Research Preliminary Report

Christopher Delaney

August 9, 2012

Introduction

Preliminary research for the project consisted mainly of acquiring general background knowledge of the field of Lie Theory through reading textbooks and articles, before moving onto studying in more depth the paper "Representations of the Lie Algebra $\mathfrak{sl}(2, \mathbb{C})$ over Polynomial Spaces" by former undergraduate mathematics student Matthew Tupper, and expanding upon it. The two main textbooks consulted were Naïve Lie Theory by Stillwell and Introduction to Lie Algebras by Erdmann and Wildon. Secondary textbooks consulted include Symmetry, Representations, and Invariants by Goodman and Wallach, and Lectures on $\mathfrak{sl}(2, \mathbb{C})$-Modules by Mazorchuk. Results of the background research are summarized below.

Preliminary Research

In informal terms, a Lie group is a group wherein the product and inverse operations are smooth, i.e. where the group is also a smooth manifold. Recall that, algebraically, a group is a set of elements together with a product operation that has an identity, is associative, and under which every element has an inverse. Geometrically, a group describes the set of symmetries of a geometric object. So a continuous or Lie group must describe continuous symmetries, such as the simple example of the symmetries of a circle. Any rotation of the circle is a symmetry, and the set of these is the set of plane rotations $\mathbf{S}^1$:

$$\mathbf{S}^1 = e^{i\theta} = \cos\theta + i\sin\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \forall\theta \in \mathbb{R}.$$

As you may have seen, $\mathbf{S}^1$ forms the unit circle in the complex plane (recall the parametrization of the circle by $x = \cos\theta,\ y = \sin\theta$). The product of any two rotations is a rotation, and every rotation clearly has an inverse rotation, so it makes sense to call $\mathbf{S}^1$ a group. A rotation can be infinitesimally small, so it also makes sense to call it continuous, or Lie.

$\mathbf{S}^1$ is a member of an important family of Lie groups called the Special Orthogonal Groups, or $SO(n)$. These are the *generalized rotation groups*, whose elements are equivalent to "rotations" in n-dimensional space in the following sense: a rotation is an orientation-preserving isometry that in general fixes only the origin, and any orientation-preserving isometry fixing only the origin is a rotation. Thus we may define rotations in n dimensions in these terms, since the generalization of length using norms and inner products gives isometry a meaning there. So to be in some $SO(n)$, transformations must preserve the length of all vectors in $\mathbb{R}^n$, or (it can be shown) equivalently, preserve the inner product.

Since matrix multiplication naturally incurs the inner product of row and column vectors, it turns out that a matrix A describes a rotation if and only if $AA^T = I$ and $\det(A) = 1$ (where I is the identity matrix and where the second condition ensures that A also preserves orientation). These $SO(n)$ are examples of *Classical Lie Groups*. The remaining families of Classical Lie Groups are $SU(n)$, essentially "rotations" of the space $\mathbb{C}^n$ of complex vectors, $Sp(n)$, "rotations" of the space $\mathbb{H}^n$ of quaternion vectors, and $GL(n, \mathbb{C})$ and $SL(n, \mathbb{C})$, the General and Special Linear groups, the groups of $n \times n$ matrices that are invertible and that have determinant 1, respectively.

A related concept to Lie groups is that of Lie algebras. Just as curves and surfaces have tangent lines and tangent planes at a point, so too can curved higher-dimensional spaces have tangent spaces of the same dimension. A Lie group is generally a curved space, and its tangent space at the identity is called its *Lie algebra*. To return to our previous example of $\mathbf{S}^1 = SO(2)$, we know that

$$\mathbf{S}^1 = SO(2) = \{A \in M_{2\times 2}(\mathbb{R})\colon AA^T = I \text{ and } \det(A) = 1\}$$

. Now take any path in $SO(2)$ passing through the identity, $A(t)$, such that $A(0) = I$. This must satisfy $A(t)A(t)^T = I$, and in differentiating this with the product rule we find that

$$A'(t)A(t)^T + A(t)A'(t)^T = \mathbf{0}$$

and then setting t=0 we obtain

$$A'(0) + A'(0)^T = \mathbf{0}.$$

So, since at t=0 these paths pass through the identity, we see that the tangent vectors to $SO(2)$ at the identity have the form $X + X^T = \mathbf{0}$ for $X \in M_{2\times 2}(\mathbb{R})$ (and it can be shown that every vector of this form is a tangent vector). This is equivalent to having the form $x + \bar{x} = 0$ for $x \in \mathbb{C}$, which means that x must be pure imaginary. Geometrically this means that the tangent to the circle $SO(2)$ is the pure imaginary line $\mathbb{R}i$, which makes sense (technically the tangent line should be $1 + \mathbb{R}i$, but the *essential* coordinate of the line is the imaginary one).

