Topics in Algebra: Quantum Groups

Taught by B. Brubaker - Notes by H. Twiss

February 28, 2020

Introduction to Quantum Groups.

Quantum groups are roughly deformations of Lie groups (or more generally reductive algebraic groups).

Example (A deformation of an algebra)**.** Let $k[X,Y]$ be a two variable polynomial ring over a field k with $XY = YX$. A deformation of this algebra is $k_q[X,Y] = \langle X, Y \mid qXY = YX\rangle$ where q[1] is a fixed parameter or a nonzero complex number.[2] As $q \to 1$ notice that $k_q[X,Y] \to k[X,Y]$.

Let's now look at an example of a complex Lie group.

Example (Complex Lie group)**.** A prototypical example is

$$\mathrm{SL}(2,\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d, \in \mathbb{C}, ad - bc = 1 \right\}.$$

Note this group can be made into a differentiable manifold where the group operations are smooth.

A complex Lie group is a group which is also a complex differentiable manifold with some regularity conditions. In particular, a complex Lie group G is a complex differentiable manifold that is also a group such that the group maps (inversion and multiplication) are smooth with respect to the differential structure of the manifold.

We'd like to start off by understanding the representations of matrix groups G.[3] That is, understand the homomorphisms $\rho : G \to \mathrm{GL}(V) \cong \mathrm{GL}(n,\mathbb{C})$ where V is an n-dimensional complex vector space. To do this, to each matrix group (more generally to each Lie group) we associate its Lie algebra.[4]

[1] We often think of q as a grading to extract finer information about $k[X,Y]$.

[2] We will see in the following why this is a good example of a deformation.

[3] For the concerned reader, all matrix groups are Lie groups, but the converse is not true.

[4] The Lie algebra arises from the Lie group by some exponential map.

Example (Complex Lie algebra)**.** The Lie algebra associated to $\mathrm{SL}(2,\mathbb{C})$ is

$$\mathfrak{sl}(2,\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{C}, a+d=0 \right\},$$

that is the set of complex 2×2 traceless matrices. If $[X,Y] = XY - YX$ is the standard bracket operation, then this Lie algebra can be expressed as follows:

$$\langle E,F,H \mid [H,E] = 2E, [H,F] = -2F, [E,F] = H\rangle,$$

where

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Returning to the general setting, let $\mathfrak{g}$ be the Lie algebra of G. If $\mathrm{Rep}(G)$ is the set of representations of G, then there is a natural map

$$\mathrm{Rep}(G) \to \mathrm{Rep}(\mathfrak{g}) \qquad \Pi \mapsto \pi(x) = \frac{d}{dt}\left(\Pi(e^{tx})\right)\Big|_{t=0}.$$

If the group has nice topological properties, then this map has an inverse.

Theorem. If G is connected and simply connected then there exists a correspondence between representations of G and representations of its Lie algebra.

We would like to pass from the Lie algebra to an associated object: the universal enveloping algebra of the Lie algebra. The universal enveloping algebra is a quotient of an object called the tensor algebra of the Lie algebra. Explicitly, the tensor algebra $T(\mathfrak{g})$ of $\mathfrak{g}$ is defined by

$$T(\mathfrak{g}) := \bigoplus_{i=0}^{\infty} \mathfrak{g}^{\otimes i} = k \oplus \mathfrak{g} \oplus (\mathfrak{g}\otimes\mathfrak{g}) \oplus \cdots.$$

The universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ of $\mathfrak{g}$ is then defined by

$$\mathcal{U}(\mathfrak{g}) := T(\mathfrak{g})/([x,y] - (x\otimes y) + (y \otimes x)),^{5}$$

where the ideal is generated by relations of the form $[x,y] - (x\otimes y) + (y\otimes x)$ for all $x, y \in \mathfrak{g}$.

[5] Observe $x\otimes y$ and $y \otimes x$ are elements of $\mathfrak{g}^{\otimes 2}$ inside $T(\mathfrak{g})$.

Example (Universal enveloping algebra)**.** The universal enveloping algebra of the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ can be expressed as

$$\mathfrak{sl}(2,\mathbb{C}) = \langle e, f, h \mid he - eh = 2e, hf - fh = -2f, ef - fe = h\rangle.^{6}$$

It is a general fact that the universal enveloping algebra contains all the representations of the Lie algebra and has a center which acts by scalars on irreducible representations. The universal enveloping algebra is the algebra we would like to deform to get a quantum group.[7]

Example (A quantum group)**.** The deformation of the universal enveloping algebra of $\mathfrak{sl}(2,\mathbb{C})$ is an example. Explicitly,

$$\mathcal{U}_q(\mathfrak{sl}(2,\mathbb{C})) = \langle E, F, K, K^{-1} \mid {}_{KFK^{-1}=q^{-2}F,[E,F]=EF-FE=(K-K^{-1})/(q-q^{-1})}^{KK^{-1}=1,K^{-1}K=1,KEK^{-1}=q^2E,}\rangle.^{8}$$

History of quantum groups: The term was coined by V. Drinfeld in his 1986 ICM lecture of the same title. He was attempting to formalize aspects of mathematical physics - exactly solvable lattice models and the quantum inverse scattering method. There was a tool known as the quantum Yang-Baxter equation which, in certain cases, could be used to solve the lattice model. Drinfeld realized that certain non-commutative Hopf algebras would produce quantum Yang-Baxter equations. These particular Hopf algebras are precisely quantum groups.

Partition Functions and Lattice Models.

We're going to start with an application of quantum groups. The application comes from statistical mechanics: interpret macroscopic behavior of a system from microscopic interactions (say of atoms).[9] Our first setting is integrable 2-dimensional lattice models.[10]

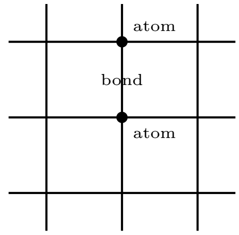
[6] We are suppressing tensor notation in the relations.

[7] A quantum group is not a group in the algebraic sense, it is an algebra.

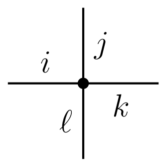
[8] It is not apparent from this form that setting $q = 1$ yields the universal enveloping algebra for $\mathfrak{sl}(2,\mathbb{C})$. We could write the quantum group in a different way to make this more apparent at the sacrifice of adding relations. We will also eventually answer the question of why this example of a quantum group is natural.

[9] Be skeptical in general, but in specific cases this works.

[10] In the following, the notation 2-dimensional lattice models will be clear, but the notion of "integrable" will not be explained.

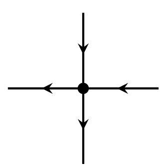


These models look like rectangular grids of finite size, where we think the vertices as atoms and each vertex only interacts with its closest neighbors (i.e., only those vertices which share an edge). We think of the edges as bonds between atoms and we decorate them with elements of a finite set of size m to indicate various possible interactions (here we denote bonds as i, j, k, and ℓ).



Let's provide an example of an integrable 2-dimensional lattice model that we will use throughout the end of this discussion.

Example (6-vertex model)**.** We will decorate every edge with an arrow either pointing in or out of a vertex. We require that every vertex has two adjacent edges pointing in and two adjacent edges pointing out. There are $\binom{4}{2} = 6$ ways to decorate the edges of any vertex, and this is why we call it the 6-vertex model. For example, a vertex may have the following decoration:



This example does arise from physics. Ice has a crystalline structure in which oxygen atoms (vertices) arrange themselves in a lattice and hydrogen atoms are closer to one oxygen atom in the lattice than its neighbors (bonds).

Given any vertex, there is a function $v \mapsto E_{i,j}^{k,\ell}(v)$ which assigns v to the energy at v depending on the decorations (i, j, k, and ℓ). This function usually depends on the location of v in the grid, but we will assume otherwise. The goal is to infer global information from local functions. For example, we like to compute quantities such as the total energy

$$\sum_{\substack{\text{admissible}\\ \text{configurations}}} \sum_{\substack{v \text{ in}\\ \text{grid}}} E_{i,j}^{k,\ell}(v).$$

The probability that atoms arrange themselves into these configurations is inversely proportional to the energy.[11] Precisely, the probability is given by $e^{-\beta\cdot\text{Energy(state)}}$ where $\beta = 1/kT$, k is Boltzmann's constant, and T is temperature.

We would like an algebraic interpretation of the total probability (in this setting it's often called the partition function[12]). It is defined as

$$Z := \sum_{\substack{\text{admissible}\\ \text{configurations}}} e^{-\beta\cdot\text{Energy(state)}} = \sum_{\substack{\text{admissible}\\ \text{configurations}}} \prod_{\substack{v \text{ in}\\ \text{grid}}} e^{-\beta E_{i,j}^{k,\ell}(v)} = \sum_{\substack{\text{admissible}\\ \text{configurations}}} \prod_{\substack{v \text{ in}\\ \text{grid}}} R_{i,j}^{k,\ell}(v),$$

where $R_{i,j}^{k,\ell}(v) := e^{-\beta E_{i,j}^{k,\ell}(v)}$ is the Boltzmann weight of v.

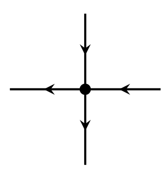
Let us give an example of why it's important to compute Z.

Example. If Q is a physical quantity then the average value of Q is defined by

$$\langle Q \rangle := \frac{\sum_{\substack{\text{admissible}\\ \text{configurations}}} Q e^{-\beta\cdot\text{Energy(state)}}}{Z}.$$

If $Q = E$, the energy of a system, then it's well-known that $\langle E \rangle = kT^2 \cdot \frac{\partial}{\partial T} \ln Z$. So, if we known Z we can compute $\langle E \rangle$ which is global information of a system.

If Z can be explicitly computed then the model is said to be exactly solvable. Let's now give an example of Boltzmann weights. We will write directions such as (NE, SE, etc.) to stand for a vertex with its surrounding edges decorated such that the directions indicate which edges point inward. For example we would write NE for

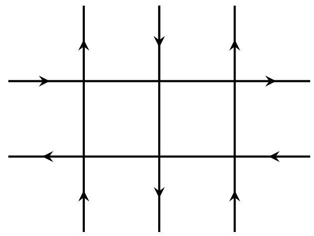


Example (Boltzmann weights)**.** Consider the following assignment of Boltzmann weights

[11] This is due to atoms arranging themselves in the lowest possible energy states.
[12] This has nothing to do with integer partitions.

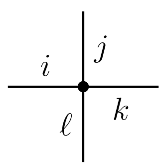
Directions	SE	NW	NE	SW	NS	EW
$R_{i,j}^{k,\ell}$	1	1	λ	λ	$1-q\lambda$	$1-q^{-1}\lambda$

Here λ is an arbitrary complex parameter and q is a nonzero parameter. and consider the following 2-dimensional lattice model

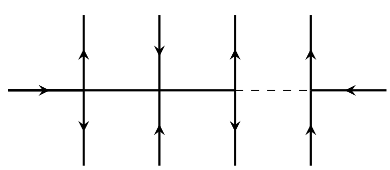


Computing Z for the diagram above means we need to find all possible fillings satisfying two arrows pointing in and two arrows pointing out at each vertex, calculating the Boltzmann weight of each, and then summing them.[13]

This leads us to understanding an algebraic interpretation of Z which is usually easier to compute. Given m decorations, we define an abstract m-dimensional complex vector space V generated by $v_1, \ldots, v_m$. We encode the $R_{i,j}^{k,\ell}$ as matrix coefficients of an endomorphism of $V \otimes V$. We call this endomorphism R. For example, in the 6-vertex model $\dim(V) = 2$ and R is a 4×4 matrix (because $\dim(V \otimes V) = 4$) with 6 nonzero entries where each nonzero entry corresponds to a Boltzmann weight. This suggests the following interpretation: when you see a vertex



think of it as $v_i \otimes v_j \mapsto R_{i,j}^{k,\ell}(v) \cdot (v_k \otimes v_\ell)$ where the mapping is given by R. With this interpretation in mind we can solve for Z one row at a time. Consider the arbitrary 2-dimensional lattice model with N columns.



[13]This gets messy very quickly.

We can think of the left most edge as an element of V and the top set of edges as an element of $V^{\otimes N}$. If T is a row of the partition function, then we can interpret T as belonging to $\mathrm{End}(V \otimes (V^{\otimes N}))$. If we record elements of $V \otimes (V^{\otimes N})$ as $v_0 \otimes v_1 \otimes v_2 \otimes \cdots \otimes v_N$, then we have the following theorem:

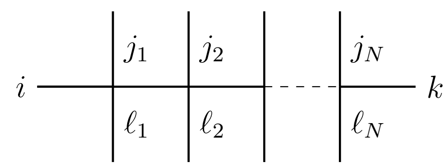
Theorem.

$$T = R_{(0,1)} R_{(0,2)} R_{(0,3)} \cdots R_{(0,N)}$$

where $R_{(i,j)}$ means apply R to $v_i \otimes v_j$ and apply the identity everywhere else.

