Exactly solved models of statistical mechanics

Valerio Toledano Laredo

Notes from a graduate cours given at Northeastern University in Spring 2011

Contents

CHAPTER 1

The Enveloping, Affine and Loop Algebras of SL(2)

1. The Lie Algebra $\mathfrak{sl}_2$

1.1. Definition. The Lie Algebra $\mathfrak{sl}_2$ can be presented as the set of trace 0, 2×2 matrices:

$$\mathfrak{sl}_2 = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \middle| \, a, b, c \in \mathbb{C} \right\}$$

Then

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

form a basis of $\mathfrak{sl}_2$. Their lie algebra structure then is given by taking the commutator in multiplication and yields the identities $[h, e] = 2e$, $[h, f] = -2f$, and $[e, f] = h$.

1.2. Bilinear Form. Now, there exists a bilinear form $(\cdot, \cdot)$ on $\mathfrak{sl}_2$ given by $(x, y) = \mathrm{tr}_{\mathbb{C}^2}(xy)$. This form is symmetric, non-degenerate and invariant under the bracket: $([x, y], z) = -(y, [x, z])$. For example:

$$(e, e) = \mathrm{tr}\ e^2 = 0 = (f, f) \tag{1.1}$$

$$(e, f) = \mathrm{tr}\ ef = 1 = (f, e) \tag{1.2}$$

$$(h, h) = \mathrm{tr}\ h^2 = 2 \tag{1.3}$$

$$(h, e) = \mathrm{tr}\ he = 0 = (hf) \tag{1.4}$$

so the dual basis given via $(\cdot, \cdot)$ has f dual to e, and $\frac{h}{2}$ dual to h.

1.3. Casmiur Operator and Finite Dimensional Representations. The next three results are elementary and will not be proven. Let $\mathfrak{g}$ be a Lie Algebra with basis x^i and dual basis x_i for $i \in I$ some indexing set. Let V be a vector space and let $\rho : \mathfrak{g} \to \mathrm{End}\ (V)$ be a representation of $\mathfrak{g}$. Then the Casmiur operator of ρ is

$$C_{(\rho)} := \sum_i \rho(x^i)\rho(x_i) = \sum_i x^i x_i = x^i x_i$$

where the first equality comes from abuse of notation and the second comes from Einstein's summation convention:

1.3.1. *Proposition.* If ρ is an irreducible representation of $\mathfrak{g}$ and if $C_{(rho)} = 0$ $\forall x \in \mathfrak{g}$ then $C_{(\rho)}$ acts as a scalar on V.

1.3.2. *Proposition.* All finite dimensional representations of $\mathfrak{sl}_2$ are completely reducible.

1.3.3. *Proposition.* For all $n \in \mathbb{N}$ there exists a unique representation V_n of $\mathfrak{sl}_2$ of dimension $n+1$ given by $V_n \cong \mathbb{C}_n[x, y]$ (- the set of polynomials in two variables the sum of whose degree's is n) with

$$e \mapsto x\frac{\partial}{\partial y} \quad f \mapsto y\frac{\partial}{\partial x} \quad h \mapsto x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}$$

It is a simple check that with the basis $x^i y^{n-i}$ for $i = 0 \dots n$ this is a $n+1$ dimensional representation of $\mathfrak{sl}_2$. For Example, $V_1 = < x, y > = \mathbb{C}^2$.

1.4. Character Formula. From here on out, $\mathfrak{g} = \mathfrak{sl}_2$. Let V be a representation of V. Then V is the direct sum of eigen spaces of h: $V = \oplus_{\lambda \in \mathbb{C}} V_{[\lambda]}$ where $V_{[\lambda]}$ is the λ eigen space of h aka the weight space of λ and $\mathbb{C}$ is understood to be the 1-dimensional dual vector space of h, ie $h^* = \mathbb{C}h$.

Now, the character formula is $\mathrm{ch}(V) = \sum_\lambda \dim\, V_{[\lambda]} e^\lambda \in \mathbb{Z}h^*$ where e^λ is a formal symbol chosen for the fact that $e^{\lambda_1} e^{\lambda_2} = e^{\lambda_1 + \lambda_2}$. It can be shown that for an exact sequence of $\mathfrak{g}$-modules

$$0 \to U \to V \to W \to 0$$

the splitting of the sequence (true for any sequence of vector spaces) gives us $\mathrm{ch}(V) = \mathrm{ch}(U) + \mathrm{ch}(W)$. It can also be shown that if U, V are $\mathfrak{g}$-modules that $\mathrm{ch}(U \otimes V) = \mathrm{ch}(U) \cdot \mathrm{ch}(V)$.

Now, let $V = V_n$ be as above and notice that $V_{[n-2i]}$ is generated by $x^{n-1} y^i$ and so is 1-dimensional. Then for the weight space decomposition $V = V_{[n]} \oplus V_{[n-2]} \oplus \dots \oplus V_{[-n]}$ give us that

$$\mathrm{ch}(V) = \sum_{i=0}^{n} \mathrm{ch}(V_{[n-2i]}) = e^n + e^{n-2} + \dots + e^{-n} = \frac{e^{n+1} - e^{-(n+1)}}{e^1 - e^{-1}}$$

Our last result for finite dimensional representations of g is the Clebsch-Gordau Rule:

$$V_m \otimes V_n \cong V_{m+n} \oplus V_{m+n-2} \oplus \dots \oplus V_{|m-n|}$$

As an example of the above we can compute $V_m \oplus V_1 = V_{m+1} \oplus V_{m-1}$

2. The Affine Lie Algebra $\widehat{\mathfrak{sl}_2}$

2.1. The Loop Algebra of $\mathfrak{sl}_2$. Let $\mathfrak{g} = \mathfrak{sl}_2$. Then the loop algebra of $\mathfrak{g}$ is the Laurant series in z with $\mathfrak{g}$ coefficients: $L_\mathfrak{g} := \mathfrak{g}[x, x^{-1}]$. For $f(z), g(z) \in L_\mathfrak{g}$, the bracket $[f, g](z) = [f(z), g(z)]$ defines a Lie Algebra structure on $L_\mathfrak{g}$. Using the notation $x(n) = x \otimes z^n$ for $x \in \mathfrak{g}$ and $n \in \mathbb{Z}$ we can rewrite the bracket as $[x(n), y(m)] = xz^n \cdot yz^m - yz^m \cdot xz^n = [x, y] \otimes z^{n+m} = [x, y](m+n)$.

Now, let $\mathfrak{g} = \mathbb{C}f \oplus \mathbb{C}h \oplus \mathbb{C}e$ by the standard decomposition with $[h, e] = 2e$, $[hf] = -2f$. Then if we look at the action of $\mathrm{ad}(h) = [h, \cdot]$ on $\mathfrak{g}$ then $\mathbb{C}f, \mathbb{C}h$ and $\mathbb{C}e$ are eigen spaces of weights -2, 0 and 2 respectively. Similarly, the loop space decomposes as $L_{\mathfrak{g}} = \mathbb{C}f[z, z^-1] \oplus \mathbb{C}h[z, z^-1] \oplus \mathbb{C}e[z, z^-1]$ with $\mathrm{ad}(h)$ again acting on these spaces as -2,0, 2 eigen spaces respectively. However, unlike in the above case, each of these eigenspaces are infinite dimensional.

2.2. The Extension of $L_{\mathfrak{g}}$ by a Derivation. First, let d be the derivation of $L_{\mathfrak{g}}$ given by $dx(n) = nx(n)$ (ie $d = z\frac{\partial}{\partial z}$). Then d satisfies the Libnetz Rule: $d[f, g] = [df, g] + [f, dg]$ since

$$d[x(n), y(m)] = (n+m)[x(n), y(m)] = n[x(n), y(m)]+m[x(n), y(m)] = [dx(n), y(m)]+[x(n), dy(m)]$$

and d commutes with addition. Now, let $\tilde{\mathfrak{g}} := L_{\mathfrak{g}} \rtimes \mathbb{C}d$ be defined to be the Lie Algebra that is $L_{\mathfrak{g}} \otimes \mathbb{C}d$ as a vector space with $L_{\mathfrak{g}} \subseteq \tilde{\mathfrak{g}}$ as a subalgebra in the obvious way and the bracket on $L_{\mathfrak{g}}$ extended by the relation $[d, d] := 0$ and $[d, f] := d(f)$ and bilinearity.

2.2.1. *Exercise.* For a general Lie algebra $\mathfrak{g}$, show that the algebra $\tilde{\mathfrak{g}}$ construed from $\mathfrak{g}$ as above is a Lie algebra.

Lets look at decompositions of $\mathfrak{g} = \mathfrak{sl}_2$. First, $\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}_- \oplus \mathbb{C}h \oplus \mathbb{C}d \oplus \tilde{\mathfrak{g}}_+$ where $\tilde{\mathfrak{g}}_+ = z\mathfrak{g}[z] \oplus \mathbb{C}e[z]$ and $\tilde{\mathfrak{g}}_- = z^{-1}\mathfrak{g}[z^{-1}] \oplus \mathbb{C}f[z^{-1}]$

Now, $\mathbb{C}e[z, z^{-1}] = \oplus_n \mathbb{C}e(n)$. Notice that each $\mathbb{C}e(n)$ is an eigenspace with eigenvalue 2 w.r.t. h and and eigenspace with eigenvalue n w.r.t. d. Similar results hold for $\mathbb{C}f[z, z^{-1}]$ and $\mathbb{C}h[z, z^{-1}]$ so we get the following decomposition of $\tilde{\mathfrak{g}}$:

$$\tilde{\mathfrak{g}} = \bigoplus_{t\in\{-2,0,2\},\ n\in\mathbb{Z}} \tilde{\mathfrak{g}}_{(t,n)}$$

where $\tilde{\mathfrak{g}}_{(t,n)}$ is a eigenspace t w.r.t. h and n w.r.t. d. Note that each $\tilde{\mathfrak{g}}_{(t,n)}$ is finite dimensional (in fact 1-dimensional unless $t = n = 0$ in which case $\tilde{\mathfrak{g}}_{(t,n)} = \mathbb{C}h \oplus \mathbb{C}d$).

Now, since $L_{\mathfrak{g}} \subseteq \tilde{\mathfrak{g}}$, any representation of $\tilde{\mathfrak{g}}$ restricts to a representation of $L_{\mathfrak{g}}$ in the natural way but there exists representations of $L_{\mathfrak{g}}$ which do not extend to a representation of $\tilde{\mathfrak{g}}$.

2.2.2. *Exercise.* Show that the representation V_a given in Definition 1.1 does not extend to a representation of $\tilde{\mathfrak{g}}$.

