LAWS OF TRIGONOMETRY ON $SU(3)$

HELMER ASLAKSEN

ABSTRACT. The orbit space of congruence classes of triangles in $SU(3)$ has dimension 8. Each corner is given by a pair of tangent vectors $(X, Y)$, and we consider the 8 functions $\text{tr} X^2$, $i \text{tr} X^3$, $\text{tr} Y^2$, $i \text{tr} Y^3$, $\text{tr} XY$, $i \text{tr} X^2 Y$, $i \text{tr} XY^2$ and $\text{tr} X^2 Y^2$ which are invariant under the full isometry group of $SU(3)$. We show that these 8 corner invariants determine the isometry class of the triangle. We give relations (laws of trigonometry) between the invariants at the different corners, enabling us to determine the invariants at the remaining corners, including the values of the remaining side and angles, if we know one set of corner invariants. The invariants that only depend on one tangent vector we will call side invariants, while those that depend on two tangent vectors will be called angular invariants. For each triangle we then have 6 side invariants and 12 angular invariants. Hence we need $18 - 8 = 10$ laws of trigonometry. If we restrict to $SU(2)$, we get the cosine laws of spherical trigonometry. The basic tool for deriving these laws is a formula expressing $\text{tr}(\exp X \exp Y)$ in terms of the corner invariants.

1. INTRODUCTION

Given a triangle in $\mathbb{R}^2$, we associate to each corner the s.a.s. data (side, angle, side) at that corner. This determines the congruence class of the triangle, and knowing the s.a.s. data at one corner, we can use the laws of trigonometry to determine the s.a.s. data at the remaining corners. Another way of looking at this is to say that to each triangle we associate 6 invariants, the 3 sides and the 3 angles. The s.a.s. congruence axiom tells us that the space of congruence classes only depends on 3 invariants. Hence, there must be $6 - 3 = 3$ relations between these 6 invariants given by, for instance, the 3 cosine laws or one cosine law and 2 sine laws.

There are classical generalizations of this to $S^2$ and $H^2$ (spherical and hyperbolic trigonometry). The generalization to the other simply connected constant curvature spaces, i.e., $\mathbb{R}^n$, $S^n$ and $H^n$, is immediate, since given any triangle $\Delta$ in $\mathbb{R}^n$ (resp. $S^n$, $H^n$), we can find a totally geodesic submanifold $N$ with $\Delta \subset N$ and $N$ isometric to $\mathbb{R}^2$ (resp. $S^2$, $H^2$).
The classical geometries ($\mathbb{R}^n$, $S^n$, and $H^n$) can be characterized by being 2-point homogeneous and satisfying the s.a.s. (side, angle, side) congruence condition. If we try to do trigonometry on the other 2-point homogeneous spaces, we will therefore need more invariants than the size of the angles and the length of the sides.

The trigonometry of $\mathbb{P}^n(C)$ was studied by Wilhelm Balschke and Hans Terheggen [BT] in 1939, after partial results by J. L. Coolidge [Co] in 1921. (Since any triangle in $\mathbb{P}^n(k)$ lies in a totally geodesic submanifold isometric to $\mathbb{P}^2(k)$, it is sufficient to study the trigonometry of $\mathbb{P}^2(k)$.) A different approach was taken by P. A. Sirokov [Sir], who obtained more complete results. Sirokov died in 1944, but the results were found among his papers, and published in 1957 by A. P. Sirokov, A. Z. Petrov and B. A. Rozenfeld. Using H. C. Wang's classification of the 2-point homogeneous spaces from 1952 [Wa], B. A. Rozenfeld [Ro] generalized Sirokov's results to all the compact 2-point homogeneous spaces, i.e., quaternionic projective space and the Cayley projective plane. In 1986, Wu-Yi Hsiang [Hs] independently developed the trigonometry of all the 2-point homogeneous spaces, using a unified geometric approach which applies equally well to the noncompact case. In 1987, Ulrich Brehm [Br] modified Hsiang's results, using an approach similar to [BT].

Hsiang's approach is to associate to each corner a set of four invariants, the length of the two adjacent sides and two angular invariants. He shows that this "s.a.s. data" determines the congruence class of the triangle, and since we get a total of nine invariants (three sides and six angular invariants) we need $9 - 4 = 5$ laws of trigonometry relating the invariants at the different corners.

The 2-point homogeneous spaces are precisely the symmetric spaces of rank 1. The rank of a symmetric space can be thought of as the minimal number of invariants for the isometry class of a pair of points. It is therefore natural to start looking at symmetric spaces of rank $> 1$. When we went from the constant curvature case to the other 2-point homogeneous spaces, we had to include more angular invariants than the size of the angle. When going to rank $> 1$, we will also need more invariants for each side. Given a symmetric space of rank $n$, we would like to find a minimal complete set of invariants corresponding to each corner, and a minimal set of relations between the invariants at the different corners.

Among the most fundamental symmetric spaces are those obtained by putting a bi-invariant metric on a compact simple Lie group of rank $n$. In this paper, we will study the trigonometry of the rank 2 group $SU(3)$.

This paper is a shortened version of my Ph.D. thesis, written under the supervision of Professor Wu-Yi Hsiang. I am grateful to him and Hsueh-Ling Huynh for many helpful suggestions.

