SL(2)

This might, perhaps, become lecture notes for a serious group theory course someday. I tried them in 2010, and added a few things concerning Heisenberg groups (in 2013) and concerning non-compact groups (in 2020). The illustrations have been taken from the classics.

GROUPS

At least apocryphically, Wigner once claimed that if it cannot be understood in terms of two by two matrices, it is not worth the trouble.[1] Presumably Wigner was thinking of group theory here. He certainly knew about it.

The set of all two by two matrices of determinant one,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} , \qquad ad - bc = 1 , \tag{1}$$

forms a *group*. The inverse of the given matrix can be written down by inspection,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} . \tag{2}$$

It takes such a simple form because the determinant is assumed to equal unity.

Matrix groups consisting of matrices of determinant one are called special linear, but which particular group we obtain depends on the nature of the matrix elements:

a, b, c, d	Name of group
Complex	$SL(2, \mathbf{C})$
$d = \bar{a}, c = -\bar{b}$	$SU(2)$
$d = \bar{a}, c = \bar{b}$	$SU(1,1)$
Real	$SL(2, \mathbf{R})$
Integers	$SL(2, \mathbf{Z})$
Integers modulo N	$SL(2, \mathbf{Z}_N)$

Here S is for special, L is for linear, and U is for unitary.

The *order* $|G|$ of a group G is the number of its elements. If the order is continuously ∞ and if in addition some technical conditions are obeyed (they are for all matrix groups) the group is a *Lie group*. Examples include $SL(2, \mathbf{C}), SL(2, \mathbf{R})$, $SU(1,1)$, and $SU(2)$. The Lie group $SU(2)$ is known from quantum mechanics courses.

[1] If Wigner never said that, I can instead quote my teacher Arne Kihlberg.

Ex: Prove that $SU(1,1)$ is a group.

The order of $SL(2,\mathbf{Z})$ is countably ∞. This group is called the modular group. Clearly $SL(2,\mathbf{Z}) \in SL(2,\mathbf{R}) \in SL(2,\mathbf{C})$.

Ex: Prove that $SL(2,\mathbf{Z})$ is a group.

$\mathbf{Z}, \mathbf{R}, \mathbf{C}$ are *rings*—they consist of elements that can be added and multiplied, in such a way that they form a commutative (*abelian*) group under addition. $\mathbf{R}, \mathbf{C}$ are *fields* (Swedish *kropp*), meaning rings such that, if the identity element under addition is excluded, they also form abelian groups under multiplication. The integers were the original role model for rings, while the real (or rational) numbers were the original role model for the definition of fields. If you remember this it is easy to check if a given algebraic structure forms a ring, or a field.

Note that in the formula (2) for the inverse of a given matrix we need only the additive inverse of the entries. This is why we require that they belong to a ring, but not necessarily to a field.

Two integers are equal modulo N if they differ by a multiple of the integer N,

$$m = n \bmod N \quad \Leftrightarrow \quad m = n + kN \ . \tag{3}$$

Integers modulo N form a field if and only if $N = p$ is a prime number. This follows from a theorem in number theory: if m and N have largest common divisor d, then one can find integers r and s such that

$$rm + sN = d \ . \tag{4}$$

If $N = p$ then $d = 1$ for all non-zero m, and r is the multiplicative inverse of m modulo p. All finite fields are known (and there are some others). Note that this is not needed for $SL(2,\mathbf{Z}_N)$ to form a group.

Ex: Compute the order of $SL(2,\mathbf{Z}_p)$.

Matrices of determinant 1 preserve volume. This is often expressed by saying that they leave a certain tensor invariant. In this case it is the epsilon-tensor. The special linear group $SL(n)$ consists of N by N matrices such that

$$g^{b_1}{}_{a_1} g^{b_2}{}_{a_2} \dots g^{b_N}{}_{a_N} \epsilon_{b_1 b_2 \dots b_N} = \det g \epsilon_{a_1 a_2 \dots a_N} = \epsilon_{a_1 a_2 \dots a_N} \ . \tag{5}$$

Other groups are defined by the requirement that they leave some other tensor invariant. Of particular interest are tensors that define a non-degenerate quadratic form, that is expressions like

$$x^a g_{ab} y^a \ . \tag{6}$$

It is assumed that the matrix g_{ab} has an inverse g^{ab}, and it is convenient to bring it to a standard form (eg. $g_{ab} = \delta_{ab}$ if g_{ab} is symmetric).

Carrying on this theme we obtain a list of subgroups of the general linear group of invertible N by N matrices:

Invariant tensor	Property	Group
$\epsilon_{a_1 a_2 \dots a_N}$	anti-symmetric	special linear
ω_{ab}	anti-symmetric	symplectic
δ_{ab}	symmetric, pos. definite	orthogonal
η_{ab}	Minkowski metric	Lorentz

This defines the groups $Sp(N)$, $SO(N)$, and $SO(1, N-1)$ as subgroups of $SL(N)$. Note that the symplectic group $Sp(2) = SL(2, \mathbf{R})$. Isomorpisms between matrix groups typically happen only for low values of N.

Ex: Find the general form of matrices belonging to $SO(2) \in SL(2, \mathbf{R})$ and $SO(1,1) \in SL(2, \mathbf{R})$.

Ex: How do you define $SU(2)$ and $SU(1,1)$ from this point of view?

Why did we not consider the more general group $GL(2)$, defined such that $ad - bc \neq 0$? The answer is that this is not so interesting, since

$$g \in GL(2) \Rightarrow g = (\text{diagonal matrix}) \times (\text{matrix in } SL(2)) \ . \tag{7}$$

Because of this decomposition, once we understand SL we understand GL too. The diagonal matrices form the *centre* of GL, i.e. they commute with everything.

The group acts on itself in three ways, left action $g \to g_1 g$, right action $g \to g g_1^{-1}$, and conjugation $g \to g_1 g g_1^{-1}$. Note that the inverse appears on the right because

$$g \to g g_1^{-1} \to g g_1^{-1} g_2^{-1} = g(g_2 g_1)^{-1} \ . \tag{8}$$

The action of the group on itself defines a *representation* of the group. In general a representation is a set of transformations of a set of objects which respects the multiplication table of the group. All matrix groups have a *defining representation* as matrices acting on some vector space, and all groups have a *trivial* (but not faithful) representation as the number 1.

If we choose a fixed group element g_1, then the set of all elements that can be written on the form $g g_1 g^{-1}$ for some $g \in G$ is called a *conjugacy class*. Any group can be partitioned into conjugacy classes, and the elements in a given conjugacy class have in a sense very similar properties. Compare Euler's theorem, according to which the rotation group can be partitioned into conjugacy classes such that each element in a given conjugacy class can be described as a rotation through some fixed angle about some unspecified fixed axis in space. The unit element is always a conjugacy class by itself.

We would like to divide $SL(2, \mathbf{R})$ into conjugacy classes. To do this it is helpful to know that any matrix M_1 can be brought into a standard *Jordan form* by means of conjugation, $M_1 \to M M_1 M^{-1}$. The catch is that Jordan's theorem assumes that an arbitrary complex matrix M can be used, while we may want to restrict ourselves to real matrices. Anyway, given a matrix M_1, the first step is to solve the characteristic equation

$$\det(M_1 - \lambda) = 0 \ . \tag{9}$$

For n by n matrices this is an nth order polynomial, and it has n complex roots. If they are all different the matrix can be brought to diagonal form by means of conjugation (but the matrix M will typically not be a rotation matrix, or even a real matrix). If some eigenvalues are equal there may be entries equal to 1 just above the diagonal. Eg, if three eigenvalues are equal, the Jordan blocks

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} , \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} , \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} , \tag{10}$$

can appear, and cannot be transformed into each other. The information about the spectrum (the possible eigenvalues) can be summarised by the n independent quantities $\det M_1, \mathrm{Tr} M_1, \ldots, \mathrm{Tr} M_1^{n-1}$. Clearly their values cannot be changed by conjugation.

For 2 by 2 matrices the possible eigenvalues are determined by the determinant and the trace of the matrix. If the determinant equals one the eigenvalues must be $(\lambda_1, \lambda_2) = (re^{i\theta}, e^{-i\theta}/r)$. But for an $SL(2, \mathbf{R})$ matrix the trace must be real, so we are stuck with the two possibilites $(r, 1/r)$ or $(e^{i\theta}, e^{-i\theta})$—and evidently a real matrix with the second spectrum cannot be diagonalised using conjugation with matrices belonging to $SL(2, \mathbf{R})$. However, when the smoke clears, one finds the following representatives for the conjugacy classes of $SL(2, \mathbf{R})$:

$$-2 < \mathrm{Tr} g < 2 \ : \ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} , \quad 0 \leq \theta < 2\pi \tag{11}$$

$$2 < \mathrm{Tr} g \ : \ \begin{pmatrix} e^u & 0 \\ 0 & e^{-u} \end{pmatrix} , \quad 0 \leq u < \infty \tag{12}$$

$$\mathrm{Tr} g < -2 \ : \ -\begin{pmatrix} e^v & 0 \\ 0 & e^{-v} \end{pmatrix} , \quad 0 \leq v < \infty \tag{13}$$

$$\mathrm{Tr} g = 2 \ : \ \begin{pmatrix} 1 & \sigma \\ 0 & 1 \end{pmatrix} , \quad \sigma = -1, 0, 1 \tag{14}$$

$$\mathrm{Tr} g = -2 \ : \ \begin{pmatrix} -1 & \sigma \\ 0 & -1 \end{pmatrix} , \quad \sigma = -1, 0, 1 \tag{15}$$

So it is almost true that the conjugacy classes are labelled by the trace.

Ex: Work out the conjugacy classes of $SU(2)$ in the same way.

Ex: Use transformations in $SL(2, \mathbf{R})$ to diagonalise an arbitrary $SO(1,1)$ matrix.

Given a subgroup $H \in G$ one can divide the group G into *left cosets* of the form gH, where H is any element in the subgroup H and g is some fixed element of G. (Right cosets also exist.) Each group element belongs to exactly one left coset. If the group G has finite order it follows that the order of the subgroup must divide the order of the group (Lagrange's theorem). An *invariant* or *normal subgroup* is a subgroup $H \in G$ such that

$$gHg^{-1} = H \text{ for all } g \in G \tag{16}$$

(i.e. $ghg^{-1} \in H$ for all $h \in H$). A subgroup is normal if it consists of entire conjugacy classes of G.

Ex: If H is an invariant subgroup, prove that the set of all left cosets gH forms a group in itself.

The resulting group is denoted G/H, and it is somehow a simpler group than the one we started out with. If the group G has finite order $|G|$, it follows that $|G/H| = |G|/|H|$.