Now then, what is an example of of a $\tilde{\mathfrak{g}}$ representation? Let V be a finite dimensional $\mathfrak{g}$ module under ρ. Then $V(z) = V[z, z^{-1}]$ (=: LV) is acted on by $L\mathfrak{g}$ by $x(n) \mapsto Xz^n = \rho(x)z^n$ and is also acted on by d by $d \mapsto D = z\frac{d}{dz}$. Then

$$[D, X(n)](a_0 + a_1 z + \ldots + a_k z^k) = \left(z\frac{d}{dz}Xz^n - Xz^n z\frac{d}{dz}\right)(a_0 + a_1 z + \ldots + a_k z^k) =$$

$$nXa_0+(n+1)Xa_1z^{n+1}+\ldots+(n+k)Xa_kz^{n+k}-(Xa_1z^{n+1}+2Xa_2z^{n+2}+\ldots+ka_kz^{n+k}) =$$

$$= nXa_0 + ((n+1)Xa_1z^{n+1} - Xa_1z^{n+1}) + \ldots + ((n+k)Xa_kz^{n+k} - ka_kz^{n+k}))$$

$$= nXa_0 + nXa_1z^{n+1} + \ldots + nXa_kz^{n+k} =$$

$$nX(z)(a_0 + a_1 z + \ldots + a_k z^k)$$

so $[D, X(n)] = nX(n)$ as required and so this defines an action of $\tilde{\mathfrak{g}}$ on LV. Therefore any finitely generated $\mathfrak{g}$ module gives us a representation of $\tilde{\mathfrak{g}}$.

3. The Affine Kac-Moody Algebra $\hat{\mathfrak{g}}$

From physics, we want to study the algebra $\tilde{\mathfrak{g}}$ perturbed by a central extension to what is called the Affine Kac-Moody Algebra.

3.1. Central Extensions. Given a Lie Algebra L, a central extension of L is a lie algebra $\hat{L}$ such that

$$0 \to \mathfrak{a} \to \hat{L} \to L \to 0$$

where $\mathfrak{a} \in Z(\hat{L})$ is an abelian subgroup of $\hat{L}$.

3.1.1. *Example:* The Heisenberg Algebra $\mathscr{H}$ is the algebra generated by p, q, z such that $[p, q] = z$, $[z, p] = 0 = [z, q]$. z is a central element of $\mathscr{H}$ and $\mathscr{H}$ is a central extension of $\mathbb{C}^2$:

$$0 \to \mathbb{C}z \to \mathscr{H} \to \mathbb{C}^2 \to 0$$

Where $\mathbb{C}^2$ is generated by $\bar{p}, \bar{q}$ s.t. $[\bar{p}, \bar{q}] = z = 0$. It can be seen that $\mathscr{H}$ is the algebra generated by $p = \frac{d}{dx}$, $z = 1$ and $q = x$.

Now, as a vector space a central extension is simply $\hat{L} = L \oplus \mathfrak{a}$ since short exact sequences of vector spaces split. For $a, a' \in \mathfrak{a}$, $x, y \in L$ the lie algebra structure is given by $[a, a']\hat{} = 0$ since $\mathfrak{a}$ is abelian, $[a, x]\hat{} = 0$ since $\mathfrak{a}$ is central and $[x, y]\hat{} = [x, y] + B(x, y)$ where $B(x, y) \in \mathfrak{a}$.

Since $[\cdot, \cdot]\hat{}$ is skew symmetric, $B : L \wedge L \to \mathfrak{a}$. We call B a 2-cocylce on L with values in $\mathfrak{a}$ if

$$B([x, y], z) + B([y, z], x) + B([z, x], y) = 0$$

3.1.2. *Exercise.* B defines a Lie algebra structure on $L \oplus \mathfrak{a}$.

Assume that B is a skew symmetric 2-cocycle. Then we have a Lie Algebra structure on $L \oplus \mathfrak{a}$ given by a central extension of L by $\mathfrak{a}$.

Now, let $j : L \to \hat{L} \cong L \oplus \mathfrak{a}$ be a map into $\hat{L}$ where $j(x) = x \oplus A(x)$ and $A(x) : L \to \mathfrak{a}$.

3.1.3. *Exercise.* Show that $B'(x, y) = B(x, y) + A([x, y])$ is a 2-cocycle.

Similarly, we will call such a map $A([x, y])$ a 2-coboudry.

3.1.4. *Exercise.* Show that a 2-coboundary is always a 2-cocycle.

Then we define $H^2(L, \mathfrak{a})$ to be the 2-cocylces modulo the 2-coboundries.

3.1.5. *Claim:* $H^2(L, \mathfrak{a})$ classifies the central extension of L by $\mathfrak{a}$..

3.1.6. *Definition.* Given two central extensions, an isomorphism of central extensions is a map ϕ such that the following commutes:

$$\begin{array}{ccccccccc} 0 & \to & \mathfrak{a} & \to & \hat{L} & \to & L & \to & 0 \\ & & \| & & \downarrow \phi & & \| & & \\ 0 & \to & \mathfrak{a} & \to & \hat{L} & \to & L & \to & 0 \end{array}$$

We have seen above that to each central extension of L by $\mathfrak{a}$ we can assign an element of $H^2(L, \mathfrak{a})$ given by the image of the map B.

3.1.7. *Exercise.* Show this map is an isomorphism..

3.2. Affine Kac-Moody Algebra. We now come to the title of this section: $\widehat{L\mathfrak{g}}$ is the central extension of $L\mathfrak{g}$ by $\mathbb{C}$ with $B(x(n), y(m)) := n\delta_{n+m,0}(x, y)$. We begin by showing that B is a 2-cocylce. First, for$x(n), y(m) \in \tilde{\mathfrak{g}}$, $B(x(n), y(m))$ is skew symmetric since since it is only nonzero if $m = -n$. But if $m = -n$ then for $B(x(n), y(m)) = n(x, y) = -(-n(y, x)) = B(y(m), x(n))$ as required. The other property of 2-cocycles is satisfied as follows: for $x(n), y(m), z(p) \in \tilde{\mathfrak{g}}$,

$$B([x(n), y(m)], z(p)) + B([y(m), z(p)], x(n)) + B([z(p), x(n)], y(m)) =$$

$$B([x, y](n + m), z(p)) + B([y, z](m + p), x(n)) + B([z, x](n + p), y(m)) =$$

$$= (n+m)\delta_{n+m+p,0}([x, y], z)+(m+p)\delta_{m+p+n,0}([y, z], x)+(n+p)\delta_{n+p+m,0}([z, x], y)$$

In this last equation, clearly if $n + m + p \neq 0$ then the above is 0. Assume $n + m + p = 0$. Then, since $(\cdot, \cdot)$ is symmetric and satisfies $([x, y], z) = -(y, [x, z])$,we have the identity's $([y, z], x) = -(z, [y, x]) = (z, [x, y]) = ([x, y], z)$ and similarly $([z, x], y) = ([x, y], z)$ so the above becomes

$$= (n + m)([x, y], z) + (m + p)([y, z], x) + (n + p)([z, x], y)$$

$$= (n+m)([x, y], z)+(m+p)([x, y], z)+(n+p)([x, y], z) = 2(n+m+p)([x, y], z) = 0$$

by the assumption that $n + m + p = 0$.

Then $\widehat{L\mathfrak{g}} = L\mathfrak{g} \oplus \mathbb{C}$ as a vector space with $[x(n), c]\hat{} = 0$, $[x(n), y(n)]\hat{} = [x, y](n+m)+n\delta_{n+m,0}(x, y)c$. Let d be a derivation with $d(x(n)) = nx(n)$, $d(c) = 0$. Then d is a derivation of $\widehat{L\mathfrak{g}}$: (by linearity of the bracket and the fact that c is central, we only have to check for $g_1, g_2 \in L\mathfrak{g}$)

$$d[x(n), y(m)]\hat{} = d([x, y](n + m)) + d(n\delta_{n+m,0}(x, y)c) = d([x, y](n + m))$$

$$= (n + m)[x, y](n + m)) = [dx(n), y(m)] + [x(n), dy(m)]$$

so d acts on $\widetilde{L\mathfrak{g}}$ and so we can define the Kac-Moody Algebra $\widetilde{\mathfrak{g}} := \widetilde{L\mathfrak{g}} \rtimes \mathbb{C}d$ with $[d, f]\tilde{}= d(f)$.

4. Inner Products

We will now consider the extension of the bilinear forms on $\mathfrak{g}$ to our new Lie Algebras. First, there is a natural extension of any form $(\cdot, \cdot) : \mathfrak{g} \otimes \mathfrak{g} \to \mathbb{C}$ to $< \cdot, \cdot >: L\mathfrak{g} \otimes L\mathfrak{g} \to \mathbb{C}[z, z^{-1}]$ by letting $< x(m), y(n) >= (x, y)z^{n+m}$ and extending linearly. Taking $(\cdot, \cdot)$ to be the inner product on $\mathfrak{g}$ we can define the following inner product on $L\mathfrak{g}$, namely, $(p, q) =\text{Res}< p, q > /z$.

4.0.1. *Exercise.* Prove the following:

1. $(x(m), y(n)) = \delta_{m+n,0}(x, y)$
2. On $L\mathfrak{g}$, $(\cdot, \cdot)$ is a nondegenerate bilinear form.
3. $(\cdot, \cdot)$ is invariant under the bracket

Then

$$(x(n), y(m)) = \text{Res}\frac{< x(n), y(m) >}{z} = \text{Res}(x, y)z^{n+m-1} = (x, y)\delta_{m+n,0}$$

since $\mathrm{Res} z^{n+m-1} \neq 0$ iff $n+m-1=-1$ by definition. Its clear than that $(\cdot,\cdot)$ is a nondegenerate bilinear form: if $p,q,r \in L\mathfrak{g}$ are of the form $p = \ldots + p_{-1}z^{-1} + p_0 + p_1 z + \ldots$

$$(c_1p+c_2r,q) = \mathrm{Res}\frac{< c_1p+c_2r,q >}{z} = \mathrm{Res}\frac{< \sum_i c_1p_iz^i + c_1r_iz^i, \sum_i q_iz^i >}{z} =$$

$$\mathrm{Res}\left(\frac{\sum_i\sum_j < c_1p_iz^i + c_1r_iz^i, q_jz^j >}{z}\right) = \mathrm{Res}\left(\frac{\sum_i\sum_j (c_1p_i + c_1r_i, q_j)z^{i+j}}{z}\right)$$

$$= \mathrm{Res}\left(\sum_i\sum_j \delta_{i+j,0}(c_1p_i + c_2r_i, q_j)\right) = \left(\sum_i (c_1p_i + c_2r_i, q_{-i})\right) = c_1\sum_i (p_i,q_{-i}) + c_2\sum_i (r_i,q_{-i})$$

$$= c_1(p,q) + c_2(p,q)$$

where the last equality can be gotten by reversing the beginning of the argument. Linearity in the second term is the same computation. To see that $L\mathfrak{g}$ is nondegenerate simply notice that for any $p \in L\mathfrak{g}$, if $p_k \neq 0$ for some k then there exists a y such that $(p_k,y) \neq 0$ by the nondegeneracy of the inner product on $\mathfrak{g}$ and

$$(p, y(-k)) = \sum_i \delta_{i-k,0}(p_i,y) = (p_k,y) \neq 0$$

Finally, let's check that $(\cdot,\cdot)$ is invariant under the bracket:

$$([x(n),y(m)],w(k)) = ([x,y](n+m),w(k)) = \delta_{n+m+k,0}([x,y],w) =$$

$$-\delta_{n+m+k,0}(y,[x,z]) = -(y(m),[x(n),z(k)])$$

Therefore $(\cdot,\cdot)$ is indeed an inner product on $L\mathfrak{g}$.

