Chapter 6: The Elekes-Sharir-Guth-Katz Framework

Adam Sheffer

April 23, 2015

> *"By the way, in case of something unexpected happens to me (car accident, plane crash, a brick on the top of my skull) I definitely ask you to publish anything we have, at your will." / György Elekes, in an email to Micha Sharir, a few years before he passed away.*

Guth and Katz's seminal work on the number of distinct distances [4] was based on several novel ideas. One can regard their proof of as consisting of four main tools:

(i) A reduction from the distinct distances problem to a problem about line intersections in $\mathbb{R}^3$. This part is usually referred to as the *Elekes-Sharir framework*. We will refer to it as the *Elekes-Sharir-Guth-Katz framework*, for reasons discussed below.

(ii) The introduction of *polynomial partitioning*, as discussed in Chapter 3.

(iii) Applying 19th century analytic geometry tools that are related to ruled surfaces, such as *flecnode polynomials* and the Cayley-Salmon theorem concerning these polynomials [8].

(iv) A polynomial interpolation technique. We saw an example of this technique in Chapter 5.

The goal of this chapter is to introduce the Elekes-Sharir-Guth-Katz framework.

1 Distinct distances between points on two lines

Before presenting the Elekes-Sharir-Guth-Katz framework, we begin with a different distinct distances problem, which can be easily reduced to an incidence problem.

In a *bipartite* distinct distances problem we have two point sets $\mathcal{P}_1$ and $\mathcal{P}_2$, and we are only interested in the number $D(\mathcal{P}_1, \mathcal{P}_2)$ of distinct distances between pairs from $\mathcal{P}_1 \times \mathcal{P}_2$. That is,

$$D(\mathcal{P}_1, \mathcal{P}_2) = \Big|\{|pq| \ : \ p \in \mathcal{P}_1, q \in \mathcal{P}_2\}\Big|$$

(where $|uv|$ denotes the (Euclidean) length of the straight segment uv).

We consider a planar bipartite problem where $\mathcal{P}_1$ is a set of m points that lie on a line ℓ_1 and $\mathcal{P}_2$ is a set of n points that lie on a line ℓ_2. Without loss of generality, we assume that $n \geq m$. When the two lines are either parallel or orthogonal, the points can be arranged so that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$. Such constructions are illustrated in Figure 1.

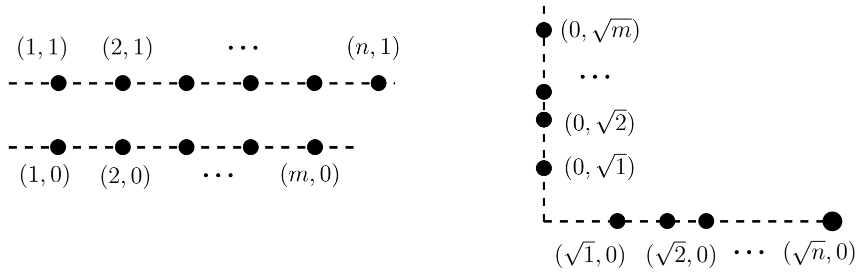

Figure 1: When the lines are either parallel or orthogonal, the points can be arranged so that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$.

On the other hand, when the two lines are neither parallel nor orthogonal, the current best construction for minimizing the number of distances yields $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta\left(n^2/\sqrt{\log n}\right)$ (by Elekes [2]). In this construction, we take ℓ_1 to be the x-axis and ℓ_2 to be the line $Z(y-x)$. We set

$$\mathcal{P}_1 = \{(i,0) : 1 \leq i \leq n\} \quad \text{and} \quad \mathcal{P}_2 = \{(i,i) : 1 \leq i \leq m\}.$$

To calculate $D(\mathcal{P}_1, \mathcal{P}_2)$ we recall a theorem that was already stated in Chapter 1.

Theorem 1.1. (Landau and Ramanujan [1, 6]) *The number of positive integers smaller than n that are the sum of two squares is $\Theta(n/\sqrt{\log n})$.*

In our case, every distance between a point of $\mathcal{P}_1$ and a point of $\mathcal{P}_2$ is of the form $\sqrt{d_x^2 + d_y^2}$, where both d_x and d_y are integers between zero and n. By Theorem 1.1, we immediately obtain $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta\left(n^2/\sqrt{\log n}\right)$.