2. The invariants of $SU(3)$

The Lie algebra of $SU(3)$ is denoted by $su(3)$, and the (standard) maximal torus by $T$. We will use the identification $T, SU(3) = (I_g)_* su(3)$, and the
bi-invariant metric induced by \( \langle X, Y \rangle_\epsilon = -\frac{1}{2} \text{tr}(XY) \). (The reason for the \( \frac{1}{2} \) factor is to make \( SU(2) \subset SU(3) \) isometric to \( S^3(1) \).) Since the metric is bi-invariant, the geodesics are left- and right-translations of one-parameter subgroups. Since we have chosen to identify \( T_gSU(3) = (I_g)_* \text{su}(3) \), we will write geodesics in the form

\[
c(t) = g \exp tX, \quad \text{where } g \in SU(3), \quad X \in \text{su}(3).
\]

Consider a triangle \( \Delta = (A, B, C) \). Since we are only interested in the isometry class of \( \Delta \), we can assume that \( C = I_3 \).

\[
\begin{align*}
B &= \exp Z = \exp X \exp Y \\
C &= I_3 \\
A &= \exp X \\
X, \ Y, \ Z &\in \text{su}(3)
\end{align*}
\]

**Figure 1**

It is well known that \( I^0(SU(3)) \), the identity component of the isometry group, is given by \( I^0(SU(3)) = SU(3) \times SU(3) \), where \((g_1, g_2)(h) = g_1 h g_2^{-1}\). For the full isometry group, we must include the outer automorphism \( c(g) = \overline{g} \) and the geodesic symmetry at the identity \( s(g) = g^{-1} \). This gives [Lo, vol. 2, p. 152]

\[
I(SU(3)) = (SU(3) \times SU(3)) \ltimes (\{I, c\} \times \{I, s\})
\]

The group of isometries fixing \( I_3 \) is \( I(SU(3), I_3) = SU(3) \ltimes (\{I, c\} \times \{I, s\}) \), with connected component \( I^0(SU(3), I_3) = SU(3) \). The induced action of \( I(SU(3), I_3) \) on \( \text{su}(3) \oplus \text{su}(3) \) is given by

\[
2 \text{Ad}(g)(X, Z) = (gXg^{-1}, gZg^{-1}), \\
c_*(X, Z) = (\overline{X}, \overline{Z}), \\
s_*(X, Z) = (-X, -Z)
\]

We now observe that the set of congruence classes of triangles in \( SU(3) \) is the same as the orbit space of the induced action of \( I(SU(3), I_3) = SU(3) \ltimes (\{I, c\} \times \{I, s\}) \) on \( \text{su}(3) \oplus \text{su}(3) \).

For simplicity, we will first consider the 2Ad-action of \( I^0(SU(3), I_3) = SU(3) \) on \( \text{su}(3) \oplus \text{su}(3) \).

**Lemma 1.** The dimension of the orbit space \( 2\text{su}(3)/2 \text{Ad}_{SU(3)} \) is 8.
Proof. Let \((X, Z) \in \mathfrak{su}(3) \oplus \mathfrak{su}(3)\). We can assume that \(X\) is diagonal.

\[
X = \begin{pmatrix}
ia & 0 & 0 \\
0 & ib & 0 \\
0 & 0 & -(a+b)
\end{pmatrix}, \quad Z = \begin{pmatrix}
ir & \alpha & \beta \\
-\bar{\alpha} & is & \gamma \\
-\bar{\beta} & -\bar{\gamma} & -i(r+s)
\end{pmatrix}.
\]

If \(X\) is regular, \(Z_{SU(3)}(X) = T\) and

\[
\text{Ad} \left( \begin{pmatrix}
u & 0 & 0 \\
0 & v & 0 \\
0 & 0 & \bar{uv}
\end{pmatrix} \right) (Z) = \begin{pmatrix}
ir & u\bar{\alpha} & u^2\gamma \bar{\beta} \\
-\bar{\alpha} & is & uu^2\gamma \\
-\bar{\beta} & -\bar{\gamma} & -i(r+s)
\end{pmatrix}.
\]

It follows that the principal isotropy subgroup of \(2 \text{Ad}\) is \(Z_3 = Z(SU(3))\), so the orbits are 8-dimensional and the dimension of the principal part of the orbit space is \(16 - 8 = 8\).

Since \(SU(3)\) is compact, it is possible to find a set of \(SU(3)\)-invariants on \(\mathfrak{su}(3) \oplus \mathfrak{su}(3)\) which will separate orbits. We will call such a set a **complete set of invariants**. Since the dimension of the orbit space is 8, we would like to find a minimal, complete set consisting of 8 invariants. Let \(\mathcal{R}[2\mathfrak{su}(3)]^{SU(3)}\) denote the polynomial \(SU(3)\)-invariants on \(\mathfrak{su}(3) \oplus \mathfrak{su}(3)\). A set \(\{f_1, \ldots, f_k\} \subseteq \mathcal{R}[2\mathfrak{su}(3)]^{SU(3)}\) is called a **basis** if every \(f \in \mathcal{R}[2\mathfrak{su}(3)]^{SU(3)}\) can be expressed polynomially in \(\{f_1, \ldots, f_k\}\). We will call \(\{f_1, \ldots, f_k\}\) a **minimal basis** if it is a basis and none of the \(f_i\)'s can be expressed polynomially in the other \(f_j\)'s. There might still be polynomial relations (syzygies) between the invariants in a minimal basis, we are just assuming that we cannot solve for any of them as a polynomial in the others.

**Lemma 2.**

\[
\mathcal{R}[2\mathfrak{su}(3)]^{SU(3)} = \{ \text{tr} f(X, Z) | f \text{ is a monomial in } X, Z \in \mathfrak{su}(3) \}.
\]

**Proof.** From [Sib] we know that the unitary invariants of \(X\) and \(Z\) are the affine invariants of \(\{X, X^*, Z, Z^*\}\) and that the affine invariants are given by traces of monomials in \(\{x, x^*, z, z^*\}\). But in our case, \(X^* = -X\), and \(Z^* = -Z\), so the lemma follows.