The centre of a group is always an invariant subgroup. A group that does not have any invariant subgroups at all is called a *simple group*. GL is not simple. Actually $SL(2)$ is not simple either, since it has $\pm\mathbf{1}$ as a non-trivial centre. The groups

$$SL(2, \mathbf{C})/Z_2 \ , \quad SL(2, \mathbf{R})/Z_2 \ , \quad SU(2)/Z_2 \ , \tag{17}$$

etc, are simple groups. Here Z_2 denotes the cyclic group represented by the two elements ± 1. From quantum mechanics courses you know that $SU(2)/Z_2$ is really the same group as $SO(3)$. Frequently one sees the notation

$$PSL(2, \mathbf{C}) = SL(2, \mathbf{C})/Z_2 \ , \quad PSL(2, \mathbf{R}) = SL(2, \mathbf{R})/Z_2 \ . \tag{18}$$

The P stands for 'projective', meaning that a factor has been ignored somehow.

The projective groups are not only simple, they are often the physical relevant ones. This is so in quantum mechanics. The group that transforms between the physical states of a qubit is not $U(2)$, it is $PSU(2)$. The point is that, although one starts with a description of physical states using a two complex dimensional vector space on which the matrices act, it is decided that vectors differing by an overall phase represent the same physical states. So the group that acts on the states is $U(2)/I(2) = SU(2)/Z_2 = PSU(2)$, where $I(2)$ is the center of $U(2)$.

<u>Literature</u>: For rather more information than you need, consult I. N. Herstein, *Topics in Algebra*, 2nd ed. 1975.

The simple group $PSL(2, \mathbf{C})$ can be faithfully represented by the *Möbius transformations*

$$z \to z' = \frac{az + b}{cz + d} \ , \quad ad - bc \neq 0 \ . \tag{19}$$

This is a one-to-one mapping of the complex plane including the point at ∞ to itself.

Ex: Check that this is a representation of $SL(2, \mathbf{C})$.

To get the complete picture it is convenient to recall that the extended complex plane can be viewed as a sphere, with ∞ as an ordinary point at its South Pole. This construction, which is known as the *Riemann sphere* or (in arbitrary dimension) as the *stereographic projection*, works as follows: We start at the sphere end, and define the sphere by

$$X^2 + Y^2 + Z^2 = 1 \ . \tag{20}$$

We project from the South Pole, at $(X, Y, Z) = (0, 0, -1)$, to an Argand plane going through the equator and coordinatised by the complex number z. Geometrically one finds the unique straight line going through the South Pole and one other point of the sphere, and assigns to it the complex number defining the point where this line goes through the complex plane. In equations

$$z = \frac{X + iY}{Z + 1} \ , \tag{21}$$

or going the other way

$$X + iY = \frac{2z}{1 + |z|^2} \ , \quad Z = \frac{1 - |z|^2}{1 + |z|^2} \ . \tag{22}$$

This gives a one-to-one map between the complex plane and the sphere minus one point—the South Pole, which in fact corresponds to the ‘extra’ point ∞.

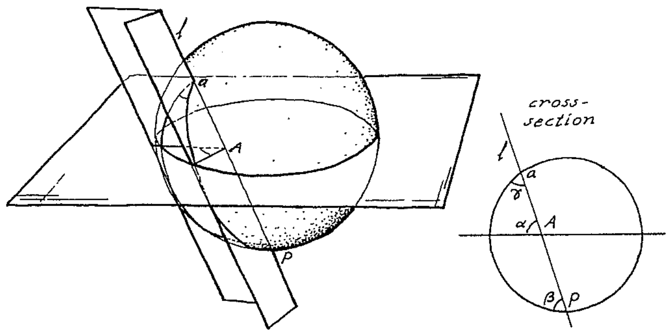


Figure 1: Why the stereographic projection is conformal. The picture was drawn by Sören Holst for his PhD thesis *Horizons and Time Machines*.

Ex: Let $z = \tan\frac{\theta}{2}e^{i\phi}$. What is (X, Y, Z)?

The stereographic projection is such that circles on the plane correspond to circles on the sphere, and straight lines on the plane correspond to circles passing through the South Pole on the sphere. Note that every straight line passes through ∞, so if ∞ is included straight lines are circles too. Angles between curves are preserved by the projection. In general maps that preserve angles are called *conformal*.

It is useful to have the Riemann sphere (or Bloch sphere, if we use the language of quantum information theory) in mind when thinking about what goes on in the extended complex plane.

Return to Möbius transformations. They have the following key properties:

1. They preserve angles between curves, and they are the most general one-to-one conformal mappings on the sphere.
2. They take circles to circles (where it is understood that a straight line counts as a circle too).
3. There is a unique Möbius transformation mapping any three points to any other set of three points.

I assume that this is at least vaguely familiar to you. Any transformation given by an analytic function is angle preserving, it is the one-to-one requirement that singles out the Möbius transformations.

Ex: Show explicitly that we can map any three points to $(0, 1, \infty)$.

We will be particularly interested in the case when a, b, c, d are real, and $ad - bc = 1$, that is when we have a representation of $SL(2, \mathbf{R})$. Obviously such Möbius transformations map the real line to itself. Moreover

$$i \rightarrow \frac{ai + b}{ci + d} = \frac{i(ad - bc) + ac + bd}{c^2 + d^2} = \frac{i + ac + bd}{c^2 + d^2} \ . \tag{23}$$

This has a positive imaginary part. By continuity it follows that the upper half plane is mapped into itself by $PSL(2, \mathbf{R})$.

Alternatively, let

$$z \rightarrow \frac{az + b}{\bar{b}z + \bar{a}} \ , \quad |a|^2 - |b|^2 = 1 \ . \tag{24}$$

We now have a representation of $SU(1, 1)$.

Ex: Check that these Möbius transformations take the unit circle into the unit circle, and the unit disk into the unit disk.

But we can find a Möbius transformation that takes the real line into the unit circle, namely

$$z \rightarrow \frac{z - i}{-iz + 1} \ . \tag{25}$$

This also takes the upper half plane into the unit disk. But if we first take the upper half plane into the unit disk, then apply a Möbius transformation taking the unit disk into itself, and finally take the unit disk back into the upper half plane, the result must be a Möbius transformation that takes the upper half plane into itself. In this way we establish a one-to-one correspondence between transformations in $PSU(1, 1)$ and transformations in $PSL(2, \mathbf{R})$. At the level of matrix groups a calculation shows that

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \in SL(2, \mathbf{R}) \ . \tag{26}$$

This establishes the isomorphism

$$SU(1, 1) \approx SL(2, \mathbf{R}) \ . \tag{27}$$

They are the same group in different guises.

Note that $SL(2, \mathbf{R})$ divides $\mathbf{C} \cup \infty$ into three *orbits*: the extended real line $\mathbf{R} \cup \infty$, the upper half plane, and the lower half plane, on each of which the group acts *transitively*, meaning that one can go from any point to any other point by means of a transformation in the group. The group $PSL(2, \mathbf{R})$ also acts *freely* on the orbits, meaning that for each group element there is a point that is moved by the transformation.

Let us now revisit the question of the conjugacy classes of $SL(2, \mathbf{R})$. The idea is that different conjugacy classes correspond to essentially different kinds of transformations. To understand a transformation it is natural to begin by asking for its fixed points. The fixed points of a Möbius transformation are determined by

$$z = \frac{az + b}{cz + d} \quad \Leftrightarrow \quad cz^2 + (d - a)z - b = 0 \ . \tag{28}$$

This is a quadratic equation and it will have two possibly coinciding roots. Hence every Möbius transformation has exactly two possibly coinciding fixed points, except for the identity $z' = z$ which leaves everything fixed. The case $c = 0$ may appear to be an exception too, but in fact it is not—if $c = 0$ one fixed point is ∞ and the other sits somewhere in the complex plane. If we assume that $c \neq 0$ we find the fixed points

$$z = \frac{1}{2c}\left(a - d \pm \sqrt{(a + d)^2 - 4}\right) \ . \tag{29}$$

Ex: Prove that a Möbius transformation is uniquely determined by its action on three points.

We now assume that a, b, c, d are real. Note that $a + d$ is the trace of the matrix. We then find three qualitively different cases:

1. $|\mathrm{Tr} g| > 2$: Two real fixed points.
2. $|\mathrm{Tr} g| < 2$: Two complex fixed points, one in each half plane.
3. $|\mathrm{Tr} g| = 2$: One real fixed point.

There are only three cases here because we are looking at $PSL(2, \mathbf{R}) = SL(2, \mathbf{R})/Z_2$, rather than at $SL(2, \mathbf{R})$ itself. The cases are called respectively

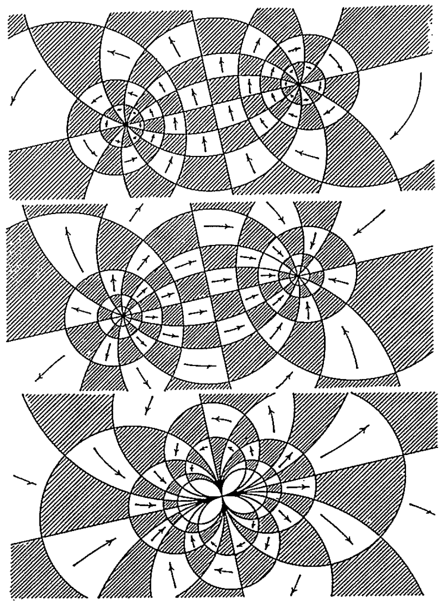


Figure 2: Elliptic, hyperbolic, and parabolic Möbius transformation. The picture illustrates Lester Ford's book on *Automorphic functions.*

hyperbolic, *elliptic*, and *parabolic* Möbius transformations. For $SL(2, \mathbf{C})$ we would also have to consider the case of complex trace.

Note that we could equally well study the conjugacy classes of $PSU(1,1)$, in which case the fixed points of a hyperbolic transformation lie on the unit circle, etc. We will pass freely between the pictures offered by these two groups.

To draw pictures of an elliptic Möbius transformation, place the fixed points at 0 and ∞ and use the $SU(1,1)$ picture. The transformation is $z' = e^{2i\phi}$, so the flow lines are those of a rotation in the unit disk. For the hyperbolic case, use the same fixed points and the $SL(2, \mathbf{R})$ picture. The transformation is $z' = e^{2u}z$, and you will see a set of straight flow lines diverging from the origin. The parabolic case is a bit trickier. If you place the fixed point at ∞ the transformation is $z' = z + b$. To see what goes on close to the fixed point, use a stereographic projection to the Riemann sphere. Having done this one can transform everything to illustrate the $SU(1,1)$ case, which is preferable since it is easier to get a feeling for the unit disk than for the upper half plane.

Ex: Prove that every hyperbolic Möbius transformation coming from $SU(1,1)$ contains exactly one flow line which is a circle or a straight line meeting the unit circle in right angles.

It is absolutely forbidden to do a calculation while doing this exercise.

