \tag{5.4}$$

Let $x = \cos\mathbf{a}$, $y = \cos\mathbf{b}$ and $z = \cos\mathbf{c}$. Fix edge $\mathbf{a}$. Let edges **b, c** vary while keeping the circumradius R constant.



*The circumcircle and edge **a** remain the same. As edge **b** and edge **c** vary, the point facing **a** slides on the circumcircle.*

Figure 5.67

As R and $x = \cos\mathbf{a}$ remain constant, equation 5.4 gives:

$$\frac{\tan^2 R}{2(1-x)} = \frac{(1-z)(1-y)}{1+2xyz-x^2-y^2-z^2} = \text{constant}. \tag{5.5}$$

Therefore, after differentiating both sides with respect to variable z, we get

$$0 = \frac{d}{dz}\frac{(1-z)(1-y)}{1+2xyz-x^2-y^2-z^2}. \tag{5.6}$$

$$0 = \frac{F_z G - F G_z}{G^2}, \tag{5.7}$$

where

$$F(y,z) = (1-y)(1-z) \text{ and } G(y,z) = 1+2xyz-x^2-y^2-z^2. \tag{5.8}$$

So, we have

$$F_z G = F G_z. \tag{5.9}$$

Computing the partial derivatives of F and G with respect to Z,

$$F_z = (z-1)\frac{dy}{dz} + (y-1). \tag{5.10}$$

$$G_z = (2xz-y)\frac{dy}{dz} + (2xy-2z). \tag{5.11}$$

$$F_z G = (1+2xyz-x^2-y^2-z^2)(z-1)\frac{dy}{dz} + (1+2xyz-x^2-y^2-z^2)(y-1). \tag{5.12}$$

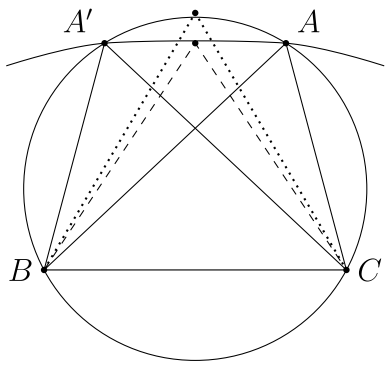
$$G_z F = (1-y)(1-z)(2xz-2y)\frac{dy}{dz} + (1-y)(1-z)(2xy-2z). \tag{5.13}$$

By equating $F_z G = F G_z$ (equation 5.9), we get

$$y = z \text{ } or \text{ } y+z = x+1. \tag{5.14}$$

Remark 5.2.4. We have reached the conclusion that, while edge **a** and R remain the same, edge **b** increases *if and only if* edge **c** decreases.

Let $\overset{\frown}{\triangle}\, ABC$ be the triangle with edge **a,b,c**, **a** $=\overset{\frown}{BC}$, **b** $=\overset{\frown}{AC}$ and **c** $=\overset{\frown}{AB}$. Let A' be a point on the circumcircle of $\overset{\frown}{\triangle}\, ABC$ such that A and A' are on the same side of $\overset{\frown}{BC}$ and $\overset{\frown}{A'B}$ and $\overset{\frown}{AC}$ are of the same length. We will refer to A' as *the symmetric image* of A on the circumcircle of $\overset{\frown}{\triangle}\, ABC$.



A' is the symmetric image of A on the same circle. The shorter arc between A and A' on Lexell's Circle lies inside the circumcircle of $\overset{\frown}{\triangle}$ ABC. The dotted triangle has the largest area, while the other three have the same area.

Figure 5.68

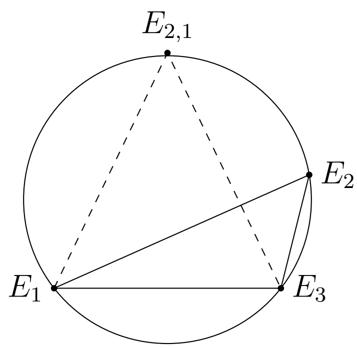
The Lexell Circle (*see Definition* 2.2.10) of $\overset{\frown}{\triangle}$ ABC with base $\overset{\frown}{BC}$ intersect with the circumcircle of $\overset{\frown}{\triangle}$ ABC at two points A and A'. When $\overset{\frown}{\triangle}$ ABC is isosceles with $\mid \overset{\frown}{AB} \mid = \mid \overset{\frown}{AC} \mid$ A and A' coincide. The shorter arc of Lexell circle between A and A' lies inside the circumcircle of $\overset{\frown}{\triangle}$ ABC. We have obtained the following result:

Corollary 5.2.5. *Fixing $\overset{\frown}{BC}$ and the circumradius and letting A vary, the area of $\overset{\frown}{\triangle}$ ABC occurs precisely when $\overset{\frown}{\triangle}$ ABC is isosceles with $\mid \overset{\frown}{AB} \mid = \mid \overset{\frown}{AC} \mid$.*

Definition 5.2.6. Using the same settings, we call this isosceles triangle *the Lexell's triangle* with base $\overset{\frown}{BC}$ for the fixed circumradius.