The question now is, can we extend such an inner product to be an inner product on $\tilde{\mathfrak{g}}$? and the answer is no: Suppose we could, then since c is central, if $x,y \in \tilde{\mathfrak{g}}$, $(c,[x,y]) = -([x,c],y) = 0$ by invariance of inner product with respect to brackets. Therefore c is perpendicular to anything that is a bracket; but $[\tilde{\mathfrak{g}},\tilde{\mathfrak{g}}] = \tilde{\mathfrak{g}}$ so $c \perp [\tilde{\mathfrak{g}},\tilde{\mathfrak{g}}]$ implies that $c \perp \tilde{\mathfrak{g}}$ so $(\cdot,\cdot)$ cannot be nondegenerate. Therefore we cannot extend an inner product on $\mathfrak{g}$ to an inner product on $\tilde{\mathfrak{g}}$.

We can, however, extend it to an inner product on $\hat{\mathfrak{g}}$ by the following: $(p,q) :=\mathrm{Res}< p,q > /z$ as before for $p,q \in L\mathfrak{g}$ and $(c,\tilde{\mathfrak{g}}) := 0$ as required by the argument above so let $(c,d) := 1$ so that c isn't perpendicular to everything in $\hat{\mathfrak{g}}$. Finally, since for $m \neq 0$,

$$(d,x(m)) = \frac{(d,[d,x(m)])}{m} = -\frac{([d,d],x(m))}{m} = 0$$

and for $m=0$, $x,y \in \mathfrak{g}$,

$$(d,[x,y]) = ([d,x],y) = (0,y) = 0$$

so $(\mathfrak{g},L\mathfrak{g})$ must be 0; we also let $(d,d)=0$. This form is nondegenerate by construction.

4.0.2. *Exercise.* Show that the above construction is a non-degenerate, invariant, bilinear form on $\hat{\mathfrak{g}}$.

4.1. Structure and Presentation of $\hat{\mathfrak{g}}$. To determent he structure of $\hat{\mathfrak{g}}$ it is useful to look at is center which we will denote $\hat{\mathfrak{h}}$. Then $\hat{\mathfrak{h}} := \mathbb{C}h \oplus \mathbb{C}c \oplus \mathbb{C}d \subseteq \hat{\mathfrak{g}}$. Then $\hat{\mathfrak{g}} = \hat{\mathfrak{h}} \oplus_\lambda \hat{\mathfrak{g}}_\lambda$ for all λ that are joint eigen values of $\hat{\mathfrak{h}}$, ie eigen values of all h, c and d. We will use this too say that $\hat{\mathfrak{g}}$ "looks like" $\mathfrak{sl}_3$.

5. Root Systems

Fo now, let $\mathfrak{g} = \mathfrak{sl}_n$ and let $\mathfrak{h} \subseteq \mathfrak{g}$ be the abelian subalgebra of diagonal matrices. Then $\mathfrak{g}$ decomposed under the action of $\mathfrak{h}$. Now, we call $x \in \mathfrak{g}$ a eigen vector of $\mathfrak{h}$ if there exists $\lambda \in h^*$ such that for all $h \in \mathfrak{h}$, $[h, x] = \lambda(h)x$. Now, if E_{kl} is the matrix with a 1 in the k'th row and l'th column, it is a simple calculation to see that

$$[E_{ii}, E_{kl}] = (\delta_{ik} - \delta_{il})E_{kl}$$

So, if h is a diagonal matrix and $\theta_k \in \mathfrak{h}$ is the Kronecker delta wrt E_{kk} then $[h, E_{kl}] = (\theta_k(h) - \theta_l(h))E_{kl}$ so the joint eigenvalues of $\mathfrak{g}$ wrt $\mathfrak{h}$ are $\{\theta_k - \theta_l\}_{k \neq l}$. Then the eigen vectors corresponding to these eigen values are precisely the E_{kl} and

5.0.1. *Proposition/Defintion.*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{k \neq l} \mathbb{C}E_{kl}$$

where E_{kl} is the weight space of $\theta_k - \theta_l$. From here on we will denote the set of nonzero roots of $\mathfrak{g}$ by $\Phi = \{\theta_k - \theta_l\}_{k \neq l} \subseteq \mathfrak{h}^* \backslash \{0\}$, the set of positive roots $\{\theta_k - \theta_l\}_{k<l}$ by Φ_+ and the set of negative roots by $\Phi_- = -\Phi_+ = \{\theta_k - \theta_l\}_{k>l}$. We note that Φ_+ is closed under addition and that $\Phi = \Phi_+ \sqcup \Phi_-$. Finally, the simple roots $\Delta \subset \Phi_+$ are the set of indecomposable element of Φ_+.

Let $\eta_\pm = \bigoplus_{\alpha \in \Phi_\pm} \mathfrak{g}_\alpha \subset \mathfrak{g}$ be the nilpotent upper and lower subalgebra, ie, given a matrix representation of $\mathfrak{g}$ we would have in block form

$$\begin{pmatrix} \ddots & & \eta_+ \\ & \eta & \\ \eta_- & & \ddots \end{pmatrix}$$

Then it is clear that just like breaks up as the direct sum of its center and its nilpotent parts as $\mathfrak{sl}_2 = \mathbb{C}f \oplus \mathbb{C}h \oplus \mathbb{C}e$, in general $\mathfrak{g} = \eta_- \oplus \mathfrak{h} \oplus \eta_+$.

6. Decomposition of $\mathfrak{g}$ to Copies of $\mathfrak{sl}_2$

6.0.2. *Definition.* We can define an inner product on $\mathfrak{g}$ the same way as before, by $(x, y) = \mathrm{tr}(\mathrm{ad}_x \mathrm{ad}_y)$. Since this inner product is nondegenerate it induces an isomorphism between $\mathfrak{h}$ and its duel space. Restricting this isomorphism to the roots we get a map $\Phi \ni \alpha \mapsto t_\alpha \in \mathfrak{h}$ such that $\alpha(t) = (t_\alpha, t)\ \forall t \in \mathfrak{h}$. Let $h_\alpha := \frac{2t_\alpha}{(\alpha,\alpha)_{\neq 0}} \in \mathfrak{h}$ denote the dual root to α.

6.1. Theorem. (Serre) $\mathfrak{g}$ decomposes into a (possibly infinite) number of copies of $\mathfrak{sl}_2$ and is presented as $\{e_i, f_i, h_i\}$ subject to

$$[h_i, h_j] = 0 \tag{6.1}$$

$$[h_i, e_j] = a_{ij} e_j \tag{6.2}$$

$$[h_i, f_j] = -a_{ij} f_j \tag{6.3}$$

$$[e_i, f_j] = \delta_{ij} h_i \tag{6.4}$$

$$\mathrm{ad}(e_i)^{1-a_{ij}} e_j = 0 = \mathrm{ad}(f_i)^{1-a_{ij}} f_j \text{ for } i \neq j \tag{6.5}$$

Proof: Let $\alpha \in \Phi$. The $\forall x \in \mathfrak{g}_\alpha$ and $\forall y \in \mathfrak{g}_{-\alpha}$, $[x, y] = (x, y)t_\alpha$ so there exists $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that $[e_\alpha, f_\alpha] = h_\alpha$. Now,

$$[h_\alpha, e_\alpha] = \alpha(h_\alpha) e_\alpha = \left(\frac{2\alpha(t_\alpha)}{(\alpha, \alpha)}\right) e_\alpha = \left(\frac{2(\alpha, \alpha)}{(\alpha, \alpha)}\right) e_\alpha = 2e_\alpha$$

and similarly, $[h_\alpha, f_\alpha] = -2f_\alpha$ so by sending $f \mapsto f_\alpha$, $e \mapsto e_\alpha$ and $h \mapsto h_\alpha$ we have an embedding $\mathfrak{sl}_2 \hookrightarrow \mathfrak{g}$. Then for each positive root we have a copy of $\mathfrak{sl}_2$ denoted $\mathfrak{sl}_2^\alpha \subseteq \mathfrak{g}$.

Now, for $\alpha = \theta_k - \theta_l \in \Phi$, let $e_\alpha = E_{kl}$, $f_\alpha = E_{lk}$ and $h_\alpha = E_{kk} - E_{ll}$; in addition for any simple root $\alpha_i \in \Delta$, let $e_i = e_{\alpha_i}$, $f_i = f_{\alpha_i}$, $h_i = h_{\alpha_i}$. Then

$$[h_i, e_j] = \alpha_j(h_i) e_j = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} e_i = a_{ij} e_i$$

where $a_{ij} \in \mathbb{Z}$. In fact, $a_{ii} = 2$ and $a_{ij} \leq 0$ for $i \neq j$. For example, for $\mathfrak{sl}_n$,

$$(a_{ij}) = \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots \\ 0 & & -1 & 2 \end{pmatrix}$$

We have now shown that $\mathfrak{g}$ is generated by $\{e_i, f_i, h_i\}$ where for each i $f_i e_i, h_i$ generate a copy of $\mathfrak{sl}_2$. We will now show that the relations hold. (1)-(3) are obvious, as for (4), if $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$ then $[x, y] \in \mathfrak{g}_{\alpha-\beta}$ so $[e_i, f_j] \in \mathfrak{g}_{\alpha_i - \alpha_j} = 0$ since all roots of $\mathfrak{g}$ are either strictly positive or strictly negative combination of simple roots. If $i = j$ then $[e_i, f_i] = h_i$ by the relations for $\mathfrak{sl}_2$.

For (5), assume again the $i \neq j$ (the case $i = j$ is trivial). Then by (3) and (4), $f_j \in V_{-\alpha_{ij}}(\mathfrak{sl}_2^\alpha)$ so f is sitting in a representation of $\mathfrak{sl}_2$, (4) give us that f is the highest weight vector in this representation so $-f_j$ is the lowest wight vector in this representation. But this is the same as saying that $\mathrm{ad}(f_i)^{1-a_{ij}} f_j = 0$. A similarly argument holds e_j.

Therefore $\mathfrak{g}$ decomposes into the direct sum of copies of $\mathfrak{sl}_2$ indexed by the simple roots.