One-Row Partition Functions.

Consider the more general 2-dimensional lattice model with N columns:



Let $T \in \mathrm{End}(V \otimes (V^{\otimes N}))$ be the associated $(m^{N+1}) \times (m^{N+1})$ matrix (called a transfer matrix). In particular,

$$v_i \otimes (v_{j_1} \otimes \cdots \otimes v_{j_N}) \mapsto \cdots + T_{i,j_s}^{k,\ell_s} \cdot v_k \otimes (v_{\ell_1} \otimes \cdots \otimes v_{\ell_N}) + \cdots .$$

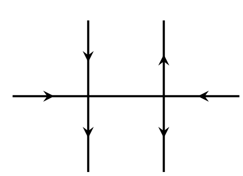
Then we have the following:

Theorem.

$$T = R_{(0,1)} R_{(0,2)} \cdots R_{(0,N)}.$$

Lets see an example of this theorem in the 6-vertex model for a one-row partition function.

Example (Computing T in the 6-vertex model)**.** We would like to compute T of the following 2-dimensional lattice model:



There is only one admissible configuration (can you find it?) so our sum only has one term, and by the definition of Z (in this case $Z = T$), we have $T = \mathrm{wt}(\mathrm{NW})\mathrm{wt}(\mathrm{EW})$. If we label the orientations $\uparrow$ and $\leftarrow$ with a $+$ and the orientations $\downarrow$ and $\rightarrow$ with a $-$, we can order the rows and columns of R by $++$, $+-$, $-+$, and $--$, and write

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda & 1-q\lambda & 0 \\ 0 & 1-q^{-1}\lambda & \lambda & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.^{14}$$

Therefore $T = \mathrm{wt}(\mathrm{NW})\mathrm{wt}(\mathrm{EW}) = 1 \cdot 1 - q^{-1}\lambda = 1 - q^{-1}\lambda$, and $T^{+,-,-}_{-,-,+}$ is precisely the coefficient of $T \in \mathrm{End}(V \otimes (V \otimes V))$ corresponding to this partition function. With our fixed ordering, $(-,-,+)$ is the 4-th element of the order and $(+,-,-)$ is the 6-th element so we are computing the $(4,6)$-entry of T. Let's verify the theorem for this coefficient. Now $R_{(0,1)}$ acts by R on the first two copies of V in $V \otimes (V \otimes V)$ and the identity on the third, and under the same basis we may write

$$R_{(0,1)} = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix}.$$

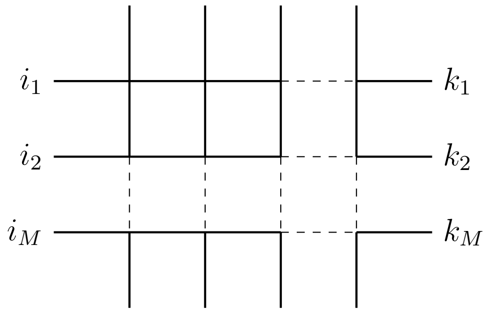
On the other hand, $R_{(0,2)}$ acts by R on the first and third copy of V and the identity on the second. In the same basis we have a block-like matrix:

$$R_{(0,2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1-q\lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 1-q^{-1}\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 1-q^{-1}\lambda & 0 & 0 \\ 0 & 0 & 1-q\lambda & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we multiply the 4-th row of $R_{(0,1)}$ by the 6-th column of $R_{(0,2)}$ we get $1 - q^{-1}\lambda$ as claimed.

For a generalized 2-dimensional lattice model,

[14] We implicitly choose an ordering of the basis, but all claims are independent of this choice.



where $\underline{j}$ and $\underline{\ell}$ will stand for the N-tuple of decorations of edges on the top and bottom row respectively and $\underline{r}^{(1)}, \ldots, \underline{r}^{(M-1)}$ are the N-tuples of decorations of edges for the intermediary rows. The partition function Z may be described as

$$Z = \sum_{\underline{r}^{(M-1)}} \sum_{\underline{r}^{(1)}} T^{k_1 \underline{r}^{(1)}}_{i_1, \underline{j}} \cdot T^{k_2, \underline{r}^{(2)}}_{i_2, \underline{r}^{(1)}} \cdots T^{k_M \underline{\ell}}_{i_M, \underline{r}^{(M-1)}}.$$ [15]

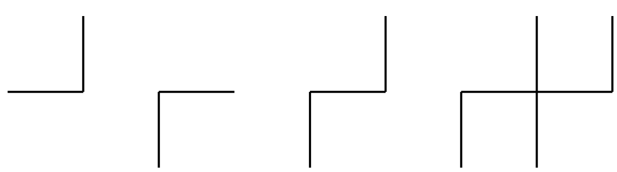
If we assume $i_s = k_s$ for $1 \leq s \leq M$, then we may identify them so our grid becomes a cylinder[16] and we may write the partition function as

$$Z = \left(\mathrm{trace}_V (T)^M\right)^{\underline{\ell}}_{\underline{j}}.$$

where $\mathrm{trace}_V(T)^M \in \mathrm{End}(V^{\otimes N})$ is the partial trace of T. Moreover, if we also have $\underline{j} = \underline{\ell}$ then we may make identifications such that our grid is a torus[17] and our partition function becomes

$$Z = \mathrm{trace}_{V^{\otimes N}} \left(\mathrm{trace}_V (T)^M\right).$$ [18]

Assuming toroidal boundary conditions, if we can find the largest eigenvalue of $\mathrm{trace}_V(T)$, let's call it k_N, then as $M \to \infty$ $Z \sim k_N^M$ asymptotically. With cylindrical boundary conditions we are able to model states as non-intersecting lattice paths as follows: whenever we see a $\uparrow$ or $\rightarrow$ we put a path and otherwise we do nothing. Some examples of states are pictured below.



[15] This is almost like matrix multiplication.
[16] This case is referred to as cylindrical boundary conditions.
[17] This is referred to as toroidal boundary conditions.
[18] The exponent inside the parentheses is just indicating matrix multiplication M times.

Toroidal boundary conditions imply a path never ends which further implies conservation of up arrows in rows. Hence $\mathrm{trace}_V(T)$ breaks up into blocks according to the number of up arrows (i.e., paths) from the bottom row. This lets us determine the eigenvalues.[19] We would like to understand a different method of computing the eigenvalues by using quantum groups.

Quantum Yang-Baxter Equation and Commuting Transfer Matrices.

Recall that for toroidal boundary conditions,

$$Z = \mathrm{trace}_{V^{\otimes N}} \left(\mathrm{trace}_V(T)^M \right).$$

Set $A = \mathrm{trace}_V(T)$, we also call A a transfer matrix.[20] We want to analyze A in particular settings. In the case of the 6-vertex model with toroidal boundary conditions we can completely understand $A \in \mathrm{End}(V \otimes V)$. Here is a sketch of the argument originally developed by Lieb in 1967 and generalized by Sutherland in the same year:

1. Model states as lattice paths for transfer matrices.
2. This splits A into blocks according to the number of paths.[21]
3. Use clever trick to diagonalize blocks.[22]

As a result, we may write $A = PA_{\mathrm{diag}}P^{-1}$ where P is the matrix of eigenvectors and A_{diag} is the matrix of distinct eigenvalues.

In general, the 6-vertex model has six free parameters (the weights). It is a reasonable assumption that the weights are symmetric upon reversal of arrows (e.g., upon reversal of arrows NS $\mapsto$ EW). Assuming this additional constraint, our model is reduced to three free parameters, and physicists call this setting the three parameter field-free model. In the field-free setting it so happens that P is independent of one of the three parameters. Label this independent parameter λ_i if we are considering

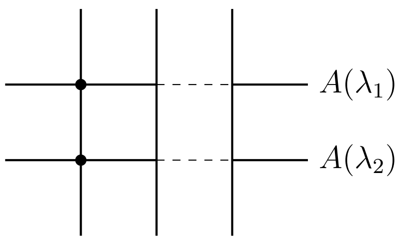
[19] The interested reader could consult Chapter 8 of Baxter's book *Exactly Solved Models in Statistical Mechanics*.

[20] Notice that A also depends on what row of our lattice we are considering since this tells us which copy of V to collapse when taking the trace.

[21] This is where the toroidal boundary conditions are used critically.

[22] The Bethe Ansatz method is used here, and it turns out that the eigenvalues of A are all distinct.

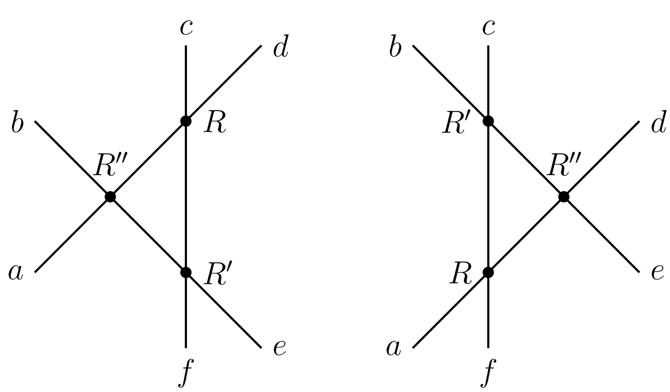
the transfer matrix associated to the i-th row of our lattice (we do this purely to distinguish between transfer matrices in different rows). Write $A_{\text{diag}}(\lambda_1)$ and $A_{\text{diag}}(\lambda_2)$ to emphasises that A_{diag} depends on this parameter (and similarly for A). These matrices are diagonal so they commute implying $A(\lambda_1)A(\lambda_2) = A(\lambda_2)A(\lambda_1)$ (i.e., transfer matrices commute). Pictorially, for any choice of $\underline{j}$ and $\underline{\ell}$, the partition functions corresponding to the 2-dimensional lattice models below are the same under the constraints mentioned above:



This fact should come as a surprise because the weights corresponding to the vertices emphasised in the left-most lattice are affected by different boundary conditions, but are invariant under interchanging rows.

A natural question to ask is if this process is reversible. That is, can one get from commuting transfer matrices to a determination of all the eigenvalues and eigenvectors of A? The answer is surprisingly yes under arbitrary boundary conditions.[23] The goal now is to determine sufficient conditions under which transfer matrices commute. The following theorem of Yang and Baxter accomplishes this:

Theorem (Quantum Yang-Baxter equation)**.** A sufficient condition for transfer matrices to commute is an $R'' \in \mathrm{End}(V \otimes V)$ such that the partition functions corresponding to the 2-dimensional lattice models below are equal

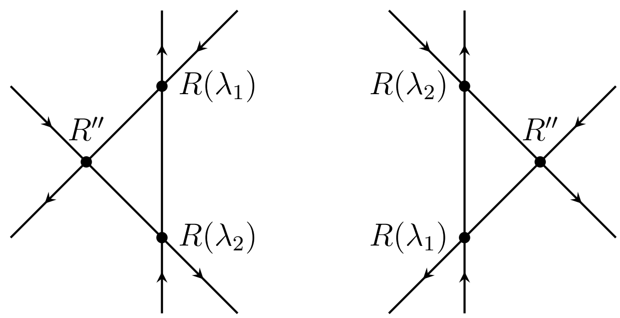


[23]The interested reader can consult section 9.5 of Baxter's book *Exactly Solved Models in Statistical Mechanics* where he answers this question in the setting of the 6-vertex model.

for any choice of decorations a, b, c, d, e, and f.

In the 6-vertex model, there are two decorations for each edge, so there are 2^6 equations and R'' is a 4×4 matrix with 6 unknowns. Let's see an example in this setting.

Example (Quantum Yang-Baxter equation in the 6-vertex model)**.** We want to find a matrix of weights R'' such that the partition functions corresponding to the 2-dimensional lattice models



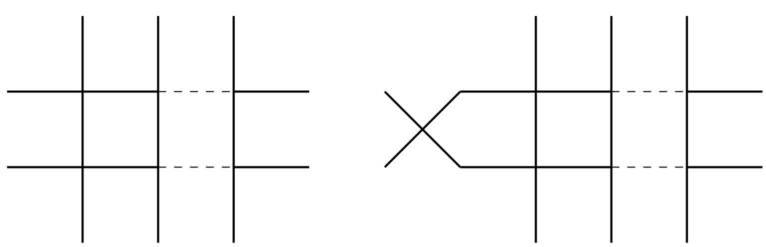
are equal. If we rotate the cross (only) in the left 2-dimensional lattice $\pi/4$ radians counter-clockwise, we may express its partition function as

$$R''(\mathrm{NW})\mathrm{SE}(\lambda_1)\mathrm{SW}(\lambda_2) + R''(\mathrm{EW})\mathrm{EW}(\lambda_1)\mathrm{NS}(\lambda_2),$$

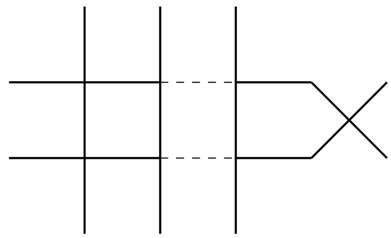
since there are only two admissible configurations (can you find them?).