This brings us to the subject of *non-Euclidean geometry*. Euclid formulated a number of axioms and postulates concerning points and straight lines. Paraphrasing slightly, the important ones say that through every pair of points there passes a unique straight line, and two lines meet in at most one point. The famous fifth postulate says that through any point not lying on a line there passes exactly one line which does not meet the given line. Such lines are called parallel. After many centuries of attempts to prove rather than assume the fifth postulate, it was suggested that one may formulate a consistent geometry in which all other axioms and postulates hold, but in which, given a line and a point not on this line, an infinite number of lines parallel with the given line pass through the given point.

Beltrami, and then Poincaré, offered a model which proves that non-Euclidean geometry is logically consistent, assuming only that Euclid's geometry itself is consistent. In this model space is the unit disk, its boundary not included. A point is a point in the unit disk. A line is a an arc of a circle, or a segment of a straight line, meeting the boundary of the disk in right angles. In this model it is easy to prove that two points determine a unique line, and that two lines meet in at most one point, but the parallel postulate is indeed modified in the manner suggested. The upper half plane can replace the unit disk in the model, if one prefers.

Ex: Prove that the sum of the angles of a triangle is less than $180°$.

The group $PSU(1,1)$ is very relevant here. Elliptic transformations work like rotations of the Poincaré disk, while hyperbolic transformations work like translations—to each line there corresponds a unique hyperbolic transformation taking the line to itself. Moreover one can find a pair of translations such that any point in the disk can be transformed into any other point. This is clearly very similar to what happens in Euclidean geometry. We can also try to define a notion of distance by insisting that the distance between any pair of points remains constant under all such transformations. It turns out

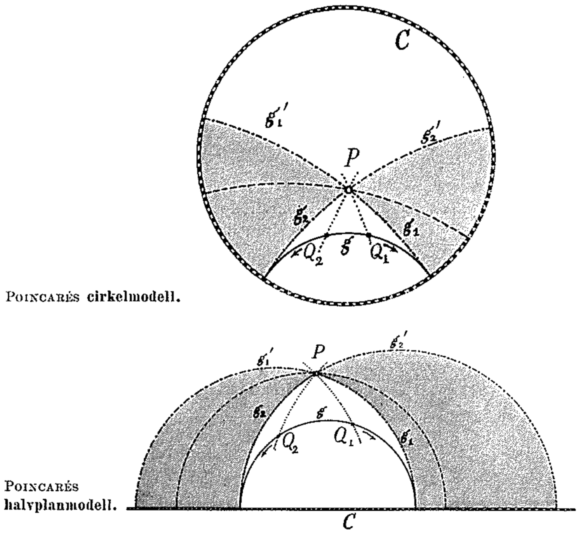


Figure 3: The modified fifth postulate in the two Poincaré models. The picture was drawn by K.-G. Hagstroem, to illustrate Marcel Riesz' book *En åskådlig bild av den icke-euklidiska geometrien.*

that there is one and only one way to do this. According to this definition the distance between a point on the unit circle and a point in the interior of the disk is always infinite. Also the sum of the angles of a triangle can be shown to depend on the area of the triangle in a simple way.

Literature: The standard reference is chapter one of L. R. Ford, *Automorphic functions*, 1929.

It is well known that $SU(2)/Z_2 \approx SO(3)$. This may lead one to suspect that $SU(1,1)$ is related to the Lorentz group $SO(1,2)$ in some way—at least they are similar in that both are three dimensional Lie groups. To explore this, introduce the symmetric matrix

$$\mathbf{x} = \begin{pmatrix} t+x & y \\ y & t-x \end{pmatrix} . \tag{30}$$

The parametrisation looks a bit odd, but we get a simple expression for the determinant,

$$\det x = t^2 - x^2 - y^2 . \tag{31}$$

This is the length squared of a vector in a three dimensional Minkowski space. Now choose a matrix $g \in SL(2, \mathbf{R})$, and compute

$$\mathbf{x} \to \mathbf{x}' = \begin{pmatrix} t'+x' & y' \\ y' & t'-x' \end{pmatrix} = g\mathbf{x}g^{\mathrm{T}} . \tag{32}$$

Note that we use the transpose on the right hand side. This is similar to our treatment of quadratic forms; the point here is that we want $\mathbf{x}'$ to be a symmetric matrix too. Clearly

$$t'^2 - x'^2 - y'^2 = \det x' = \det x = t^2 - x^2 - y^2 . \tag{33}$$

This is the defining property of a *Lorentz transformation*, so the vector (t, x, y) has been mapped to the vector (t', x', y') by a Lorentz transformation. Since the dimensions of the groups match, we have established the isomorphism

$$SL(2, \mathbf{R})/Z_2 = SO_0(1,2) . \tag{34}$$

The subscript is just fine print: We deal with the connected component of the Lorentz group, which means that we cannot change the sign of t.

Ex: Work out the explicit 3 by 3 matrix that acts on Minkowski space, as a function of the parameters in an $SL(2, \mathbf{R})$ matrix.

We can play the same game in four dimensions. Let $\mathbf{x}$ be a general Hermitean matrix

$$\mathbf{x} = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} . \tag{35}$$

Again the determinant is the length squared of a vector in Minkowski space, now of four dimensions. Pick a matrix $g \in SL(2,\mathbf{C})$, and compute

$$\mathbf{x} \rightarrow \mathbf{x}' = g\mathbf{x}g^\dagger . \tag{36}$$

This is again a Hermitean matrix, and again its determinant is preserved, so we have established the isomorphism

$$SL(2,\mathbf{C})/Z_2 = SO_0(1,3) . \tag{37}$$

The isomorphism $SU(2)/Z_2 = SO(3)$ is a special case of this, which arises if we restrict ourselves to transformations that preserve the trace of $\mathbf{x}$. The story ends here. There are no such isomorphisms for the higher dimensional Lorentz groups.

But why did the isomorphism happen in the first place? To understand it, we must understand how Minkowski space is divided into orbits under Lorentz transformations. If we pick a future pointing timelike vector of unit length as a *fiducial vector* on which we act with the Lorentz group, the resulting orbit consists of the hyperboloid

$$t^2 - x^2 - y^2 = 1 , \quad t > 0 . \tag{38}$$

This is a spacelike surface in Minkowski space. If the fiducial vector is spacelike the orbit is a timelike hyperboloid. If it is null (lightlike), the orbit is the forwards or backwards lightcone. The group acts freely on each of these orbits. Note that all points on a given orbit are at the same distance from the origin; in this sense the orbits behave like the orbits of the rotation group in ordinary space.

Rather than considering the lightcone as such, let us consider the set of all null rays through the origin. Again $SO_0(1,2)$ acts freely on this set. But the set of all null rays is clearly in one-to-one correspondence to a circle, which should ring a bell, since $PSU(1,1)$ also acts freely on a circle.

To describe the set of all null rays, note that a symmetric 2 by 2 matrix of determinant zero can be written as

$$x^{AB} = \pm u^A u^B \tag{39}$$

for a two dimensional vector $\mathbf{u}$.

Ex: Prove this. Prove a similar statement for an Hermitean matrix.

In this connection the two component vector is called a *spinor*. When $SL(2,\mathbf{R})$ acts on a matrix $\mathbf{x}$ of this form it acts separately on each spinor,

$$\mathbf{x} \rightarrow g\mathbf{x}g^{\mathrm{T}} \quad \Leftrightarrow \quad \mathbf{u} \rightarrow g\mathbf{u} \ . \tag{40}$$

Since we are interested in null rays rather than null vectors we identify matrices $\mathbf{x}$ that differ by an overall factor. Hence we will identify spinors differing by overall factors too. Momentarily closing our eyes to the fact that the second component may equal zero, we can label the set of all spinors up to a factor by a single real number x,

$$\left(\begin{array}{c} x_0 \\ x_1 \end{array} \right) \sim \frac{1}{x_1} \left(\begin{array}{c} x_0 \\ x_1 \end{array} \right) = \left(\begin{array}{c} \frac{x_0}{x_1} \\ 1 \end{array} \right) \equiv \left(\begin{array}{c} x \\ 1 \end{array} \right) \ . \tag{41}$$

Evidently

$$\left(\begin{array}{c} x_0 \\ x_1 \end{array} \right) \rightarrow \pm \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \left(\begin{array}{c} x_0 \\ x_1 \end{array} \right) \quad \Leftrightarrow \quad x \rightarrow \frac{ax+b}{cx+d} \ . \tag{42}$$

This is how real Möbius transformations act on null rays.

Ex: Show that spinors with the second component equal to zero correspond to the point ∞ on the real line.

We have divided the set of all non-zero spinors into equivalence classes, in such a way that all spinors that differ by a common factor are regarded as identical. This set is known as the *real projective line*, and can be identified with the set $\mathbf{R} \cup \infty$ on which real Möbius transformations act—or via an overall Möbius transformation to the unit circle. A rotation in Minkowski space corresponds to an elliptic Möbius transformation, and a boost to a hyperbolic Möbius transformation. The set of all null rays in 2+1 dimensions can be regarded as a real projective line.

We can treat complex spinors in the same way. We then end up with the *complex projective line*, which is identical to the Riemann sphere on which the group $SL(2, \mathbf{C})$ acts through Möbius transformations. The set of complex spinors up to overall factors can be identified with the set of null rays in a four dimensional Minkowski space, so this gives rise to the isomorphism $PSL(2, \mathbf{C}) \approx SO_0(1, 3)$. In higher dimensions this trick fails: the set of all null rays is always a sphere, but the only spheres that are projective lines (over some number fields) are $\mathbf{S}^1$ and $\mathbf{S}^2$—and two more examples if you bring in something called quaternions and octonions, but we do not go into this here.

Return to 2+1 dimensions. There are other orbits in Minkowski space, notably the spacelike hyperboloids $t^2 - x^2 - z^2 = 1$. They are not so different from spheres in Euclidean space. In particular, any point on such a hyperboloid can be taken to any other point by means of a Lorentz transformation, just as any point on a sphere can be taken to any other by means of a rotation. Rotations preserve all distance relations in Euclidean space, and Lorentz transformations preserve all distance relations in Minkowski space. This means that, whatever impression one may get by drawing a picture of the hyperboloid, all its points are equivalent. There is no special point anywhere. Moreover the group of transformations that preserve distances on the hyperboloid is isomorphic to $SU(1, 1)$, the group that—we claimed—preserves distances as defined in non-Euclidean geometry.