In order to reduce expressions of the form \(\text{tr} f(X, Z)\), we will use the polarized Cayley-Hamilton Theorem, first used by Dubnov [Du1] in 1935 and later by Rivlin [Ri] in 1955. (For the \(n \times n\) case see [Le].) The Cayley-Hamilton Theorem for \(M(3, \mathbb{C})\) can be written as

\[
M^3 - M^2 \text{tr } M + \frac{1}{2} M [(\text{tr } M)^2 - \text{tr } M^2] - I [\frac{1}{3} \text{ tr } M^3 - \frac{1}{2} \text{ tr } M^2 \text{ tr } M + \frac{1}{6} (\text{tr } M)^3] = 0.
\]

Polarizing this, we get the multilinear version.
Lemma 3 (The polarized Cayley-Hamilton Theorem).

\[ MNP + MPN + NMP + RPM + PNM + NPM - M[\text{tr} NP - \text{tr} N \text{tr} P] \\
- N[\text{tr} MP - \text{tr} M \text{tr} P] - P[\text{tr} MN - \text{tr} M \text{tr} N] - (NP + PN) \text{tr} M \\
- (PM + PM) \text{tr} N - (MN + NM) \text{tr} P \\
- I[\text{tr} M \text{tr} N \text{tr} P - \text{tr} M \text{tr} NP - \text{tr} N \text{tr} MP - \text{tr} P \text{tr} MN \\
+ \text{tr} MNP + \text{tr} PNM] = 0. \]

Using the polarized Cayley-Hamilton Theorem, Dubnov [Du2] showed (see also Rivlin [Ri]):

Lemma 4. A basis for traces of polynomials in two arbitrary \( 3 \times 3 \) matrices is given by the 12 invariants

\[ \text{tr} X, \text{tr} X^2, \text{tr} X^3, \text{tr} Z, \text{tr} Z^2, \text{tr} Z^3, \text{tr} XZ, \]
\[ \text{tr} X^2Z, \text{tr} XZ^2, \text{tr} X^2Z^2, \text{tr} XYZ, \text{tr} XZ^2 \]
\[ \text{tr} XZ^2X^2, \text{tr} XZ^2Z^2. \]

Furthermore, \( \text{tr} XZ^2X^2 \) and \( \text{tr} XZ^2Z^2 \) are the roots of a quadratic equation with coefficients expressible in terms of the 10 other invariants.

Remark. This quadratic syzygy is very complicated. Dubnov only indicates how to get it, and does not give an explicit formula.

For skew-Hermitian, \( \text{tr} = 0 \) matrices, this can be reduced further.

Lemma 5.

\[ \text{tr} XZ^2X^2 = \frac{\text{tr} XZ^2Z^2}{2}, \]
\[ \text{Re} \text{tr} XZ^2Z^2 = \frac{1}{2} [\text{tr} XZ \text{tr} X^2Z + \text{tr} X^2Z \text{tr} XZ^2 - \frac{1}{3} \text{tr} X^3 \text{tr} Z^3]. \]

Proof. The first statement is immediate since \( X \) and \( Z \) are skew-Hermitian. For the second statement, we use the polarized Cayley-Hamilton Theorem. If we set \( M = N = X \) and \( P = Z \), multiply on the right by \( XZ^2 \) and take the trace, we get

\[ 2 \text{tr} XZ^2X^2 + 2 \text{tr} XZ^2X^2 + 2 \text{tr} X^3Z^3 \\
- 2 \text{tr} X^2Z^2 \text{tr} XZ - \text{tr} X^2 \text{tr} XZ^3 - 2 \text{tr} X^2Z \text{tr} XZ^2 = 0. \]

Similarly

\[ \text{tr} XZ^3 = \frac{1}{2} \text{tr} XZ \text{tr} Z^2, \]
\[ \text{tr} X^3Z^3 = \frac{1}{2} \text{tr} X^2 \text{tr} XZ \text{tr} Z^2 + \frac{1}{4} \text{tr} X^2 \text{tr} XZ \text{tr} Z^3. \]

Substituting (4) and (3) into (2), the formula in Lemma 5 follows. \( \square \)

Observe that \( \text{tr} X^2, \text{tr} Z^2, \text{tr} XZ \) and \( \text{tr} X^2Z^2 \) are real-valued, \( \text{tr} X^3, \text{tr} Z^3, \text{tr} X^2Z \) and \( \text{tr} XZ^2 \) are purely imaginary and \( \text{tr} XZ^2X^2 \) is complex.

Combining Lemmas 4 and 5, we get the following proposition.
Proposition 1.

\[ \text{tr} X^2, i \text{tr} X^3, \text{tr} Z^2, i \text{tr} Z^3, \text{tr} XZ, i \text{tr} X^2Z, \text{tr} XZ^2, \text{tr} X^2Z^2 \]

and \( \text{Im} \text{tr} XZX^2Z^2 \) form a basis for \( \mathbb{R}[\text{su}(3)]^{SU(3)} \).

Let us denote the 8 first invariants in Proposition 1 by \( \{I_1, \ldots, I_8\} \). Since the dimension of the orbit space is 8, we are interested in the extent to which \( \{I_1, \ldots, I_8\} \) determine the orbits. Dubnov's quadratic syzygy suggests that they in general will only determine pairs of orbits. We will give a direct proof of this.

**Proposition 2.** \( \{I_1, \ldots, I_8\} \) will in general only determine pairs of orbits. Specifically, consider the two pairs \((X, Z_1)\) and \((X, Z_2)\) with

\[
X = \begin{pmatrix} ia & 0 & 0 \\ 0 & ib & 0 \\ 0 & 0 & -(a+b) \end{pmatrix},
\]

\[
Z_1 = \begin{pmatrix} ic & r & p+iq \\ -r & id & s \\ -p+iq & -s & -i(c+d) \end{pmatrix},
\]

\[
Z_2 = \begin{pmatrix} ic & r & -p+iq \\ -r & id & s \\ p+iq & -s & i(c+d) \end{pmatrix},
\]

where \( X \) is regular, \( r, s > 0 \), and \( p \neq 0 \). \( \{I_1, \ldots, I_8\} \) will agree on the two orbits containing \((X, Z_1)\) and \((X, Z_2)\) respectively. Otherwise \( \{I_1, \ldots, I_8\} \) will separate orbits.