We now derive the current best lower bound for this problem, derived by Sharir, Sheffer, and Solymosi [9]. Notice that there remains a huge gap between the current best upper and lower bounds.

Theorem 1.2. *Let $\mathcal{P}_1$ be a set of n points on a line ℓ_1 and let $\mathcal{P}_2$ be a set of m points on a line ℓ_2, both in $\mathbb{R}^2$, such that ℓ_1 and ℓ_2 are neither parallel nor orthogonal. Then*

$$D(\mathcal{P}_1, \mathcal{P}_2) = \Omega\left(\min\left\{n^{2/3}m^{2/3}, n^2, m^2\right\}\right).$$

Proof. We rotate the plane so that ℓ_1 become the x-axis and translate the plane so that $\ell_1 \cap \ell_2$ is the origin. Let s denote the slope of ℓ_2 after the rotation. We set $D = D(\mathcal{P}_1, \mathcal{P}_2)$ and denote the D distinct distances in $\mathcal{P}_1 \times \mathcal{P}_2$ as $\delta_1, \ldots, \delta_D$. Let Q be the set of quadruples (a,p,b,q) where $a, b \in \mathcal{P}_1$ and $p, q \in \mathcal{P}_2$, such that $|ap| = |bq| > 0$ and $ap \neq bq$ (the two segments are allowed to share at most one endpoint). An example is depicted in Figure 2.

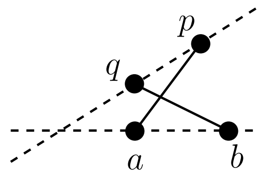

Figure 2: In a quadruple $(a,p,b,q) \in Q$, we have $|ap| = |bq| > 0$.

The quadruples are ordered, so that (a,p,b,q) and (b,q,a,p) are considered as two distinct elements of Q. We prove the theorem by double counting $|Q|$.

Let $E_i = \{(a,p) \in \mathcal{P}_1 \times \mathcal{P}_2 \mid |ap| = \delta_i\}$, for $i = 1,\ldots,D$. Notice that $\sum_{i=1}^{D} |E_i| = mn$. We have

$$|Q| = 2\sum_{i=1}^{D} \binom{|E_i|}{2} \geq \sum_{i=1}^{D} (|E_i| - 1)^2 .$$

By the Cauchy-Schwarz inequality,

$$\left(\sum_{i=1}^{D}(|E_i| - 1)\right)^2 \leq \left(\sum_{i=1}^{D} (|E_i| - 1)^2\right)\left(\sum_{i=1}^{D} 1\right) = D\sum_{i=1}^{D} (|E_i| - 1)^2 .$$

By combining the above, we get

$$|Q| \geq \sum_{i=1}^{D} (|E_i| - 1)^2 \geq \frac{1}{D}\left(\sum_{i=1}^{D}(|E_i| - 1)\right)^2 = \frac{(mn-D)^2}{D}. \tag{1}$$

Consider a quadruple (a,p,b,q) where $a,b \in \mathcal{P}_1$ and $p,q \in \mathcal{P}_2$. Write $a = (a_x,0), b = (b_x,0), p = (p_x, sp_x)$, and $q = (q_x, sq_x)$. Recall that this quadruple is in Q if and only if $|ap| = |bq|$, or

$$(a_x - p_x)^2 + s^2p_x^2 = (b_x - q_x)^2 + s^2q_x^2. \tag{2}$$

We apply a reduction to a parametric plane. For every pair of points $p,q \in \mathcal{P}_2$, we define a corresponding point $v_{pq} = (p_x,q_x) \in \mathbb{R}^2$. We denote the set of these $\Theta(m^2)$ points as $\mathcal{P}'$. For every pair of points $a,b \in \mathcal{P}_1$, we define a corresponding hyperbola γ_{ab} given by the equation

$$a_x^2 - b_x^2 - 2a_xx + 2b_xy + x^2(1+s^2) - y^2(1+s^2) = 0.$$

We denote the set of these $\Theta(n^2)$ hyperbolas as $\mathcal{H}$. Notice that (2) is satisfied if and only if the point v_{pq} is incident to the hyperbola γ_{ab}. That is, to obtain an upper bound for $|Q|$, it suffices to obtain an upper bound for $I(\mathcal{P}',\mathcal{H})$.