So this is how one may calculate the Lie algebra of a given Lie group. The Lie algebra captures much of the structure of the Lie group, and since the tangent space is flat and not curved it is much easier to work with. For one, it forms a vector space. Not only that, but it satisfies the extra structure of the *Lie bracket*, an operation that captures much of the generally non-commutative content of the Lie group product. In brief, the Lie bracket for the Lie algebras of matrix Lie groups is $[A, B] = AB - BA$ for A, B elements of the Lie algebra. It is even possible to return to the Lie group from the associated Lie algebra in many cases; it turns out the exponential function maps the tangent space into the Lie group. Applied to the $SO(2)$ example above, this makes sense, as the paths $A(t) = e^{it}$ are precisely the paths mapped from $\mathbb{R}i$ into $SO(2)$ such that $A(0) = I$.

More generally, as an algebraic structure, a Lie algebra is a vector space L together with a map $[\ ,\ ]\colon L \times L \to L$ such that $[x, x] = 0\ \forall x \in L$ and $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0\ \forall x, y, z \in L$. The latter condition is known as the *Jacobi Identity*, and it may seem obscure at first, but it turns out to be essential in preserving the structure of the Lie group to which the algebra is tangent.

A further shadow of the Lie group, if you will, can be found in the concept of a *representation.* A representation of L is a vector space V together with a Lie algebra homomorphism $\varphi\colon L \to End(V)$, the space of endomorphisms or linear transformations on the vector space V, or sometimes more concretely the space of matrices representing those transformations in a given basis. Recall that a homomorphism is a map that preserves operations; thus a vector space homomorphism is precisely a linear transformation, and a Lie algebra homomorphism must also preserve the Lie bracket.

A representation can be thought of as representing a more abstract structure (the Lie algebra) by a more definitive one (the space of linear transformations on a vector space). A common representation is the *adjoint representation*, described by

$$ad\colon L \to End(L),\ (ad(x))(y) = [x, y]\ \forall x, y \in L$$

where you recall that $[x, y]$ is the Lie bracket of x and y.

An equivalent structure to a representation is that of a *Lie module.* A Lie module for a given Lie algebra L is a vector space V together with a map $L \times V \to V,\ (x, v) \mapsto x \cdot v$ for $x \in L$ and $v \in V$ such that $x \cdot v = (\varphi(x))(v)$ for some representation φ. Representations and modules are just different ways of seeing the same structure; the choice of seeing a given case as one or the other will depend only on which is more convenient.

A final concept to define before moving onto the research at hand is that of a *weight.* A weight is really a generalized eigenvalue; that is, where a vector that is affected in a scalar fashion by a linear transformation is said to be an eigenvector of the transformation, a vector that is affected in a scalar fashion by a whole set of linear maps is said to be a weight vector of the set. The linear maps will each have a different eigenvalue with respect to the weight vector, so we create a function λ that sends each map in the set to its eigenvalue with respect to the weight vector. This λ is called the weight.

We now move onto more specific and new research.

Advanced Research

Tupper's paper deals with a specific representation of a Lie algebra, that is the *oscillator representation* of $\mathfrak{sl}(2, \mathbb{C})$, the Lie algebra to the Special Linear group mentioned above. Recall that the Lie group $SL(2, \mathbb{C})$ is defined by

$$SL(2, \mathbb{C}) = \{X \in M_{2\times 2}(\mathbb{C})\colon \det(X) = 1\},$$

and using the formula $\det(e^A) = e^{Tr(A)}$ together with the fact that the exponential function maps the tangent space into the Lie group, we see that

$$1 = \det(e^x) = e^{Tr(x)} \Leftrightarrow Tr(x) = 0.$$

So we see that

$$\mathfrak{sl}(2, \mathbb{C}) = \{x \in M_{2\times 2}(\mathbb{C})\colon Tr(x) = 0\}$$

