Lie Theory Through Examples

John Baez

Lecture 2

1 A_3

Now let's draw the A_3 lattice. To do this, we must start with some general theory. You see, last time we just *guessed* that the best way to draw the A_2 lattice gave a hexagonal pattern. But we need to be clearer about the rules of the game. If someone hands us a compact simple Lie group K, we can find a maximal torus $T \subseteq K$. Then we can work out the Lie algebra of this maximal torus, $\mathfrak{t}$. We can then work out the lattice L, which is the kernel of

$$\begin{array}{rcl} e\colon \mathfrak{t} & \to & T \\ x & \mapsto & \exp(2\pi x). \end{array}$$

But to *draw* this lattice we should know a way to measure angles and distances in $\mathfrak{t}$.

Luckily, any compact simple Lie group K has a god-given inner product on its Lie algebra. This inner product restricts to give an inner product on the Lie algebra of the maximal torus. So, there's a *god-given best way* to measure distances and angles in $\mathfrak{t}$. This is why it makes sense to say, for example, that the lattice for A_2 is hexagonal.

Here's how this god-given inner product works. It's an important idea, so let's start out being very general: let $\mathfrak{g}$ be any Lie algebra. Then each element $x \in \mathfrak{g}$ gives a linear operator from $\mathfrak{g}$ to itself, namely the operator 'bracketing with x'. This is usually called $\mathrm{ad}(x)$:

$$\begin{array}{rcl} \mathrm{ad}(x)\colon \mathfrak{g} & \to & \mathfrak{g} \\ y & \mapsto & [x,y] \end{array}$$

Why is it called 'ad'? That's short for 'adjoint', since it defines a representation of the Lie algebra $\mathfrak{g}$ on itself, and people call this representation **adjoint representation**. But why do they call it this? It seems to have nothing to do with the adjoint of an operator — I don't know why the word 'adjoint' gets used these two ways, just like the word 'conjugate'. So, I need to do a little historical research someday.

Anyway, we can try to cook up an inner product on $\mathfrak{g}$ as follows:

$$k(x,y) = \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)).$$

This is called the **Killing form**. Whenever $\mathfrak{g}$ is finite-dimensional, this expression makes sense: 'tr' stands for the trace of an operator, which is just the sum of its diagonal entries when we write it as a matrix in any basis. The Killing form is clearly linear in each argument:

$$k(x, \alpha y + \beta z) = \alpha k(x,y) + \beta k(x,z)$$

and similarly for the first argument. Thanks to the cyclic property of the trace, we also have

$$k(x,y) = k(y,x).$$

So, if $\mathfrak{g}$ is a *real* Lie algebra, the Killing form automatically has all the properties of an inner product except positive definiteness. If $\mathfrak{g}$ is *complex*, the Killing form won't be an inner product, since an inner product on a *complex* vector space should be linear in one slot and conjugate-linear in the other.

In fact, we have:

Theorem 1 *If $\mathfrak{g}$ is a finite-dimensional real Lie algebra, the Killing form k is negative definite if and only if $\mathfrak{g}$ is a direct sum of Lie algebras of compact simple Lie groups. By **negative definite**, we mean that for every nonzero $x \in \mathfrak{g}$*

$$k(x,x) < 0.$$

Theorem 2 *If $\mathfrak{g}$ is a finite-dimensional complex Lie algebra, the Killing form k is a nondegenerate bilinear form if and only if $\mathfrak{g}$ is a direct sum of Lie algebras of complex simple Lie groups. By **nondegenerate**, we mean that for every nonzero $x \in \mathfrak{g}$ there exists $y \in \mathfrak{g}$ with*

$$k(x,y) \neq 0.$$

So: for any compact simple group K the *negative* of the Killing form is an inner product on the Lie algebra $\mathfrak{k}$.

Now we should work out some examples. It's not too painful for $\mathfrak{su}(2)$:

Exercise 1 *Work out the Killing form (x,y) where x and y range over this basis of $\mathfrak{su}(2)$:*

$$i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

*Fans of quantum physics will know that σ_1, σ_2 and σ_3 are called **Pauli matrices**.*

But what's the Killing form like on $\mathfrak{k} = \mathfrak{su}(3)$, and why does the lattice $L \subseteq \mathfrak{t} \subseteq \mathfrak{k}$ look hexagonal in this case? And what happens for $\mathfrak{su}(4)$? Though I'm sure it would be good for my soul, I'm too lazy to compute the Killing form even for $\mathfrak{su}(3)$. So, I'll cheat and use a theorem to compute it *up to a constant factor.*

Theorem 3 *Suppose K is a compact simple Lie group and suppose K has an irreducible representation on the real vector space V. Then any two inner products on V that are both invariant under the action of K are proportional.*