**Proof:** Consider the pair \((X, Z)\). Since we are only interested in the orbit of \((X, Z)\), we can assume \( X \) is diagonal,

\[
X = \begin{pmatrix} ia & 0 & 0 \\ 0 & ib & 0 \\ 0 & 0 & -i(a+b) \end{pmatrix}, \quad a \geq b \geq -(a+b).
\]

Now

\[
\text{Ad} \left( \begin{pmatrix} u & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & \overline{uv} \end{pmatrix} \right) = \begin{pmatrix} ic & \alpha & \beta \\ -\overline{\alpha} & id & \gamma \\ -\overline{\beta} -\overline{\gamma} & -(c+d) \end{pmatrix}
\]

(5)

so by choosing \( u \) and \( v \) properly, we can bring \( Z \) into the form

\[
Z = \begin{pmatrix} ic & r & p+iq \\ -r & id & s \\ -p+iq & -s & -i(c+d) \end{pmatrix}, \quad r, s \geq 0.
\]

Every orbit contains a pair in the above form, and using (5), it is easy to see that if \( X \) is regular and \( r, s > 0 \), then the orbit contains only one such pair.

Writing out the formulas for \( \{I_1, \ldots, I_8\} \) at \((X, Z)\) we find that knowing the values of \( \{I_1, \ldots, I_8\} \) determines \( a, b, c, d, r, s, |p+iq| \) and \( q \). But knowing \( |p+iq| \) and \( q \) will only determine \( p \) up to sign.
The proof that the 8 invariants determine the orbit when \( r \) or \( s = 0 \) or \( X \) is singular is similar. \( \Box \)

**Proposition 3.** The nine invariants \( \{I_1, \ldots, I_9, \text{Im} \, \text{tr} \, XZ^2Z^2\} \) for a minimal complete basis.

**Proof.** In view of Propositions 1 and 2, all we need to do is to show that \( \text{Im} \, \text{tr} \, XZ^2Z^2 \) separates the two orbits described in Proposition 2, but

\[
\text{tr} \, XZ_1^2Z_1^2 = -|p + iq|^2 [(b^3 + 3ab^2 + 3a^2b + 2a^3)c + (b^3 + 2ab^2 + 2a^2b + a^3)d] \\
- (p + q)rs(a-b)(2ia^2 + 5iab + 2ib^2) + \text{terms not involving } p \text{ or } q,
\]

so replacing \( p \) by \(-p\) will change the value. \( \Box \)

3. The invariants of the full isometry group

Let us now return to our geometric problem. We need to consider the orbits of the full isometry group. It turns out that four \( SU(3) \) orbits are joined together to form one \( I(SU(3), I_3) \) orbit, and in particular the two orbits described in Proposition 2 will be identified.

Unfortunately, the \( I^0(SU(3), I_3) \) invariants described in the previous section are not absolute \( I(SU(3), I_3) \) invariants. To remedy this, we define the character

\[
\chi: I(SU(3)) = (SU(3) \times SU(3)) \ltimes \langle \{I, c\} \times \{I, s\} \rangle \to \mathbb{R},
\]

\[
\chi(g_1, g_2, c^j, s^j) = (-1)^{l+j}.
\]

It is easy to see that \( \text{tr} \, X^2, \text{tr} \, Z^2, \text{tr} \, XZ, \text{tr} \, X^2Z^2 \) will be \( I(SU(3), I_3) \) invariants, while \( i \text{tr} \, X^3, i \text{tr} \, Z^3, i \text{tr} \, X^2Z, i \text{tr} \, XZ^2 \) will be \( I(SU(3), I_3) \) semi-invariants, i.e.,

\[
P(g(X, Z)) = \chi(g)P(X, Z).
\]

We can now state our first main result.

**Theorem 1.** Let \( (X, Z), (A, B) \in su(3) \oplus su(3) \). Then \( (X, Z) \) and \( (A, B) \) will lie on the same orbit under \( I(SU(3), I_3) \leftrightarrow \)

\[
(1) \quad (\text{tr} \, X^2, \text{tr} \, Z^2, \text{tr} \, XZ, \text{tr} \, X^2Z^2) = (\text{tr} \, A^2, \text{tr} \, B^2, \text{tr} \, AB, \text{tr} \, A^2B^2)
\]

and

\[
(2+) \quad (i \text{tr} \, X^3, i \text{tr} \, Z^3, i \text{tr} \, X^2Z, i \text{tr} \, XZ^2) = (i \text{tr} \, A^3, i \text{tr} \, B^3, i \text{tr} \, A^2B, i \text{tr} \, AB^2)
\]

or

\[
(2-) \quad (i \text{tr} \, X^3, i \text{tr} \, Z^3, i \text{tr} \, X^2Z, i \text{tr} \, XZ^2) = -(i \text{tr} \, A^3, i \text{tr} \, B^3, i \text{tr} \, A^2B, i \text{tr} \, AB^2).
\]

**Proof.** If \( (X, Z) \) and \( (A, B) \) lie on the same orbit, the invariants satisfy (1) and (2+) or (2−). Assume now that \( (X, Z) \) and \( (A, B) \) satisfy (1) and (2+).
Because of Proposition 1, we only need to show that the two \( I^0(SU(3), I_3) \) orbits containing \((X, Z_1)\) and \((X, Z_2)\), respectively, where

\[
X = \begin{pmatrix}
ia & 0 & 0 \\
0 & ib & 0 \\
0 & 0 & -(a + b)
\end{pmatrix},
\]

\[
Z_1 = \begin{pmatrix}
ic & r & p + iq \\
-r & id & s \\
-p + iq & -s & -i(c + d)
\end{pmatrix},
\]

\[
Z_2 = \begin{pmatrix}
ic & r & -p + iq \\
-r & id & s \\
p + iq & -s & -i(c + d)
\end{pmatrix}
\]

with \( X \) regular, \( r, s > 0 \) and \( p \neq 0 \), will become one \( I(SU(3), I_3) \) orbit. This follows from

\[
\text{Ad} \left( \begin{pmatrix}
-1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & -1
\end{pmatrix} \right) \circ c_* \circ s_*(X) = X,
\]

\[
\text{Ad} \left( \begin{pmatrix}
-1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & -1
\end{pmatrix} \right) \circ c_* \circ s_*(Z_1) = Z_2.
\]