6.2. The Decomposition For $\hat{g}$. Lets return to the case $\mathfrak{g} = \mathfrak{sl}_2$. Recall that $\hat{\mathfrak{h}} = \mathfrak{h} \oplus \mathbb{C}c \oplus \mathbb{C}d$ and let $\Phi = \{\alpha, -\alpha\}$ with $\alpha \in \mathfrak{h}^*$, $\alpha(h) = 2$ be the root system of $\mathfrak{g}$. Then the above decomposition yields

$$\hat{\mathfrak{g}} = \hat{h} \bigoplus_{(\alpha,n) \in \Phi \times \mathbb{Z}} \hat{\mathfrak{g}}(\alpha, n) \bigoplus_{n \in \mathbb{Z}^\times} \hat{\mathfrak{g}}(0, n)$$

A bit of explanation is needed. The above roots come from the affine root system $\hat{\Phi} = \Phi \times \mathbb{Z}\delta \sqcup \{0\} \times \mathbb{Z}$ where $\delta \in \hat{h}$ is a linear form dual to d, ie $\delta(d) = 1$, $\delta(c) = 0$, $\delta(h) = 0$. One way to think of this decomposition is as an infinite "matrix"

$$graphichere$$

where each box is a copy of $\mathfrak{sl}_2$ (or more generally $\mathfrak{sl}_n$). We call the elements labeled + positive roots and denote them $\hat{\Phi}_+$ and those labeled − negative roots and denote them $\hat{\Phi}_-$. Then $\hat{\Phi}_+ = \{\Phi \sqcup 0\} \times \mathbb{N}^*\delta \sqcup \{\alpha\} \times \{0 \cdot \delta\}$ and $\hat{\Phi}_-$.

The simple roots of $\hat{\Phi}_+$ come in two flavors. Looking at the $n = 0$ and $n = 1$ boxes we have simple roots $\alpha_1 = (\alpha, 0)$ and $\alpha_0 = (-\alpha, 1)$. Outside of these boxes we have similar simple roots: for all $n \geq 0$, $(\alpha, n) = n(-\alpha, 1) + (n + 1)(\alpha, 0)$, $(-\alpha, n) = n(-\alpha, 1) + (n - 1)(\alpha, 0)$ and $(0, n) = n(-\alpha, 1) + n(\alpha, 0)$.

Now, recall that we have a form $(\cdot, \cdot)$ on $\hat{\mathfrak{g}}$ and that

$$(\cdot, \cdot)|_{\hat{\mathfrak{h}}} = (\cdot, \cdot)_{\mathfrak{h}} \oplus \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

For $\hat{\alpha} \in \hat{\mathfrak{h}}^* \to t_{\hat{\alpha}} \in \hat{\mathfrak{h}} \rightsquigarrow h_{\hat{\alpha}} = \frac{2t_{\hat{\alpha}}}{(\hat{\alpha}, \hat{\alpha})}$ so for $(\alpha, n) = \pm\alpha + n\delta \rightsquigarrow \pm h + nc$ and

$$(\hat{\alpha}, \hat{\alpha}) = \left\{ \begin{matrix} 2 & \text{Real Roots } \pm \alpha + n\delta \\ 0 & \text{Imaginary Roots } n\delta \end{matrix} \right\}$$

Now, the simple roots are real. Let $h_1 = h$, $h_0 = -h + c$ and let $\alpha_1 = (\alpha, 0)$, $e_1 = e$, $f_1 = f$ and $h_1 = h$; finally, let $\alpha_0 = (-\alpha, 1)$, $e_0 = f(1)$, $f_0 = e(-1)$ and $h_0 = -h + c$.

6.2.1. *Exercise.* Check e_0, f_0 and h_0 form an $\mathfrak{sl}_2$ triple.

Finally, we can calculate the affine Cartan matrix to be

$$a_{ij} = \frac{2(\hat{\alpha}, \hat{\alpha})}{(\hat{\alpha}, \hat{\alpha})} = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} = \hat{A}$$

The affine Cartan matrix.

6.3. Theorem. $\tilde{\mathfrak{g}}$ is presented on $\{e_i, f_i, h_i\}_{i=0,1}$ with

$$[h_i, h_j] = 0 \tag{6.6}$$
$$[h_i, e_j] = a_{ij}e_j \tag{6.7}$$
$$[h_i, f_j] = -a_{ij}f_j \tag{6.8}$$
$$[e_i, f_j] = \delta_{ij}h_i \tag{6.9}$$
$$\mathrm{ad}(e_i)^{1-a_{ij}}e_j = 0 = \mathrm{ad}(f_i)^{1-a_{ij}}f_j \tag{6.10}$$

Note, this is not $\hat{\mathfrak{g}}$ because this only has a 2 dim Cartan subalgebra, indeed $\hat{\mathfrak{g}}$ is presented as above but with $\{e_i, f_i, h_i, d\}_{i=0,1}$ and the additional relations

$$[d, h_i] = 0 \tag{6.11}$$
$$[d, e_1] = [d, f_1] = 0 \tag{6.12}$$
$$[d, e_0] = e_0 \tag{6.13}$$
$$[d, f_0] = -f_0 \tag{6.14}$$

7. The Enveloping Algebra of a Lie Algebra

The enveloping algebra $U\mathfrak{g}$ is defined to be the the unique solution to the following universal problem:

$U\mathfrak{g}$ is an associative unitial algebra with a linear map $\rho : \mathfrak{g} \to U\mathfrak{g}$ such that $\rho(x)\rho(y)-\rho(y)\rho(x) = \rho([x,y])$ and if U is another such algebra there exists a unique unitial algebra map $\pi : \mathscr{U}(\mathfrak{g}) \to U$ such that

$$\begin{tikzcd} & U\mathfrak{g} \arrow[d, dotted, "p"] \\ \mathfrak{g} \arrow[ur, "\rho"] \arrow[r, "\rho'"'] & U \end{tikzcd}$$

7.1. Proposition: $U\mathfrak{g}$ exists and is unique up to isomorphism.

Proof: Uniqueness is clear from the universal property. Define $T_{\mathfrak{g}} := \bigoplus_k \mathfrak{g}^{\otimes k}$ to be the tensor algebra. Now, let $\tilde{\rho} : \mathfrak{g} \hookrightarrow T_{\mathfrak{g}}$ as an element of degree 1 and let $U\mathfrak{g} := T\mathfrak{g}/x \otimes y - y \otimes x - [x,y]$ be the result of modding out $T_{\mathfrak{g}}$ by the algebra generated by the Lie bracket. Then $\rho : \mathfrak{g} \to U\mathfrak{g}$ is simply $\tilde{\rho}$ factoring through the quotient. Clearly, this satisfies the requirement that

$$\rho(x)\rho(y) - \rho(y)\rho(x) = x \otimes y - y \otimes x = [x,y] = \rho([x,y])$$

For any other associative unital algebra U such that $\rho' : \mathfrak{g} \to U$ we can define a map $p : U\mathfrak{g} \to U$ in the logical way by $p : x_1 \otimes \ldots, \otimes x_k \mapsto p(x_1) \ldots p(x_k)$. Therefore $U\mathfrak{g}$ is a universal enveloping algebra and so by uniqueness is *the* universal enveloping algebra.

$\square$

7.2. Properties of $U\mathfrak{g}$. Any representation of $\mathfrak{g}$ extends to a representation of $U\mathfrak{g}$. Indeed, if V is a $\mathfrak{g}$ module with representation π we have by the universal property the following diagram:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\pi"] \arrow[dr, "\rho"'] & \mathrm{End} \\ & U\mathfrak{g} \arrow[u, "\exists\ \tilde{\pi}"'] \end{tikzcd}$$

So V is a $U\mathfrak{g}$ via p.

Similarly, any representation of $U\mathfrak{g}$ corresponds to a representation of $\mathfrak{g}$ by composition:

$$\begin{tikzcd} U\mathfrak{g} \arrow[r, "\tilde{\pi}"] \arrow[dr, "\rho"'] & \mathrm{End} \\ & \mathfrak{g} \arrow[u, "\exists\ \tilde{\pi}\circ\rho"'] \end{tikzcd}$$

Combining the two facts above we see that $\{\text{Reps of } \mathfrak{g}\} \equiv \{\text{Reps of } U\mathfrak{g}\}$.

Now, $T\mathfrak{g} = \bigoplus_k \mathfrak{g}^{\otimes k}$ is $\mathbb{N}$ graded but the ideal $I = \{x \otimes y - y \otimes x - [x,y]\}$ is not homogeneous so $U\mathfrak{g}$ doesn't inherit the grading from $T\mathfrak{g}$. However, $U\mathfrak{g}$ is filtered since the standard filtration of $T\mathfrak{g}$ descends to the filtration $U\mathfrak{g} = \cup_{n\in\mathbb{N}} U\mathfrak{g}_n$ where $U\mathfrak{g}_n = \{\rho(x_1) \ldots \rho(x_n) | x_i \in \mathfrak{g}\}$.

7.3. Lemma: The graded of $U\mathfrak{g}$ given by $\mathrm{Gr}(U\mathfrak{g}) := \oplus_n U\mathfrak{g}_n/U\mathfrak{g}_{n-1}$ is commutative.

Proof: This is a simple check: $\rho(x)\rho(y) = \rho(y)\rho(x) + \rho([x,y])$ in $U\mathfrak{g}$ so in $\mathrm{Gr}(U\mathfrak{g})$, $\rho(x)\rho(y) = \rho(y)\rho(x)$. □

7.4. Corollary: If $S\mathfrak{g} := \oplus_{n\geq 0} S^n\mathfrak{g}$ is the symmetric algebra of $\mathfrak{g}$ then $S\mathfrak{g} \xrightarrow{\sigma} \mathrm{Gr}(U\mathfrak{g})$ by $S_0 \ni x \mapsto \rho(x) \in\mathrm{Gr}(U\mathfrak{g})$ and $S \ni x_1 \dots x_k \mapsto \rho(x_1)\dots\rho(x_k) \in\mathrm{Gr}(U\mathfrak{g})$

7.5. Theorem: (PBW). The map $\sigma : S\mathfrak{g} \to \mathrm{Gr}(U\mathfrak{g})$ is in fact an isomorphism of graded algebras for any Lie algebras $\mathfrak{g}$ over any field. The proof can be found elsewhere.

7.6. Corollary's of PBW:.

(1) $\dim U\mathfrak{g}_n/Ug_{n-1} = \dim S^n\mathfrak{g}$
(2) the defining map $\rho : \mathfrak{g} \to U\mathfrak{g}$ is an embedding. This justifies writing x instead of $\rho(x)$ in $U\mathfrak{g}$ by abuse of notation.
(3) If $\{x_a\}_{a\in\Lambda}$ is a basis of $\mathfrak{g}$ then the lexicographically ordered monomials in x_a form a basis of $U\mathfrak{g}$.

7.6.1. *Exercise.* Prove the above.

(4) Suppose that as a vector space $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ where $\mathfrak{g}_i$ are Lie subalgebras of $\mathfrak{g}$, then by universality

$$\begin{tikzcd} U\mathfrak{g}_i \arrow[r, hook, "\pi_i"] & U\mathfrak{g} \\ \mathfrak{g}_i \arrow[u, "\rho_i"] \arrow[r, hook] & \mathfrak{g} \arrow[u, "\rho"'] \end{tikzcd}$$

so $U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \to U\mathfrak{g}$ is a linear map of bimodules given by $\pi : g_1 \otimes g_1 \mapsto \pi_1(g_1) \otimes \pi_2(g_2)$. We claim that this is an isomorphism.