Proof: The idea here is that if we have two inner products on V, say $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle'$, then we can write

$$\langle x, y \rangle' = \langle x, Ty \rangle$$

for some unique operator $T \colon V \to V$. If both inner products are invariant under the action of K on V, T must commute with the action of K. But now suppose V is an irreducible representation! Then by Schur's Lemma every operator $T \colon V \to V$ that commutes with the action of K must be a multiple of the identity. So,

$$\langle x, y \rangle' = c \langle x, y \rangle$$

for some constant c. ∎

The group $\mathrm{SU}(n)$ acts on $\mathfrak{su}(n)$ by conjugation. This is also called the **adjoint representation**, but now with a capital A:

$$\mathrm{Ad}(g)x = gxg^{-1} \qquad g \in \mathrm{SU}(n), x \in \mathfrak{su}(n).$$

Exercise 2 *Check that the adjoint representation* Ad *of* $\mathrm{SU}(n)$ *on* $\mathfrak{su}(n)$ *is irreducible. (Hint: it's enough to show that the representation* ad *is irreducible.)*

Corollary 1 *Suppose $\langle \cdot, \cdot \rangle$ is any inner product on $\mathfrak{su}(n)$ that is invariant under the adjoint action of* $\mathrm{SU}(n)$*. Then this inner product is proportional to the Killing form.*

This is handy because there's an easy way to get such an inner product:

$$\langle x, y \rangle = -\mathrm{tr}(xy)$$

where now we're just taking the usual trace of the product of matrices $x, y \in \mathfrak{su}(n)$! This is easy to compute.

So now we're in business. For starters, consider the lattice L in $\mathfrak{t} \subseteq \mathfrak{su}(3)$. As we saw last time, L consists of integer linear combinations of these two vectors:

$$B_1 = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Let's work out their inner products.

$$\begin{aligned} \langle B_1, B_1 \rangle &= -\mathrm{tr} \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= -\mathrm{tr} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= 2 \end{aligned}$$

An incredibly similar calculation shows

$$\langle B_2, B_2 \rangle = 2.$$

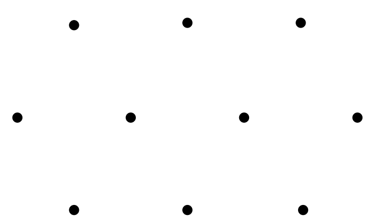
We also have

$$\begin{aligned} \langle B_1, B_2 \rangle &= -\mathrm{tr} \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \\ &= -\mathrm{tr} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= -1 \end{aligned}$$

So: both B_1 and B_2 have the same length, and the angle between them is

$$\arccos\left(\frac{\langle B_1, B_2 \rangle}{\|B_1\| \|B_2\|}\right) = \arccos\left(-\frac{1}{2}\right) = 120°,$$

or if you prefer, $\pi/3$ radians. So indeed, the lattice of their integer linear combinations looks hexagonal:



Now let's do A_3. Here we have $K = \mathrm{SU}(4)$. As before, we get a maximal torus consisting of diagonal matrices:

$$T = \{ \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix} : \ a, b, c, d \in \mathrm{U}(1), abcd = 1 \}.$$

This is a 3-dimensional torus, since we've got 4 numbers but one equation. So, its Lie algebra is 3-dimensional:

$$\mathfrak{t} = \{ \begin{pmatrix} ia & 0 & 0 & 0 \\ 0 & ib & 0 & 0 \\ 0 & 0 & ic & 0 \\ 0 & 0 & 0 & id \end{pmatrix} : \ a, b, c, d \in \mathbb{R}, a + b + c + d = 0 \}.$$

And it's easy to work out the A_3 lattice. It's a lot like the A_2 lattice, only bigger:

$$L = \ker(e) = \{ \begin{pmatrix} ia & 0 & 0 & 0 \\ 0 & ib & 0 & 0 \\ 0 & 0 & ic & 0 \\ 0 & 0 & 0 & id \end{pmatrix} : \ a, b, c, d \in \mathbb{Z}, a + b + c + d = 0 \}.$$

But how do we *draw* this lattice? It's sitting in the 3d space $\mathfrak{t}$, so we should be able to draw it. But points in this space correspond to 4-tuples (a, b, c, d) with $a + b + c + d = 0$, and it's a bit tough (though not impossible) to draw stuff in 4 dimensions.

One method is to find a set of **generators** for our lattice — vectors whose integer linear combinations give everything in the lattice. We can work out their lengths and the angles between these generators. Then we can find vectors with the same lengths and angles between them in $\mathbb{R}^3$, and draw those... and their integer linear combinations.