If \((X, Z)\) and \((A, B)\) satisfy (1) and (2−), then \((X, Z)\) and \((s_*(A), s_*(B))\) satisfy (1) and (2+). This completes the proof. \(\Box\)

This can be made more explicit as follows:

**Corollary 1.** The four pairs

1. \[
\begin{pmatrix}
ia & 0 & 0 \\
0 & ib & 0 \\
0 & 0 & -(a + b)
\end{pmatrix},
\]

2. \[
\begin{pmatrix}
ia & 0 & 0 \\
0 & ib & 0 \\
0 & 0 & -(a + b)
\end{pmatrix},
\]

3. \[
\begin{pmatrix}
i(a + b) & 0 & 0 \\
0 & -ib & 0 \\
0 & 0 & -ia
\end{pmatrix},
\]

4. \[
\begin{pmatrix}
i(a + b) & 0 & 0 \\
0 & -ib & 0 \\
0 & 0 & -ia
\end{pmatrix},
\]

where \( X \) is regular, \( r, s > 0 \), and \( p \neq 0 \), lie on different \( I^0(SU(3), I_3) \) orbits, but on the same \( I(SU(3), I_3) \) orbit.

**Proof.** In the proof of Theorem 1, we showed that 1 and 2 lie on the same \( I(SU(3), I_3) \) orbit, but 3 is obtained from 1 by applying

\[
\text{Ad} \left( \begin{pmatrix}
0 & 0 & 1 \\
0 & -1 & 0 \\
1 & 0 & 0
\end{pmatrix} \right) \circ s_*,
\]

so the four orbits are joined into one. \(\Box\)
4. A Trace Formula

In §3, we associated to the corner at $C = I_3$ a set of 8 invariants $\{I_1, \ldots, I_8\}$ expressed in terms of the two tangent vectors $X$ and $Z$ which make up the corner at $C = I_3$ (cf. Figure 1). We will call these the corner invariants at $C$. Similarly, the corner at $A$ is made up of the two tangent vectors $(l_A)_*(-X)$ and $(l_A)_*(-Y)$, and the corner at $B$ of $(l_B)_*(-Y)$ and $(l_B)_*(-Z)$, so after left-translating the triangle by $A^{-1}$ and $B^{-1}$ respectively, we get two new sets of corner invariants.

We will call the invariants that only depend on one tangent vector side invariants, while the ones that involve two tangent vectors will be called angular invariants. This gives four angular invariants for each corner and two side invariants for each side, for a total of 18 invariants. Since the triangle is determined by the eight corner invariants at any corner, we would like to find 10 relations (laws of trigonometry) between the invariants at different corners.

The tangent vectors $X$, $Y$ and $Z$ are related by

$$\exp X \exp Y = \exp Z.$$  \hspace{1cm} (6)

We need to express this relation in terms of the corner invariants described in §3. First recall the following simple fact. Assume that $X \in M(3, \mathbb{C})$ is diagonalizable. Then there is an $A \in \text{Gl}(3, \mathbb{C})$ such that

$$X = A \begin{pmatrix} \lambda_1(X) & 0 & 0 \\ 0 & \lambda_2(X) & 0 \\ 0 & 0 & \lambda_3(X) \end{pmatrix} A^{-1} = ADA^{-1},$$

where $\lambda_i(X), \ i = 1, 2, 3$, denotes the eigenvalues of $X$. Then

$$\text{tr}(\exp tX) = \text{tr}(\exp t(ADA^{-1})) = \text{tr}(A(\exp tD)A^{-1})$$

$$= \text{tr}(\exp tD) = \sum_{i=1}^{3} e^{\lambda_i(X)t}.$$  \hspace{1cm} (7)

Taking the trace of (6) and applying (7) to the right-hand side, we get

$$\text{tr}(\exp X \exp Y) = \sum_{i=1}^{3} e^{\lambda_i(Z)}. $$

We would therefore like to have a formula expressing $\text{tr}(\exp X \exp Y)$ in terms of the corner invariants of $-X$ and $Y$. We will first give a generalization of (7), expressing $\text{tr}(U \exp X)$ in terms of traces of products of $U$ and $X$ (Theorem 2). We will then deduce a formula for $\text{tr}(\exp X \exp Y)$ (Theorem 3) from which the laws of trigonometry will follow easily.

Theorem 2. 1. Let $X \in M(n, \mathbb{C})$ be diagonalizable with distinct eigenvalues $\lambda_i$. Let $U \in M(n, \mathbb{C})$. Then

$$\text{tr}(U \exp tX) = \sum_{i,j=1}^{n} a_{ij} \text{tr}(X^{j-1}U)e^{\lambda_i t}.$$  \hspace{1cm} (8)
where
\[ a_{ij} = (-1)^j s_{n-j}(\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_n)/ \prod_{k \in \{1, \ldots, n\} - \{i\}} (\lambda_i - \lambda_k) \]
and \( s_k \) is the sum over all products of \( k \) distinct \( \lambda \)'s.