Proof: Choose an ordered basis $\{x^i_a\}_{a\in\Lambda^i}$ of $\mathfrak{g}_i$. Then $\{\{x^1_a\}_{a\in\Lambda^1}, \{x^2_b\}_{b\in\Lambda^2}\}$ is an ordered basis of $\mathfrak{g}$ (if we put $\Lambda_1 < \Lambda_2$). Then any element $x^1_a \dots x^1_{a'} \otimes x^2_b \dots x^2_{b'}$ in the basis of $U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$ corresponding to the bases above correspond to the lexicographically ordered monomial $x^1_a \dots x^1_{a'} x^2_b \dots x^2_{b'}$ in the basis for $U\mathfrak{g}$. This correspondence is clearly a bijection. □

7.7. Remark: Assume that the field K has characteristic 0. Then in addition to the map $S\mathfrak{g} \to\mathrm{Gr}(U\mathfrak{g})$ there exists a map $S\mathfrak{g} \to U\mathfrak{g}$ given by the composition $S\mathfrak{g} \hookrightarrow T\mathfrak{g} \to U\mathfrak{g}$ where $x_1 \dots x_k \mapsto \frac{1}{m}\sum_{\sigma\in\sigma_m} x_{\sigma(1)} \dots x_{\sigma(m)}$. This map is an isomorphism of filtered vector spaces.

8. Representations of $\hat{h}$

Recall that $\hat{h} = \mathbb{C}h \oplus \mathbb{C}c \oplus \mathbb{C}d$.

8.1. Definition:

(1) A representation of $\hat{g}$ is diagonalizable if $V = \otimes_{\lambda\in\hat{h}^*} V_\lambda$ where $V_\lambda = \{v \in V | tv = \lambda(t)v, \forall t \in \hat{h}\}$
(2) A diagonalizable representation is called integrable if e_i and f_i act nilpotently, ie $\forall v \in V$, $\exists n \in \mathbb{N}$ such that $e_i^n v = f_i^n v = 0$.

For $i = 0, 1$, let $g^i :=< e_i, f_i, h_i >\cong \mathfrak{sl}_2$

8.2. Proposition: If V is integrable, then as a $\mathfrak{g}_i$ module V decomposes into a possibly infinite direct sum of finite dimensional irreducible modules invariant under the action of $\hat{h}$. ie, $V = \oplus_{a \in \mathbb{C}^*} V_a^i$ as a $\mathfrak{g}^i$ module.

8.3. Corollary: V integrates to a representation of the group $\mathrm{sl}_2^{(i)}$ with Lie algebra $\mathfrak{g}_i$.

Proof: Let $v \in V[\lambda]$ and let $\gamma_v = \mathfrak{g}^i v$. Then

$$e_i f_i^k v = e_i f_i f_i^{k-1} v = h_i f_i^{k-1} v + f_i e_i f_i^{k-1} v = \sum_{j=0} f_i^j h_i f_i^{k-1-j} v + f_i^k e_i v$$

$$= < \lambda - (k-1-j)\alpha_i, h_i >$$

since $[t, f_i] = -\alpha_i(t) f_i$. But this is

$$= k(< \lambda, h_i > -k + 1) f_i^{k-1} v + f_i^k e_i v$$

so if $\sum_{k,m \geq 0} f_i^k e_i^m v$ is invariant under $\mathfrak{g}_i$ and $\hat{h}$ then we can calculate the effect of applying h_i to an element of v in the same way as we did for e_i above. By local nilpotence, γ_v is finite dimensional so all $v \in V$ are contained in a finite dimensional $\mathfrak{g}^i$ submoduel γ_v invariant under $\hat{h}$. By complete reducibility, γ is then the direct sum of finite dimensional $\mathfrak{sl}_2$ moduels. The proof follows. □

Note: In general, V is not complete reducible as a $\hat{\mathfrak{g}}$ module.

8.4. Category $\mathscr{O}$. A representation V is in Category $\mathscr{O}$ if

(1) V is diagonal, ie $V = \oplus_{\lambda \in h^*} V_\lambda$
(2) V_λ is finite dimensional
(3) There exist $\lambda_1, \ldots, \lambda_r \in h^*$ such that the nonzero eigen values of V denoted $P(V) := \{\lambda | V_\lambda \neq 0\} \subseteq \cup_{i=1}^n \{\mu \in h^* | \mu \leq \lambda_i\}$ where $\mu \leq \lambda$ means $\lambda - \mu$ is the sum of positive roots.

8.4.1. *Example:* Let V be a highest weight representation of $\lambda \in \hat{\mathfrak{h}}^*$, ie there $\exists v \in V_\lambda$ such that $e_i v = 0$ and $U\hat{\mathfrak{g}} v = V$. Then since $\hat{\mathfrak{g}} = \hat{\mathfrak{n}}_- \oplus \hat{\mathfrak{h}} \oplus \hat{\mathfrak{n}}_+$ by the PBW Theorem, $U\hat{\mathfrak{g}} = U\hat{\mathfrak{n}}_- \otimes U\hat{\mathfrak{h}} \otimes U\hat{\mathfrak{n}}_+$. Now, since $U\hat{\mathfrak{n}}_+$ consists of powers of e_i's, $U\hat{\mathfrak{n}}_+ v = 0$ and since v is an eigen vector of $U\hat{\mathfrak{h}}$, $U\hat{\mathfrak{g}} v = U\hat{\mathfrak{n}}_- v$.

8.4.2. *Exercise.* Show that V is in catagory $\mathscr{O}$.

9. Verma Modules

9.1. Definition: M_λ is a Verma Module of highest weight λ iff every module of highest weight λ is a quotient of M_λ.

9.2. Proposition:

(1) For any $\lambda \in \hat{\mathfrak{h}}^*$ there exists a unique Verma module of highest weight λ.
(2) M_λ is a free $U\hat{\mathfrak{n}}$ module of rank 1.
(3) M_λ contains a unique maximal proper submodule.

Proof:

(1) Uniqueness is clear from the definition. For existence, let $I_\lambda \subseteq U\hat{\mathfrak{g}}$ be the left ideal generated by $\hat{n}_+$ and $(t-\lambda(t).1)$ for $t \in \hat{\mathfrak{h}}$ and set $M_\lambda = U\hat{\mathfrak{g}}/I_\lambda$. The M_λ is killed by $\hat{\mathfrak{n}}_+$ and $t \in \hat{\mathfrak{h}}$ acts on M_λ by $t.1 = \lambda(t)1$.

Furthermore, any mod. of highest weight representation will be given by

(2) By PBW, $U\hat{\mathfrak{g}} = U\hat{\mathfrak{n}}_{\text{-}} \otimes U\hat{\mathfrak{h}} \otimes U\hat{\mathfrak{n}}_+$ so when we mod out by I_λ, we kill $\hat{\mathfrak{n}}_+$ and $\hat{\mathfrak{h}}$ is identified with the scalars so $U\hat{\mathfrak{g}}/I_\lambda \cong U\hat{\mathfrak{n}}_{\text{-}}$ as a left $U\hat{\mathfrak{n}}_{\text{-}}$ module.

(3) Let $M' \subseteq M_\lambda$ be a submodule of M_λ. We will show that M' is proper iff $M'[\lambda] = 0$. First, if $M'[\lambda] \neq 0$ then $v_\lambda \in M'$; but then $M_\lambda = U\hat{\mathfrak{g}}v_\lambda \subseteq M'$ so $M' = M_\lambda$. If $M'[\lambda] = 0$ then M' is clearly proper as it does not contain v_λ.

Now, let $M' = \sum_{M'' \subsetneq M_\lambda} M''$ be the sum of all proper submodules, then $M'[\lambda] = \sum_{M'' \subsetneq M_\lambda} M''[\lambda] = 0$ so $M' \subsetneq M_\lambda$ is the unique proper submodule of M_λ. □

9.3. Corollary: There exists a unique, irreducible highest weight modules of weight λ denoted $L(\lambda)$.

Proof: Let $L(\lambda) = M_\lambda/M'_\lambda$. Then $L(\lambda)$ is clearly irreducible. It is unique since if $\pi_i : M_\lambda \to L_i(\lambda)$ then $\ker \pi_i$ is a maximal proper submodule of M_λ. Then by part 3 of the the proposition $\ker \pi_1 = \ker \pi_2$ so $L_1(\lambda) = M_\lambda / \ker \pi_1 = L_2(\lambda)$. □

10. Integrable Modules

Recall that a $\hat{\mathfrak{g}}$ module is integrable iff h_i acts diagonally and f_i, e_i act nilpotently.

10.1. Proposition: L_λ is integrable iff $\lambda(h_i) \in \mathbb{Z}_+$, $i = 0, 1$.

Proof: $\Rightarrow$) $h_i v_\lambda = \lambda(h_i)v_\lambda$ and v_λ lies in a finite dimensional $\mathfrak{sl}_2^{(i)}$ module so $\lambda(h_i) \in \mathbb{Z}$ by the representation theory of $\mathfrak{sl}_2$. Since $e_i v_\lambda = 0$, $\lambda(h_i) \in \mathbb{Z}_{\geq 0}$. **********************

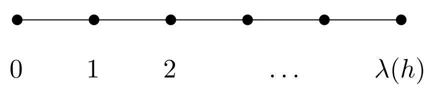
$\Leftarrow$) We showed above that $e_i f_i^k v = k(\lambda(h_i) - k - 1)f_i^{k-1}v$ so, in particular, $e_i f_i^{\lambda(h_i)+1} v_\lambda = 0$. Moreover, for $j \neq i$, $e_j f_i^{\lambda(h_i)+1} = 0$ since e_i and f_j commute. This then implies that the submodule generated by $f_i^{\lambda(h_i)+1}$ is proper. Now, for $U\hat{\mathfrak{g}} = U\hat{\mathfrak{n}}_{\text{-}} \otimes U\hat{\mathfrak{h}} \otimes U\hat{\mathfrak{n}}_+$, $U\hat{\mathfrak{n}}_{\text{-}}$ decreases the weight of a vector, $U\hat{\mathfrak{h}}$ scales a vector and $U\hat{\mathfrak{n}}_+$ increases the weight so it is impossible to "capture" anything "above" $f_i^{\lambda(h_i)+1}v_\lambda$. But $L(\lambda)$ has no proper submodules so $f_i^{\lambda(h_i)+1}v_\lambda = 0$. It follows that the f_i's are locally nilpotent since they are locally nilpotent on V_λ and are locally nilpotent on $\hat{\mathfrak{g}}$:

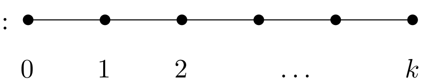
For $m >> 0$ and $x_i \in \mathfrak{n}_{\text{-}}$, $f_i^m x_1 \ldots x_k v_\lambda =ad(f)^{m_1}(x_i) \cdot f^{m_2} v_\lambda$. By picking m large enough we force m_1 or m_2 to be large so, since $f^{m_2}v_\lambda = 0$ for some m_2 and $\text{ad}(f_i)^{1-a_{ij}}(f_j) = 0$ by the serre relations f_i acts nilpotently. □

10.2. Integrable Highest Weight Modules for $\hat{\mathfrak{sl}}_2$. Finally, we will look at highest weight representations for $\hat{\mathfrak{g}}$. The eigenvalues of $\hat{\mathfrak{h}}$ are of the form $\hat{\lambda} = h\omega = \lambda + k\Lambda_0 + n\delta$, where $\lambda \in \mathfrak{h}^*$, $\Lambda_0(c) = 1$, $\Lambda_0(h \notin \mathbb{C}c) = 0$, and $\delta(d) = 1$, $\delta(h \notin \mathbb{C}d) = 0$.