Why might the quantum Yang-Baxter equation guarantee transfer matrices commute? The argument is as follows.

Proof sketch. For simplicity, assume cylindrical boundary conditions, and consider the 2-dimensional lattice models (we're suppressing decorations here):



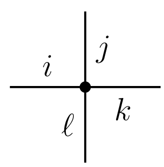
Denote the partition function of the left model by Z_l, the partition function with interchanged rows by Z_l', the partition function of the right model by Z_r, and let wt stand for the weight of the cross in the right model. Then $Z_r = \mathrm{wt} \cdot Z_l$. By repeatedly applying the quantum Yang-Baxter equation, the partition function of the right model is invariant under shifting the cross across columns given we interchange the weights of the column we pass.[24] Moving the cross across the entire lattice produces the model



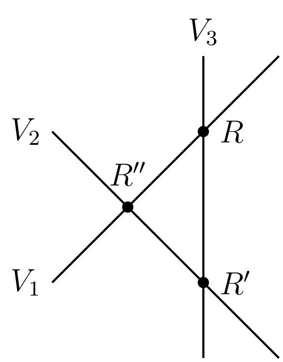
where the vertical rows are interchanged and $Z_r = Z_l' \cdot \text{wt}$. Hence $\text{wt} \cdot Z_l = Z_l' \cdot \text{wt}$ implying the transfer matrices commute. $\square$

Understanding Solutions to the Quantum Yang-Baxter Equation.

Recall that when you see a vertex



we think of it as $v_i \otimes v_j \mapsto R_{i,j}^{k,\ell}(v) \cdot (v_k \otimes v_\ell)$ where the mapping is given by R. In other words, we associate an element $R \in \text{End}(V \otimes V)$ to this picture. Suppressing decorations and labeling lines, consider



[24]This type of argument is sometimes called a "train argument".

where V_1, V_2, and V_3 are vector spaces labeling the lines. We can think of this picture as an element of $\mathrm{End}(V_1 \otimes V_2 \otimes V_3)$ where the segments of the edges closest to V_1, V_2, and V_3 are inputs and when we pass a vertex moving left the inputs are acted upon by matrices R'', R, and R' respectively. Moreover, the matrices only act on the vector spaces associated to the lines which intersect at their vertex (so R'' acts on V_1 and V_2 and acts as the identity on V_3). The quantum Yang-Baxter equation then reads

$$R''RR' = R'RR''.$$

Notice that this is a purely algebraic statement, so we have an algebraic interpretation of the quantum Yang-Baxter equation.[25]

We now understand that solutions to the quantum Yang-Baxter equation gives rise to commuting transfer matrices (by a train argument) and these commuting transfer matrices solve[26] the partition function Z by a method of Baxter. We would now like to answer the following questions about the quantum Yang-Baxter equation:

1. Are there solutions to the quantum Yang-Baxter equation?
2. If there are solutions, can we provide a source for lots of solutions?
3. What does the quantum Yang-Baxter equation have to do with quantum groups?

The answer to (1) is a yes and we would like to illustrate it with an example in the 6-vertex model.

Example (Solutions to quantum Yang-Baxter equation for the 6-vertex model)**.** In our running example we have $R(\lambda_1, q)$ and $R(\lambda_2, q)$. The solution to the quantum Yang-Baxter Equation is also of this form: $R'' = R''(\lambda_3, q)$ where λ_3 satisfies the follow equation

$$\lambda_3 - \lambda_1 + \lambda_2 + \lambda_1\lambda_2\lambda_3 = (q + q^{-1})\lambda_3\lambda_2.$$

Label the weights of SW, NW, NE, SE, EW, and NS by a_1, b_1, a_2, b_2, c_1, and c_2 (in the field-free setting $a_1 = a_2$, $b_1 = b_2$, and $c_1 = c_2$) and define invariants

$$\Delta_1 = \frac{a_1a_2 + b_1b_2 - c_1c_2}{2a_1b_1} \qquad \Delta_2 = \frac{a_1a_2 + b_1b_2 - c_1c_2}{2a_2b_2}$$

(in the field-free setting $\Delta_1 = \Delta_2$). Then we have a theorem:

[25]The interested reader can find an algebraic proof of the quantum Yang-Baxter equation in 7.5 of Chari and Pressley's *A Guide to Quantum Groups*.

[26]By solve we mean described in closed form using familiar functions.

Theorem (Brubaker-Bump-Friedberg)**.** In the 6-vertex model, there exists a solution R'' to the quantum Yang-Baxter equation if and only if $\Delta_1(R) = \Delta_1(R')$ and $\Delta_2(R) = \Delta_2(R')$. In particular, R'' can be described in terms of R and R'.

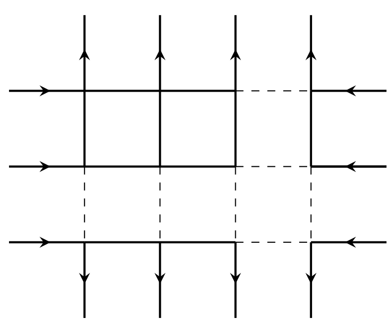
If we are not in the 6-vertex model there is still a general recipe for obtaining solutions to the quantum Yang-Baxter equation. We outline this recipe and in it answer questions (2) and (3). The following is technical. In the general setting, representations of quantum groups give rise to solutions of the quantum Yang-Baxter equation. Quantum groups themselves are special examples of Hopf algebras. To be precise, quantum groups are quasi-triangular Hopf algebras. That is, a pair $(H, \mathcal{R})$ consisting of a Hopf algebra H and a an element $\mathcal{R} \in H \otimes H$ obeying nice properties including an "arbitrary quantum Yang-Baxter equation" $\mathcal{R}_{12}\mathcal{R}_{13}\mathcal{R}_{23} = \mathcal{R}_{23}\mathcal{R}_{13}\mathcal{R}_{12}$. To obtain solutions of the quantum Yang-Baxter equation, one takes a representation (ρ, V) of H and a quantum group $(H, \mathcal{R})$. It so happens that $(\rho \otimes \rho)(\mathcal{R})$ give an "honest" matrix solution to the quantum Yang-Baxter equation. In particular, the quantum group $\mathcal{U}_q(\mathfrak{sl}(2, \mathbb{C}))$ gives rise to solutions to the quantum Yang-Baxter equation for the 6-vertex model.

We'd now like to give a simple in-depth example of computing the entire partition function.

Example (Ice)**.** In ice, the energy of every oxygen atom is the same. Therefore we my assume all weights are 1. This means we may express our partition function as

$$Z = \sum_{\substack{\text{admissible} \\ \text{configurations}}} 1.$$

That is, Z counts the fillings of the lattice given boundary conditions. Let $M = N$ and consider the 2-dimensional lattice model



The admissible fillings of this lattice are in bijections with $N \times N$ alternating sign matrices (matrices with entries in $0, 1, -1$ such that nonzero entries in rows and columns alternate $1, -1, 1 \ldots$ and entries in rows and columns sum to 1). Given a vertex in an admissible configuration of the lattice, assign a 0 to the corresponding entry in a $N \times N$ matrix unless the weight is EW in which case assign a 1 or NS in which case assign a -1. It is a theorem of Zeilberger, Kuperberg, Stroganov, and Okada (all independently) that the number of $N \times N$ alternating sign matrices is given by the formula

$$\frac{1!4!7! \cdots (3N-2)!}{N!(N+1)! \cdots (2N-1)!}.$$

Zeilberger's proof establishes shows these matrices and another combinatorical object are equinumerous but he does not establish a bijection. Kuperberg's, Stroganov's, and Okada's proof use quantum groups, which are more elegant.

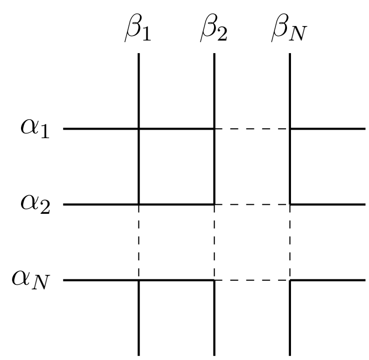
Stroganov's and Okada's Proof and Generalizations.

We would like to discuss the proof strategy that Stroganov and Okada used to count the number of $N \times N$ alternating sign matrices[27].

In the field-free setting define Boltzmann weights

$$\begin{aligned} a &= qx - q^{-1}x^{-1}, \\ b &= x - x^{-1}, \\ c &= q - q^{-1}, \end{aligned}$$

where x and q are free parameters. Now let $\alpha_1, \ldots, \alpha_N$ and $\beta_1, \ldots, \beta_N$ be free parameters and label the rows and columns of the following 2-dimensional lattice model with them as pictured below (notice the lattice has N rows and N columns):



[27]The interested reader can consult *Six-Vertex, Loop, and Tiling Models: Integrability and Combinatorics* by Zinn-Justin for a deeper discussion.

Assign weights as in the field-free setting (after decorating edges), but set $x = \alpha_i/\beta_j$ if the vertex is in the row α_i and column β_j. Stroganov and Okada realized that if we set $q = e^{i\pi/3}$, $\underline{\alpha} = (q^{-1}, \ldots, q^{-1})$ and $\underline{\beta} = (1, \ldots, 1)$, then $Z_N(\underline{\alpha}, \underline{\beta})$ is a complex multiple of the number of $N \times N$ alternating sign matrices. So, we need to compute $Z_N(\underline{\alpha}, \underline{\beta})$. It is a claim that $Z_N(\underline{\alpha}, \underline{\beta})$ is characterized by the following properties

1. $Z_N(\underline{\alpha}, \underline{\beta})$ is symmetric in the entries of $\underline{\alpha}$.
2. $Z_N(\underline{\alpha}, \underline{\beta})$ is symmetric in the entries of $\underline{\beta}$.
3. For each $1 \le i \le N$, $\alpha_i^{N-1} Z_N(\underline{\alpha}, \underline{\beta})$ is a polynomial of degree at most $N-1$ in α_i^2.
4. For each $1 \le i \le N$, $\beta_i^{N-1} Z_N(\underline{\alpha}, \underline{\beta})$ is a polynomial of degree at most $N-1$ in β_i^2.
5. There is a relationship between $Z_N(\underline{\alpha}, \underline{\beta})$ and $Z_{N-1}(\alpha_2, \ldots, \alpha_N, \beta_2, \ldots, \beta_N)$ if $\alpha_1 = \beta_1$. Explicitly,

$$\begin{aligned} Z_N(\underline{\alpha}, \underline{\beta}) =& (q - q^{-1}) \prod_{i=2}^{N} \left(q\frac{\alpha_1}{\alpha_i} - q^{-1}\frac{\alpha_i}{\alpha_1} \right) \prod_{j=2}^{N} \left(q\frac{\beta_j}{\alpha_1} - q^{-1}\frac{\alpha_1}{\beta_j} \right) \\ & \cdot Z_{N-1}(\alpha_2, \ldots, \alpha_N; \beta_2, \ldots, \beta_N). \end{aligned}$$

Property (1) is proved as follows: Show $Z_N(\underline{\alpha}, \underline{\beta})$ has an associated solution to the quantum Yang-Baxter equation by showing $\Delta(x, q)$ is independent of x.[28] Then use a train argument to illustrate that $Z_N(\underline{\alpha}, \underline{\beta})$ is invariant under interchanging rows of the model (i.e., is symmetric in the entries of $\underline{\alpha}$).[29] (2) is proved by rotating the lattice and using an analogous argument. (3) and (4) we encourage the reader to prove themselves. We exclude the proof of (5).

Izergin found an expression for a function with these properties by using q-binomial coefficients and determinants in them. When we set $q = e^{i\pi/3}$, $\underline{\alpha} = (q^{-1}, \ldots, q^{-1})$, and $\underline{\beta} = (1, \ldots, 1)$, the result is easily seen to be

$$\frac{1!4!7! \cdots (3N-2)!}{N!(N+1)! \cdots (2N-1)!}.$$

[28] Notice we're using the result of Brubaker, Bump, and Friedberg.

[29] We don't assume toroidal boundary conditions here so we need to do a little more work to show the weights we interchange are the same as we apply the train argument.

One can use similar techniques to evaluate a more general class of 6-vertex 2-dimensional lattice models with specified boundary conditions. Say the lattice has r rows and ℓ columns. The top boundary conditions are a sequence of up and down arrows and we encode this sequence by an integer partition λ with nonnegative distinct parts such that $\lambda+\rho$ has distinct parts where $\rho = (r-1, r-2, \ldots, 1, 0)$. Given $\lambda+\rho$, we put an up arrow at column i if i is in a part of $\lambda+\rho$ and put a down arrow otherwise. Let's see an example.