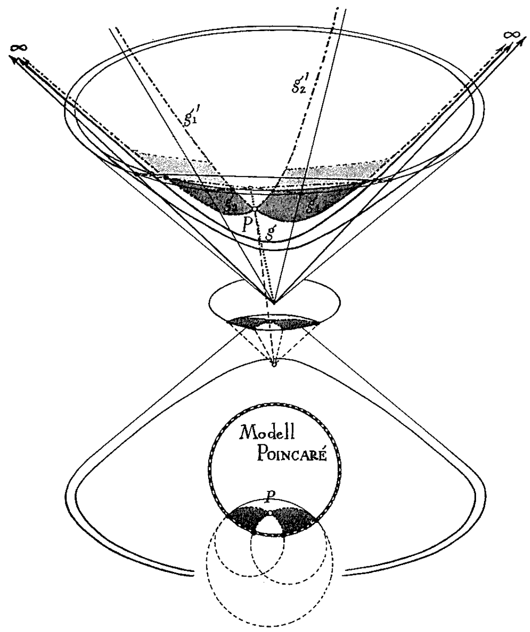
In effect we now have a second model of non-Euclidean geometry. The relation to the Poincaré disk model is given by a stereographic projection in Minkowski space. To set it up, consider the hyperboloid

$$T^2 - X^2 - Y^2 = 1 \ , \quad T > 0 \ , \tag{43}$$

in Minkowski space (we change notation slightly here), place a unit disk at $T = 0$ and centered at the origin, and project between the disk and the hyperboloid by means of straight lines from the point $(-1, 0, 0)$. This will give the formulæ

$$z = \frac{X + iY}{T + 1} \ , \tag{44}$$

or going the other way



Hyperboloidmodellen. Övriga modeller som projektioner av denna.

Figure 4: The hyperboloid and the Poincaré model. The picture is again by K.-G. Hagstroem, in Marcel Riesz' *En åskådlig bild av den icke-euklidiska geometrien.*

$$X + iY = \frac{2z}{1-|z|^2} \ , \quad T = \frac{1+|z|^2}{1-|z|^2} \ . \tag{45}$$

Note that we have to assume that $|z|^2 < 1$ for this to make sense.

We can now give an elegant interpretation of the 'lines' of the Poincaré model. Remember that a 'straight line' on a sphere, that is a *geodesic* on the sphere, is a Great Circle—the intersection of the sphere with a plane through the center of the sphere. Similarly a geodesic on the hyperboloid will be the intersection of the hyperboloid with a plane through the origin of Minkowski space, of the form

$$x_0T + x_1X + x_2Y = 0 \ , \tag{46}$$

where (x_0, x_1, x_3) must be a spacelike vector if the plane is going to intersect the hyperboloid at all.

Ex: Prove that a geodesic on the hyperboloid, so defined, corresponds to a line (a segment of a circle meeting the unit circle at right angles) on the Poincaré disk.

Clearly our definitions imply that geodesics on the hyperboloid transform among themselves under the Lorentz group. The same must therefore be true for the lines in the Poincaré model.

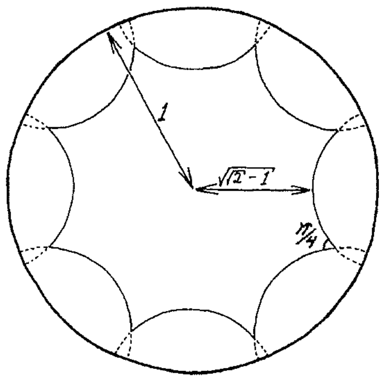


Figure 5: A regular octagon, of the size needed to create a smooth pretzel. The picture was drawn by Sören Holst for a paper in *Classical and Quantum Gravity.*

Using differential geometry one can show that the Riemann curvature scalar on the hyperboloid is constant and negative, while the Ricci tensor vanishes. Spaces of constant negative curvature plays an important role in mathematics. As an example, take a square in an Euclidean plane and identify opposite sides to form a flat torus; note that the vertices become a single point—of no special significance—on the torus. A similar trick is impossible in the Poincaré disk, because the sum of the angles of the square (or any quadrilateral) is less than 2π. Trying to identify opposite sides would lead to a torus with a non-smooth singular point. On the other hand, take an octagon in the flat plane. The sum of the angles at the vertices exceeds 2π, so we cannot identify opposite sides and obtain some smooth closed surface. However, in the Poincaré disk we can shrink the angles by letting the octagon grow, so there will be some critical size for which the operation of identifying opposite sides will succeed. The result turns out to have the topology of a pretzel, and a metric of constant negative curvature by construction.

Ex: Why did I not consider a hexagon?

The symmetry groups of the two planes—flat and negatively curved—actually played their roles here. The opposite sides of the square are identified by means of a translation taking one into the other. The same is true for the octagon: there is a unique (hyperbolic) Möbius transformation that is used to identify a pair of opposing edges—as a glance first at Fig. 2 and then at Fig. 5 should make clear.

Taking polygons with $4g$ edges in the Poincaré disk, and identifying the edges with suitable Möbius transformations, we obtain closed surfaces with constant negative curvature and $g > 1$ 'holes' in them. They are often referred to as *Riemann surfaces* of *genus* g. The torus and the sphere complete the list of Riemann surfaces.

Literature: A lovely perspective on all this is in R. Penrose: Relativistic symmetry groups, in A. O. Barut: *Group theory in non-linear problems*, 1974.

THE GROUP MANIFOLD OF $SL(2, \mathbf{R})$

The set of all elements of a continuous group form a space known as its *group manifold*. Let us see what this is for the group of all 2 by 2 matrices of determinant one. Any real 2 by 2 matrix can be parametrised as

$$g = \begin{pmatrix} U+Y & X+V \\ X-V & U-Y \end{pmatrix} . \tag{47}$$

This will be an element of $SL(2, \mathbf{R})$ if and only if

$$\det g = -X^2 - Y^2 + U^2 + V^2 = 1 . \tag{48}$$

The story is similar to that of $SU(2)$, but there is a significant difference too: the group manifold of $SU(2)$ is a *compact* space (in fact a 3-sphere), while that of $SL(2, \mathbf{R})$ is *non-compact*.

As we know, the group acts on itself according to

$$g \to g_{\mathrm{L}} g g_{\mathrm{R}}^{-1} . \tag{49}$$

Since we may choose g_{L} and g_{R} independently this is a rather large group of transformations, needing $3 + 3 = 6$ parameters for its description. For a general Lie group G, it would be the group $G \times G$. Note we can use left action $g \to g_L g$ to go from any point in the group manifold to any other point; the group acts transitively on itself.

A fundamental idea is that one can use a suitably large set of transformations of a space to define its geometry, say because it may be the case that there is an essentially unique way of measuring distances that is preserved by the transformations. This idea works splendidly here. We begin by defining the *Maurer-Cartan form*

$$g^{-1} dg . \tag{50}$$

This is a differential one-form defined on the group manifold, or if you like it defines a covariant vector field on the group.

Ex: Write this out explicitly using the coordinates X, Y, U, V.

The Maurer-Cartan form is invariant under left action by any fixed group element g_L;

$$g^{-1}dg \to (g_L g)^{-1}d(g_L g) = g^{-1}g_L^{-1}g_L dg = g^{-1}dg \ . \tag{51}$$

We can go on to define a group invariant metric

$$ds^2 = -\frac{1}{2}\text{Tr } g^{-1}dgg^{-1}dg \ . \tag{52}$$

Using the coordinates X, Y, U, V this is

$$ds^2 = -dX^2 - dY^2 + dU^2 + dV^2 \ . \tag{53}$$

To derive this we must remember that $d(-X^2 - Y^2 + U^2 + V^2) = 0$.

Ex: Prove that the group metric (52) is invariant also under right action, so that the isometry group of the group manifold G is $G \times G$.

From eq. (48) we already know that we can regard the group manifold of $SL(2, \mathbf{R})$ as a three dimensional hypersurface sitting inside a four dimensional space coordinatised by the *embedding coordinates* X, Y, U, V. We have now been led to a natural metric on this space, which in turn gives rise to a natural metrix on the group manifold. When viewed in this way the group manifold of $SL(2, \mathbf{R})$ is often referred to as 2+1 dimensional *anti-de Sitter space*.

Ex: Repeat all the above considerations for the group $SU(2)$.

In one special sense the story ends here—the only spheres that are group manifolds are the one and three dimensional spheres $\mathbf{S}^1$ and $\mathbf{S}^3$, and the only Lorentzian space that is a group manifold is 2+1 dimensional anti-de Sitter space. Although every Lie group has its group manifold, only very special spaces are group manifolds.

To understand anti-de Sitter space in a more intimate way, we introduce a set of three genuine coordinates (t, ρ, ϕ) through

$$(X, Y, U, V) = \left(\frac{2\rho\cos\phi}{1-\rho^2}, \ \frac{2\rho\sin\phi}{1-\rho^2}, \ \frac{1+\rho^2}{1-\rho^2}\cos t, \ \frac{1+\rho^2}{1-\rho^2}\sin t \right) \ . \tag{54}$$

The new coordinates range over

$$0 \leq t < 2\pi \ , \qquad 0 < \rho < 1 \ , \qquad 0 \leq \phi < 2\pi \ . \tag{55}$$

The coordinates (ρ, ϕ) are polar coordinates on a unit disk. The origin of the coordinate system coincides with the unit element of the group. The metric (52) now takes the form

$$ds^2 = \left(\frac{1+\rho^2}{1-\rho^2} \right)^2 dt^2 - dl^2 \ , \tag{56}$$

$$dl^2 = \frac{4}{(1-\rho^2)^2}(d\rho^2 + \rho^2 d\phi^2) \ . \tag{57}$$

To understand this, consider a surface at constant t. This is evidently an infinite spacelike surface, and the metric dl^2 on this surface is precisely the metric on the Poincaré disk. So far this is not all that different from Minkowski space—we have a stack, not of flat planes, but of planes with constant negative curvature. However, there is the peculiar feature that the time coordinate is periodic. Thus the topology of the group manifold is $\mathbf{S}^1 \times \mathbf{R}^2$ rather than the $\mathbf{R}^3$ topology enjoyed by Minkowski space. There is no grand father paradox, because there are no grandfathers on a group manifold.

The coordinates (t, ρ, ϕ) are called *sausage coordinates*, because they display anti-de Sitter space as a salami sliced by infinitely thin Poincaré disks.

Ex: Check that the spatial metric dl^2 really is the metric that one would obtain from the hyperboloid model of non-Euclidean geometry.

With this understanding we can revisit the question of conjugacy classes of $SL(2, \mathbf{R})$. Recall that a conjugacy class consists of all group elements that can be transformed into each other using $g \to g_1 g g_1^{-1}$ for some group element g_1. Since the same element occurs on the left and on the right, this is a proper subgroup of the full isometry group. We know that different conjugacy classes have different traces, and we also know that

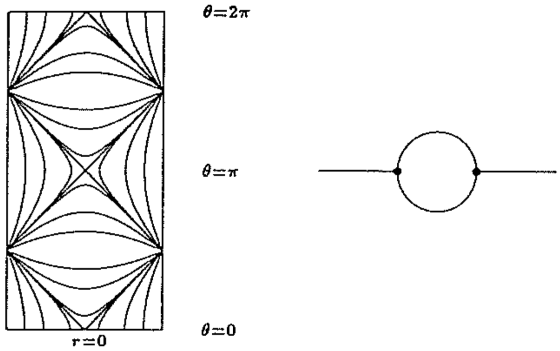


Figure 6: The conjugacy classes of $SL(2,\mathbf{R})$. The picture occurs in Joakim Hallin's PhD thesis *Some Aspects of Non-Perturbative Quantum Field Theory.* What we see on the left is a conformal diagram of anti-de Sitter space.

$$\mathrm{Tr}g = 2U = 2\frac{1+\rho^2}{1-\rho^2}\cos t \ . \tag{58}$$

Unless $U = \pm 1$ there is only one conjugacy class at the given value of U. But the equation that defines the group manifold can be rewritten

$$X^2 + Y^2 - V^2 = U^2 - 1 = -(1-U^2) \ . \tag{59}$$

For $|U| < 1 \Leftrightarrow |\mathrm{Tr}g| < 2$ this is a spacelike hyperboloid. For $|U| > 1$ it is a timelike hyperboloid. For $|U| = 1$ it is a cone, and naturally splits into three pieces, consisting of the vertex, a forwards cone, and a backwards cone.