Lemma 5.2.7. *With fixed circumradius the regular triangles have the largest area.*

Proof. (Lemma 5.2.7) Let $\overset{\frown}{\triangle}$ $E_1E_2E_3$ be an arbitrary spherical triangle with fixed circumradius R. Let e_1, e_2, e_3 denote the edges facing points E_1, E_2, E_3 respectively. If $\overset{\frown}{\triangle}$ $E_1E_2E_3$ not regular, then it has *a longest edge* and *a shortest edge.* Say e_3 is longest, and e_1 is the shortest. Fix edge e_2. We get *the Lexell's triangle* with base e_2. Let's denote this isosceles *Lexell's triangle* by $\widehat{\triangle}\ E_{1,1}E_{2,1}E_{3,1}$ and denote edge $\widehat{E_{1,1}E_{2,1}}$, $\widehat{E_{2,1}E_{3,1}}$, $\widehat{E_{1,1}E_{3,1}}$ by $e_{3,1}$, $e_{1,1}$, $e_{2,1}$ respectively. By Corollary 5.2.5, $\widehat{\triangle}\ E_{1,1}E_{2,1}E_{3,1}$ has bigger area than $\widehat{\triangle}\ E_1E_2E_3$.



The dashed triangle is the Lexell's triangle with base e_2. Two vertices and one edge remain unchanged $E_{1,1} = E_1$, $E_{3,1} = E_3$, and $e_2 = e_{2,1}$.

Figure 5.69

Remark 5.2.8. $e_2 = e_{2,1}$. $E_2 \neq E_{2,1}$ but $E_1 = E_{1,1}$ and $E_{3,1} = E_3$. The second subindex indicate the number of step we are at.

If $\widehat{\triangle}\ E_{1,1}E_{2,1}E_{3,1}$ is a regular triangle, then we are done. Otherwise, $\widehat{\triangle}\ E_{1,1}E_{2,1}E_{3,1}$ must have a shortest edge and a longest edge. Then we will repeat the above procedure and get an isosceles triangle of bigger area denoted by $\widehat{\triangle}\ E_{1,2}E_{2,2}E_{3,2}$. If $\widehat{\triangle}\ E_{1,2}E_{2,2}E_{3,2}$ happens to be regular, then we are done; otherwise, repeat to get $\widehat{\triangle}\ E_{1,3}E_{2,3}E_{3,3}$.

Remark 5.2.9. The recursive procedure described above terminates at step m if $\widehat{\triangle}\ E_{1,m}E_{2,m}E_{3,m}$ happens to be regular. However, it may never terminate. By *remark* 5.2.4, every step of the transformation to the next Lexell's triangle of larger area, the longest edge of the triangle will decrease, and the shortest edge of the triangle will increase. The reader may apply a rigourous "ϵ-δ" argument to show this but it is evident to see: as m goes to ∞ the area of $\overset{\frown}{\triangle}\, E_{1,m}E_{2,m}E_{3,m}$ *infinitesimally increases* to the area of a regular triangle of the same circumradius.

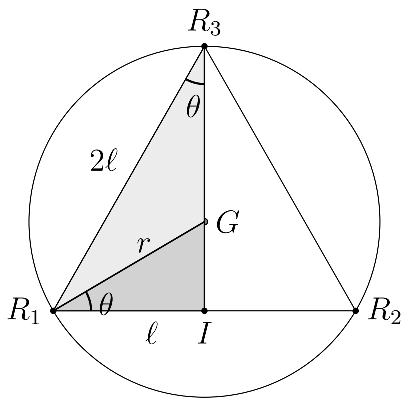
Therefore, with the same circumradius, a regular triangle has the largest area.