Example (Specifying boundary conditions using integer partitions)**.** If $\lambda = (2, 2, 0)$ then $\rho = (2, 1, 0)$ and $\lambda + \rho = (4, 3, 0)$. The top boundary conditions are then given by the the sequence of arrows

$$\uparrow \qquad \uparrow \qquad \downarrow \qquad \downarrow \qquad \uparrow$$

We encode the bottom boundary analogously. Given a matrix with top boundary $\lambda + \rho_r$, bottom boundary $\mu + \rho_\ell$, left boundary all right right arrows, and right boundary all left arrows, we have a theorem of Brubaker, Bump, and Friedberg:

Theorem (Brubaker-Bump-Friedberg)**.** If $\Delta = 0$, then we may write the partition function as

$$Z_{\lambda+\rho_r/\mu+\rho_\ell} = Z_{\rho_r/\rho_\ell} \cdot S_{\lambda/\mu}\left(\frac{b_2^{(1)}}{a_1^{(1)}}, \ldots, \frac{b_2^{(r-\ell)}}{a_1^{(r-\ell)}}\right),$$

where Z_{ρ_r/ρ_ℓ} is a computable partition function depending on ρ_r and ρ_ℓ, $S_{\lambda/\mu}$ is a skew Schur polynomial depending on λ and μ, and the $a_1^{(i)}$'s and $b_2^{(j)}$'s are weights.

Relations Among Symmetric Functions.

In the previous section we discovered that $Z_N(\underline{\alpha}, \underline{\beta})$ was symmetric in the entries of $\underline{\alpha}$ and $\underline{\beta}$. More generally, we would like to know which symmetric functions are representable as partition functions of a 2-dimensional lattice model (not necessarily in the 6-vertex setting). A natural question to ask is why we care about which symmetric functions are representable as partition functions. The punchline is that partition functions satisfy many functorial properties so we can often prove a lot of identities about symmetric functions if we represent them as partition functions.

Recall that given a partition λ with r nonnegative parts and $\rho = (r-1, r-2, \ldots, 1, 0)$, $\lambda+\rho$ is a partition with distinct parts which encodes boundary conditions for a 2-dimensional lattice model. In the 6-vertex model, if we declare all the left boundary conditions to point to the right, all the right boundary conditions to point to the left, and all the bottom boundary conditions to point down, then we have an assignment $\lambda \mapsto Z_\lambda$ with with weights given by

$$\begin{aligned} a_1 &= 1 \quad & a_2 &= x_i, \\ b_1 &= 0 \quad & b_2 &= x_i, \\ c_1 &= x_i \quad & c_2 &= 1, \end{aligned}$$

where i is an index for the row number. In this setting we have the following theorem:

Theorem. Let $\underline{x} = (x_1, \ldots, x_r) \in \mathbb{C}^r$. Then

$$Z_\lambda(\underline{x}) = \underline{x}^\rho \cdot S_\lambda(\underline{x}),$$

where $\underline{x}^\rho = (x_1^{r-1}, x_2^{r-2}, \ldots, x_{r-1}, x_r^0)$ and $S_\lambda(\underline{x})$ is the Schur polynomial corresponding to λ.

What do we mean when we say $S_\lambda(\underline{x})$ is the Schur polynomial corresponding to λ? If we have a Young diagram corresponding to the partition λ, we may fill the boxes of the diagram with elements from the alphabet $\{1, 2, \ldots, r\}$ to make it a semi-standard Young tableaux.[30] If T is a semi-standard Young tableaux, then the weight $\mathrm{wt}(T)$ of T is defined to by the r-tuple where the i-th entry consists of the number of boxes filled with letter i in the alphabet. Let's see an example.

Example (Semi-standard Young tableaux)**.** The Young diagram corresponding to $\lambda = (6, 4, 2)$ is

.

Our alphabet is $\{1, 2, 3\}$ and a filling of this Young diagram which produces a semi-standard Young tableaux T is

1	1	2	2	2	3
2	2	3	3		
3	3				

.

Then $\mathrm{wt}(T) = (2, 5, 5)$.

[30] This means the row entries weakly increase and the column entries strictly increase.

Letting SSYT(λ) stand for the set of all semi-standard Young tableaux with Young diagram corresponding to λ, the Schur polynomial $S_\lambda(\underline{x})$ corresponding to λ is defined by

$$S_\lambda(\underline{x}) = \sum_{T \in \mathrm{SSYT}(\lambda)} \underline{x}^{\mathrm{wt}(T)}.$$

There is an interesting inner product (called the Hall inner product[31]), and the Schur polynomials form an orthonormal basis with respect to this inner product. Moreover, in representation theory, irreducible and finite dimensional (necessarily polynomial) representations of $\mathrm{GL}_n(\mathbb{C})$ are indexed by partitions, and the characters of matrices with eigenvalues $x_1, \dots, x_n$ are given by $S_\lambda(\underline{x})$. Pictorially,

$$\mathrm{GL}_n(\mathbb{C}) \xrightarrow{\rho_\lambda} \mathrm{End}(V) \xrightarrow{\mathrm{Tr}} \mathbb{C} \qquad \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_r \end{pmatrix} \mapsto S_\lambda(\underline{x}).$$

As an aside, there are some other interesting polynomials which have connections to symmetric polynomials. We will speak somewhat loosely in the following. The Macdonald polynomials $P_\lambda(\underline{x}; q, t)$ are polynomials dependent on two parameters q and t. If we set $q = 0$, then we obtain the Hall-Littlewood polynomials $P_\lambda(\underline{x}; t)$, and setting $t = 0$ gives the Schur polynomials $S_\lambda(\underline{x})$. If we take the Macdonald polynomials, set $t = q^a$ and perform a limiting procedure $q \to 1$, we obtain the Jack polynomials. One can define the so-called non-symmetric Macdonald polynomials $E_\alpha(\underline{x}; q, t)$ where α is a composition. If we "average over all permutations of α," we get the usual Macdonald polynomials $P_\lambda(\underline{x}; q, t)$.

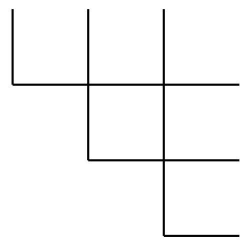
In 2019, Berodin and Wheeler produced a 2-dimensional lattice model with quantum Yang-Baxter equations whose partition function was $E_\alpha(\underline{x}; q, t)$ with $\underline{x} \in \mathbb{C}^n$. The decorations on an edge come from the quantum group module $\mathcal{U}_q(\hat{\mathfrak{sl}}_{n+1})$.[32] In particular, the decorating set is unbounded but it comes from the quantum group module above and so the associated quantum Yang-Baxter equation has a solution!

In fact, there is a more general recipe which takes a symmetric function and breaks it into non-symmetric pieces. If we assume toroidal boundary conditions recall that we can think of decorations (in the 6-vertex model) as paths. If we color the starting and ending points of the paths with the same set and permute them then the symmetric function breaks into pieces according to which permutations give rise to paths with the same start and end color. Let's see an example.

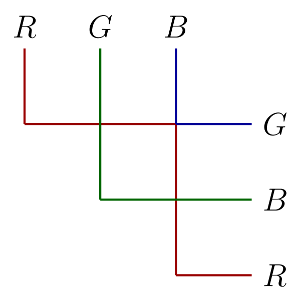
[31] It is a symmetric bilinear form on the ring of symmetric functions.

[32] Here $\hat{\mathfrak{sl}}_{n+1}$ can be thought of as an affine version of $\mathfrak{sl}_{n+1}$.

Example (Breaking symmetric functions according to path colorings)**.** If we have the following path diagram



Then the RGB coloring below corresponds to a non-symmetric piece (notice the right labeling has been permuted):



The important fact here is that if a path emanates from a coloring at the top it needs to terminate at the corresponding coloring on the right.

There are two other interesting properties of Schur polynomials that we can deduce using lattice models. The first is that if we have a lattice model with top boundary conditions given by $\lambda + \rho$ and remove the top row this gives rise to a branching rule. In particular, we have the property

$$S_\lambda(\underline{x}) = \sum_\mu c_\mu \cdot S_\mu(\underline{x}),$$

where μ ranges over the partitions with one less part than λ and c_μ is a constant depending on μ. In the same setting we can also glue certain lattices together. If $\lambda + \rho$ are the top boundary conditions for a lattice model, let $\tilde{\lambda} + \rho$ be the boundary conditions such that if we flip the $\tilde{\lambda} + \rho$ upside down its bottom row decorations match with the decorations of $\lambda + \rho$. If Z denotes the partition function for the new lattice model then we have

$$Z = \sum_\mu Z_\mu(\underline{x}) Z_{\tilde{\mu}}(\underline{x}),$$

where μ ranges over all partitions corresponding to the gluing. This identity is called the dual Cauchy identity.

Algebras, Coalgebras, and Bialgebras.

Let A be an algebra over k. In the remainder of these notes, by an algebra we always mean and associative algebra with unit. That is, a k-vector space A with an associative multiplication, by which we mean a k-linear map $m : A \otimes A \to A$, and a unit element 1_A with $m(a, 1_A) = m(1_A, a)$ for all $a \in A$.[33] In terms of diagrams, associativity of the algebra means that the following diagram commutes:

$$\begin{tikzcd} & A \otimes A \otimes A \arrow[dl, "m\otimes\mathrm{id}"'] \arrow[dr, "\mathrm{id}\otimes m"] & \\ A \otimes A \arrow[dr, "m"'] & & A \otimes A \arrow[dl, "m"] \\ & A & \end{tikzcd}$$

For any $a \in A$ there exists a map

$$\eta_a : k \to A \qquad \lambda \mapsto \lambda a.$$

So $\eta_a(1) = a$ and if there exists a unit 1_A, we have $\eta(1) := \eta_{1_A}(1) = 1_A$. In diagrams, A possessing a unit is equivalent to a map η such that the following diagrams commute

$$\begin{tikzcd} A \otimes A \arrow[dr, "m"] & \\ k \otimes A \arrow[u, "\eta\otimes\mathrm{id}"] \arrow[r, "\sim"] & A \end{tikzcd} \qquad \begin{tikzcd} A \otimes A \arrow[dr, "m"] & \\ A \otimes k \arrow[u, "\mathrm{id}\otimes\eta"] \arrow[r, "\sim"] & A \end{tikzcd}$$

With this in mind we can identify the multiplicative unit in the algebra with the multiplicative unit in k. We will call the unit 1. Let us give some examples of algebras.

Example (Algebras)**.** We have two primary examples of interest.

- Let k be a field and G be any finite group. Then the group algebra $k[G]$ is defined as the k-vector space with basis $\{e_g\}_{g\in G}$ with multiplication $e_g e_h = e_{gh}$ where gh denotes multiplication in G.
- Let V be a vector space. Then the tensor algebra $T(V) = \bigoplus_{i=0}^{\infty} V^{\otimes i}$ (which we have already encountered) can be made into an algebra. If we think of the elements of $T(V)$ as linear combinations of finite strings of vectors then the multiplication operation is concatenation of strings.

[33]We usually suppress the multiplication notation.

We would now like to define the dual notation of an algebra, namely a coalgebra. A coalgebra C over a field k is a k-vector space with an operation $\Delta : C \to C \otimes C$ called the coproduct. All our coalgebras will be assumed to be coassociative and with counit. In other words, the following diagram commutes

$$\begin{tikzcd}
& C \otimes C \otimes C & \\
C \otimes C \arrow[ur, "\Delta\otimes\mathrm{id}"] & & C \otimes C \arrow[ul, "\mathrm{id}\otimes\Delta"'] \\
& C \arrow[ul, "\Delta"] \arrow[ur, "\Delta"'] &
\end{tikzcd}$$

and there exists a map $\epsilon : C \to k$, called the counit, such that the two diagrams below commute

$$\begin{tikzcd}
C \otimes C \arrow[d, "\epsilon\otimes\mathrm{id}"'] & \\
k \otimes C \arrow[r, "\sim"] & C \arrow[ul, "\Delta"']
\end{tikzcd}
\qquad
\begin{tikzcd}
C \otimes C \arrow[d, "\mathrm{id}\otimes\epsilon"'] & \\
C \otimes k \arrow[r, "\sim"] & C \arrow[ul, "\Delta"']
\end{tikzcd}$$

If we label the first copy in $C \otimes C$ by $C_{(1)}$ and the second copy by $C_{(2)}$ then we may describe the comultiplication map Δ as

$$\Delta : C \to C_{(1)} \otimes C_{(2)} \qquad c \mapsto \sum_{i=1}^{n} c^i_{(1)} \otimes c^i_{(2)}.$$

This is often too verbose in the literature. Comultiplication will usually be written as $\Delta(c) = \sum_{i=1}^{n} c^i_{(1)} \otimes c^i_{(2)}$ or $\Delta(c) = c_{(1)} \otimes c_{(2)}$ where the latter notation is horrible but indicates that the structure is completely linear so we only need to check on each term. The two latter notations introduced are called "Sweedler Notation". Lets give some examples of coalgebras.