It can be easily verified that, since $s \neq 0$, no hyperbola in $\mathcal{H}$ is degenerate (i.e., the product of two lines). Moreover, the same hyperbola of $\mathcal{H}$ cannot be obtained by two different pairs of points $a,b \in \mathcal{P}_1$.

By definition, two (non-degenerate) hyperbolas intersect in at most four points. Thus, there is no $K_{5,2}$ in the incidence graph of $\mathcal{P}' \times \mathcal{H}$. This is a rather bad restriction, but we can improve it by noting that the roles of ℓ_1 and ℓ_2 (and of $\mathcal{P}_1$ and $\mathcal{P}_2$) are fully symmetric, and can be interchanged; that is, we can regard pairs of distinct points of $\mathcal{P}_1$ as points in the plane, and pairs of distinct points of $\mathcal{P}_2$ as hyperbolas, defined in a fully symmetric manner. In particular, this symmetry implies that the incidence graph of $\mathcal{P}' \times \mathcal{H}$ contains no copy of $K_{2,5}$. We apply our point-curve incidence bound from Chapter 3 (Theorem 2.1) on $\mathcal{H} \times \mathcal{P}'$ with $s = 2$ and $t = 5$, to obtain

$$|Q| = O\left(|\mathcal{P}'|^{2/3}|\mathcal{H}|^{2/3} + |\mathcal{P}'| + |\mathcal{H}|\right) = O\left(m^{4/3}n^{4/3} + m^2 + n^2\right).$$

Combining this bound with the lower bound in (1) implies the assertion of the theorem. □

Recently, Pach and de Zeeuw [7] generalized the above technique from lines to general algebraic curves.

2 The Framework

Elekes and Sharir used to think about the distinct distances problem, and around the turn of the millennium Elekes communicated to Sharir the basics of a reduction from this problem to a problem of bounding the number of intersections in a set of helices in $\mathbb{R}^3$. Later on, Elekes also sent Sharir the quote from the top of this document.

After Elekes' death, Sharir worked on publishing these ideas. He simplified the reduction so that it resulted in a problem of bounding the number of intersections in a set of parabolas in $\mathbb{R}^3$, and used tools from Guth and Katz's joints paper [3] to obtain some initial results for this incidence problem. Publishing the reduction, thereby exposing it for the first time to the general community, proved to be a good idea, because hardly any time had passed before Guth and Katz managed to apply it to get their almost tight bound for the distinct distances problem. Guth and Katz further simplified the reduction so that it has now resulted in a problem concerning intersections between lines in $\mathbb{R}^3$.

Due to this turn of events, we refer to the reduction as the Elekes-Sharir-Guth-Katz framework (or *ESGK framework*, for short). We now describe this reduction.

Consider a set $\mathcal{P}$ of n points in the plane, and let x denote the number of distinct distances that are determined by pairs of points from $\mathcal{P}$. As in the previous section, the reduction revolves around the set

$$Q = \left\{(a, p, b, q) \in \mathcal{P}^4 : |ap| = |bq| > 0\right\}.$$

The quadruples in Q are ordered in the sense that (a, p, b, q), (b, q, a, p), (p, a, q, b), and the other possible permutations are all considered as distinct elements of Q. In a quadruple $(a, p, b, q) \in Q$, the segments ap and bq are allowed to share vertices, though we do not allow a quadruple where both $a = b$ and $p = q$ (the case where $a = q$ and $b = p$ is allowed). As in the previous section, the reduction is based on double counting $|Q|$, and we begin by deriving a lower bound. We denote the set of (nonzero) distinct distances that are determined by $\mathcal{P} \times \mathcal{P}$ as $\delta_1, \ldots, \delta_x$. Also, for $1 \le i \le x$, we set

$$E_i = \left\{(p, q) \in \mathcal{P}^2 : |pq| = \delta_i\right\}.$$

We consider (p,q) and (q,p) as two distinct pairs in E_i. Notice that $\sum_{i=1}^{x} |E_i| = n^2 - n$ since every ordered pair of distinct points of $\mathcal{P} \times \mathcal{P}$ is contained in a unique set E_i. By applying the Cauchy-Schwarz inequality, we have

$$|Q| = \sum_{i=1}^{x} 2\binom{|E_i|}{2} > \sum_{i=1}^{x} (|E_i| - 1)^2 \geq \frac{1}{x}\left(\sum_{i=1}^{x} |E_i| - 1\right)^2 = \frac{(n^2 - n - x)^2}{x}. \qquad (3)$$

It remains to derive an upper bound for $|Q|$. Specifically, if we manage to derive the bound $|Q| = O(n^3 \log n)$, combining it with (3) would immediately imply $x = \Omega(n/\log n)$.