Now, $cV_\lambda = k(c)V_\lambda$ so $c|_V = k\mathrm{id}$ since c is central. Then $c\lambda v_\lambda = \lambda c v_\lambda = k\lambda v_\lambda$. If V is integrable, then $\hat{\lambda}(h_1) \in \mathbb{Z}_+$ by the work done previously so $\hat{\lambda}(h_1) = \lambda(h)$ and $\mathbb{Z}_+ \ni \hat{\lambda}(h_0) = \hat{\lambda}(-h + c) = -\lambda(h) + k$. Now, $L(\hat{\lambda})$ is integrable iff $\lambda(h) \in \mathbb{Z}_+$ and $\lambda(h) \leq k$.

Fix k in $\mathbb{N}$. Then, the irreducible representations of $\mathfrak{sl}_2$ and $\hat{\mathfrak{sl}}_2$ are given by

$\mathfrak{sl}_2$: 

$\hat{\mathfrak{sl}}_2$: 

So there are a finite number of irreducible representations for any fixed k.

11. Solutions to Exercises

(Exercise 2.2.1) For a general Lie algebra $\mathfrak{g}$, show that $\tilde{\mathfrak{g}}$ is a Lie algebra.

Proof: Clearly $L_{\mathfrak{g}}$ is a Lie Algebra since all the properties of the bracket on $L_{\mathfrak{g}}$ follow directly from the linearity of the bracket on $\mathfrak{g}$ and the definition $[f,g](z) = [f(z), g(z)]$. Now, Let $f, g \in L_g$ and let $c_1, c_2 \in \mathbb{C}$. Then

$$[f + c_1 d, g + c_2 d] = [f, g] + [f, c_2 d] + [c_1 d, g] + [c_1 d, c_2 d] = [f, g] - c_2 d(f) + c_1 d(g)$$

First, this is bilinear since $[\cdot, \cdot]$ is and d is linear. It's alternating since

$$[g+c_2 d, f+c_1 d] = [g, f]-c_1 d(g)+c_2 d(f) = -([f, g]+-c_2 d(f)+c_1 d(g)) = -[f+c_1 d, g+c_2 d]$$

Finally,

$$[d, [f, g]] + [f, [g, d]] + [g, [d, f] = [d(f), g]) + [f, d(g)] + [f, -d(g)] + [g, d(f)] = 0$$

and

$$[d, [d, f]] + [d, [f, d]] + [f, [d, d]] = [d, d(f)] + [d, -d(f)] + [f, 0] = 0$$

and by definition

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0 = [d, [d, d]]$$

So since the bracket is bilinear,

$$[f + c_1 d, [g + c_2 d, h + c_3 d]] + [g + c_2 d, [h + c_3 d, f + c_1 d] + [h + c_3 d, [f + c_1 d, g + c_2 d]]$$

Can be rewritten as the sum of factors of one of the above four forms and so is 0 as required for the Jacobi Identity. □

(Exercise 2.2.2) The representation V of $L\mathfrak{g}$ given by Definition 1.1 does not extend to a representation of $\tilde{\mathfrak{g}}$.

Proof: If $V \cong \mathbb{C}^2$ is the representation given in Definition 1.1 then V is a $L_{\mathfrak{g}}$ module by $\rho_a : L_{\mathfrak{g}} \to$End$(V)$ st $g \mapsto g(a)$ for some $a \in \mathbb{C}$. Assume that this representation could be extended to one of $\tilde{\mathfrak{g}}$. Then we would must have a matrix $\rho_a(d) = D$ such that if $x \mapsto X$, $[D, X(n)] = nx(n)$. But

$$[D, X(n)] = DX(n)-X(n)D = DXa^n-Xa^nD = a^n(DX-XD) = a^n[D, X] \neq nXa^n$$

since $[D, X]$ cannot equal nX for all n. Therefore V does not extend to a representation of $\tilde{\mathfrak{g}}$.

(Exercise 3.1.2) Show that a 2-cocylce defines a Lie algebra structure on the central extension of L by $\mathfrak{a}$.

Proof: First,

$$[a_1 + x_1, a_2 + x_2]\hat{} = [a_1, a_2]\hat{} + [a_1, x_2]\hat{} + [x_1, a_2]\hat{} + [x_1, x_2]\hat{}$$

so, since all these terms are 0 except the last, it suffices to check that $[x_1, x_2]\hat{} = [x_1, x_2] + B(x_1, x_2)$ satisfies the properties of the bracket. The bilinearity and alternating properties follow from the definition of the wedge product. The Jacobi identity holds by the definition of a 2-cocylce since

$$[x, [y, z]\hat{}]\hat{} = [x.[y, z] + B(y, z)]\hat{} = [x, [y, z]]\hat{} + [x, B(y, z)]\hat{} =$$
$$[x, [y, z]] + B(x, [y, z])$$

since $[x, B(y, z)] = 0$. So

$$[x, [y, z]\hat{}]\hat{} + [y, [z, x]\hat{}]\hat{} + [z, [x, y]\hat{}]\hat{} =$$
$$[x, [y, z]] + B(x, [y, z]) + [y, [z, x]] + B(y, [z, x]) + [z, [x, y]] + B(z, [x, y]) = 0$$

and B defines a Lie Algebra structure on $L \oplus \mathfrak{a}$. Furthermore, the short exact sequence of vector spaces

$$1 \to \mathfrak{a} \to L \oplus \mathfrak{a} \to L \to 1$$

is an ses of Lie algebras since both of the maps are Lie algebra morphisms: $[a, a] = 0 \mapsto 0 = [a, a]\hat{}$ and $[x, y]\hat{} = [x, y] + B(x, y) \mapsto [x, y]$ since $B(x, y) \in \mathfrak{a}$.

(Exercise 3.1.3) Show that $B'(x, y) = B(x, y) + A([x, y])$ is a 2-cocycle.

Proof: First, $B' : L \wedge L \to \mathfrak{a}$ since B does and A is linear and $[x, y]$ is skew symmetric. Then fact that B' is a 2-cocycle is a simple computation:

$$B'([x, y], z) + B'([y, z], x) + B'([z, x], y)$$
$$= B([x, y], z) + A([[x, y], z]) + B([y, z], x) + A([[y, z], x]) + B([z, x], y) + A([[z, x], y])$$
$$= B([x, y], z) + B([y, z], x) + B([z, x], y) + A([[x, y], z] + [[y, z], x] + [[z, x], y]) = 0 + A(0) = 0$$

(Exercise 3.1.4) Show that a 2-coboundary is always a 2-cocycle.

Proof: First, $A([x, y]) = A(-[y, x]) = -A([y, x])$ so $A : L \wedge L \to \mathfrak{a}$. Now, by linearity,

$$A([[x, y], z]) + A([[y, z], x]) + A([[z, x], y]) = A([[x, y], z] + [[y, z], x] + [[z, x], y]) = A(0) = 0$$

so $A([x,y])$ is a 2-cocycle.

(Exercise 3.1.7) Show the map taking an isomorphism class of central extension of L by $\mathfrak{a}$ to its corresponding element in the homology $H^2(L,\mathfrak{a})$ defined above is an isomorphism.

Proof: Since central extension is defined by a 2-cocycle this map is clearly onto; since every map $B : L \wedge L \to \mathfrak{a}$ defines a Lie algebra structure on $L \oplus \mathfrak{a}$ we simple need to check that two central extension are isomorphic iff they differ by a coboundary. One way is obvious, assume that $B'(x,y) = B(x,y) + A([x,y])$. Then $\phi : L \oplus \mathfrak{a} \to L \oplus \mathfrak{a}$ by $\phi : x \oplus a \mapsto x \oplus (a + A(x))$.

Indeed this is a vector space isomorphism:

Injective: $x \oplus (a+A(x)) = y \oplus (b+A(y))$ implies that $x = y$ and so $a+A(x) = b + A(x)$ which implies that $a = b$.

Surjective: $x \oplus -A(x) \mapsto x$ and $0 \oplus a \mapsto a$ so it is surjective on a basis of $L \oplus \mathfrak{a}$. It is a Lie algebra isomorphism since

$$\phi([x,y]_B) = \phi([x,y] \oplus B(x,y)) = [x,y] \oplus B(x,y) + A([x,y]) = [x,y]_{B'}$$

It is clear that this map respects the commutative diagram below:

$$\begin{array}{ccccccccc} 0 & \to & \mathfrak{a} & \to & \hat{L} & \to & L & \to & 0 \\ & & \| & & \downarrow \phi & & \| & & \\ 0 & \to & \mathfrak{a} & \to & \hat{L} & \to & L & \to & 0 \end{array}$$

Therefore if two maps differ by a 2-coboundary they yield isomorphic central extensions and so the map from $H^1(L,\mathfrak{a}) \to$isomorphism classes of central extensions of L by $\mathfrak{a}$ is well defined and surjective. We will now see that it has an inverse.

Now, assume that two maps $B, B' : L \wedge L \to \mathfrak{a}$ are isomorphic via a map $\phi : L \oplus \mathfrak{a} \to L \oplus \mathfrak{a}$ such that the above commutes. Then by the commutativity of the diagram $\phi(0 \oplus a) = 0 \oplus a$ for $a \in \mathfrak{a}$ and $\phi(x \oplus 0) = x \oplus \hat{\phi(x)}$ for all $x \in L$, $\phi : L \to \mathfrak{a}$. In general then

$$\phi(x \oplus a) = \phi(x \oplus 0) + (0 \oplus a) = x \oplus \hat{\phi}(x) + 0 \oplus a = x \oplus (\hat{\phi}(x) + a)$$

Now, since ϕ is a Lie algebra homomorphism, by letting $x := x \oplus 0$ and $a := 0 \oplus a$ by abuse of notation we get

$$\phi([x,y]_B) = \phi([x,y]+B(x,y)) = [x,y]+\hat{\phi}([x,y] \oplus B(x,y)) = [x,y]+\hat{\phi}([x,y])+B(x,y)$$

and

$$[\phi(x), \phi(y)]_{B'} = [x + \hat{\phi}(x), y + \hat{\phi}(y)]_{B'} = [x,y]_{B'} = [x,y] + B'(x,y)$$

where the second equality on the last row comes from the fact that $\hat{\phi}(x)$ is central. But since ϕ is a Lie algebra homomorphism these must be equal so

$$[x,y] + \hat{\phi}([x,y]) + B(x,y) = [x,y] + B'(x,y)$$

and $B'(x,y) = \hat{\phi}([x,y]) + B(x,y)$ where $\hat{\phi} : L \to \mathfrak{a}$. Therefore B and B' differ by a 2-coboundary and so there is a map sending isomorphism classes of central extensions of L by $\mathfrak{a}$ to $H^2(L,\mathfrak{a})$. It is clear that this map is the inverse of the one constructed above.