We should work out these lengths and angles using minus the Killing form. But we'll use this inner product, which is proportional, and easier to compute:

$$\langle x, y \rangle = -\mathrm{tr}(xy).$$

Here are some nice generators for the A_3 lattice:

$$B_1 = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}$$

I could do these calculations, but they're so easy I'll let you:

Exercise 3 *Show that*

$$\langle B_1, B_1 \rangle = \langle B_2, B_2 \rangle = \langle B_3, B_3 \rangle = -2$$
$$\langle B_1, B_2 \rangle = \langle B_2, B_3 \rangle = -1$$
$$\langle B_1, B_3 \rangle = 0$$

This is a lot like what we got for A_2, except that now B_1 and B_3 have vanishing inner product.

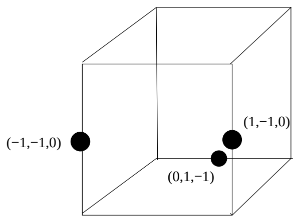
In short: all of B_1, B_2, B_3 have length$^2 = 2$, the inner product of consecutive ones is -1, and the inner product of nonconsecutive ones is 0. Or even more simply put: **they're all the same length, each is at a 120° angle to the next one, but apart from that they're orthogonal.** We'll see this pattern extends to A_n for higher n.

Now, how do we draw them? We just find vectors in $\mathbb{R}^3$ that meet the above description, and draw those! These will do quite nicely:

$$\begin{aligned} b_1 &= (1,-1,0) \\ b_2 &= (0,1,-1) \\ b_3 &= (-1,-1,0) \end{aligned}$$

Cute trick, huh? Note that b_1 and b_2 look just like B_1 and B_2, but written as vectors instead of diagonal matrices, and without the annoying factor of 'i'. But we cleverly choose b_3 so that has the same length as the rest, is orthogonal to b_1, and has dot product -1 with b_2.

Here's what they look like:



They're midpoints of some edges of the cube whose corners are $(\pm 1, \pm 1, \pm 1)$. We get a prettier picture if draw more points in the lattice generated by these guys. It's easy to see that this lattice contains these 12 points:

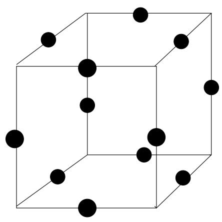
$$\begin{aligned} &(\pm 1, \pm 1, 0) \\ &(0, \pm 1, \pm 1) \\ &(\pm 1, 0, \pm 1). \end{aligned}$$

If you think about it a minute, you'll see this lattice is

$$\mathrm{A}_3 = \{(a,b,c)\colon\ a,b,c \in \mathbb{Z},\ a+b+c \text{ is even}\}.$$

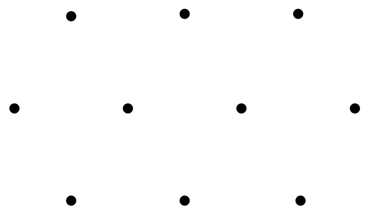
I'm calling this A_3 because it's one picture of the A_3 latttice.

What does this lattice look like? It's easy to see that apart from the origin itself, the points closest to the origin are the 12 points I just listed. A cube has 12 edges, so these are *all* the midpoints of the edges of cube whose corners are $(\pm 1, \pm 1, \pm 1)$. They look like this:



To get the whole lattice we just keep repeating this picture!

It may not be obvious, but the resulting lattice also shows up naturally when we stack oranges in a triangular pyramid. To convince yourself of that, first notice that the A_2 lattice sits in the A_3 lattice, because we can fit $\mathfrak{su}(3)$ inside $\mathfrak{su}(4)$. The A_3 lattice looks like this:



and that's our first layer of oranges. Then there are lots more layers. Each orange touches 6 oranges in the same layer, but also 3 in the layer above and 3 in the layer below. That's a total of 12, which is right — because we know the origin, and thus *every* point in the A_3 lattice, has 12 nearest neighbors.

We also get A_3 lattice when we stack oranges in a *square* pyramid! To see that, use this picture:

$$A_3 = \{(a, b, c) \colon \ a, b, c \in \mathbb{Z}, \ a + b + c \text{ is even}\}.$$

Now the points whose z coordinate is zero are the centers of our first layer of oranges. These points

$$\{(a, b, 0) \colon \ a, b \in \mathbb{Z}, \ a + b \text{ is even}\}$$

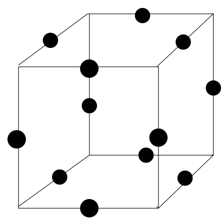
form a square lattice — at least if you turn your head 45°. Each orange touches 4 oranges in the same layer, but also 4 in the layer above:

$$\{(a, b, 1) \colon \ a, b \in \mathbb{Z}, \ a + b \text{ is odd}\}$$

and 4 in the layer below:

$$\{(a, b, -1) \colon \ a, b \in \mathbb{Z}, \ a + b \text{ is odd}\}$$

For example, the nearest neighbors of the 'origin orange' are centered here:



So, the A_3 lattice is very natural when you're trying to pack equal-sized balls in 3 dimensions. In fact, it's optimal if you're trying to pack them as densely as possible! You get a density of $\pi/\sqrt{18}$, or about 79%. There are other equally good ways, but none better.