Before we give some examples of coalgebras we would like to note that given two algebras A_1 and A_2 (or coalgebras C_1 and C_2) we can construct a new algebra (or coalgebra) where the underlying vector space is $V_1 \otimes V_2$ (or $C_1 \otimes C_2$) with multiplication map $m((a_1 \otimes a_2), (a'_1 \otimes a'_2)) = m_1(a_1, a'_1) \otimes m_2(a_2, a'_2)$ (or $\Delta(c_1 \otimes c_2) = \Delta_1(c_1) \otimes \Delta_2(c_2)$) where m_1 and m_2 are the multiplications on A_1 and A_2 respectively (or Δ_1 and Δ_2 are the comultiplications on C_1 and C_2 respectively). The units (or counits) are obvious.

Example (Coalgebras)**.** We will be primarily concerned with two examples of coalgebras.

- The algebra $k[G]$ can also be realized as a coalgebra with comultiplication defined on the basis by $\Delta(e_g) = e_g \otimes e_g$ for all $g \in G$ and counit $\epsilon(e_g) = 1$ for all $g \in G$.

- The tensor algebra $T(V)$ can also be realized as a coalgebra in two ways. The first is to define Δ by

$$\Delta(v_1 \otimes \cdots \otimes v_\ell) = \sum_{j=0}^{\ell} (v_0 \otimes \cdots \otimes v_j) \otimes (v_{j+1} \otimes \cdots \otimes v_\ell),$$

  where $v_0 = v_\ell = 1$, and then extend linearly. The second is to first define Δ by $\Delta(v) = v \otimes 1 + 1 \otimes v$ where we consider $v \otimes 1, 1 \otimes v \in T(V) \otimes V$ for all $v \in T^1(V) = T(V)$[34]. We then extend this definition recursively by $\Delta(v_1 \otimes v_2) = \Delta(v_1) \otimes \Delta(v_2)$ where $v_1 \otimes v_2 \in T^2(v)$.

These two coalgebras are not the same, but they do have the same unit ϵ given by $\epsilon(v) = 0$ for all $v \in V$ and $\epsilon(m) = m$ for all $m \in k$.[35]

We also have the notion of a biaglebra. A bialgebra B over a field k is both an algebra over k and a coalgebra over k in which Δ and ϵ are algebra maps (i.e., m and η are coalgebra maps) such that the following conditions are satisfied

1. $\Delta \circ m = (m \otimes m) \circ (\Delta \otimes \Delta)$.
2. Δ preserves the identity.
3. $\epsilon \circ m = m \circ (\epsilon \otimes \epsilon)$.
4. ϵ preserves the identity.

There is a commutative diagram describing each of these properties and we encourage the interested reader to formulate these properties in terms of commutative diagrams. Let's now see some examples of Bialgebras.

Example (Bialgebras)**.** Below are our two primarily examples of bialgebras.

- The group algebra $k[G]$ can be considered as a bialgebra by checking properties $(1)-(4)$ directly.

- The second coalgebra structure on $T(V)$ makes it into a bialgebra while the first one does not.

[34] We are being purposely verbose here.
[35] In other words, ϵ is just projection onto the $V^{\otimes 0} = k$ factor.

Hopf Algebras.

We will state some more facts about bialgebras. In general, given any finite set S one may form the k-vector space with basis $\{v_s\}_{s\in S}$ and define a coalgebra structure on it by $\Delta(v_s) = v_s \otimes v_s$. In particular, for the group algebra $k[G]$, this give a bialgebra.[36]

We would like to illustrate a nice property of bialgebras. Given a bialgebra B and two left B-modules V and W, $V \otimes W$ has a natural left $(B \otimes B)$-module structure. If we compose with the coalgebra map $\Delta : B \to B \otimes B$ we get a natural B-algebra structure on $V \otimes W$ defined by

$$b(v \otimes w) := \Delta(b)(v \otimes w) \quad (\text{for } v \otimes w \in V \otimes W),$$

where $\Delta(b)(v \otimes w)$ acts by the natural left $(B \otimes B)$-module structure on $V \otimes W$. We now introduce Hopf algebras. A Hopf algebra H over k is a bialgebra over k with a map $s : H \to H$, called the antipode map, such that $m \circ (s \otimes \mathrm{id}) \circ \Delta = \eta \circ \epsilon = m \circ (\mathrm{id} \otimes s) \circ \Delta$. Equivalently, the following diagram commutes:

$$\begin{tikzcd}
H \arrow[r, "\eta\circ\epsilon"] \arrow[d, "\Delta"'] & H \\
H \otimes H \arrow[r, "\mathrm{id}\otimes s", "s\otimes\mathrm{id}"'] & H \otimes H \arrow[u, "m"']
\end{tikzcd}$$

Let's give some examples of Hopf algebras.

Example. In our running example we will have the following Hopf algebras:

- We can make the bialgebra $k[G]$ into a Hopf algebra by defining the antipode map s by $e_g \mapsto e_{g^{-1}}$ on the basis and then extending linearly.[37]
- We can make the bialgebra $T(V)$ into a Hopf algebra by defining $s(v) = -v$ in $T^1(V)$ and then extending by the universal property of tensor algebras[38].

The commutative diagram above can also be thought of in the following way. It says that we can give $\mathrm{End}(H)$ an associative algebra structure by where given $f, g \in \mathrm{End}(H)$, we define their product $f * g$ by the map $m \circ (f \otimes g) \circ \Delta$. In this algebra, $\eta \circ \epsilon$ is the identity.

The antipode map s also has the following properties:

[36]Elements $x \in B$ of a bialgebra such that $\Delta(x) = x \otimes x$ are called grouplike.
[37]The antipode map s can be thought of as a sort of inverse, but generally $s^2 \neq \mathrm{id}$.
[38]We are not extending linearly.

1. s is an antialgebra map. In other words, $s(hg) = s(g)s(h)$ for all $g, h \in H$, and $s(1_H) = 1_H$.

2. s is an anticoalgebra map. That is, $(s \otimes s) \circ \Delta = \epsilon$.

3. s is unique if it exists.

We would like to state some other properties of Hopf algebras. We say H is commutative if the algebra multiplication is commutative. In other words, the following diagram commutes:

$$\begin{tikzcd} H \otimes H \arrow[r, "\tau"] \arrow[d, "m"'] & H \otimes H \arrow[dl, "m"] \\ H & \end{tikzcd}$$

In this diagram $\tau : H \otimes H \to H \otimes H$ is the map defined by $h \otimes g \mapsto g \otimes h$ for all $h, g \in H$.[39] We say H is cocommutative if $\tau \circ \Delta = \Delta$. This may be phrased as saying that the following diagram commutes:

$$\begin{tikzcd} H \otimes H \arrow[r, "\tau"] & H \otimes H \\ H \arrow[u, "\Delta"] \arrow[ur, "\Delta"'] & \end{tikzcd}$$

We also have the following properties of Hopf algebras:

1. If H is commutative and cocommutative then $s^2 = \mathrm{id}$.

2. If H cocommutative then then H-structure on $V \otimes W$ (where V and W are left H-modules) is not isomorphic to that of $W \otimes V$.

For quantum group we would like to relax the condition that $V \otimes W \cong W \otimes V$ as left H-modules (i.e., relax cocommutativity). Let's give some examples of these properties.

Example. If G is abelian, $k[G]$ as a Hopf algebra is commutative; $k[G]$ is always cocommutative. As a Hopf algebra, $T(V)$ is commutative if and only if $\dim(V) \leq 1$; $T(V)$ is always cocommutative.

[39] It is sometimes called the swap map or interchange map.

Now recall the definition of a Lie algebra. If $\mathfrak{g}$ is a Lie algebra, it is a vector space with a bracket operation $[\cdot,\cdot] : \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$. We can form the tensor algebra $T(\mathfrak{g})$ and quotient by the ideal generated by all relations of the form $[g_1, g_2] - (g_1 \otimes g_2) + (g_2 \otimes g_1)$ to get the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$[40] as explained previously.

We end with three comments of Majid:

1. Any theorem, true for both group algebras and universal enveloping algebras, is true for all cocommutative Hopf algebras.

2. Sweedler has a book with a lot of results on the classification of finite dimensional Hopf algebras, but there remain many open questions.

3. There is very little supply of noncommutative and noncocommutative Hopf algebras until Drinfeld and Jimbo gave examples examples of $\mathcal{U}_q(\mathfrak{g})$ for nice Lie algebras $\mathfrak{g}$.

Left Modules and Dual Hopf Algebras.

Given to left H-modules V and W (considered as algebras) recall that $V \otimes W$ inherits a left H-module structure, and if H is cocommutative[41] $V \otimes W \cong W \otimes V$ as left H-modules. In general, $V \otimes W$ and $W \otimes V$ may be quite different as left H-modules.[42]

If M is a left H-module, consider the dual $M^* = \mathrm{Hom}_k(M, k)$. We can define a left action[43] on M^* given by

$$(hf)(m) := f(s(h)m) \quad \text{(for all } h \in H \text{ and } m \in M).$$

If s is invertible, then we can define a second left action by

$$(hf)(m) := f(s^{-1}(h)m) \quad \text{(for all } h \in H \text{ and } m \in M).$$

These two actions do not give rise to isomorphic H-modules.

We would now like to give an example of another Hopf algebra.

[40] This algebra contains all the of representation theory of $\mathfrak{g}$.

[41] As an aside, $s^2 = \mathrm{id}$ if H is commutative or cocommutative, and s is bijective if H is finite dimensional.

[42] When we define quantum groups, we want to limit how bad this difference in left H-module structure can be.

[43] For a general algebra A we get a right action on M^*.

Example. For reasons which will become clear, define $\mathcal{U}_q(\mathfrak{b}_+)$ by

$$\mathcal{U}_q(\mathfrak{b}_+) := \langle X, K, K^{-1} \mid KK^{-1} = 1 = K^{-1}K, KX = qXK\rangle^{44},$$

and define maps Δ, ϵ, and s by

$$\Delta(X) = (X \otimes 1) + (K \otimes X), \quad \Delta(K) = K \otimes K, \quad \Delta(K^{-1}) = K^{-1} \otimes K^{-1},$$
$$\epsilon(X) = 0, \quad \epsilon(K) = \epsilon(K^{-1}) = 1,$$
$$s(X) = -K^{-1}X, \quad s(K) = K^{-1}, \quad s(K^{-1}) = K.$$

It is a fact that this set of relations and definitions extends to define an infinite dimensional[45] noncommutative noncocommutative Hopf algebra. We sketch a proof of this fact below.

Proof sketch. Extend Δ and ϵ and check that this gives algebra maps. Then extend s as an antialgebra map and check the axioms for s on generators. Then the axioms for s on products will follow automatically because Δ and ϵ are well-defined multiplicatively. □

It can also be checked that for any $u \in \mathcal{U}_q(\mathfrak{b}_+)$ we have $s^2(u) = K^{-1}uK$ so s is bijective and hence invertible.

It is best to think of $\mathcal{U}_q(\mathfrak{b}_+)$ as a q-deformation of the universal enveloping algebra $\mathcal{U}(\mathfrak{b}_+)$ where $\mathfrak{b}_+$ is the Lie algebra of the Boreal subgroup of $\mathrm{SL}(2,\mathbb{C})$. For the reader unfamiliar with Borel subgroups, the Boreal subgroup B of $\mathrm{SL}(2,\mathbb{C})$ is

$$B = \left\{ \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \mid a, b \in \mathbb{C} \right\}.$$

The Lie algebra $\mathfrak{b}_+$ has as generators

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

In the deformation $D \to K, K^{-1}$ and $X \to X$.

In the setting of finite dimensional Hopf algebras (which we are rarely in) we have a dual space H^{*}[46] and a canonical isomorphism between $(H^*)^*$ and H. We also have a natural map $\langle \cdot, \cdot \rangle : H^* \otimes H \to k$, called the evaluation map, defined by

$$\langle \phi, v \rangle := \phi(v).$$

[44] Here q is an arbitrary element of k.

[45] The set $\{x^n\}_{n \in N}$ can be extended to a basis.

[46] This is the vector space dual.

Given H we can formulate all the axioms for the Hopf algebra structure of H^* in terms of the evaluation map. Indeed, the Hopf algebra axioms for H^* are equivalent to the following identities

1. $\langle \phi \circ \psi, a \rangle = \langle \phi \circ \psi, \Delta(a) \rangle$.

2. $\langle 1, a \rangle = \epsilon(a)$.

3. $\langle \Delta \circ \phi, a \otimes b \rangle = \langle \phi, m(a \otimes b) \rangle$.