$\square$

Let $\overset{\frown}{\triangle}\, R_{max}$ be a regular triangle with the same circumcircle of $\overset{\frown}{\triangle}\, T_{max}$. The area of $\overset{\frown}{\triangle}\, R_{max}$ cannot be less than the average $\frac{1}{2N-4}$. Since $\overset{\frown}{\triangle}\, T_{max}$ and $\overset{\frown}{\triangle}\, R_{max}$ are of the same size, $\overset{\frown}{\triangle}\, \Omega > \overset{\frown}{\triangle}\, R_{max}$. This proves the following result:

Corollary 5.2.10. *The circumradius of the largest triangle in $\mathscr{M}_N$ is bigger than or equal to a regular triangle of size $\frac{1}{3(2N-4)}$.*

The circumradius of a regular triangle with area $\frac{1}{3(2N-4)}$ can be computed. In Figure 5.70 is a regular triangle with area $\frac{1}{2N-4}$. G is the circumcenter and I is the mid-point of an edge.



I is the middle point of edge $\overset{\frown}{R_1R_2}$. $\overset{\frown}{R_3I}$ is perpendicular to $\overset{\frown}{R_1R_2}$. In the Traditional Right triangle $\overset{\frown}{\triangle}\ R_1GI$, $\measuredangle GIR_1 = \pi/2$, $\measuredangle R_1GI = \pi/3$ and $\measuredangle GR_1I = \frac{1}{6(2N-4)} + \pi/6$.

Figure 5.70

Since

$$area(\overset{\frown}{\triangle}\ R_1R_2R_3) = 3\measuredangle R_3R_1R_2 - \pi = \frac{1}{2N-4}, \tag{5.15}$$

$$\measuredangle R_3R_1R_2 = 2\measuredangle GR_1I = \frac{1}{3(2N-4)} + \pi/3. \tag{5.16}$$

$$\measuredangle GR_1I = \frac{1}{6(2N-4)} + \pi/6. \tag{5.17}$$

$\overset{\frown}{\triangle}\ R_1GI$ is a *Traditional Right triangle* with $\measuredangle GIR_1 = \pi/2$. For computation simplicity, let $\theta = \frac{1}{6(2N-4)} + \pi/6$ and $|\ \overset{\frown}{R_1R_2}\ | = 2\ell$ and $r = |R_1G|$ (*as indicated in Figure* 5.70).

In triangle $\overset{\frown}{\triangle}\ R_1IR_3$, the *Spherical Rules of Sines* (*Theorem* 2.4),

$$\frac{\sin(2\ell)}{\sin(\pi/2)} = \frac{\sin\ell}{\sin\theta}, \tag{5.18}$$

which simplifies to

$$\cos\ell = \frac{1}{2}(\sin\theta)^{-1}. \tag{5.19}$$

In triangle $\overset{\frown}{\triangle} R_1IG$, applying the *Spherical Rules of Sines* (*Theorem* 2.4) gives:

$$\frac{\sin r}{\sin(\pi/2)} = \frac{\sin \ell}{\sin(\pi/3)}, \tag{5.20}$$

which can be rewritten as

$$1 - \cos^2 r = \frac{4}{3}(1 - \cos^2 \ell). \tag{5.21}$$

Plugging equation 5.19 into equation 5.21, we eventually get

$$\cos^2 r = \frac{1}{3}\frac{1}{\sin^2 \theta} - \frac{1}{3} = \frac{1}{3}\cot^2 \theta \tag{5.22}$$

or

$$\cos r = \frac{\sqrt{3}}{3}\cot \theta. \tag{5.23}$$

So we know that the distance between $\mathscr{M}_N$ and its next deep hole, i.e the separation of $\mathscr{M}_{N+1}$ $\delta(\mathscr{M}_{N+1})$, is bigger than r.

Corollary 5.2.11. (Memphie's Lower Bound of $\delta(\mathscr{M}_N)$)

$$\delta(\mathscr{M}_{N+1}) \geq \arccos(\frac{\sqrt{3}}{3}\cot(\frac{\pi}{6} + \frac{1}{12N - 24})) \ , \tag{5.24}$$

or equivalently

$$\delta(\mathscr{M}_N) \geq \arccos(\frac{\sqrt{3}}{3}\cot(\frac{\pi}{6} + \frac{1}{12N - 36})) \ . \tag{5.25}$$

Remark 5.2.12. *The convergence rate of the lower bounds is about* $0.62\sqrt{N}$ *for large* N.

5.3 Spherical Cap Discrepancy of $\mathscr{M}$

Counting is one of the hardest topics in mathematics. Now given a well-spaced spherical sequence of length N and an arbitrary spherical cap or hight t,, we would like to estimate how many points of the sequence is inside the cap. In this particular case, the sequence of length N we are looking at is the finite truncation of the infinite sequence $\mathscr{M}$. With the upper and lower bounds for $\delta(\mathscr{M}_N)$ obtained in the previous section, we know that if the height of a spherical cap is between

$$\frac{\sqrt{3}}{3}\cot(\frac{\pi}{6}+\frac{1}{12N-36}) \tag{5.1}$$

and

$$\frac{1}{2}\cot^2(\frac{N\pi}{6N-12})-\frac{1}{2} \tag{5.2}$$

then, for sure, this cap contains a point of $\mathscr{M}_N$. However, if height of this cap happens to be taller than

$$\frac{\sqrt{3}}{3}\cot(\frac{\pi}{6}+\frac{1}{12N-24}), \tag{5.3}$$

then it is possible that it contains no point in $\mathscr{M}_N$. We have the following bounds.