Ex: Redraw Fig. 6 for the group $PSL(2,\mathbf{R})$.

At constant V the conjugacy class has a metric given by

$$ds^2 = -dX^2 - dY^2 + dU^2 \quad \text{or} \quad dl^2 = dX^2 + dY^2 - dU^2 \ . \tag{60}$$

(I am willing to change the overall sign in order to get a positive definite rather than a negative definite metric on a spacelike surface.)

Let us now make more intimate use of the fact that $SL(2,\mathbf{R})$ is a Lie group. The unit element sits at $(X,Y,U,V) = (0,0,1,0)$. Close to the unit element we find that

$$g \approx \begin{pmatrix} 1+x_2 & x_1+x_0 \\ x_1-x_0 & 1-x_2 \end{pmatrix} = \mathbf{1} + x_0\gamma_0 + x_1\gamma_1 + x_2\gamma_2 \ , \tag{61}$$

where

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \gamma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \gamma_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{62}$$

The determinant equals 1 to first order in the parameters x_i. The matrices γ_i form a basis for the set of all real traceless two by two matrices, and obey

$$[\gamma_1, \gamma_2] = -2\gamma_0 \ , \qquad [\gamma_2, \gamma_0] = 2\gamma_1 \ , \qquad [\gamma_0, \gamma_1] = 2\gamma_2 \ . \tag{63}$$

This defines the *Lie algebra* of $SL(2, \mathbf{R})$: a vector space equipped with an anti-symmetric product, which can be thought of as the *tangent space* of the group at the origin.

Ex: Write down the commutator algebra enjoyed by the i times the Pauli matrices, $i\sigma_i$. This is the real Lie algebra of $SU(2)$. What is the key difference?

Let us now consider *one parameter subgroups* of $SL(2, \mathbf{R})$. An example is

$$g(\sigma) = e^{\sigma\gamma_0} = \mathbf{1} + \sigma\gamma_0 - \frac{\sigma^2}{2!}\mathbf{1} - \frac{\sigma^3}{3!}\gamma_0 + \ldots = \begin{pmatrix} \cos\sigma & \sin\sigma \\ -\sin\sigma & \cos\sigma \end{pmatrix} \ . \tag{64}$$

The calculation is very simple because $\gamma_0^2 = -\mathbf{1}$. In the sausage coordinates this is a curve given by

$$(t, \rho, \phi) = (\sigma, 0, 0) \ . \tag{65}$$

In fact this is a timelike *geodesic* through the origin (i.e. through the unit element of the group).

This did not happen by accident. For any Lie group we can write the one parameter subgroups

$$g(\sigma) = e^{\sigma\tilde{g}} \ , \tag{66}$$

where $\tilde{g}$ is any element of the Lie algebra of the group. In the present case this must be a traceless two by two matrix, to ensure that the determinant of g equals unity. Since we assume that the group is a matrix group it is clear that $g(0)$ is its unit element, so this defines some curve through there. At the unit element this curve has the tangent vector

$$\lim_{\sigma \to 0} \frac{dg}{d\sigma} = \tilde{g} \ , \tag{67}$$

which shows that the Lie algebra indeed can be regarded as the vector space spanned by all tangent vectors at the origin. We know that, if we start from a given point in a direction determined by a given tangent vector, we will obtain a unique geodesic from these data. It is a theorem that, in a Lie group, the curves defined by the one parameter subgroups are indeed geodesics through the unit element in the group manifold.

If we want the geodesics through any other given point, we can use left action to move the geodesics we have already.

Ex: For $SU(2)$ physicists usually include, for good reasons, an imaginary factor i in the exponent of $g(\sigma)$. But since $SU(2)$ has a real Lie algebra, there is a reason for not doing it as well. What reason?

To get all the geodesics through the unit element, observe that

$$\begin{aligned} g(\sigma) = e^{x_0\gamma_0 + x_1\gamma_1 + x_2\gamma_2} = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \right) + \\ + \left(\begin{array}{cc} x_2 & x_1 + x_0 \\ x_1 - x_0 & -x_2 \end{array} \right) \left(1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots \right) \ . \end{aligned} \tag{68}$$

For timelike geodesics, choose $(x_0, x_1, x_2) = \sigma(\hat{x}_0, \hat{x}_1, \hat{x}_2)$ so that

$$x^2 \equiv x_0^2 - x_1^2 - x_2^2 = \sigma^2(\hat{x}_0^2 - \hat{x}_1^2 - \hat{x}_2^2) = \sigma^2 \ . \tag{69}$$

Thus $(\hat{x}_0, \hat{x}_1, \hat{x}_2)$ is a timelike unit vector in a Minkowski space. Then

$$g(\sigma) = \left(\begin{array}{cc} \cos\sigma + \hat{x}_2 \sin\sigma & (\hat{x}_1 + \hat{x}_0)\sin\sigma \\ (\hat{x}_1 - \hat{x}_0)\sin\sigma & \cos\sigma - \hat{x}_2\sin\sigma \end{array} \right) \ . \tag{70}$$

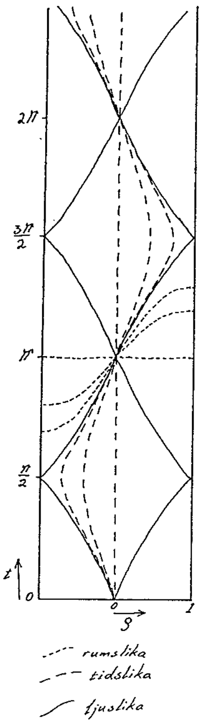


Figure 7: Geodesics on a two dimensional slice ($Y = 0$) of $SL(2, \mathbf{R})$, in sausage coordinates. The picture was drawn by Sören Holst for his Master's Thesis *Om existensen av slutna tidslika kurvor konstruerade med hjälp av massor i 2+1 dimensioner.*

Comparing to eq. (47) we see that these curves can be written as

$$(X(\sigma), Y(\sigma), U(\sigma), V(\sigma)) = (\hat{x}_1 \sin\sigma, \hat{x}_2 \sin\sigma, \cos\sigma, \hat{x}_0 \sin\sigma) \ . \tag{71}$$

The rather more complicated expression for the geodesics in sausage coordinates can be worked out from here. One interesting fact is that the entire family of timelike geodesics will refocus at the point $(X, Y, U, V) = (0, 0, -1, 0)$.

Spacelike geodesics are obtained similarly; unlike the timelike ones they are not periodic in σ. To get the null geodesics, set

$$(x_0, x_1, x_2) = \sigma(1, \cos\phi, \sin\phi) \quad \Rightarrow \quad x^2 = 0 \ . \tag{72}$$

Then

$$g(\sigma) = \left(\begin{array}{cc} 1 + \sigma \sin\phi & \sigma(\cos\phi + 1) \\ \sigma(\cos\phi - 1) & 1 - \sigma\sin\phi \end{array} \right) \ . \tag{73}$$

Hence

$$(X(\sigma), Y(\sigma), U(\sigma), V(\sigma)) = (\sigma\cos\phi, \sigma\sin\phi, 1, \sigma) \ , \tag{74}$$

confirming that $U = 1$ is the light cone from the unit element.

With a full understanding of the topology, geometry, isometry group, and geodesics of the group manifold, we are satisfied. Moreover all that we did generalises straightforwardly to any matrix Lie group. The only thing that changes—usually dramatically—is the length of the calculations.

Literature: For the geometry of 2+1 dimensional anti-de Sitter space, see S. Holst, *Horizons and Time Machines*, 2000.

HEISENBERG GROUPS

We will now examine a finite group in some detail. It will take awhile, but it will transpire that 2 by 2 matrices play a key role here as well. We begin with a *presentation* of the group, which means that we introduce a few symbols, declare what relations are satisfied by them, and finally insist that the group consists of all possible *words* that can be written down using these symbols subject to these relations. We obtain the *Heisenberg group* $H(N)$ from the presentation

$$ZX = \omega XZ \quad \Leftrightarrow \quad ZXZ^{-1}X^{-1} = \omega \tag{75}$$

$$\omega X = X\omega \qquad\qquad \omega Z = Z\omega \tag{76}$$

$$X^N = Z^N = \omega^N = \mathbf{1} \ . \tag{77}$$

Using these relations a general group element be written as $\omega^k X^i Z^j$, where i, j, k are integers modulo N. The order of the group is clearly N^3. It is a *nilpotent* group, meaning that the set of group elements g that can be written in the form $g = g_1 g_2 g_1^{-1} g_2^{-1}$ forms an Abelian subgroup. In a sense nilpotent groups are as close to Abelian as they can be, without actually being so. At the other end of the scale, a group is *simple* if every element of the group can be written in this way.

Ex: Show that this definition of "simple" agrees with that given in the first lecture.

We will want a matrix representation of our group. This is easily obtained because every nilpotent group can be represented by *upper triangular* matrices. For the Heisenberg group, 3 by 3 matrices will do. Consider

$$g(k,i,j) = \begin{pmatrix} 1 & j & k \\ 0 & 1 & i \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{78}$$

Setting, in turn, $(i,j,k) = (1,0,0)$, $(0,1,0)$, $(0,0,1)$ we define three matrices X, Z, ω obeying eqs. (75 - 76). To ensure that eqs. (77) hold we assume that

i, j, k are integers modulo N. Thus the "parameters" on which the matrices depend belong to the ring of integers modulo N. This is sometimes taken as the defining representation of the Heisenberg group. It suggests, correctly as it turns out, that the case when N is a prime number is particularly nice because in this case, and in this case only, the ring of integers modulo N is a field. Every non-zero integer has a multiplicative inverse in arithmetic modulo a prime number. Other variants of the Heisenberg group can be obtained by choosing the parameters in the defining representation in different ways, in particular we obtain a Lie group by choosing them to be real numbers.

The representation in terms of 3 by 3 matrices is not particularly useful because it is not unitary. To obtain a unitary representation we begin by fixing a faithful representation of the Abelian subgroup generated by ω. The obvious choice is

$$\omega = e^{\frac{2\pi i}{N}} \ . \tag{79}$$

Irreducible representations of Abelian groups are always one-dimensional so the only choice here was that we could have chosen any primitive root of unity, that is to say any root of unity for which $\omega^k = 1$ only if $k = N$.