4. $\langle \phi, 1 \rangle = \epsilon(\phi)$.

5. $\langle s \circ \phi, a \rangle = \langle \phi, s(a) \rangle$.

We encourage the reader to pair the identities with the corresponding Hopf algebra axioms. In general (i.e., in the infinite dimensional setting), we only know $(H \otimes H)^* \cong H^* \otimes H^*$. We say that two Hopf algebras H and H' are dually paired if there exists a nondegenerate pairing $\langle \cdot, \cdot \rangle : H \otimes H \to k$[47] satisfying the five properties above. In general, H may have several dual pairings.

Examples of Dual Hopf Algebras.

We would like to give several examples of dual pairings.

Example (Dual pairings)**.**

- As we have already said, if H is a finite dimensional Hopf algebra it has a dual Hopf algebra H^* which is automatically dually paired to H.

- Consider the Hopf algebra $k[G]$. It is dually paired with the Hopf algebra $k(G)$ where the underling algebra structure of $k(G)$ is the algebra of functions on G with pointwise multiplication.

- If $\mathfrak{g}$ is a finite dimensional complex semisimple Lie algebra (think $\mathfrak{sl}_2$) with Lie group G, then the dual of $\mathcal{U}(\mathfrak{g})$ is the coordinate algebra $\mathbb{C}[G]$ for G over $\mathbb{C}$ defined by
$$\mathbb{C}[G] = \mathbb{C}[x_{i,j}]_{1 \le i \le j \le n} / (p(\underline{x}))$$

[47] This is not necessarily an evaluation map.

where $\underline{x} = (x_{i,j})_{1 \le i \le j \le n}$ and $(p(\underline{x}))$ is the ideal generated by the polynomial equations which give embeddings of G into $\mathrm{Mat}_n(\mathbb{C})$. Its Hopf algebra structure is given on generators by

$$\Delta(x_{i,j}) = \sum_{k=1}^{n} (x_{i,k} \otimes x_{k,j}) \quad \text{and} \quad \epsilon(x_{i,j}) = \delta_{i,j}$$

where $\delta_{i,j}$ is the Kronecker delta. We omit the definition of s because it is more complicated to define and dependent on cofactors of the matrix $[x_{i,j}]$. The pairing between these Hopf algebras is given by extending the pairing

$$\langle \alpha, x_{i,j} \rangle = \rho(\alpha)_{i,j} \quad (\text{for all } \alpha \in \mathfrak{g}),$$

where $\rho : \mathfrak{g} \to \mathrm{Mat}_n(\mathbb{C})$ is the defining representation, to $\mathcal{U}(\mathfrak{g})$.

It is a general fact that $\mathcal{U}_q(\mathfrak{b}_+)$ is self-dual. We would like to explore this fact, but first let us provide reasoning for why duality is important.

Hopf duality sets up a duality between left H-modules and right H-comodules.[48] Moreover, there are several constructions of $\mathcal{U}_q(\mathfrak{g})$ using duality. We state two below:

- There is a natural action of the Hopf algebra $\mathrm{SL}_q(2)$[49] on the quantum plane[50], and the dual is $\mathcal{U}_q(\mathfrak{sl}_2)$.
- There is a construction of Drinfeld where given a Hopf algebra H and a dual H^* of H, one can form a Hopf algebra $D(H, H^*)$. For finite dimensional H and H^*, $D(H, H^*)$ is guaranteed to be a quasi-triangular[51] Hopf algebra.

Let's being our discussion of $\mathcal{U}_q(\mathfrak{b}_+)$ with a couple of facts.

Proposition. $\mathcal{U}_q(\mathfrak{b}_+)$ has $\{X^m K^n\}_{\substack{m \ge 0 \\ n \in \mathbb{Z}}}$ as a k-basis.

Proof sketch. Show $\{X^m K^n\}_{\substack{m \ge 0 \\ n \in \mathbb{Z}}}$ is a spanning set by checking that monomials are stable under multiplication by any element of $\mathcal{U}_q(\mathfrak{b}_+)$ (we can check this on generators). For linear independence, consider the commutative ring $R := k[A, B, B^{-1}]$. with basis $\{A^m B^n\}_{\substack{m \ge 0 \\ n \in \mathbb{Z}}}$. We have endomorphisms $f, g : R \to R$ defined on basis elements by

$$f(A^m B^n) = A^{m+1} B^n \quad \text{and} \quad g(A^m B^n) = q^m A^m B^{n+1}.$$

[48]This is to be defined.
[49]We have not discussed this Hopf algebra yet.
[50]Think of the affine plane with the relation $xy = qyx$.
[51]That is, it gives rise to an abstract quantum Yang-Baxter equation.

Observe g has an inverse g^{-1} given by

$$g^{-1}(A^m B^n) = q^{-m} A^m B^{n-1}.$$

Check $g \circ f = q(f \circ g)$. This implies that we have a map $\mathcal{U}_q(\mathfrak{b}_+) \to \mathrm{End}_k(R)$ defined on generators by $X \mapsto f$, $K \mapsto g$, and $K^{-1} \mapsto g^{-1}$. Now check f, g, and g^{-1} are linearly independent in $\mathrm{End}_k(R)$ implying X, K, and K^{-1} are linearly independent in $\mathcal{U}_q(\mathfrak{b}_+)$. $\square$

The second fact is as follows:

Proposition. In $\mathcal{U}_q(\mathfrak{b}_+)$ we have the identity

$$\Delta(X^m) = \sum_{r=0}^{m} \begin{bmatrix} r \\ m \end{bmatrix}_q X^{m-r}(K^r \otimes X^r),$$

where

$$\begin{bmatrix} r \\ m \end{bmatrix}_q = \frac{[m]_q!}{[r]_q![m-r]_q!},$$

with $[r]_q! = [r]_q[r]_{q-1} \cdots [1]_q$ and $[r]_q = (1-q^r)/(1-q)$.[52] We also make the convention that

$$\begin{bmatrix} m \\ m \end{bmatrix}_q = \begin{bmatrix} 0 \\ m \end{bmatrix}_q = 1.$$

Proof sketch. Recall that Δ is defined by extending it on basis elements multiplicatively with the base condition $\Delta(X) = (X \otimes 1) + (K \otimes X)$. Then apply the q-binomial formula to summands and use induction:

$$(A+B)^n = \sum_{m=0}^{n} \left[\frac{n}{m}\right]_q A^m B^{n-m} \quad (\text{if } qAB = BA).$$

$\square$

With these two properties we can show $\mathcal{U}_q(\mathfrak{b}_+)$ is self-dual. Indeed, we have the following proposition and proof sketch due to Majid:

Proposition. The identities

$$\langle K, K \rangle = q, \quad \langle X, X \rangle = 1, \quad \text{and} \quad \langle X, K \rangle = \langle X, K \rangle = 1,$$

uniquely determines non-degenerate pairing between $\mathcal{U}_q(\mathfrak{b}_+)$ and itself satisfying properties (1)-(5) of a dual pairing.

[52] We say $\begin{bmatrix} r \\ m \end{bmatrix}_q$ is a q-binomial coefficient.

Proof sketch. We're going to construct an extension of the pairing by assuming some functions satisfy the pairing properties, define the pairing using this extension, and then recheck that it satisfies all the properties. For $m \geq 0$ and $n \in \mathbb{Z}$, define functions

$$f_{m,n}(u) = \langle X^m K^n, u\rangle \quad (\text{for all } u \in \mathcal{U}_q(\mathfrak{b}_+)),$$

and assume they satisfy the pairing properties as well as the identities in the proposition. First evaluate $f_{m,n}$ on X and K. It can be checked that

$$f_{m,n}(K) = q^n \delta_{m,0} \quad \text{and} \quad f_{m,n}(X) = \delta_{m,1}$$

where δ is the Kronecker delta. Now show for any $u, u' \in \mathcal{U}_q(\mathfrak{b}_+)$, that

$$f_{m,n}(uu') = \sum_{r=0}^{m} \begin{bmatrix} r \\ m \end{bmatrix}_q f_{m-r,n+r}(u) f_{r,n}(u').$$

Then extend the pairing by the identity above (which uniquely determines it). Now go back and check all the pairing properties are satisfied on the basis elements $\{X^m K^n\}_{\substack{m\geq 0 \\ n\in\mathbb{Z}}}$. $\square$

Quasitriangular Hopf Algebras.

We are going to define quantum groups in this section, but to do that we first need to introduce quasitriangular Hopf algebras.

Recall that if B is a bialgebra, then given two left B-algebras V and W we can form a left B-algebra $V \otimes W$. In category-theoretic language we say that left B-modules with left B-module homomorphisms form a left monodial[53] category. Since comultiplication is associative, the operation $(V, W) \mapsto V \otimes W$ is associative. So we really have a associative left monodial category. Also recall that $V \otimes W$ and $W \otimes V$ may have very different structures as left B-modules. Moreover, if B is cocommutative then $\tau : V \otimes W \to W \otimes V$ gives a left B-module isomorphism between $V \otimes W$ and $W \otimes V$. Now suppose that $V \otimes W \cong W \otimes V$ via some (not necessarily τ) left B-module isomorphism. More generally, suppose we have a natural family of left B-module isomorphisms $t_{V,W} : V \otimes W \to W \otimes V$, for all B-modules V and W, that is compatible with associativity.[54] By this we mean that the following diagram (and a dual one with the subscripts interchanged) is commutative

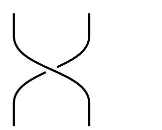
[53]We say monodial since the operation taking V and W to $V \otimes W$ defines a monoid on the space of B-modules.

[54]This is more general since we can think of $\tau : V \otimes W \to W \otimes V$ as a family of isomorphisms if we vary V and W across all left B-modules.

$$\begin{tikzcd}
V \otimes (W \otimes X) \arrow[r, "t_{V,W\otimes X}"] & (W \otimes X) \otimes V \\
(V \otimes W) \otimes X \arrow[u] \arrow[d, "t_{V,W}\otimes \mathrm{id}"'] & W \otimes (X \otimes V) \arrow[u] \\
(W \otimes V) \otimes X \arrow[r] & W \otimes (V \otimes X) \arrow[u, "\mathrm{id}\otimes t_{V,X}"']
\end{tikzcd}$$

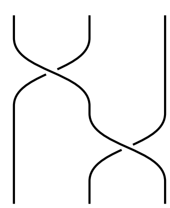
where all the unlabeled arrows are the obvious associative isomorphisms on tensor products of modules. This more general structure, in the category theoretic language, is called a braided left monodial category. We give an example to illustrate why we call this category braided.

Example (Braided monodial category)**.** The prototypical example of a braided monodial category comes from braids on n-strings which justifies the name. Let the objects of this category be elements of $\mathbb{Z}_{\geq 0}$ and the morphisms be braids between them.[55] For example, the braid



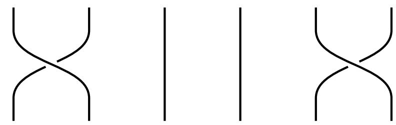
is an example of a morphism between the object 3. Composition of braids is given by stacking. For example composing the braids

produces the braid



The monodial operation is given by concatenating strings. For example, concatenation of the two braids above (instead of composition) produces the braid

[55]By the definition of a braid this means there are only morphisms between m and n if $m = n$.



We encourage the interested reader to deduce what the isomorphism t_{b_1,b_2} is for arbitrary braids b_1 and b_2.

If $t_{V,W} : V \otimes W \to W \otimes V$ is a braiding[56], then consider the map $t_{B,B} B \otimes B \to B \otimes B$ and let $R = t_{B,B}(1 \otimes 1)$. Conversely given any $R = \sum_{i=1}^{n} b^i_{(1)} \otimes b^i_{(2)} \in B \otimes B$, we can construct a family of morphisms

$$t^{(R)}_{V,W} : V \otimes W \to W \otimes V \qquad v \otimes w \mapsto \sum_{i=1}^{n} ((b^i_{(1)} w) \otimes (b^i_{(2)} v)).$$

In fact, this family of morphisms is a braiding if and only if R satisfies the following three conditions

1. R is invertible.
2. $\tau \circ \Delta = R \Delta R^{-1}$, where $R \Delta R^{-1}$ takes h to $R \Delta(h) R^{-1}$ for all $h \in H$.
3. $(\Delta \otimes \mathrm{id})(R) = R_{(1,3)} R_{(2,3)}$ and $(\mathrm{id} \otimes \Delta)(R) = R_{(1,3)} R_{(1,2)}$ where $R_{(i,j)}$ is the image of R under the algebra morphism $\phi_{(i,j)} : B \otimes B \to B \otimes B \otimes B$ which acts by id on the i and j copies of B inside $B \otimes B \otimes B$ and trivially on the other copy.

Moreover, there is a bijective correspondence between R satisfying these properties and braidings.