Kepler guessed this way back in 1611 — so it's called the **Kepler conjecture**. It's really hard to prove! In 1997, Thomas Hales published a plan for a computer-aided proof. Later Hales gave a full proof in a series of papers totaling over 250 pages. It uses ideas like global optimization, linear programming, and interval arithmetic. The files containing the computer code and data for for the proof are about 3 gigabytes in size.

Not surprisingly, people had trouble checking Hales' proof. He submitted it to *Annals of Mathematics*, and in in 2003, it was reported that the paper would be published with a note stating that parts of the paper had not been checked, even though 12 referees had worked on it for more than four years! However, Hales objected. When the proof finally appeared in 2005, the publication contained no such note.

- Thomas C. Hales, A proof of the Kepler conjecture, *Ann. Math.* **162** (2005), 1065–1185.

For more information try Hales' website and this popular history:

- George G. Szpiro, *Kepler's Conjecture: How Some of the Greatest Minds in History Helped Solve One of the Oldest Math Problems in the World*, Wiley, New York, 2003.

But right now, we should think more about what the A_3 lattice looks like. If we connect the dots in the picture above, we get a solid called a **cuboctahedron** with squares and equilateral triangles as faces. It's called a cuboctahedron because its vertices are the midpoints of the 12 edges of a cube, but also they're the midpoints of the 12 edges of a regular octahedron!

In other words, you can get a cuboctahedron by taking a cube and chopping off the corners so much that you chop halfway down each edge. But, you can also get a cuboctahedron by doing the same thing to an octahedron. So, it's a kind of 'hybrid' of a cube and an octahedron.

The cube and regular octahedron are both **Platonic solids**, or **regular polyhedra**. In other words: they're convex polyhedra whose faces are all identical regular polygons, with the same number of faces meeting at each vertex. There are 5 Platonic solids:

- the tetrahedron — 4 vertices, 6 edges, 4 triangular faces
- the octahedron — 6 vertices, 12 edges, 8 triangular faces
- the cube — 8 vertices, 12 edges, 6 square faces
- the dodecahedron — 20 vertices, 30 edges, 12 pentagonal faces
- the icosahedron — 12 vertices, 30 edges, 20 triangular faces

The cube and the octahedron are **dual**. If you put a dot in the middle of each face of a cube, these dots will be the vertices of an octahedron. Conversely, if you put a dot in the middle of each face of an octahedron, these dots will be the vertices of a cube! Similarly, the dodecahedron and icosahedron are dual. The tetrahedron is its own dual.

Dual Platonic solids have the same number of edges. Even better, if you put a dot in the middle of each edge of a Platonic solid, you get the same shape as if you put a dot in the middle of each edge of its dual! For the cube and the octahedron, these dots are the vertices of a cuboctahedron. For the dodecahedron and icosahedron, you get the vertices of an **icosidodecahedron**. For the tetrahedron, you get the vertices of an octahedron. All these quaint-sounding facts will turn out to be important as we dig deeper into group theory.

We've seen the A_3 lattice is full of overlapping cuboctahedra, each being a kind of 'hybrid' of a cube and an octahedron. There's another important way Platonic solids show up in the A_3 lattice. We can't fill space with equal-sized tetrahedra, and we can't do it with octahedra. But if we use both, we can fill space in a pattern that Buckminster Fuller called an **octet truss** — you see it used in architecture sometimes, because it's really rigid. And, the vertices of an octet truss form an A_3 lattice!

The A_3 lattice is truly ubiquitous. It arises in many ways:

1. Stack balls in a triangular pyramid. Their centers form an A_3 lattice.
2. Stack balls in a square pyramid. Their centers form an A_3 lattice.
3. Take a cubic lattice and draw a dot at the center of each cube and the midpoint of each edge. These dots form an A_3 lattice.
4. Take a cubic lattice and draw a dot at each corner and the midpoint of each face. These dots form an A_3 lattice.
5. Take a cubic lattice, color the cubes alternately red and black in a 3d checkerboard pattern, and draw a dot at the center of each red cube. These dots form an A_3 lattice.

Construction 4 is popular in crystallography; this is why crystallographers call the A_3 lattice a **face-centered cubic** or **fcc** lattice. Construction 5 generalizes to n dimensions and gives what is called the **$\mathbf{D}_n$ lattice**. As we shall see, it is a special feature of 3 dimensions that $\mathrm{D}_3 \cong \mathrm{A}_3$.

Exercise 4 *Prove that constructions 4 and 5 both give an* A_3 *lattice.*

Exercise 5 *Draw pictures of all 5 constructions and convince yourself without using formulas that they all give the same lattice, up to rotation and rescaling.*