The standard basis for this Lie algebra is

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The oscillator representation represents $\mathfrak{sl}(2,\mathbb{C})$ via the endomorphisms of the space of polynomials in one variable, and we can define it as $\rho\colon \mathfrak{sl}(2,\mathbb{C}) \to End(C[x])$ by defining its action on the basis vectors of $\mathfrak{sl}(2,\mathbb{C})$:

$$\rho(h) = x\frac{d}{dx} + \frac{1}{2}, \qquad \rho(e) = \frac{i}{2}x^2, \qquad \rho(f) = \frac{i}{2}\frac{d^2}{dx^2}$$

To be a representation, this must preserve the Lie bracket, and it is easily checked that it does. Hence we have made $C[x]$ an $\mathfrak{sl}(2,\mathbb{C})$-module.

A useful process is that of decomposition, where one breaks up a module into irreducible submodules in such a way that the module can be described as a direct sum of its submodules. Tupper desired to accomplish such a decomposition of $C[x,y]$, the space of polynomials in two variables, and to do so he used the tensor product to essentially make two copies of the oscillator representation so that it may be applied to $C[x,y]$. The new two-variable oscillator then becomes

$$\sigma(h) = x\frac{d}{dx} + y\frac{d}{dy} + 1, \qquad \sigma(e) = \frac{i}{2}(x^2+y^2), \qquad \sigma(f) = \frac{i}{2}\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2}\right).$$

We can then proceed to use lowest-weight spaces to decompose $C[x,y]$. A lowest-weight h-eigenvector of a module is a vector in that module with the least weight (i.e. with the lowest possible eigenvalue). This turns out to be equivalent to being a vector killed by the action of f (easily proven). Tupper locates lowest-weight h-eigenvectors $b_n^+ = (x+iy)^n$ and $b_n^- = (x-iy)^n$, defines the associated lowest-weight subspaces of $C[x,y]$ as

$$V_n^+ = Span\{(\sigma(e))^s(b_n^+)\}_{s\geq 0}, \qquad V_n^- = Span\{(\sigma(e))^s(b_n^-)\}_{s\geq 0}$$

and then proceeds to prove that the V_n are irreducible submodules and ultimately that

$$C[x,y] = \bigoplus_{n\geq 0} V_n^+ \oplus \bigoplus_{n\geq 0} V_n^-.$$

In an attempt to extend these results, we ask whether it is possible to form a similar decomposition of $C[x,y,z]$, or, more generally, of $C[x_1,x_2,...,x_n]$ for any n. To inform ourselves further on the general case, we look at $C[x,y,z]$ as an $\mathfrak{sl}(2,\mathbb{C})$-module in more detail and discuss and refute various hypotheses.

It is a simple matter to view $C[x,y,z]$ as a tensor product of $C[x]$ with itself three times, and equally simple to extend the tensor product of the oscillator representation to polynomials of three or indeed of n variables (the proof of this for $n=3$ is just simple, if long, calculation; an alternate proof for any n using induction was also produced). Thus our new representation $\psi\colon\ \mathfrak{sl}(2,\mathbb{C})\to End(C[x,y,z])$ is defined by

$$\psi(h) = x\frac{d}{dx} + y\frac{d}{dy} + z\frac{d}{dz} + \frac{3}{2}$$

$$\psi(e) = \frac{i}{2}(x^2+y^2+z^2)$$

$$\psi(f) = \frac{i}{2}\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2}\right).$$

It was initially thought that perhaps since the two-variable case required h-eigenvectors that were homogeneous of degree 1 as the base for its decomposing lowest-weight subspaces, then the three-variable case would require h-eigenvectors starting homogeneous of degree 2. After much calculation and deliberation it was soon realized that this was not so, and, after discarding an intuitive search approach in favour of a more methodical one, the search for h-eigenvectors of degree 1 in $C[x, y, z]$ proceeded as follows.