We may now define a quasitriangular Hopf algebra over k. A quasitriangular Hopf algebra over k (also known as a quantum group over k) is a pair (H, R) consisting of a Hopf algebra H over k and an element $R \in H \otimes H$ such that the three properties above are satisfied.[57]

Quasitriangular Hopf Algebras and Quantum Yang-Baxter Equations.

We would like to state and sketch the proof of a lemma about quasitriangular Hopf algebras.

[56] By this we mean the family of maps $t_{V,W} : V \otimes W \to W \otimes V$ gives rise to a braided left monodial category of left B-modules and left B-module homomorphisms.

[57] We will not go into detail about why we call this structure quasitriangular.

Lemma. If (H, R) is a quasitriangular Hopf algebra, then we have the following:

- $(\epsilon \otimes \mathrm{id})(R) = (\mathrm{id} \otimes \epsilon)(R) = 1$, $(s \otimes \mathrm{id})(R) = R^{-1}$, and $(\mathrm{id} \otimes s)(R^{-1}) = R$.
- $(H, \tau(R^{-1}))$ is a quasitriangular Hopf algebra.
- In $H \otimes H \otimes H$ we have the abstract quantum Yang-Baxter equation

$$R_{(1,2)}R_{(1,3)}R_{(2,3)} = R_{(2,3)}R_{(1,3)}R_{(1,2)}.$$

Proof sketch. We will sketch each property individually.

1. We will show $(e \otimes \mathrm{id})(R) = 1$; the other cases are handled similarly. By the properties of quasitriangular Hopf algebras we have

$$(\Delta \otimes \mathrm{id})(R) = R_{(1,3)}R_{(2,3)} \quad \text{and} \quad (\epsilon \otimes \mathrm{id}) \circ \Delta = \mathrm{id}.$$

   Apply $(\epsilon \otimes \mathrm{id}) \circ \mathrm{id}$ to both side of the first of the two identities above. The right-hand side is R. The left-hand side is

$$((\epsilon \otimes \mathrm{id}) \circ \mathrm{id})(R_{(1,3)}R_{(1,2)}) = (\epsilon \otimes \mathrm{id})(R)\epsilon(1)R.$$

   Now use the fact that R is invertible.

2. This is straightforward to check.

3. We know

$$R_{(1,2)}R_{(1,3)}R_{(2,3)} = R_{(1,2)}(\Delta \otimes \mathrm{id})(R) \quad \text{and} \quad R_{(1,2)}\Delta R^{-1}_{(1,2)}$$

   where $R_{(1,2)}\Delta R^{-1}_{(1,2)}$ sends h to $R_{(1,2)}\Delta(h)R^{-1}_{(1,2)}$ for all $h \in H$. Therefore

$$\begin{aligned} R_{(1,2)}R_{(1,3)}R_{(2,3)} &= R_{(1,2)}(\Delta \otimes \mathrm{id})(R) \\ &= (\tau \otimes \Delta \otimes \mathrm{id})(R)R_{(1,2)} \\ &= ((\tau \otimes \mathrm{id}) \circ (\Delta \otimes \mathrm{id}))(R)R_{(1,2)} \\ &= (\tau \otimes \mathrm{id})(R_{(1,3)}R_{(2,3)})R_{(1,2)} \\ &= R_{(2,3)}R_{(1,3)}R_{(1,2)}. \end{aligned}$$

   Observe that in the second to last line τ flipping the subscripts of the $R_{(i,j)}$.

□

Observe that this connects quasitriangular Hopf algebras and quantum Yang-Baxter equations! We have now shown that a quasitriangular Hopf algebra gives rise to an abstract quantum Yang-Baxter equation. In fact, we get an honest quantum Yang-Baxter equation by choosing a representation (ρ, V) for H and then applying $\rho \otimes \rho$ to R. That is, the matrix $(\rho \otimes \rho)(R)$ satisfies a quantum Yang-Baxter equation.

An Investigation of $\mathcal{U}_q(\mathfrak{sl}_2)$.

While we have defined quasitriangular Hopf algebras and stated some of their properties we have not yet given any examples. In the following we would like to do an in-depth investigation.

Consider, as previously, the algebra

$$\mathcal{U}_q(\mathfrak{sl}(2,k)) = \langle E, F, K, K^{-1} \mid {}_{KFK^{-1}=q^{-2}F, [E,F]=EF-FE=(K-K^{-1})/(q-q^{-1})}^{KK^{-1}=1, K^{-1}K=1, KEK^{-1}=q^2E,} \rangle.$$

This definition makes sense over any characteristic 0 field k with $q \neq 0, 1, -1$ (for example take $k = \mathbb{C}$ and $q \in \mathbb{C}^*$ not a unit). We would like to answer the following questions about $\mathcal{U}_q(\mathfrak{sl}(2,k))$:

1. Does $\mathcal{U}_q(\mathfrak{sl}(2,k))$ have a Hopf algebra structure?
2. Is $\mathcal{U}_q(\mathfrak{sl}(2,k))$ really a q-deformation of $\mathcal{U}(\mathfrak{sl}(2,k))$?
3. What are the finite dimensional left $\mathcal{U}_q(\mathfrak{sl}(2,k))$-modules and how do they compare to the left $\mathcal{U}(\mathfrak{sl}(2,k))$-modules?[58]
4. Is $\mathcal{U}_q(\mathfrak{sl}(2,k))$ quasitriangular and can we determine a method to compute $R \in H \otimes H$?[59]
5. Can we construct $\mathcal{U}_q(\mathfrak{sl}(2,k))$ by some natural process?

We will answer these questions over the next few sections.

To answer the first, we do get a Hopf algebra structure. The idea here mimics that of defining a Hopf algebra structure on $\mathcal{U}_q(\mathfrak{b}_+)$. We define Δ and ϵ on generators and extended multiplicatively, define s on generators and extend as an antialgebra, then check all relations. In particular,

$$\begin{gathered}
\Delta(E) = (E \otimes 1) + (K \otimes E), \quad \Delta(F) = (F \otimes K^{-1}) + (1 \otimes F), \\
\Delta(K) = K \otimes K, \quad \Delta(K^{-1}) = K^{-1} \otimes K^{-1}, \\
\epsilon(E) = \epsilon(F) = 0, \quad \epsilon(K) = \epsilon(K^{-1}) = 1, \\
s(E) = -K^{-1}E, \quad s(F) = -FK, \quad s(K) = K^{-1}, \quad s(K^{-1}) = K.
\end{gathered}$$

The answer to this question is "sort of". We can present $\mathcal{U}(\mathfrak{sl}_2)$ as

$$\mathcal{U}(\mathfrak{sl}_2) = \langle X, Y, H \mid [X,Y] = H, [H,X] = 2X, [H,Y] = -2Y \rangle.^{60}$$

[58] This will come down to two cases, either q is a root of unity or it is not.

[59] This will turn out to be no in general because R will be an infinite sum and hence not in the tensor product.

[60] Here X is similar to E, Y is similar to F, and H is similar to K and K^{-1} in $\mathcal{U}_q(\mathfrak{sl}_2)$.

We would hope that $\mathcal{U}_q(\mathfrak{sl}_2) \to \mathcal{U}(\mathfrak{sl}_2)$ as $q \to 1$, but this is broken because $q - q^{-1}$ is in the denominator of a relation for our presentation of $\mathcal{U}_q(\mathfrak{sl}_2)$. Here's how we fix this issue. It can be shown that $\mathcal{U}_q(\mathfrak{sl}_2)$ is isomorphic (as an algebra) to another algebra $\mathcal{U}'$ defined by

$$\mathcal{U}'_q := \langle E, F, K, K^{-1}, L \mid {}_{[L,E]=q(EK+K^{-1}E),[L,F]=-q^{-1}(FK+K^{-1}F)}^{R(\mathcal{U}_q(\mathfrak{sl}_2)),[E,F]=L,(q-q^{-1})L=K-K^{-1},} \rangle$$

where $R(\mathcal{U}_q(\mathfrak{sl}_2))$ is representing all the relations for $\mathcal{U}_q(\mathfrak{sl}_2)$.[61] The isomorphism $\varphi : \mathcal{U}' \to \mathcal{U}_q(\mathfrak{sl}_2)$ is defined by $E \mapsto E$, $F \mapsto F$, $K \mapsto K$, $K^{-1} \mapsto K^{-1}$, and $L \mapsto [E, F]$. Now $\mathcal{U}'_q$ at $q = 1$ satisfies $\mathcal{U}'_1 \cong \mathcal{U}(\mathfrak{sl}_2)[K]/(K^2 - 1)$ so we get a projection onto $\mathcal{U}_q(\mathfrak{sl}_2)$ via a map defined by $E \mapsto X$, $F \mapsto Y$, $K \mapsto 1$, and $L \mapsto H$.

The Representation Theory of $\mathcal{U}(\mathfrak{sl}_2)$.

We'd like to understand the representation theory of $\mathcal{U}(\mathfrak{sl}_2)$. We will need the following preliminary result before we dig into the representation theory:

Lemma. We have the two identities:

$$X^p H^q = (H - 2p)^q X^p \quad \text{and} \quad Y^p H^q = (H + 2p)^q Y^p.$$

Proof sketch. Use induction. □

A basis theorem will also be needed:

Theorem. $\{X^i Y^j H^k\}_{i,j,k \in \mathbb{Z}_{\geq 0}}$ is a k-basis for $\mathcal{U}(\mathfrak{sl}_2)$.

Proof sketch. This is in the same spirit as proving a k-basis for $\mathcal{U}_q(\mathfrak{b}_+)$ done earlier. □

Finally, we will need a result about the center of $\mathcal{U}(\mathfrak{sl}_2)$:

Theorem. $C := XY + YX + \frac{H^2}{2}$ is in the center of $\mathcal{U}(\mathfrak{sl}_2)$. In fact, C generates the center.

Proof sketch. All that needs to be checked is that brackets with C vanish, and checking this on generators is enough. The second statement follows from a theorem of Harish and Chandra. □

[61] We don't use this algebra because there are far more relations to check.

We would like to use these results to determine all the finite dimensional left $\mathcal{U}(\mathfrak{sl}_2)$-modules (throughout we assume $k = \mathbb{C}$ for simplicity). The theory of highest weight vectors and highest weight modules will be useful here. We recast the main definitions in terms of $\mathcal{U}(\mathfrak{sl}_2)$ for brevity. If V is a finite dimensional left $\mathcal{U}(\mathfrak{sl}_2)$-module, a nonzero vectors $v \in V$ is said to be of weight $\lambda \in \mathbb{C}$ if $Hv = \lambda v$. It is called a highest weight vector if in addition $Xv = 0$. We say V is a highest weight representation of $\mathcal{U}(\mathfrak{sl}_2)$ if it is generated by a highest weight vector.

We now state and sketch the proof of a proposition:

Proposition. Every finite dimensional left $\mathcal{U}(\mathfrak{sl}_2)$-module V has a highest weight vector.

Proof sketch. H has some eigenvector w with an eigenvalue say α.[62] If $xw = 0$ we are done. If not, consider the sequence $\{X^n w\}_{n\geq 0}$. By the first result, $X^n w$ is an eigenvector of H with eigenvalue $\alpha + 2n$. But v is finite dimensional so there can only be finitely many distinct eigenvalues so there exists an n such that $X^n w \neq 0$ and $X^m w = 0$ for $m > n$. This implies $X^{n+1} w$ is a highest weight vector. □

We can now state the main structure theorem for finite dimensional left $\mathcal{U}(\mathfrak{sl}_2)$-modules.

Theorem (Structure Theorem for finite dimensional left $\mathcal{U}(\mathfrak{sl}_2)$-modules)**.** The finite dimensional simple left $\mathcal{U}(\mathfrak{sl}_2)$-modules (up to isomorphism) are indexed by nonnegative integers n, call them $V(n)$, of dimension $n+1$ with highest weight vector v_n of weight n, and with the weight vectors of weights $n, n-2, \ldots, -n$ forming a $\mathbb{C}$-basis for $V(n)$. In particular, the action of X raises the weight of a weight vector by 2, the action of Y on a weight vector lowers the weight by 2, and C acts on the $V(n)$ by scalars where the scalar is $n(n+2)/2$.[63]

We note that the actions of X and Y do not undo one another. For example, if V is a finite dimensional simple left $\mathcal{U}(\mathfrak{sl}_2)$-module, with highest weight vector v, then $V = \langle v \rangle = \langle XYv \rangle$, but $XYv = v$ need not occur. However, the weight spaces are 1-dimensional so XYv and v differ by a scalar. Moreover, with this theorem $\{v_p = Y^p v_n / p!\}_{0\leq p\leq n}$ is a $\mathbb{C}$-basis for $V(n)$.

It is natural to ask the following question: given two finite dimensional simple left $\mathcal{U}(\mathfrak{sl}_2)$-modules $V(n)$ and $V(m)$, what is the decomposition of $V(n) \otimes V(m)$? Luckily, in this setting the question is not so difficult to answer as we have the following theorem:

[62] This relies on the fact that $k = \mathbb{C}$ is algebraically closed. The proof does hold in the general setting however.