Therefore the central extensions $L \oplus \mathfrak{a}$ are in one to one correspondence with elements of $H^2(L,\mathfrak{a})$.

(Exercise 4.0.1) Let $(p,q) :=$Res$< p,q > /z$. Prove the following:

(1) $(x(m), y(n)) = \delta_{m+n,0}(x,y)$
(2) On $L\mathfrak{g}$, $(\cdot,\cdot)$ is a nondegenerate bilinear form.
(3) $(\cdot,\cdot)$ is invariant under the bracket

Proof:

(1)
$$(x(n), y(m)) = \text{Res}\frac{< x(n), y(m) >}{z} = \text{Res}(x,y)z^{n+m-1} = (x,y)\delta_{m+n,0}$$
since Res$z^{n+m-1} \neq 0$ iff $n+m-1 = -1$ by definition.

(2) If $p,q,r \in L\mathfrak{g}$ are of the form $p = \ldots + p_{-1}z^{-1} + p_0 + p_1 z + \ldots$
$$(c_1 p + c_2 r, q) = \text{Res}\frac{< c_1 p + c_2 r, q >}{z} = \text{Res}\frac{< \sum_i c_1 p_i z^i + c_1 r_i z^i, \sum_i q_i z^i >}{z} =$$
$$\text{Res}\left(\frac{\sum_i \sum_j < c_1 p_i z^i + c_1 r_i z^i, q_j z^j >}{z}\right) = \text{Res}\left(\frac{\sum_i \sum_j (c_1 p_i + c_1 r_i, q_j) z^{i+j}}{z}\right)$$
$$= \text{Res}\left(\sum_i \sum_j \delta_{i+j,0}(c_1 p_i + c_2 r_i, q_j)\right) = \left(\sum_i (c_1 p_i + c_2 r_i, q_{-i})\right) = c_1 \sum_i (p_i, q_{-i}) + c_2 \sum_i (r_i, q_{-i})$$
$$= c_1(p,q) + c_2(p,q)$$
where the last equality can be gotten by reversing the beginning of the argument. Linearity in the second term is the same computation. To see that $L\mathfrak{g}$ is nondegenerate simply notice that for any $p \in L\mathfrak{g}$, if $p_k \neq 0$ for some k then there exists a y such that $(p_k, y) \neq 0$ by the nondegeneracy of the inner product on $\mathfrak{g}$ and
$$(p, y(-k)) = \sum_i \delta_{i-k,0}(p_i, y) = (p_k, y) \neq 0$$

(3) Let's check that $(\cdot,\cdot)$ is invariant under the bracket:
$$([x(n), y(m)], w(k)) = ([x,y](n+m), w(k)) = \delta_{n+m+k,0}([x,y], w) =$$
$$-\delta_{n+m+k,0}(y, [x,z]) = -(y(m), [x(n), z(k)])$$
Therefore $(\cdot,\cdot)$ is indeed an inner product on $L\mathfrak{g}$.

(Exercise 4.0.2) Show that $(p, q) :=\mathrm{Res}_{z=0}\frac{<p,q>}{z}$ is a non degenerate, invariant bilinear form on $\hat{\mathfrak{g}}$.

Proof

First, $(\cdot,\cdot)$ is bilinear since for $a, b, e, f \in \hat{\mathfrak{g}}$ where $a = p_a + c_a + d_a$ for $p_a \in L\mathfrak{g}$, $c_a \in \mathbb{C}c$ and $d_a \in \mathbb{C}d$,

$$(a + b, e) = (p_a + p_b + c_a + c_b + d_a + d_b, e) =$$

$$(p_a + p_b, e) + (c_a + c_b, e) + (d_a + d_b, e) = (p_a + p_b, p_e) + (c_a + c_b, e_d) + (d_a + d_b, e_c)$$

$$= (p_a + p_b, p_e) + (c_a + c_b)e_d + (d_a + d_b)e_c =$$

$$(p_a, p_e) + (c_a, e_d) + (d_a, e_c) + (p_b, p_e) + (c_b, e_d) + (d_b, e_c) =$$

$$(a, e) + (b, e)$$

A similarly argument shows linearity in the second term. We now show that tt is invariant with respect to brackets

Since c is central and so commutes with everything and d is only nontrivial on $L\mathfrak{g}$, we have, for $e = w(k) + c + d$,

$$([x(n) + c + d, y(m) + c + d], e) = ([x(n), y(m)] + [x(n), d] + [d, y(m)], e)$$

$$= ([x(n), y(m)] + n(x, y)\delta_{n+m}c - nx(n) + my(m), w(k) + c + d)$$

$$= ([x(n), y(m)], w(k)) + n(x, y)\delta_{n+m} + (-nx(n) + my(m), w(k))$$

$$= ([x(n), y(m)], w(k)) + n(x, y)\delta_{n+m} - n(x, w)\delta_{n+k} + m(y, w)\delta_{k+m}$$

Now, by a similar computation we see that

$$(y(m) + c + d, [x(n) + c + d, e])$$

$$= (y(m), [x(m), w(k)]) + n(x, w)\delta_{n+k} + (y(m), -nx(n) + kw(k))$$

$$= (y(m), [x(m), w(k)]) + n(x, w)\delta_{n+k} + k(y, w)\delta_{m+k} - n(y, x)\delta_{m+n}$$

Since $(\cdot,\cdot)$ is invariant on $x(n), y(m), w(k)$ we only need to check that

$$n(x, y)\delta_{n+m}-n(x, w)\delta_{n+k}+m(y, w)\delta_{k+m} = -(n(x, w)\delta_{n+k}+k(y, w)\delta_{m+k}-n(y, x)\delta_{m+n})$$

But this is quiet simple since the only term for which this is no obvious is $m(y, w)\delta_{k+m} = -k(y, w)\delta_{m+k}$; but both sides are 0 unless $m = -k$ so the equality holds. Therefore the killing form extends to a invariant bilinear form on $\hat{g}$.

(Exercise 6.2.1) Check that in $\hat{g}$, e_0, f_0, h_0 form an $\mathfrak{sl}_2$ triple.

Proof: Let $e_0 = f(1)$, $f_0 = e(-1)$ and $h_0 = -h + c$. Then

$$[e_0, f_0] = [f(1), e(-1)] = [f, e] + (f, e)c = -h + c = h$$

since $[e, f] = h$ and $(f, e) = 1$ by the relations in 1.1 and 1.2. Similarly

$$[h_0, f_0] = [-h + c, e(-1)] = -[h, e(-1)] = -[h, e](0 - 1) = -2e(-1) = -2f_0$$

and

$$[h_0, e_0] = [-h + c, f(1)] = -[h, f(1)] = -[h, f](0 + 1) = 2f(1) = 2e_0$$

Therefore e_0, f_0, h_0 form an $\mathfrak{sl}_2$ triple.

(Exercise 7.6.1) Show that if $\{x_a\}_{a\in\Lambda}$ is a basis of $\mathfrak{g}$ then the lexicographically ordered monomials in x_a form a basis of $U\mathfrak{g}$.

Proof: The lexicographically ordered monomials form a basis of the symmetric algebra $S\mathfrak{g}$ by its definition and so form a basis of $\mathrm{Gr}(U\mathfrak{g})$ by PBW. First, it's clear that the LOM's of degree 1 form a basis of $U\mathfrak{g}_1$ since $T\mathfrak{g}_1 = \mathrm{Gr}U\mathfrak{g}_1 = S\mathfrak{g}_1 = \mathfrak{g}$ and $\mathrm{Gr}U\mathfrak{g}_1 \subseteq U\mathfrak{g}_1 \subseteq T\mathfrak{g}_1$.

Now,by induction the LOM's of degree $n-1$ form a basis of $U\mathfrak{g}_{n-1}$. But the LOM's of degree n form a basis of $\mathrm{Gr}(U\mathfrak{g})_n = U\mathfrak{g}_n/U\mathfrak{g}_{n-1}$. But every element of $U\mathfrak{g}_n$ can be written as an element of $U\mathfrak{g}_n/U\mathfrak{g}_{n-1}$ plus an element of $U\mathfrak{g}_{n-1}$ so if the LOM's of degree n form a basis of $U\mathfrak{g}_n/U\mathfrak{g}_{n-1}$ and the LOM's of degree $n-1$ form a basis of $U\mathfrak{g}_{n-1}$ then the union of these two bases is a basis for $U\mathfrak{g}_n$. But the union is simply the LOM's of degree n. Since $U\mathfrak{g} = \cup_{n\in\mathbb{N}} U\mathfrak{g}_n$, the union of all such bases forms a basis of $U\mathfrak{g}$, but this is simple the set of all LOM's.

(Exercise 8.4.2) Show that a highest weight representation V is in category $\mathscr{O}$

Proof: First, $V = U\hat{\mathfrak{n}}_- v$ is diagonal since for any basis element (ie lexicographically ordered monomial so element of the form $f_0^m f_1^n$), $f_0^m f_1^n \cdot v$ we have

$$h_0(f_0^m f_1^n \cdot v) = [h_0, f_0] f_0^{m-1} f_1^n \cdot v + f_0 h_0 f_0^{m-1} f_1^n \cdot v$$
$$= -2 f_0^m f_1^n \cdot v + f_0 h_0 f_0^{m-1} f_1^n \cdot v = \ldots = -2m f_0^m f_1^n \cdot v + f_0^m h_0 f_1^n \cdot v$$
$$= -2m f_0^m f_1^n \cdot v - 2n f_0^m f_1^n \cdot v + f_0^m f_1^n h_0 \cdot v$$
$$= (-2m + 2n + \lambda(h_0)) f_0^m f_1^n \cdot v = (\lambda(h_0) - m\alpha_0(h_0) - n\alpha_1(h_0)) f_0^m f_1^n \cdot v$$

where $\alpha_0 = (-\alpha, 1)$ and $\alpha_1 = (\alpha, 0)$ as before. By practicality the same computation,

$$h_1(f_0^m f_1^n \cdot v) = (2m - 2n + \lambda(h_1)) f_0^m f_1^n \cdot v = (\lambda(h_1) - m\alpha_0(h_1) - n\alpha_1(h_1)) f_0^m f_1^n \cdot v$$

Finally,

$$d(f_0^m f_1^n \cdot v) = [d, f_0] f_0^{m-1} f_1^n \cdot v + f_0 d f_0^{m-1} f_1^n \cdot v =$$
$$= -f_0^m f_1^n \cdot v + f_0 d f_0^{m-1} f_1^n \cdot v = \ldots = -m f_0^m f_1^n \cdot v + f_0^m d f_1^n \cdot v$$
$$(\lambda(d) - m) f_0^m f_1^n \cdot v = (\lambda(d) - m\alpha_0(d) - n\alpha_1(d)) f_0^m f_1^n \cdot v$$

So since $f_0^m f_1^n \cdot v$ is a basis of V of eigenvectors of $\hat{\mathfrak{g}}$, $\hat{\mathfrak{g}}$ acts diagonally. Therefore V satisfies the first condition of category $\mathscr{O}$.