From this point on the relation

$$\sum_{k=0}^{N-1} \omega^k = 0 \tag{80}$$

is regarded as a matter of course. Incidentally the question which subsets of the roots of unity sum to zero, given N, is hard.

Ex: Sort this question out for $N \leq 9$ (which is easy).

Next we turn our attention to Z. Since $Z^N = 1$ all its eigenvalues must be Nth roots of unity. The question is whether its spectrum is non-degenerate, that is whether the eigenvalues are all different. Without loss of generality we may assume that Z is given in diagonal form. Suppose ω^r is an eigenvalue and $|r\rangle$ the corresponding eigenvector. Then

$$Z|r\rangle = \omega^r|r\rangle \ . \tag{81}$$

But from eq. (75) it follows that

$$ZX|r\rangle = \omega XZ|r\rangle = \omega^{r+1}X|r\rangle \ . \tag{82}$$

Thus $X|r\rangle$ is an eigenvector with eigenvalue ω^{r+1}. In this way it follows that all the eigenvalues $\omega^r, \omega^{r+1}, \ldots, 1, \ldots, \omega^{r-1}$ must appear. Because ω is a primitive root of unity these are all distinct, which means that the spectrum of Z is non-degenerate. By means of a unitary permutation of the eigenvectors we can order them as we like. We have also shown that

$$X|r\rangle = |r+1\rangle \ . \tag{83}$$

We arrive at an essentially unique representation, the matrix elements of our operators being

$$\langle r|Z|s\rangle = \omega^s \delta_{rs} \ , \qquad \langle r|X|s\rangle = \delta_{r,s+1} \ , \qquad \langle r|X^iZ^j|s\rangle = \omega^{sj}\delta_{r,s+i} \ , \tag{84}$$

where $0 \leq r, s \leq N-1$. For $N = 2$ X and Z are the Pauli matrices σ_x and σ_z, which explains the notation. For $N = 3$ we get

$$Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix} , \qquad X = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \ . \tag{85}$$

In every dimension the matrices are *monomial*, meaning that they have only one non-zero entry in each row and in each column.

The representation that we have obtained is unique up to unitary equivalence, once we have chosen the representation of the Abelian subgroup generated by ω. One says that the representation of the entire group is *induced* by the representation chosen for this subgroup. The matrices are monomial because we start with a one-dimensional representation for the subgroup.

An interesting oddity is that

$$\det Z = \omega\omega^2 \ldots \omega^{N-1} = e^{\frac{2\pi i}{N}\frac{N(N-1)}{2}} = e^{i\pi(N-1)} = \pm 1 \ . \tag{86}$$

The plus sign applies if N is odd and the minus sign if N is even. The same is true for X. Hence the Heisenberg group is a finite subgroup of $SU(N)$ in the former case, and a subgroup of $U(N)$ in the latter. This is the first sign that even dimensions will cause some trouble with phase factors. One more sign is that if we set $N = 2$ we find that $X^2 = Z^2 = \mathbf{1}$, while $(XZ)^2 = -\mathbf{1}$.

It would have been more natural if all the cyclic subgroups had the same order. We can deal with this by defining $Y = iXZ$ so that $Y^2 = \mathbf{1}$, but then we have brought in a fourth root of unity, so that Y does not belong to the Heisenberg group as we have defined it. Keep this in mind while we proceed.

Repeated use of $ZX = \omega XZ$ leads to

$$(X^i Z^j)(X^k Z^l) = \omega^{jk-il}(X^k Z^l)(X^i Z^j) \ . \tag{87}$$

The phase factor on the right hand side is interesting if we think of i and j as labelling N^2 points in a discrete lattice. Indeed we can think of "vectors"

$$\mathbf{p}_1 = \begin{pmatrix} i \\ j \end{pmatrix} \ , \qquad \mathbf{p}_2 = \begin{pmatrix} k \\ l \end{pmatrix} . \tag{88}$$

(The reason for the quotation marks around "vectors" is that the definition of vector spaces requires that there is a field of "scalars" that can be used to form linear combinations. However, we are using the ring of integers modulo N for our scalars, and this is not a field unless N is a prime number. Ignore this remark.) We can then interpret the anti-symmetric quadratic form

$$\Omega(\mathbf{p}_1, \mathbf{p}_2) = jk - il = -\Omega(\mathbf{p}_2, \mathbf{p}_1) \tag{89}$$

as a *symplectic form* evaluated for a pair of such vectors. This notion is best described in the ordinary plane (over the real numbers) where $\Omega(\mathbf{x}_1, \mathbf{x}_2)$ gives the oriented area spanned by two vectors $\mathbf{x}_1$ and $\mathbf{x}_2$. We simply take over this definition to the discrete plane containing N^2 points that we are now working on. In classical mechanics a symplectic form is used to define Poisson brackets, and a space equipped with a symplectic form is often referred to as a *phase space* (where the word "phase" has nothing to do with the phase factors in the Heisenberg group).

Ex: Show that a symplectic form defines an oriented area on $\mathbf{R}^2$.

Our discrete phase space can be identified with the Abelian group

$$H(N)/I(N) = \mathbf{Z}_N \times \mathbf{Z}_N \ . \tag{90}$$

Here $I(N)$ is the centre of the Heisenberg group, that is to say the set of elements that commute with every other element, and $\mathbf{Z}_N$ is the cyclic group of order N. The quotient group is the Heisenberg group with all phase factors ignored.

At this point our attitude to phase factors is beginning to be distinctly schizophrenic. From one point of view they are crucial. If the phase factor ω had been set equal to 1 from the outset the group would have turned into the Abelian group $\mathbf{Z}_N \times \mathbf{Z}_N$. The representation theory would collapse, because every irreducible representation of an Abelian group is one-dimensional. However, from another point of view the phase factors are without interest because they modify only the overall phases of the Hilbert space vectors, and physically the overall phase of a Hilbert space vector is meaningless. Acting in all possible ways on a given Hilbert space vector with the Heisenberg group gives only N^2 physically distinct quantum states, so from this point of view it is natural to think of the Heisenberg group as the Abelian translation group acting on the discrete plane with N^2 elements. But then again much of the interest in this picture comes from the symplectic form, which turns the discrete plane into a discrete phase space—and the symplectic form comes from the phase factors.

Before we proceed we will introduce a new phase factor, namely

$$\tau = e^{\frac{i\pi(N+1)}{N}} = -e^{\frac{i\pi}{N}} . \tag{91}$$

Then

$$\tau^2 = \omega \ , \qquad \tau^{2N} = \tau^{N^2} = 1 \ , \qquad \tau^N = \begin{cases} +1 & N \text{ odd} \\ -1 & N \text{ even} \end{cases} . \tag{92}$$

If N is odd τ is one of the Nth roots of unity,

$$N = 2n - 1 \qquad \Rightarrow \qquad \tau = \omega^n . \tag{93}$$

In fact $n = 1/2$, the multiplicative inverse of 2, which exists in arithmetic modulo N if and only if N is odd. It is a fact that rings and fields where 2 does not have a multiplicative inverse are considerably difficult to deal with. If N is even this happens to us, and if we introduce τ into the group we are in fact modifying its definition. This is a price worth paying though, and it is not too high because we still have a group with an essentially unique unitary representation, and also because the phase factors do not matter for the action of the group on the discrete phase space.

So, whether N is odd or even we redefine the Heisenberg group as consisting of all possible phase factors τ^k, where k is an integer, together with the N^2 *displacement operators*

$$D_{ij} = \tau^{ij} X^i Z^j = \tau^{ij+2sj} \delta_{r,s+i} \ . \tag{94}$$

The matrix representation we use was given as well. It follows that

$$D_{ij} D_{kl} = \tau^{kj-il} D_{i+k,j+l} = \omega^{kj-il} D_{kl} D_{ij} \ , \qquad D_{ij}^\dagger = D_{-i,-j} \ . \tag{95}$$

Now the symplectic structure is indeed evident in every step, and moreover the operation of taking the adjoint has become quite transparent. These are considerable advantages.

Ex: Check that the first equality in eq. (95) would not have come out as nicely (with the symplectic form in evidence) without the phase factor in the definition of the displacement operators.

The Heisenberg group always gives a *unitary operator basis*, an important notion in quantum information theory. The idea is to regard the set of all matrices acting on $\mathbf{C}^N$—each of which has N^2 complex valued matrix elements—as an N^2 dimensional Hilbert space in its own right. This is possible, and the scalar product is defined as

$$(A, B) = \frac{1}{N} \mathrm{Tr} A^\dagger B \ . \tag{96}$$

The factor $1/N$ is introduced for convenience. Using the explicit representation (94) it is easy to see that all dispacement operators are traceless, with the obvious exception of D_{00}. From the group property it is then easy to check that

$$(D_{ij}, D_{kl}) = \delta_{ik} \delta_{jl} \ . \tag{97}$$

This means that the displacement operators form an orthonormal basis in operator space, and since they are also unitary this is a unitary operator basis. Using the Fierz formula

$$\sum_{i,j} (D_{ij})_{rs} (D_{ij})^*_{tu} = N \delta_{rt} \delta_{su} \tag{98}$$

we can now express an arbitary operator as a linear combination

$$A = \sum_{i,j} D_{ij} \frac{1}{N} \mathrm{Tr} D_{ij}^\dagger A \ . \tag{99}$$

Note that it was by no means obvious from the start that there exists an operator basis consisting of unitary operators only. The group manifold of $U(N)$ has only N^2 real dimensions, while the real dimension of the set of all matrices is twice that.

Ex: Derive the Fierz formula and from it eq. (99).

We now ask for the *automorphism group* of the Heisenberg group, that is to say transformations that take it to itself. More precisely we ask for the *normalizer* of the Heisenberg group within the unitary group $U(N)$, that is all unitary transformations U such that

$$UD_{ij}U^\dagger = D_{i'j'} \ . \tag{100}$$

Here the pair of integers (i', j') may differ from the pair (i, j). The set of all such unitaries necessarily forms a group, although there will be phase factors multiplying U and not determined by this equation. Evidently the Heisenberg group itself is a subgroup of its normalizer, but we ask for *outer automorphisms*, that is to say the set of all unitaries in the normalizer but not actually in the group it normalizes. The phase factor in the group law provides the clue. It will clarify matters if we collect the pairs of integers (i, j) and (k, l) into "vectors" $\mathbf{p}_1$ and $\mathbf{p}_2$. We also write $G\mathbf{p}$ for $\mathbf{p}'$. The group law is

$$D_{\mathbf{p}_1} D_{\mathbf{p}_2} = \tau^{\Omega(\mathbf{p}_1, \mathbf{p}_2)} D_{\mathbf{p}_3} \ , \qquad \mathbf{p}_3 = \mathbf{p}_1 + \mathbf{p}_2 \ , \quad \Omega(\mathbf{p}_1, \mathbf{p}_2) = jk - il \ . \tag{101}$$

But the phase factor commutes with everything, so if we apply a transformation like (100) to the group law we obtain

$$D_{G\mathbf{p}_1} D_{G\mathbf{p}_2} = \tau^{\Omega(\mathbf{p}_1, \mathbf{p}_2)} D_{G\mathbf{p}_3} \ . \tag{102}$$

This is the correct group law if and only if

$$\Omega(G\mathbf{p}_1, G\mathbf{p}_2) = \Omega(\mathbf{p}_1, \mathbf{p}_2) \quad \Leftrightarrow \quad j'k' - i'l' = jk - il \ . \tag{103}$$

Because these equalities must hold in the exponent of τ, and since τ is an Nth root of unity if N is odd and a $2N$th root of unity if N is even, they must hold for all integers taken modulo N if N is odd and modulo $2N$ if N is even.