A transformation of the plane is said to be a *rigid motion* if it preserves distances between points. The rigid motions of the plane are rotations, translations, reflections, and their combinations. A *proper rigid motion* is a rigid motion that also preserves orientation; that is, an ordered triple of points abc forms a left turn after applying the transformation if and only if it originally formed a left turn. See Figure 3 for an example.

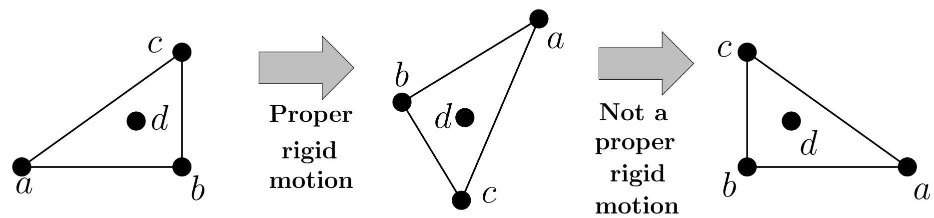

Figure 3: The second transformation is a rigid motion but not a proper one.

The proper rigid motions are exactly the transformations that are obtained by combining rotations and translations. In fact, every (planar) proper rigid motion is either a single rotation or a single translation. (That is, any combination of rotations and translations results in a single translation or in a single rotation; more details can be found in [5, Section 1.5] and in the exercises following it.)

For a pair of points $a, b \in \mathcal{P}$, consider the rotations that take a to b. The origin of such a rotation must be equidistant from a and b. In other words, the centers of these rotations must all be on the perpendicular bisector of the segment ab. Conversely, every point on the perpendicular bisector of ab is the origin of a rotation that takes a to b. See Figure 4(a) for an illustration.

Consider a quadruple $(a, p, b, q) \in Q$ and recall that by definition $|ap| = |bq|$. We can always apply a rotation that takes a to b and then rotate around the new position of a until p is taken to q. This translation followed by a rotation is a proper rigid motion taking ap to bq. To see that there is a unique proper rigid motion that takes ap to bq, we denote by ℓ_1 and ℓ_2 the perpendicular bisectors of the segments ab and pq, respectively. If ℓ_1 and ℓ_2 are parallel, then there is a unique translation taking ap to bq (and no rotations; e.g., see Figure 4(b)). Similarly, if ℓ_1 and ℓ_2 intersect, then there is a unique rotation taking ap to bq, and no translations. Specifically, the origin of this rotation is the point $\ell_1 \cap \ell_2$, and the angles of rotation from a to b and from p to q are equal because $|ap| = |bq|$ (see Figure 4(c)).

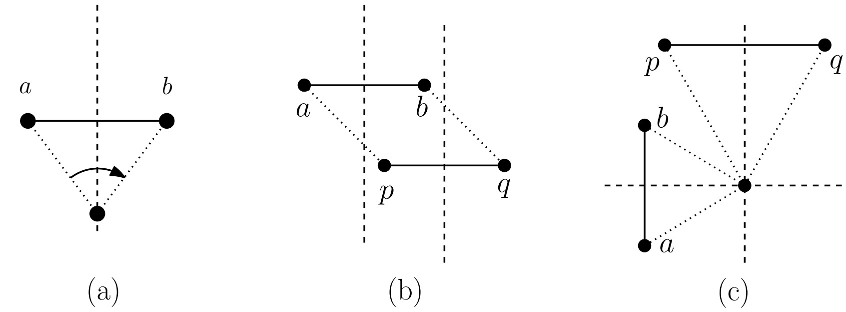

Figure 4: (a) The origin of the rotation must be on the perpendicular bisector. (b) If the perpendicular bisectors are parallel then there is a translation taking ap to bq. (c) If the perpendicular bisectors intersect, there is a rotation taking ap to bq, and its origin is the intersection point.