2. In particular, if \( S \in SU(3) \) has distinct eigenvalues, this can be written as
\[
\text{tr}(U \exp tX) = \sum_{i=1}^{3} \frac{2 \text{tr}X^2U + 2\lambda_i(X) \text{tr}XU + [2\lambda_i(X)^2 - \text{tr}X^2] \text{tr}U}{6\lambda_i(X)^2 - \text{tr}X^2} e^{\lambda_i(X)t}.
\]

3. If \( X \in su(3) \) has a double eigenvalue \( \lambda(X) \), we get
\[
\text{tr}(U \exp tX) = \left( \frac{1}{3\lambda(X)} \text{tr}XU + \frac{2}{3} \text{tr}U \right) e^{\lambda(X)t} + \left( \frac{-1}{3\lambda(X)} \text{tr}XU + \frac{1}{3} \text{tr}U \right) e^{-2\lambda(X)t}.
\]

**Proof.** For the proof of (8), we will for simplicity assume that \( n = 3 \). Writing \( X = ADA^{-1} \), we get
\[
\text{tr}((\exp tX)U) = \text{tr}(A(\exp tD)A^{-1}U) = \text{tr}((\exp tD)A^{-1}UA).
\]
Since
\[
\exp tD = \begin{pmatrix}
  e^{\lambda_1(X)t} & 0 & 0 \\
  0 & e^{\lambda_2(X)t} & 0 \\
  0 & 0 & e^{\lambda_3(X)t}
\end{pmatrix},
\]
we get
\[
\text{tr}(\exp tXU) = \sum_{i=1}^{3} \text{tr}(E_i A^{-1}UA) e^{\lambda_i(X)t},
\]
where
\[
E_1 = \begin{pmatrix}
  1 & 0 & 0 \\
  0 & 0 & 0 \\
  0 & 0 & 0
\end{pmatrix}, \quad E_2 = \begin{pmatrix}
  0 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 0
\end{pmatrix}, \quad E_3 = \begin{pmatrix}
  0 & 0 & 0 \\
  0 & 0 & 1 \\
  0 & 1 & 0
\end{pmatrix}.
\]
We can identify the set of diagonal \( 3 \times 3 \) matrices with \( \mathbb{C}^3 \). We can then think of
\[
\phi(M) = \text{tr}(MA^{-1}UA) = \text{tr}(AMA^{-1}U)
\]
as a linear functional on \( \mathbb{C}^3 \). Since we do not have any simple expression for \( A \), we cannot immediately calculate \( \phi(E_i) \). We know, however, that
\[
\phi(I) = \text{tr}(IA^{-1}UA) = \text{tr}(AA^{-1}U) = \text{tr}U,
\]
\[
\phi(D) = \text{tr}(DA^{-1}UA) = \text{tr}(ADA^{-1}U) = \text{tr}XU,
\]
\[
\phi(D^2) = \text{tr}(D^2A^{-1}UA) = \text{tr}(AD^2A^{-1}U) = \text{tr}X^2U.
\]
If $X$ is regular, we get a new basis for $\mathbb{C}^3$ by setting

$$ F_1 = I, \quad F_2 = D, \quad F_3 = D^2. $$

Now

$$ F_i = V E_i, \quad \text{where} \quad V = \begin{pmatrix}
1 & \lambda_1(X) & \lambda_1(X)^2 \\
1 & \lambda_2(X) & \lambda_2(X)^2 \\
1 & \lambda_3(X) & \lambda_3(X)^2
\end{pmatrix}, $$

so we get

$$
(12) \quad \begin{pmatrix}
\text{tr}(E_1 A^{-1} U A) \\
\text{tr}(E_2 A^{-1} U A) \\
\text{tr}(E_3 A^{-1} U A)
\end{pmatrix} = 
\begin{pmatrix}
\phi(E_1) \\
\phi(E_2) \\
\phi(E_3)
\end{pmatrix} = V^{-t} \begin{pmatrix}
\phi(F_1) \\
\phi(F_2) \\
\phi(F_3)
\end{pmatrix} = V^{-1} \begin{pmatrix}
\text{tr} U \\
\text{tr} X U \\
\text{tr} X^2 U
\end{pmatrix},
$$

where

$$ V^{-t} = \begin{pmatrix}
\frac{\lambda_2 \lambda_3}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} & \frac{-\lambda_2 - \lambda_3}{\lambda_1 \lambda_3} & \frac{1}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} \\
\frac{\lambda_1 \lambda_3}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)} & \frac{\lambda_1 - \lambda_3}{\lambda_2 \lambda_3} & \frac{1}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)} \\
\frac{\lambda_1 \lambda_2}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)} & \frac{-\lambda_1 - \lambda_2}{\lambda_3 \lambda_2} & \frac{1}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)}
\end{pmatrix}. $$

Combining (11) and (12) we get

$$
\text{tr}(U \exp tX) = (e^{\lambda_1(X)t} e^{\lambda_2(X)t} e^{\lambda_3(X)t}) V^{-1} \begin{pmatrix}
\text{tr} U \\
\text{tr} X U \\
\text{tr} X^2 U
\end{pmatrix}. $$

Multiplying this out gives (8) in the case $n = 3$. For the general case, it is sufficient to observe that $(a_{ij})$ is the inverse transpose of the Vandermonde matrix. To prove (9) we simply observe that

$$ 6\lambda_i(X)^2 - \text{tr} X^2 = \begin{cases}
(\lambda(1(X) - \lambda_2(X))(\lambda_1(X) - \lambda_3(X)), & i = 1, \\
(\lambda_2(X) - \lambda_1(X))(\lambda_2(X) - \lambda_3(X)), & i = 2, \\
(\lambda_3(X) - \lambda_1(X))(\lambda_3(X) - \lambda_2(X)), & i = 3,
\end{cases} $$

$$ 2\lambda_i(X)^2 - \text{tr} X^2 = \begin{cases}
\lambda_1(X)\lambda_3(X), & i = 1, \\
\lambda_1(X)\lambda_2(X), & i = 2, \\
\lambda_1(X)\lambda_2(X), & i = 3,
\end{cases} $$

$$ \lambda_i(X) = \begin{cases}
-(\lambda_2(X) + \lambda_3(X)), & i = 1, \\
-(\lambda_1(X) + \lambda_3(X)), & i = 2, \\
-(\lambda_1(X) + \lambda_2(X)), & i = 3.
\end{cases} $$