At this point we simplify matters by assuming that N is odd, and leave the complications in the even dimensional case to their fate.

The transformation G must leave the symplectic form invariant. The most general linear transformation of our "vectors" is

$$\begin{pmatrix} i \\ j \end{pmatrix} \rightarrow \begin{pmatrix} i' \\ j' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} \ , \tag{104}$$

where the entries of the 2 by 2 matrix are integers taken modulo N, since all our arithmetic is modulo N. One checks easily that this transformation preserves the symplectic form if and only if

$$\alpha\delta - \beta\gamma = 1 \quad \text{modulo } N \ . \tag{105}$$

To really clinch matters we should now think a bit more about how we handled our phase factors, and also ask whether there are non-linear transformations of this kind. But in fact there are none, and we already have the answer. The matrices that preserve the symplectic form—which by definition constitute the *symplectic group*—form the group $SL(2, \mathbf{Z}_N)$. The normalizer of the Heisenberg group within the group of unitaries is built from the symplectic group together with the Heisenberg group itself. The resulting group is known as the *Clifford group*.

Ex: If you can figure out exactly why Clifford's name appears here you get a free lunch.

Due to the phase factors the structure of the Clifford group is somewhat involved. If we divide out its centre we obtain the group which acts on the discrete phase space, and this has a simple description since it as a semidirect product of the symplectic group $SL(2, \mathbf{Z}_N)$ and the Abelian translation group $\mathbf{Z}_N \times \mathbf{Z}_N$. To define the terms here, a *direct product* $G \times K$ of two groups G and K consists of ordered pairs of group elements (g, k) such that

$(g_1,k_1)(g_2,k_2) = (g_1g_2,k_1k_2)$. In a *semidirect product* matters are more complicated since the group G acts on the group K. Things simplify if the group K is Abelian, say if G consists of rotations and K of translations. Every group element in the group of rotations and translations can be written as an ordered pair (r,t), with the group law

$$(r_1,t_1)(r_2,t_2) = (r_1r_2, t_1 + r_1t_2) \ . \tag{106}$$

You know this, it is just expressed in an involved way. In our case $SL(2,\mathbf{Z}_N)$ plays the role of the rotations and the direct product group $\mathbf{Z}_N \times \mathbf{Z}_N$ plays the role of the translations.

Ex: For $N=5$, pick some elements of the Clifford group at random and divide phase space into orbits under these group elements.

We have identified the symplectic group consisting of the matrices

$$G = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \ , \qquad \alpha\delta - \beta\gamma = 1 \ . \tag{107}$$

But we still do not know how it is represented by unitary matrices. Provided that the integer β has a multiplicative inverse modulo N the answer is

$$U_G = \frac{e^{i\theta}}{\sqrt{N}} \sum_{r,s=0}^{N-1} \tau^{\beta^{-1}(\delta r^2 - 2rs + \alpha s^2)} |r\rangle\langle s| \ . \tag{108}$$

To prove that this is indeed a representation we check, using our representation of the displacement operators, that

$$U_G D_{ij} U_G^{-1} = \tau^{(\alpha i + \beta j)(\gamma i + \delta j) + 2s(\gamma i + \delta j)} \delta_{r,s+\alpha i + \beta j} = D_{i',j'} \tag{109}$$

where (i',j') are given in eq. (104). This is a straightforward and not too tricky calculation. Because the displacement operators D_{ij} form a unitary operator basis the representation is thereby determined uniquely up to the so far arbitrary phase factor $e^{i\theta}$. But what happens if $\beta = 0$, or more generally if β does not have a multiplicative inverse modulo N? (This happens if β is not relatively prime to N. If N is a prime number the integers modulo N form a field, and every non-zero element has a multiplicative inverse. A certain amount of number theory necessarily creeps in at this point.) We walk around this problem by means of the decomposition

$$\left(\begin{array}{cc} \alpha & \beta \\ \gamma & \delta \end{array} \right) = \left(\begin{array}{cc} 0 & -1 \\ 1 & x \end{array} \right) \left(\begin{array}{cc} \gamma + x\alpha & \delta + x\beta \\ -\alpha & -\beta \end{array} \right) . \tag{110}$$

The integer x can always be chosen so that $\delta + x\beta$ is relatively prime to N. If $N = p$ is prime we can set $x = 0$ since either β or δ must have a multiplicative inverse in this case.

Ex: Check, in a few examples, whether the matrices U_G appearing if β^{-1} does not exist are monomial.

Choosing the so far arbitrary phase factors $e^{i\theta}$ in the expression for U_G so that we obtain a true group is difficult but doable.

Once we start analyzing the symplectic group in detail it matters a lot whether N is prime or not. If $N = p$ is a prime number the order of the group is

$$|SL(2, \mathbf{Z}_p| = p(p^2 - 1) . \tag{111}$$

In general the order is

$$|SL(2, \mathbf{Z}_N| = N^3 \prod \left(1 - \frac{1}{p^2} \right) , \tag{112}$$

where the product runs over all primes that divide N counted once each.

Ex: Work out the order for $2 \leq N \leq 12$.

The next step is to divide the group into conjugacy classes. If $N = p$ is a prime number the story is exactly analogous to the story for $SL(2, \mathbf{R})$. The conjugacy classes are uniquely labelled by the trace of the matrix, unless the trace equals ± 2 in which case there are three different conjugacy classes (with upper triangular representatives). The remaining conjugacy classes can again be divided into hyperbolic conjugacy classes having diagonal representatives, and elliptic conjugacy classes that cannot be diagonalized. Some number theory creeps in here, because when diagonalizing a two-by-two matrix we have to solve a quadratic equation and the quadratic equation may or may not have solutions in the field of integers modulo p. When N is a composite number the story is more complicated because there may be several conjugacy classes among matrices with the same trace.

Once we have divided a group up into conjugacy classes we are close to understanding it. In a finite group every group element g generates a cyclic group of some order (because if there did not exist a smallest integer k such that $g^k = \mathbf{1}$ the order of the group would have to be infinite). For small N we can get an intimate understanding because of the isomorphisms

$$SL(2, \mathbf{Z}_2) = S_3 \ , \quad PSL(2, \mathbf{Z}_3) = A_4 \ , \quad PSL(2, \mathbf{Z}_5) = A_5 \ . \tag{113}$$

Here $PSL(2, \mathbf{Z}_p) = SL(2, \mathbf{Z}_p)/\pm\mathbf{1}$ is the simple group that results by ignoring its Abelian centre, S_n is the group of all possible permutations of n objects, and A_n is the group of all even permutations of n objects. More pertinently, S_3 is the symmetry group of a triangle (including rotations through π around three axes lying within the triangle itself), A_4 is the group of rotations taking a tetrahedron to itself, and A_5 is the group of rotations taking a dodecahedron to itself.

Ex: Make cardboard models of the five Platonic solids and use them to list the conjugacy classes of their symmetry groups.

Literature: A detailed treatment (except that he gives no instructions for how to make cardboard models) is by D. M. Appleby in J. Math. Phys. **46** (2005) 052107.

REPRESENTATION THEORY

Given a group, we are interested in representing it using matrices with complex entries in every possible way. To cut this down a little, we are interested in *inequivalent irreducible* representations only. By definition a representation is reducible if the vector space on which the matrices act contains a proper subspace that is left invariant by all matrices in the representation. It can be shown that for all finite groups (and for all compact Lie groups) every representation admits a choice of basis in the vector space such that the matrices take a block diagonal form, where each block defines an irreducible representation. Moreover each *irrep*—as they are are commonly abbreviated—can be assumed to be unitary. That is

- Every representation of a finite or compact group is equivalent to a direct sum of irreducible ones.
- There is a scalar product such that a given irrep is unitary.

Two representations are said to be equivalent if the matrices in one representation can be turned into those in the other by means of a basis change in the space on which they act. The first statement implies that we can bring all the matrices of a given matrix into the form

$$U(g) = \left(\begin{array}{c|c|c|c} U^{(1)}(g) & \mathbf{0} & \dots & \mathbf{0} \\ \hline \mathbf{0} & U^{(2)}(g) & \dots & \mathbf{0} \\ \hline \vdots & \vdots & & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \dots & U^{(k)}(g) \end{array} \right) . \tag{114}$$

The second statement is clearly plausible. For any group element g in a finite group there exists an integer n such that $g^n = \mathbf{1}$, hence the eigenvalues of a matrix representing g must be roots of unity, as they would be if the matrix were unitary. To prove it, let $\langle v|w\rangle_0$ be any scalar product. Then we define a new scalar product by means of a sum over all group elements,

$$\langle v, w\rangle = \frac{1}{|G|} \sum_{g\in G} \langle U(g)v, U(g)w\rangle_0 \ . \tag{115}$$

Ex: Prove that the sum does indeed define a scalar product.

Ex: Prove that $U(g)$ is unitary with respect to the new scalar product.

Fortunately we do not have to perform this sum, it is enough to know that we lose no generality by restricting ourselves to unitary representations. We can generalise to any compact Lie group—which by definition has a group manifold of finite volume—if we replace the sum with an integral. This disposes of $SU(2)$, but at this point we still know nothing about non-compact groups such as $SL(2, \mathbf{R})$.

Note that a representation does not have to be *faithful*. Indeed the trivial representation, in which all group elements are represented by the number 1, is available for any group.

The task of finding all inequivalent irreducible representations can still be a hard problem, but for finite groups there are some general theorems to guide us. It can be shown that:

- The number of inequivalent irreducible representations of G equals the number of conjugacy classes of G.
- Let the dimensions of the distinct irreducible representations be n_i. Then

$$\sum_i n_i^2 = |G| \ . \tag{116}$$

- The dimensions n_i divide $|G|$.
- A simple group has only one one-dimensional irrep.

I should now introduce characters, and then focus on the finite group $SL(2, \mathbf{Z}_p)$, with p a prime number.

<u>Literature</u>: The idea (which was a good one) was to follow J. E. Humphrey's paper in Amer. Math, Monthly **82** (1975) 21.