By the above, we have the following equivalent definition for Q: a quadruple (a, p, b, q) is in Q if and only if there exists a proper rigid motion τ that takes ap to bq. We say that the quadruple (a, p, b, q) corresponds to τ. That is, our goal is to show that the number of quadruples from $\mathcal{P}^4$ that correspond to a proper rigid motion is $O(n^3 \log n)$. As already noted, such a bound, combined with (3), would lead to the Guth-Katz bound $x = \Omega(n/\log n)$.

We first bound the number of quadruples in Q that correspond to a translation. Given the first three points of a quadruple $(a, p, b, ?)$, there is at most one point in $\mathcal{P}$ that can complete it to a quadruple that corresponds to a translation. Thus, $O(n^3)$ quadruples in Q correspond to a translation.

Bounding the number of quadruples in Q that correspond to a rotation is more difficult. A rotation can be parameterized using three parameters — two parameters for the origin and another one for the angle of rotation. Given a rotation with origin (o_x, o_y) and an angle of α, Guth and Katz [4] parameterized it as $(o_x, o_y, \cot(\alpha/2)) \in \mathbb{R}^3$. The advantage of this parametrization is that, given a pair of points $a, b \in \mathbb{R}^2$, the set of parametrizations of the rotations that take a to b is exactly the following line in $\mathbb{R}^3$:

$$\ell_{ab} = \left(\frac{a_x + b_x}{2}, \frac{a_y + b_y}{2}, 0 \right) + t \left(\frac{b_y - a_y}{2}, \frac{a_x - b_x}{2}, 1 \right), \quad \text{for } t \in \mathbb{R}. \tag{4}$$

The proof of this property is a rather standard calculation which is not relevant to the rest of our discussion, so we postpone it to Appendix A. Notice that the projection of ℓ_{ab} on the xy-plane is the perpendicular bisector of ab (since the projection contains the midpoint of a and b, and is orthogonal to the line incident to a and b). Specifically, ℓ_{ab} is obtained by "lifting" the perpendicular bisector of ab to a line in $\mathbb{R}^3$ whose slope in the z-direction is $2/|ab|$.

Consider a quadruple $(a, p, b, q) \in \mathcal{P}^4$ and the corresponding lines ℓ_{ab} and ℓ_{pq} in $\mathbb{R}^3$. If the intersection point $p = \ell_{ab} \cap \ell_{pq}$ exists, then p is the parametrization of a rotation taking both a to b and p to q. That is, the quadruple (a, p, b, q) corresponds to a rotation (and is thus in Q) if and only if ℓ_{ab} and ℓ_{pq} intersect. For an example, see Figure 5.

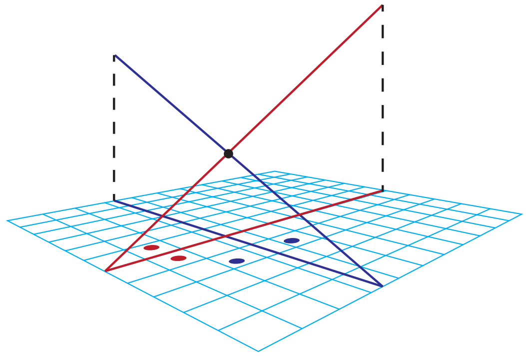

Figure 5: A quadruple of points in the plane, the two perpendicular bisectors, and their "lifting" to $\mathbb{R}^3$.

Pairs of points from $\mathcal{P}^2$ yield n^2 lines in the parametric space $\mathbb{R}^3$, and there is a bijection between quadruples of Q that correspond to rotations and pairs of intersecting lines. Thus, an upper bound of $O(n^3 \log n)$ on the number of pairs of intersecting lines would imply $|Q| = O(n^3 \log n)$, as required. There are sets of n^2 lines in $\mathbb{R}^3$ with a significantly larger number of intersecting pairs. Fortunately, as we will see in the following two chapters, the sets of lines that we deal with have additional properties that we can rely on.

References

[1] B. C. Berndt and R. A. Rankin, *Ramanujan: Letters and Commentary*, Amer. Math. Soc., Providence, RI, 1995.