To prove (10), we can either modify the proof of (8) or take a limit in (9). This completes the proof of Theorem 2. $\square$

Remark. The above proof is due to Hsueh-Ling Huynh. We would also like to outline our original proof, using the polarized Cayley-Hamilton Theorem. For simplicity we will only consider (9), i.e., $X \in su(3)$ and regular.
Consider
\[ g(X, U, t) = \text{tr}(U \exp tX), \quad X \in \text{su}(3), \ U \text{ arbitrary}. \]

Then
\[ g'(t) = \text{tr}(XU \exp tX), \quad g''(t) = \text{tr}(X^2U \exp tX). \]

Setting \( P = \exp tX \) and \( M = N = X \) in the polarized Cayley-Hamilton Theorem, we get
\[
\begin{align*}
6X^2 \exp tX &= -\exp tX \text{tr}X^2 - 2X \text{tr}(X \exp tX) \\
&= -2X^2 \text{tr}(\exp tX) - I[2 \text{tr}(X^2 \exp tX) - \text{tr}(\exp tX) \text{tr}X^2] = 0.
\end{align*}
\]

We now multiply by \( U \) and take the trace
\[
6g''(t) - \text{tr}X^2g(t) = (2 \text{tr}X^2U - \text{tr}X^2 \text{tr}U) \text{tr}(\exp tX)
+ 2 \text{tr}XU \text{tr}(X \exp tX) + 2 \text{tr}U \text{tr}(X^2 \exp tX).
\]

Using the fact that
\[
\text{tr}(X^k \exp tX) = \frac{d^k}{dt^k} \text{tr} (\exp tX) = \sum_{i=1}^{3} \lambda_i(X)^k e^{\lambda_i(X)t},
\]
we get
\[
6g''(t) - \text{tr}X^2g(t) = \sum_{i=1}^{3} [2 \text{tr}X^2U - \text{tr}X^2 \text{tr}U + 2\lambda_i \text{tr}XU + 2\lambda_i^2 \text{tr}U]e^{\lambda_i t},
\]
\[
g(X, U, 0) = \text{tr}U, \quad g'(X, U, 0) = \text{tr}XU.
\]

Solving this we get the formula in the second part of Theorem 2.

If we set \( U = \exp sY \) and use Theorem 2, we get the following formula.

**Theorem 3.** Let \( X, Y \in \text{su}(3) \). If both \( X \) and \( Y \) are regular
\[
\begin{align*}
\text{tr}(\exp tX \exp sY) &= \sum_{i,j=1}^{3} \frac{e^{\lambda_i(X)t} e^{\lambda_j(Y)s}}{6\lambda_i(X)^2 - \text{tr}X^2 \lambda_j(Y)^2 - \text{tr}Y^2}
\times [4 \text{tr}X^2Y^2 - \text{tr}X^2 \text{tr}Y^2 + 4\lambda_i(X) \text{tr}XY^2 + 4\lambda_j(Y) \text{tr}X^2Y
+ 4\lambda_i(X)\lambda_j(Y) \text{tr}XY - 2\lambda_i(X)^2 \text{tr}Y^2
- 2\lambda_j(Y)^2 \text{tr}X^2 + 12\lambda_i(X)^2\lambda_j(Y)^2].
\end{align*}
\]

5. **Deriving the laws of trigonometry**

We are now ready to derive the laws of trigonometry from combining Theorems 2 and 3 with the relation
\[
\exp Z = \exp X \exp Y.
\]
From (13) we get

\begin{align*}
(14) \quad \text{tr}(\exp X \exp Y) &= \text{tr}(\exp Z), \\
(15) \quad \text{tr}(X \exp X \exp Y) &= \text{tr}(X \exp Z), \\
(16) \quad \text{tr}(X^2 \exp X \exp Y) &= \text{tr}(X^2 \exp Z), \\
(17) \quad \text{tr}(Y \exp X \exp Y) &= \text{tr}(Y \exp Z), \\
(18) \quad \text{tr}(Y^2 \exp X \exp Y) &= \text{tr}(Y^2 \exp Z).
\end{align*}

We want to express these equations in terms of the corner invariants. On the right-hand sides we can apply Theorem 2 directly and Theorem 3 applies to the left-hand side of (14). For the right-hand sides of (15)--(18) we could either make repeated use of Theorem 2 to derive a formula for \(\text{tr}(V \exp X \exp Y)\), or we can simply observe that in (15)

\begin{equation}
(19) \quad \text{tr}(X \exp X \exp Y) = \left. \frac{\partial}{\partial t} \right|_{t=1} \text{tr}(\exp tX \exp Y),
\end{equation}

and similarly for (16)--(18).

Using Theorems 2 and 3 and (19), we can now rewrite (14)--(18) in terms of the corner invariants. For convenience, we will use the following notation:

\begin{align*}
\omega_i(X) &= \text{Im} \lambda_i(X) = -i \lambda_i(X), \\
\text{tr}(U \exp tX) &= \sum_{i=1}^3 A_i(X, U) e^{\lambda_i(X)t}, \\
A_i(X, U) &= \frac{2 \text{tr} X^2 U + 2 \lambda_i(X) \text{tr} XU + [2 \lambda_i(X)^2 - \text{tr} X^2] \text{tr} U}{6 \lambda_i(X)^2 - \text{tr} X^2}, \\
\text{tr}(\exp tX \exp sY) &= \sum_{i,j=1}^3 \frac{e^{\lambda_i(X)t}}{6 \lambda_i(X)^2 - \text{tr} X^2} \frac{e^{\lambda_j(Y)s}}{6 \lambda_j(Y)^2 - \text{tr} Y^2} B_{i,j}(X, Y), \\
B_{i,j}(X, Y) &= 4 \text{tr} X^2 Y^2 - \text{tr} X^2 \text{tr} Y^2 + 4 \lambda_i(X) \lambda_j(Y) \text{tr} XY \\
&\quad + 4 \lambda_i(Y) \lambda_j(Y) \text{tr} X^2 + 4 \lambda_i(X) \lambda_j(Y) \text{tr} Y^2 \\
&\quad - 2 \lambda_i(X)^2 \text{tr} Y^2 - 2 \lambda_j(Y)^2 \text{tr} X^2 + 12 \lambda_i(X)^2 \lambda_j(Y)^2.
\end{align*}