A NON-COMPACT LIE GROUP

The representation theory of the Lie algebra of the rotation group, or equivalently the Lie algebra of its covering group $SU(2)$, is well known from textbooks on quantum mechanics. That is a compact group, with a compact group manifold. The representation theory of the non-compact group $SO(1,2)$, or its covering groups (there are several), is different in many interesting ways.

Here is a little bit of the story: The Lie algebra of $SO(2,1) \sim SU(1,1)$ is obtained by flipping two signs in the Lie algebra of $SO(3) \sim SU(2)$. We get

$$[L_1, L_2] = iL_0 \qquad [L_2, L_0] = -iL_1 \qquad [L_0, L_1] = -iL_2 \ . \tag{117}$$

Ex: Where did the signs come from? Show that flipping one sign would have been enough.

To find all the irreducible representations we proceed just as we do for the angular momentum algebra, taking due care about signs. First we form

$$M_+ = \frac{1}{\sqrt{2}}(L_1 - iL_2) \qquad M_- = \frac{1}{\sqrt{2}}(-L_1 - iL_2) \ . \tag{118}$$

The Lie algebra for the new basis vectors becomes

$$[L_0, M_\pm] = \pm M_\pm \qquad [M_+, M_-] = L_0 \ . \tag{119}$$

These commutators are exactly those that we are used to from $SU(2)$. There is a subtle difference though, because our plan is to represent the L_i as Hermitean operators, meaning that we aim for a notion of Hermiticity such that $M_+^\dagger = -M_-$. This sign difference is crucial when we compare the unitary representations of the two groups.

Let us first study linear representations, without any concern for Hermiticity or unitarity. First we look for a maximal set of commuting operators. We take it to include L_0. There is one more, namely the Casimir operator

$$C = L_0^2 - L_1^2 - L_2^2 = 2M_-M_+ + L_0(L_0 + 1) \ . \tag{120}$$

(Note that the final expression is identical to the one we are used to from the $SU(2)$ case.) At the level of rigour where we operate we expect that the vectors in the space on which the operators act should be labelled by two quantum numbers, and we also expect that we can diagonalize L_0. Using a and b for the labels we find the faithful representation

$$\begin{aligned} L_0|a,b\rangle &= \tfrac{1}{2}(a-b)|a,b\rangle \\ M_+|a,b\rangle &= \tfrac{1}{\sqrt{2}}b|a+1,b-1\rangle \\ M_-|a,b\rangle &= \tfrac{1}{\sqrt{2}}a|a-1,b+1\rangle \qquad . \end{aligned} \tag{121}$$

It is easy to see that $j = \frac{1}{2}(a+b)$ is invariant under the action of the generators, and that

$$C|a,b\rangle = j(j+1)|a,b\rangle \ . \tag{122}$$

The eigenvalue is unaffected by the change $j \to -j-1$, and in fact this transformation connects two equivalent representations.

In addition to j one more invariant can be found. We separate out the non-integral part e_0 from the eigenvalue of L_0 by writing

$$\frac{1}{2}(a-b) = e_0 + m \ , \tag{123}$$

where m is any integer. Since the eigenvalue changes in integer steps e_0 is an invariant too. The non-integral part will affect the periodicity of the group element e^{iL_0}. If $e_0 = 0$ the group becomes $SO(3)$ or $SO(1,2)$ as the case may be. If $e_0 = -1/2$ the group is $SU(2)$ or $SU(1,1)$. The group manifold is then simply connected in the first case, and the story stops there, but in the second case e_0 can be any number. We reach the *universal covering group* by letting e_0 take an irrational value, in which case we have unwrapped the circle and the topology of the group manifold changes from $\mathbf{S}^1 \times \mathbf{R}^2$ to $\mathbf{R} \times \mathbf{R}^2$.

The nature of the linear representations hinges on whether a or b are integers. If a is a positive integer we can lower its value in integer steps, until we reach a vector for which

$$M_-|0,b\rangle = 0 \ . \tag{124}$$

The value of $a = j+m+e_0$ cannot be lowered any further, which means that the set of vectors for which a is a non-negative integer is invariant under the action of the group.

> **Ex**: a is an integer but b is not. Show that the representation is reducible (we have found an invariant subspace), but not completely reducible (the 'matrices' cannot be brought to block-diagonal form).

Assume that a is an integer but b is not. We then have an irreducible representation denoted $\mathcal{D}^+$, for which we can set

$$e_0 = -j \ , \qquad m = 0, 1, 2, \ldots \ . \tag{125}$$

Similarly, if b is a non-negative integer then we reach a vector for which

$$M_+|a, 0\rangle = 0 \ . \tag{126}$$

If a is not an integer we obtain the irreducible representation $\mathcal{D}^-$, for which we set

$$e_0 = j \ , \qquad m = 0, -1, -2, \ldots \ . \tag{127}$$

If both a and b are negative integers we do not find any invariant subspaces, but if both of them are non-negative integers something interesting happens. Then $2j = a + b \geq 0$ and we find an invariant subspace for which

$$\frac{1}{2}(a-b) = m = -j, \ldots, 0, \ldots, j \ . \tag{128}$$

This is a finite dimensional representation.

At the level of linear representations we have found three inequivalent types of irreducible representations. If neither a nor b are non-negative integers the vector $|a, b\rangle$ belongs to an infinite dimensional representation that we denote $\mathcal{D}(j, e_0)$, if exactly one of them is a non-negative integer the vector belongs to an irreducible representation denoted $\mathcal{D}^\pm(j)$ (depending on whether a or b is a non-negative integer), and if both of them are non-negative integers the vector belongs to a finite dimensional irreducible representation of dimension $2j + 1 = a + b + 1$. So far there is no distinction between $SU(2)$ and $SU(1, 1)$.

The distinction between compact and non-compact groups appears when we impose the Hermiticity properties

$$SU(2)\ :\quad M_+^\dagger = M_-\ , \qquad\qquad SU(1,1)\ :\quad M_+^\dagger = -M_-\ . \tag{129}$$

We continue to assume that L_0 is Hermitean and diagonal in the chosen basis. To define Hermitean conjugation we need to turn our vector space into a Hilbert space through the introduction of a scalar product. We will then insist that there exist normalizing factors N_m such that the vectors

$$|j,m\rangle \equiv N_m|a,b\rangle = N_m|j+m+e_0, j-m-e_0\rangle \tag{130}$$

form an orthonormal basis. That is to say that the normalizing factors have to be chosen so that

$$\langle j',m'|j,m\rangle = \delta_{j,j'}\delta_{m,m'}\ . \tag{131}$$

Insisting on these things will not only determine the normalizing factors, it will also restrict the possible values of the invariants e_0 and $j(j+1)$. In this basis we have

$$\begin{aligned} L_0|j,m\rangle &= (e_0+m)|j,m\rangle \\ M_+|j,m\rangle &= \tfrac{N_m}{2N_{m+1}}(j-m-e_0)|j,m+1\rangle \\ M_-|j,m\rangle &= \tfrac{N_m}{2N_{m-1}}(j+m+e_0)|j,m-1\rangle \qquad . \end{aligned} \tag{132}$$

The first observation is that

$$\langle j,e_0|L_0|j,e_0\rangle = m+e_0\ . \tag{133}$$

Insisting that $L_0^\dagger = L_0$ forces us to conclude that

$$\mathrm{Im}[e_0] = 0 \tag{134}$$

for both groups. So from now on e_0 is a real number.

We then impose $M_+^\dagger = -M_-$. This implies that

$$\langle j',m'|M_+|j,m\rangle = -(\langle j,m|M_-|j',m'\rangle)^*\ , \tag{135}$$

from which we quickly deduce that

$$\left|\frac{N_{m+1}}{N_m}\right|^2 = \frac{m+e_0-j}{m+1+e_0+j^*} = \frac{m+e_0+\frac{1}{2}-(j+\frac{1}{2})}{m+e_0+\frac{1}{2}+(j^*+\frac{1}{2})} \ . \tag{136}$$

This must be positive for all the allowed values of m.

> **Ex**: What happens for $SU(2)$ and how does this change the following discussion?

Suppose that $j = \sigma + i\lambda$ is complex. Then the preceding formula becomes

$$\left|\frac{N_{m+1}}{N_m}\right|^2 = \frac{m+e_0+\frac{1}{2}-(\sigma+\frac{1}{2}+i\lambda)}{m+e_0+\frac{1}{2}-(-\sigma-\frac{1}{2}+i\lambda)} \ . \tag{137}$$

The right hand side must be real. If $\lambda \neq 0$ the only possibility is

$$j = -\frac{1}{2} + i\lambda \quad \Rightarrow \quad C = -\frac{1}{4} - \lambda^2 \ . \tag{138}$$

This one parameter family of unitary representations arises as the special case $\mathcal{D}_P$ of the linear representation $\mathcal{D}(j, e_0)$, and is known as the *principal series*. We can restrict e_o to lie in the interval $[-1/2, 1/2)$ without loss of generality. Examining eq. (136) we see that we can set

$$N_m = 1 \tag{139}$$

for the principal series. There is a supplementary series as well, for which j is real. The troublesome cases for the unitarity condition (136) arise when $m = 0$ and $m = -1$. If we restrict e_0 to lie in the interval $[-1/2, 1/2)$ they lead to

$$\left|j + \frac{1}{2}\right| < \frac{1}{2} - |e_0| \quad \Rightarrow \quad C < |e_0|(|e_0| - 1) \quad \Rightarrow \quad -\frac{1}{4} \leq C \leq 0 \ . \tag{140}$$

This is the *supplementary series* $\mathcal{D}_S$.

For the two *discrete series* $\mathcal{D}^{\pm}$ the unitarity condition is easily met:

$$\mathcal{D}^+(j): \qquad e_0 = -j \ , \qquad m = 0, 1, 2, \ldots \tag{141}$$

$$\mathcal{D}^-(j): \qquad e_0 = j \ , \qquad m = 0, -1, -2 \ldots \ . \tag{142}$$

For physical applications it is interesting that we can obtain an irreducible representation with the spectrum of L_0 bounded from below. The finite dimensional representation cannot be made unitary by any choice of the normalizing factors, in fact non-compact groups never have finite dimensional unitary representations (except for the trivial representation).

For the principal series we were able to set $N_m = 1$. In the remaining cases we can choose the normalizing factors to be

$$N_m = \sqrt{\frac{(m+e_0-1-j)!}{(m+e_0+j)!}} \ , \tag{143}$$

where we interpret $z! = \Gamma(1+z)$.

We end here, admittedly a little abruptly.

Literature: A very readable description version of this story is given by A. O. Barut and C. Fronsdal, Proc. Roy. Soc. **A287** (1965) 532. The paper by Bargmann, to which they refer, gives the story in full.

THE MODULAR GROUP

In which I should show how $SL(2, \mathbf{Z})$ can be used to tesselate the upper half plane, introduce modular forms, discuss snow crystals, and more.