First, to determine lowest-weight candidates, the action of h was applied to the general form of the first-degree homogeneous polynomial raised to an arbitrary power, to determine which could be eigenvectors:

$$\begin{aligned}\psi(h)(ax + by + cz)^n &= (x\frac{d}{dx} + y\frac{d}{dy} + z\frac{d}{dz} + \frac{3}{2})(ax + by + cz)^n \\ &= n(ax + by + cz)^{n-1}[ax + by + cz] + \frac{3}{2}(ax + by + cz)^n \\ &= (n + \frac{3}{2})(ax + by + cz)^n\end{aligned}$$

And so we see that *any* homogeneous polynomial of degree 1 is an h-eigenvector for all powers n. Not all of these are of lowest weight, however; recall that being lowest-weight is equivalent to being killed by the action of f. So as before, we apply the action of f to the general h-eigenvector of degree 1:

$$\begin{aligned}\psi(f)(ax + by + cz)^n &= \frac{i}{2}\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2}\right)(ax + by + cz)^n \\ &= \frac{i}{2}(n)[\frac{d}{dx}(a(ax + by + cz)^{n-1}) + \frac{d}{dy}(b(ax + by + cz)^{n-1}) \\ &\quad + \frac{d}{dz}(c(ax + by + cz)^{n-1}] \\ &= \frac{i}{2}(n)(n-1)(ax + by + cz)^{n-2}[a^2 + b^2 + c^2]\end{aligned}$$

Setting this equal to 0 implies that the non-trivial eigenvectors will be of the form $(ax + by + cz)^n$ such that $a^2 + b^2 + c^2 = 0$. There are of course infinitely many of these, counting scalar multiples. But setting $a = \pm 1$ and $b = \pm 1$ implies that $c = \pm i\sqrt{2}z$, and not counting scalar multiples implies that the four basic h-eigenvectors are:

$$\begin{aligned}\gamma_1 &= x + y + i\sqrt{2}z \\ \gamma_2 &= x + y - i\sqrt{2}z \\ \gamma_3 &= x - y + i\sqrt{2}z \\ \gamma_4 &= x - y - i\sqrt{2}z\end{aligned}$$

So, following the pattern of the two-variable case, these γ will be the building blocks, if you will, of the lowest-weight submodules into which we hope to decompose $C[x, y, z]$.

An issue was immediately observed, however. To prove the decomposition of $C[x, y]$, it was necessary to describe a basis of Z_1, the space of homogeneous polynomials of degree 1 (in two variables), using lowest-weight vectors (and the same for Z_n for all n); this was straightforward in the two-variable case, as there were two independent lowest-weight eigenvectors of degree 1 and Z_1 had dimension 2. Starting the same way in $C[x, y, z]$, however, we notice that this Z_1 has dimension 3, and we have *four* independent lowest-weight eigenvectors available. In choosing three of them, we will have four possible bases from which to construct our Z_1. It is of course likely that all four will be similar in structure.

We would next like to be able to describe bases for our new three-variable Z_n in terms of three of our four lowest-weight h-eigenvectors, the γ. First an inquiry into the dimension of each Z_n. In each case the dimension of Z_n will be equal to the number of ways of choosing n objects from a set of three objects (x, y and z) with repetition allowed, since these are all the ways of getting terms of degree n. Basic combinatorics tells us that this result is $\binom{3+n-1}{n} = \binom{n+2}{n} = \frac{(n+2)!}{2n!}$. So Z_2 will have 6 terms, Z_3 10 terms, Z_4 15 terms, and so on.

Now, it can be proven that the product of any two lowest-weight polynomials is itself a lowest-weight polynomial (which fits with our knowledge that the γ^n are lowest-weight for any power n). In the two-variable case, this was exploited to make the submodules appear elegant; $\psi(e) = \frac{i}{2}(x^2 + y^2)$ is precisely a scalar multiple of the product of the lowest-weight eigenvectors b_1^+ and b_1^-. Seen through this light, really the bases for the Z_n in the two-variable case were simply all possible n-combination products of b_1^+ and b_1^-. It is expected that a similar principle will hold in $C[x, y, z]$; the n-combination products of three of the four γ will result in a dimension identical to that of the Z_n, and so all that remains is proving independence.

It seems clear that this will allow a decomposition of $C[x, y, z]$ similar to that of $C[x, y]$.

Ongoing Research

In addition to the aforementioned independence proof, still to be done is the organization of these products of the γ into irreducible submodules in the same vein as in the $C[x, y]$ case, to create a proper submodular decomposition. Also to be researched is the generalization of both the two- and three-variable cases into n dimensions.

This research will continue to be conducted and will provide further insight into the structures of the more complicated multi-variable polynomial spaces.