In these formulas both \(X\) and \(Y\) are assumed to be regular. But in the singular case the formulas can be modified by taking a limit as in Theorem 2, so we will only state the laws of trigonometry in the regular case. Since \(\text{tr}\) is complex valued on \(SU(3)\), we take the real and imaginary parts of (14)--(18), giving us the following 10 equations.
Theorem 4.

1. \[ \sum_{i,j=1}^{3} \cos(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y) = \sum_{k=1}^{3} \cos \omega_k(Z), \]

2. \[ \sum_{i,j=1}^{3} \sin(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y) = \sum_{k=1}^{3} \sin \omega_k(Z), \]

3. \[ \sum_{i,j=1}^{3} \cos(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_i(X) = \sum_{k=1}^{3} \cos \omega_k(Z)A_k(Z, X), \]

4. \[ \sum_{i,j=1}^{3} \sin(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_i(X) = \sum_{k=1}^{3} \sin \omega_k(Z)A_k(Z, X), \]

5. \[ \sum_{i,j=1}^{3} \cos(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_i(X)^2 = \sum_{k=1}^{3} \cos \omega_k(Z)A_k(Z, X^2), \]

6. \[ \sum_{i,j=1}^{3} \sin(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_i(X)^2 = \sum_{k=1}^{3} \sin \omega_k(Z)A_k(Z, X^2), \]

7. \[ \sum_{i,j=1}^{3} \cos(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_j(Y) = \sum_{k=1}^{3} \cos \omega_k(Z)A_k(Z, Y), \]

8. \[ \sum_{i,j=1}^{3} \sin(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_j(Y) = \sum_{k=1}^{3} \sin \omega_k(Z)A_k(Z, Y), \]

9. \[ \sum_{i,j=1}^{3} \cos(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_j(Y)^2 = \sum_{k=1}^{3} \cos \omega_k(Z)A_k(Z, Y^2), \]

10. \[ \sum_{i,j=1}^{3} \sin(\omega_i(X) + \omega_j(Y))B_{i,j}(X, Y)\lambda_j(Y)^2 = \sum_{k=1}^{3} \sin \omega_k(Z)A_k(Z, Y^2). \]

1 and 2 relate the side invariants of \(Z\) to the corner invariants of \((-X, Y)\).

3, 4, 5 and 6 relate the angular invariants of \((X, Z)\) to the corner invariants of \((-X, Y)\).

7, 8, 9 and 10 relate the angular invariants of \((-Y, -Z)\) to the corner invariants of \((-X, Y)\).

In this way we are able to determine all the 18 fundamental invariants once we know one set of eight corner invariants.

Remark. Observe that the semi-invariants (including the eigenvalues) occur in such a way that the sign ambiguities cancel.

6. Restricting to \(SU(2)\)

Assume that \(A, B \in SU(3)\) have a common eigenvector. We can then think of them as lying on a copy of \(SU(2)\) in \(SU(3)\). But \(SU(2)\) is isometric to \(S^3(1)\), so the 10 laws of trigonometry should reduce to the usual laws of spherical trigonometry.
To verify this, first observe that if \( X, Z \in \mathfrak{su}(2) \), then \( Y \) is also in \( \mathfrak{su}(2) \), since \( SU(2) \) is totally geodesic. Hence we can assume

\[
X = \begin{pmatrix} iu & \alpha & 0 \\ -\bar{\alpha} & -iu & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} iv & \beta & 0 \\ -\bar{\beta} & -iv & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} iw & \gamma & 0 \\ -\bar{\gamma} & -iw & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

Setting \( \angle(X, Z) = C \), \( \angle(-X, Y) = A \) and \( \angle(-Y, -Z) = B \), we get

\[
\lambda_1(X) = i|X|, \quad \lambda_2(X) = -i|X|, \quad \lambda_3(X) = 0, \quad \text{tr} \, XZ = -2|X||Z| \cos C \\
\text{tr} \, X^2 = -2|X|^2, \quad \text{tr} \, X^2 Z = \text{tr} \, XZ^2 = 0, \quad \text{tr} \, X^2 Z^2 = 2|X|^2|Z|^2.
\]

When substituting into 1–10, equations 2, 3, 6, 7 and 10 cancel, while 1, 5 and 9 reduce to

\[
\cos |X| \cos |Y| + \cos A \sin |X| \sin |Y| = \cos |Z|,
\]

and 4 and 8 reduce to

\[
\sin |X| \cos |Y| - \cos |X| \sin |Y| \cos A = \cos C \sin |Z|, \\
\sin |Y| \cos |X| - \cos |Y| \sin |X| \cos A = \cos B \sin |Z|.
\]

We see that (20) is the usual cosine law in spherical trigonometry, while (21) and (22) follow easily from combining (20) with the cosine laws

\[
\cos |X| \cos |Z| + \cos C \sin |X| \sin |Z| = \cos |Y|, \\
\cos |Y| \cos |Z| + \cos B \sin |Y| \sin |Z| = \cos |X|.
\]

References


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720

Current address: Department of Mathematics, National University of Singapore, Singapore 0511, Republic of Singapore