[2] G. Elekes, A note on the number of distinct distances, *Period. Math. Hung.*, **38** (1999), 173–177.

[3] L. Guth, and N. H. Katz, Algebraic methods in discrete analogs of the Kakeya problem, Advances in Mathematics **225.5** (2010), 2828–2839.

[4] L. Guth and N. H. Katz, On the Erdős distinct distances problem in the plane, *Annals of Mathematics*, **181** (2015), 155–190.

[5] J. Stillwell, *Naive lie theory*, Springer, New York, London, 2008.

[6] E. Landau, Über die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindeszahl der zu ihrer additiven Zusammensetzung erforderlichen Quadrate, *Arch. Math. Phys.* **13** (1908), 305–312.

[7] J. Pach and F. de Zeeuw, Distinct distances on algebraic curves in the plane, arXiv:1308.0177.

[8] G. Salmon, *A Treatise on the Analytic Geometry of Three Dimensions*, Vol. 2, 5th edition, Hodges, Figgis and co. Ltd., Dublin, 1915.

[9] M. Sharir, A. Sheffer, and J. Solymosi, Distinct distances on two lines, *J. Combinat. Theory Ser. A*, **120** (2013), 1732–1736.

A Lines in the parametric space $\mathbb{R}^3$

In this appendix we consider the parametrization of planar rotations that is presented in Section 2. That is, a rotation with origin point $o \in \mathbb{R}^2$ and angle α is parameterized by the point $\left(o_x, o_y, \cot\frac{\alpha}{2}\right) \in \mathbb{R}^3$. Given a pair of points $a, b \in \mathbb{R}^2$, we explain why the set of parametrizations of the rotations that take a to b is a line in $\mathbb{R}^3$. Recall that the origin of such a rotation is required to be on the perpendicular bisector of the segment ab. This situation is depicted in Figure 6, where $c = ((a_x+b_x)/2, (a_y+b_y)/2)$ is the midpoint of the segment ab and $\delta = |ab| = \sqrt{(a_x-b_x)^2+(a_y-b_y)^2}$.

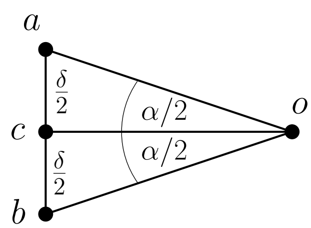

Figure 6: A rotation with origin o and angle α that takes a to b.

Notice that the slope of the perpendicular bisector of ab is $s = (a_x - b_x)/(b_y - a_y)$ and that it is incident to c. Since o is on the perpendicular bisector of ab, we have

$$o_y - \frac{a_y+b_y}{2} = s\left(o_x - \frac{a_x+b_x}{2}\right). \tag{5}$$

We assume that that $o_x \geq c_x$ and $o_y \geq c_y$, and set $d_x = o_x - c_x$ and $d_y = o_y - c_y$ (the other cases can be similarly handled). By (5), we have $d_y = sd_x$, which implies

$$|co| = \sqrt{d_x^2+d_y^2} = d_x\sqrt{1+s^2} = d_x\sqrt{\frac{(a_x-b_x)^2+(b_y-a_y)^2}{(b_y-a_y)^2}} = \frac{\delta d_x}{b_y-a_y}. \tag{6}$$

By looking at Figure 6, we notice that $|co| = \frac{\delta}{2}\cot\frac{\alpha}{2}$. Combining this with (6), we obtain

$$\frac{b_y-a_y}{2}\cdot\cot\frac{\alpha}{2} = d_x = o_x - c_x. \tag{7}$$

By (5) and (7), we have that $\left(o_x, o_y, \cot\frac{\alpha}{2}\right)$ is on the line

$$\ell_{ab} = \left(\frac{a_x+b_x}{2}, \frac{a_y+b_y}{2}, 0\right) + t\left(\frac{b_y-a_y}{2}, \frac{a_x-b_x}{2}, 1\right), \qquad \text{for } t \in \mathbb{R}.$$

Conversely, since we can choose o to be any point on the perpendicular bisector, any point on ℓ_{ab} is the parametrization of a rotation that takes a to b. Thus, ℓ_{ab} is exactly the set of parametrizations of the rotations that take a to b.