[63] This last fact may be used to prove that any finite dimensional left $\mathcal{U}(\mathfrak{sl}_2)$-module is semisimple.

Theorem (Clebsch-Gordon Formula)**.** Given two finite dimensional simple left $\mathcal{U}(\mathfrak{sl}_2)$-modules $V(n)$ and $V(m)$,

$$V(n) \otimes V(m) \cong V(n+m) \otimes V(n+m-2) \otimes \cdots \otimes V(n-m).$$

Proof sketch. Check that the dimension of both sides is $(n+1)(m+1)$. This means we are done if we can find highest weight vectors for weights $n+m-2p$ where $0 \leq p \leq m$ in $V(n) \otimes V(m)$. If v and v' are the highest weight vectors for $V(n)$ and $V(m)$ respectively, let v_p and v'_p be the corresponding $\mathbb{C}$-basis vectors. Check that

$$\sum_{i=0}^{p} \frac{(m-p+i)!(n-i)!}{(m-p)!n!}(v_i \otimes v'_{p-i})$$

for $0 \leq p \leq m$ are the desired highest weight vectors. □

In general, quantum groups (like $\mathcal{U}_q(\mathfrak{sl}_2)$) will give rise to beautiful algorithms for computing decompositions of tensor products using bases like those just discussed.[64]

We end with a definition introducing the Hopf algebra point of view. If H is a Hopf algebra and A is an algebra (over k) then we say A is a left H-module algebra (or is a left Hopf module-algebra) if (as a vector space) A has a left H-module structure and $m : A \otimes A \to A$ and $\eta : k \to A$ are H-module maps in the sense that the following two properties are satisfied:

1. $hm(a \otimes b) = \sum_{i=1}^{n}(h^i_{(1)}a)(h^i_{(2)}b)$ for all $h \in H$ and $a, b \in A$.

2. $h1 = \epsilon(h)1$ for all $h \in H$.

Indeed, we say the last map is an H-module map because the counit gives k a natural H-module structure. In general, we have the following proposition:

Proposition. Any Hopf algebra H acts on on itself (as a left H-module algebra) where for all $h, g \in H$, we define

$$h \cdot g := m((id \otimes s)(\Delta(h)(g \otimes 1))) = \sum_{i=1}^{n} h^i_{(1)} g s(h^i_{(2)}).$$[65]

As above, and in general, we will specify the action with a $\cdot$ if a Hopf algebra is acting on itself (or notation is very similar) to avoid confusion with multiplication. Let's see a few examples:

[64] These are often called crystal bases or canonical bases in select texts.

[65] If H is commutative it can be checked that this action is trivial.

Example (Hopf module-algebras)**.**

- Consider the Hopf algebra $H = k[G]$. Since $\Delta(h) = h \otimes h$ and $s(h) = h^{-1}$, we have $h \cdot g = hgh^{-1}$.

- Let $\mathfrak{g}$ be a Lie algebra and consider the Hopf algebra $H = \mathcal{U}(\mathfrak{g})$. If $h, g \in \mathfrak{g} \subset \mathcal{U}(\mathfrak{g})$, then recall $\Delta(h) = (h \otimes 1) + (1 \otimes h)$ and $s(h) = -h$ and similarly for g. Then $h \cdot g = hg - gh$ which the advanced reader will realize is the adjoint action on $\mathfrak{g}$.

If H' is dually paired to H, then we can defined a left action of H' via the pairing $\langle \cdot, \cdot \rangle$ by

$$\phi \cdot h := \sum_{i=1}^{n} h^i_{(1)} \langle \phi, h^i_{(2)} \rangle.$$

We'd now like to state and sketch the proof of a lemma:

Lemma. For any Lie algebra $\mathfrak{g}$, an algebra A is a left Hopf module-algebra over $\mathfrak{g}$ if and only if A has a left $\mathfrak{g}$-module structure on which elements act by derivations.

Proof sketch. For the forward direction, recall that given $g \in \mathfrak{g}$, $\Delta(g) = (g \otimes 1) + (1 \otimes g)$. If we require

$$g(ab) = \sum_{i=1}^{n} (g^i_{(1)} a)(g^i_{(2)} b)$$

for all $a, b \in A$, then this becomes $x(ab) = x(a)b + bx(a)$. Conversely, there is a small lemma which states that if A is an Hopf module satisfying the unit properties of a module and satisfies the multiplication properties on generators, then it defines a module-algebra. Using this lemma all that needs to be checked are the multiplication properties on multiplication of generators[66]. □

From this lemma, we have a theorem which we will not give a proof for:

Theorem. Let $\mathcal{U}(\mathfrak{sl}_2)$ act on polynomials $p \in k[x, y]$ by

$$Xp = x\frac{\partial p}{\partial y}, \quad Yp = y\frac{\partial p}{\partial x}, \quad \text{and} \quad Hp = x\frac{\partial p}{\partial x} - y\frac{\partial p}{\partial y}.$$

This makes $k[x, y]$ into a left Hopf module-algebra over $\mathcal{U}(\mathfrak{sl}_2)$. The submodules $k[x, y]_n$ consist of homogeneous polynomials of degree n and are isomorphic to the simple modules $V(n)$ discussed previously.

This concludes our discussion of the representation theory of $\mathcal{U}(\mathfrak{sl}_2)$

[66] We are using multiplication in two different ways here.

The Representation Theory of $\mathcal{U}_q(\mathfrak{sl}_2)$

This is going to be in the same spirit as the previous section, but the arguments and material will be more difficult. We are going to need a few preliminary statements:

Lemma. There exists a unique automorphism ω of $\mathcal{U}_q(\mathfrak{sl}_2)$ sending $E \mapsto F$, $F \mapsto E$, $K \mapsto K^{-1}$, and $K^{-1} \mapsto K$ such that $w^2 = \mathrm{id}$.

Proof sketch. Since the automorphism is defined on generators, it is unique if it exists. It's now just a short check to see that ω is compatible with all the relations defining $\mathcal{U}_q(\mathfrak{sl}_2)$. □

This lemma will essentially cut the workload in half as can be seen below:

Lemma. For $m \geq 0$, we have the identity

$$[E, F^m] = [m]_q F^{m-1} \frac{q^{-(m-1)}K - q^{(m-1)}K^{-1}}{q - q^{-1}}$$

where $[m]_q = \frac{q^n - q^{-n}}{q - q^{-1}}$.

Proof sketch. Use induction. □

There is also a dual identity which is easily proved with the help of ω:

Lemma. For $m \geq 0$, we have the identity

$$[E^m, F] = [m]_q E^{m-1} \frac{q^{(m-1)}K - q^{-(m-1)}K^{-1}}{q - q^{-1}}$$

where $[m]_q = \frac{q^n - q^{-n}}{q - q^{-1}}$.

Proof sketch. Apply ω to the first lemma. □

We will also need a basis theorem:

Theorem. $\mathcal{U}_q(\mathfrak{sl}_2)$ has $\{F^i K^j E^\ell\}_{\substack{i,\ell \geq 0 \\ j \in \mathbb{Z}}}$ as a k-basis.

Proof sketch. The proof is in the same spirit as for $\mathcal{U}_q(\mathfrak{b}_+)$. □

We also have a short corollary:

Corollary. $\mathcal{U}_q(\mathfrak{sl}_2)$ has no zero divisors.

Proof sketch. Consider the subalgebra $\mathcal{U}_0 \subset \mathcal{U}_q(\mathfrak{sl}_2)$ defined by

$$\mathcal{U}_0 := \langle K, K^{-1} \rangle.$$

Observe any element of $\mathcal{U}_q(\mathfrak{sl}_2)$ is expressible as a linear combination of terms of the form $F^s h E^r$ where $h \in \mathcal{U}_0$ and $r, s \geq 0$. Say $u \in \mathcal{U}_q(\mathfrak{sl}_2)$ has leading term (r, s) if $F^s h E^r$ is a term in u and if all other terms $F^{s'} h E^{r'}$ have $s' < s$ or $s = s'$ and $r' < r$. Now show that if u has leading term (s, r) an u' has leading term (p, m), then uu' has leading term $(s + p, r + m)$. $\square$

We now want to classify finite dimensional left modules M for $\mathcal{U}_q(\mathfrak{sl}_2)$. For the moment we will assume the characteristic of the base field k is not 2 and $q \in k$ is not a root of unity. We let a weight vector be an eigenvector under the action of K. If λ is an eigenvalue of M, let $M_\lambda = \{m \in M \mid Km = \lambda m\}$. The relations for $\mathcal{U}_q(\mathfrak{sl}_2)$ imply $EM_\lambda \subseteq M_{q^2\lambda}$ and $FM_\lambda \subseteq M_{q^{-2}\lambda}$. So the space

$$\bigoplus_{n \in \mathbb{Z}} M_{q^{2n}\lambda}$$

is a left submodule of M for any choice of λ. We cannot guarantee any of these summands are nonzero, but if $M_\lambda \neq 0$ for some λ and M is simple this implies

$$M = \bigoplus_{n \in \mathbb{Z}} M_{q^{2n}\lambda}$$

where all but finitely many of the summands are nonzero. We now have a proposition:

Proposition. Let M be a finite dimensional left $\mathcal{U}_q(\mathfrak{sl}_2)$-module. Then there exist positive integers r and s such that $E^r M = 0$ and $F^s M = 0$.

Proof. We will prove $E^r M = 0$ as the proof for F is analogous. Let

$$M_f = \{m \in M \mid f(K)^n m = 0 \text{ for } n >> 0\}$$

where $f \in k[x]$ is an irreducible. Observe that if M_f and M_g are nonzero, then $M_f = M_g$ if and only if f is a constant multiple of g. Since M is finite dimensional

$$M = \bigoplus_{f \text{ irreducible}} M_f$$

where all but finitely many of the summands are nonzero. Given $f \in k[x]$ with $M_f \neq 0$, define $f_i(x) = f(q^i x)$. Observe f_i is irreducible as well. Since $EK = q^{-2}KE$,

$Ef(K) = f_{-2}(K)E$ and induction on r shows $E^r f(K) = f_{-2r}(K)E^r$ for all $r > 0$. This implies $E^r M_f \subseteq M_{f_{-2r}}$. So it suffices to show that for all irreducible f there exists an r such that $M_{f_{-2r}} = 0$ for then (by the finite dimensionality of M) we can take the maximum such r and we will be done. Suppose $M_{f_{-2r}} \neq 0$ for all $r > 0$. Because M is finite dimensional there exists a positive integer s such that $M_{f_{-2r}} = M_{f_{-2s}}$ for all $r \geq s$. But then for $r \geq s$, f_{-2r} and f_{-2s} differ by a nonzero scalar. If the polynomials have a nonzero constant term then $M_{f_{-2r}} = M_{f_{-2s}} = M_x$ but K is invertible so $M_x = 0$ which is a contradiction. So the constant term is zero. But then they differ in the leading term by $q^{2(s-r)n}$ for some n, but this is never 1 because q is not a root of unity. Hence $M_{f_{-2r}} = 0$ for some r. □

This leads us to a corollary which we will consider as the penultimate statement for this section (with our choice of q):

Corollary. Let M be a finite dimensional left $\mathcal{U}_q(\mathfrak{sl}_2)$-module. Then M is a direct sum of its weight spaces and all weights are of the form $\pm q^a$ with $a \in \mathbb{Z}$

Proof. Recall the linear algebra fact that an endomorphism of a finite dimensional vector space are diagonalizable if and only if its minimal polynomial splits into linear factors each appearing with multiplicity one. With this fact in mind, we will write down the minimal polynomial for the action of K on M. By the previous proposition, there exists $s > 0$ such that $F^s M = 0$. Recall that

$$[E, F] = \frac{K - K^{-1}}{q - q^{-1}}.$$

The relation above and the fact $F^s EM = 0$ together imply $F^{s-1}(FE)M = F^{s-1}(EF - \frac{K-K^{-1}}{q-q^{-1}})M = 0$. If we keep moving powers of F out and using the relation, then we get a polynomial which annihilates all of M. In particular, the s-th iteration of this process produces the polynomial

$$h_r^{(s)} = \prod_{j=-(r-1)}^{r-1} \frac{Kq^{r-s+j} - K^{-1}q^{-(r-s+j)}}{q - q^{-1}}.$$

By induction on $0 \leq r \leq s$, $F^{s-r}h_r^{(s)}M = 0$. Taking $r = s$ and simplifying the polynomial algebraically shows

$$h_s^{(s)} = \prod_{j=-(s-1)}^{s-1} (k - q^{-j})(k + q^{-j})$$

which is a polynomial with distinct linear factors of multiplicity one. The minimal polynomial also divides it so the minimal polynomial inherits these properties and this proves the corollary. □