Corollary 5.3.1. *The spherical cap discrepancy of $\mathscr{M}_N$ is bounded below by*

$$\frac{1}{2}-\frac{\sqrt{3}}{6}\cot(\frac{\pi}{6}+\frac{1}{12N-24}). \tag{5.4}$$

5.3.1 Some Thoughts on Further Investigation

As the number of points grows to ∞, the area of each triangle will be getting very small. Meanwhile, the lengths of edges of all the triangles should eventually be "evened out". We will end our journey of spherical cap discrepancy with some ideas on how this investigation can be carried further.

Using the same notation 17 introduced in Chapter 3 *Memphis' Triangulation*, there are two types of edges: *cross edge* and *radius edge*.

Conjecture 1. (Memphis' Edge Conjecture)

The radio of the number of cross edge versus the number of radius edges goes to 1/2 as the number of points goes to infinity.

Every point in the triangulation is connected to other points by edges.

Conjecture 2. (Memphis' Angle Conjecture)

As the number of points grows to ∞*, the number of edges connected to "most" points is equal to 6.*

Conjecture 3. (Memphis' Area Conjecture)

For "most" triangles in $T \in \mathscr{M}_N$ $\lim_{N\to\infty} area(T)(2N-4) = 1$.

Of course, we need to make the description of these conjectures more precise. What does "most points" or "most triangles" even mean? Making the statements more accurate will be part of the next journey. As most of this field is still largely unexplored, there is lots of room left for creative inventions of new tools and new techniques. More adventures are awaiting for Memphis.

Bibliography

[1] *Random Number Generation and Quasi-Monte Carlo Methods*, 1992, Harald Niederreiter.

[2] *The distribution of points in a cube and the approximate evaluation of integrals*, Zh. Vychisl. Mat. i Mat. Fiz., 7 (1967), pp. 784-802.

[3] *Point sets and sequences with small discrepancy*, Monatsh. Math., 104 (1987), pp. 273-337.

[4] M. Hall, *Combinatorial Theory*, 2nd ed., John Wiley, New York, 1986.

[5] J. Dénes and A. D. Keedwell, *Latin Squares and Their Applications*, Academic Press, New York, 1974

[6] H. J. Ryser, *Combinatorial Mathematics*, Mathematical Association of America, Washington, DC, 1963.

[7] R. Lidl, H. Niederreiter, *Finite Fields*, Addision Wesley, Reading, MA, 1983.

[8] *The Right Right Triangle on the Sphere*, William Dickinson and Mohammad Salmassi, The College Mathematics Journal, Vol. 39, No. 1 (Jan., 2008), pp. 24-33

[9] *Least Action Principle Of Crystal Formation Of Dense Packing Type And Kepler's Conjecture* (2001).

[10] *Least Action Principle Of Crystal Formation Of Dense Packing Type And Kepler's Conjecture* (2001), pp. 66-67.

[11] *On the Densest Packing of Spherical Caps*, L.Fejes Toth, The American Mathematical Monthly, Vol. 56, No. 5 (May, 1949), pp. 330-331.

[12] *Sur la sphère vide. A la mémoire de Georges Voronoï*, Boris Delaunay, Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na, (1934), issue 6, pages 793–800

[13] *Irregularities of Distribution* (1987), Jośeph Beck, Cambridge University Press

[14] *Irregularities of distribution. VII.* Acta Arith. 21 (1972), Wolfgang M. Schmidt, page 45–50.

[15] *On irregularities of distribution.* Mathematika 1 (1954), Klaus F. Roth, page 73–79.

[16] *Sums of distances between points on a sphere. II. Proc. Amer. Math. Soc.*, K. B. Stolarsky. ,41:575–582, 1973.

[17] *Point sets on the sphere S2 with small spherical cap discrepancy*, C.Aistleitner, J.S.Brauchart, J.Dick, Discrete Comput. Geom. 48 (2012), no. 4, 990–1024.

[18] *Quasi–Monte Carlo rules for numerical integration over the unit sphere* S^2, J.S.Brauchart, J.Dick, Numer. Math. 121 (2012), no. 3, 473–502.