Now, the above computations show that the weight spaces are all of the form $\mu = \lambda - m\alpha_0 - n\alpha_1$ and so $\lambda - \mu = m\alpha_0 + n\alpha_1$ is a positive root and $\mu < \lambda$ for all weights μ of V. Therefore V satisfies the third condition of category $\mathscr{O}$.

Finally, V satisfies the second condition of category $\mathscr{O}$ since each weight space $\mu = \lambda - m\alpha_0 - n\alpha_1$ is one dimensional, corresponding to $f_0^m f_1^n \cdot v$. Therefore V is in category $\mathscr{O}$.

CHAPTER 2

Quantum group $U_q\mathfrak{sl}_2$

1. Gaussian integers

1.1. Let v be an indeterminate. Define for each $n \in \mathbb{Z}$:

$$[n]_v := \frac{v^n - v^{-n}}{v - v^{-1}}$$

The following properties are easily verified:

(1) For $n \geq 1$ we have:
$$[n]_v = v^{n-1} + v^{n-3} + \cdots + v^{-n+3} + v^{-n+1}$$
and $[-n]_v = -[n]_v$. Therefore, we have
$$[n]_v \in \mathbb{Z}[v, v^{-1}]$$

(2)
$$[n]_v\big|_{v=1} = n$$

(3) $[0]_v = 0$, $[1]_v = 1$ and
$$[2]_v = v + v^{-1}$$

(4) For $m, n \in \mathbb{Z}$, $[m]_v + [n]_v \neq [n+m]_v$. However
$$v^{-m}[n]_v + v^n[m]_v = [n+m]_v$$

Proof. Using the definition of the Gaussian integers we have:

$$\begin{aligned} v^{-m}[n]_v + v^n[m]_v &= \frac{v^{-m}(v^n - v^{-n})}{v - v^{-1}} + \frac{v^m(v^n - v^{-n})}{v - v^{-1}} \\ &= \frac{v^{m-n} - v^{-m-n} + v^{n+m} - v^{m-n}}{v - v^{-1}} \\ &= [n+m]_v \end{aligned}$$

□

1.2.

$$[n]_v! := [n]_v[n-1]_v \cdots [1]_v$$
$$\begin{bmatrix} a \\ n \end{bmatrix}_v := \frac{[a]_v!}{[a-n]_v![n]_v!}$$

Lemma. *For every $a, n \in \mathbb{N}$, $a \geq n$ we have:*

$$\begin{bmatrix} a+1 \\ n \end{bmatrix}_v = v^{-n} \begin{bmatrix} a \\ n \end{bmatrix}_v + v^{a-n+1} \begin{bmatrix} a \\ n-1 \end{bmatrix}_v$$

and hence $\begin{bmatrix} a \\ n \end{bmatrix}_v \in \mathbb{Z}[v, v^{-1}]$.

PROOF. We begin by computing the right–hand side :

$$\begin{aligned}
\text{R.H.S.} &= v^{-n}\begin{bmatrix} a \\ n \end{bmatrix}_v + v^{a-n+1}\begin{bmatrix} a \\ n-1 \end{bmatrix}_v \\
&= v^{-n}\frac{[a]_v!}{[a-n]_v![n]_v!} + v^{a-n+1}\frac{[a]_v!}{[n-1]_v![a-n+1]_v!} \\
&= \frac{[a]_v!}{[n]_v![a-n+1]_v!}\left(v^{-n}[a-n+1]_v + v^{a-n+1}[n]_v\right) \\
&= \frac{[a]_v!}{[n]_v![a-n+1]_v!}[a+1]_v \\
&= \begin{bmatrix} a+1 \\ n \end{bmatrix}_v
\end{aligned}$$

□

2. Definition of $U_q\mathfrak{sl}_2$

Let k be a field and $q \in k$ be a non–zero element such that $q^2 \neq 1$. Mainly we will have the following examples in mind:

(1) $k = \mathbb{C}$ and $q \in \mathbb{C} \setminus \{0, 1, -1\}$.
(2) $k = \mathbb{C}(v)$ and $q = v$.

2.1. Definition. $U_q\mathfrak{sl}_2$ is a unital associative algebra over k generated by $\{K^{\pm 1}, E, F\}$ subject to the following relations:

(QG1)

$$KK^{-1} = K^{-1}K = 1$$

(QG2)

$$KEK^{-1} = q^2E \qquad\qquad KFK^{-1} = q^{-2}F$$

(QG3)

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

It will be convenient in computations to have the following notation:

$$[K; a] := \frac{q^aK - q^{-a}K^{-1}}{q - q^{-1}}$$

Note that we have the following reformulation of the relations of $U_q\mathfrak{sl}_2$:

(1) (QG3) is equivalent to $EF - FE = [K; 0]$.
(2) (QG2) is equivalent to :

$$[K; a]F = F[K; a-2]$$
$$[K; a]E = E[K; a+2]$$

Lemma. *The following relations hold in $U_q\mathfrak{sl}_2$, for every $r, s \geq 1$:*

$$EF^s = F^sE + [s]_qF^{s-1}[K; 1-s]$$
$$FE^r = E^rF - [r]_qE^{r-1}[K; r-1]$$

Proof. We only prove the first of the two identities. The proof of the second one is exactly similar. The proof is based on induction on s. For $s = 1$ the relation $EF = FE + [K; 0]$ is precisely the relation (QG3). We proceed with the induction step:

$$\begin{aligned}
EF^{s+1} &= (EF)F^s \\
&= (FE + [K;0])F^s \\
&= FEF^s + [K;0]F^s \\
&= F(F^sE + [s]_q F^{s-1}[K;1-s]) + F^s[K;-2s] \\
&= F^{s+1}E + F^s\left([s]_q[K;1-s] + [K;-2s]\right) \\
&= F^{s+1}E + F^s\left(\frac{q^{-s}([s]_q q + q^{-s})K + q^s([s]_q q^{-1} + q^s)K^{-1}}{q-q^{-1}}\right) \\
&= F^{s+1}E + [s+1]_q[K;-s]
\end{aligned}$$

where we have used the following two special cases of $v^{-n}[m]_v + v^m[n]_v = [n+m]_v$:

$$\begin{aligned}
v^{-1}[m]_v + v^m &= [m+1]_v \\
v^{-n} + v[n]_v &= [n+1]_v
\end{aligned}$$

□

3. PBW theorem

Proposition. *The set of monomials $\mathcal{S} = \{F^sK^nE^r : r, s \in \mathbb{N}, n \in \mathbb{Z}\}$ spans $U_q\mathfrak{sl}_2$.*

Proof. It suffices to prove that the set $\mathcal{S}$ is stable under left multiplication by $\{K^{\pm}, E, F\}$. We check this using the Lemma 2.1:

(1)
$$F.F^sK^nE^r = F^{s+1}K^nE^r$$

(2)
$$K^{\pm 1}.F^sK^nE^r = q^{\mp 2s}F^sK^{n\pm 1}E^s$$

(3)
$$E.F^sK^nE^r = q^{-2n}F^sK^nE^{r+1} + [s]_qF^{s-1}[K;1-s]K^nE^r$$

□

Theorem. *The set of monomials $\mathcal{S} = \{F^sK^nE^r : r, s \in \mathbb{N}, n \in \mathbb{Z}\}$ is linearly independent in $U_q\mathfrak{sl}_2$.*

Proof. Let $V = k[X, Y, Z^{\pm 1}]$ and define operators $\rho(E), \rho(F), \rho(K^{\pm 1})$ on V by:

$$\begin{aligned}
\rho(F).Y^sZ^nX^r &= Y^{s+1}Z^nX^r \\
\rho(K^{\pm 1})Y^sZ^nX^r &= q^{\mp 2s}Y^sZ^{n\pm 1}X^r \\
\rho(E)Y^sZ^nX^r &= q^{-2n}Y^sZ^nX^{r+1} + [s]_qY^{s-1}[Z;1-s]Z^nX^r
\end{aligned}$$

where $[Z;a] := \dfrac{Zq^a - Z^{-1}q^{-a}}{q-q^{-1}}$.

Claim: The operators $\{\rho(E), \rho(F), \rho(K^{\pm 1})\}$ satisfy the relations (QG1), (QG2) and (QG3).

Assuming the claim, ρ extends to a representation of $U_q\mathfrak{sl}_2$ on V. Moreover we have:

$$\rho(F^s K^n E^r).1 = Y^s Z^n X^r$$

Therefore $\{F^s K^n E^r.1 : r, s \in \mathbb{N}, n \in \mathbb{Z}\} \subset V$ is a linearly independent set. This proves that the set $\mathcal{S}$ is linearly independent and we are done.

Proof of the claim: We verify the relations directly:

(QG1) $\rho(K)\rho(K^{-1}) = \rho(K^{-1})\rho(K) = 1$ is clear.

(QG2) We prove the relation $KEK^{-1} = q^2 E$. The proof for the case of F is absolutely similar.

$$\begin{aligned}
\rho(K)(\rho(E)(\rho(K^{-1})(Y^s Z^n X^r))) &= q^{2s}\rho(K)\rho(E)(Y^s Z^{n-1} X^r) \\
&= q^{2s}\rho(K)\left(q^{-2(n-1)} Y^s Z^{n-1} X^{r+1} + [s]_q Y^{s-1}[Z; 1-s] Z^{n-1} X^r\right) \\
&= q^{-2(n-1)} Y^s Z^n X^{r+1} + q^2 [s]_q Y^{s-1}[Z; 1-s] Z^n X^r \\
&= q^2 \rho(E)(Y^s Z^n X^r)
\end{aligned}$$

(QG3)

$$\begin{aligned}
[\rho(E), \rho(F)] Y^s Z^n X^r &= \rho(E)(\rho(F)(Y^s Z^n X^r)) - \rho(F)(\rho(E)(Y^s Z^n X^r)) \\
&= \rho(E)(Y^{s+1} Z^n X^r) - \rho(F)\left(q^{-2n} Y^s Z^n X^{r+1} + [s]_q Y^{s-1}[Z; 1-s] Z^n X^r\right) \\
&= q^{-2n} Y^{s+1} Z^n X^{r+1} + [s+1]_q Y^s [Z; -s] Z^n X^r \\
&\quad - q^{-2n} Y^{s+1} Z^n X^{r+1} - [s]_q Y^s [Z; 1-s] Z^n X^r \\
&= Y^s \left([s+1]_q [Z; -s] - [s]_q [Z; 1-s]\right) Z^n X^r \\
&= Y^s \left(\frac{q^{-2s} Z - q^{2s} Z^{-1}}{q - q^{-1}}\right) Z^n X^r \\
&= \left(\frac{K - K^{-1}}{q - q^{-1}}\right) Y^s Z^n X^r
\end{aligned}$$

$\square$

4. Representation theory of $U_q\mathfrak{sl}_2$

For notational convenience we set $\mathcal{U} = U_q\mathfrak{sl}_2$ in this section. We further assume that q is not a root of unity.

4.1.

Proposition. *If V is a finite–dimensional representation of $\mathcal{U}$ then there exist $r, s \geq 0$ such that $E^r = F^s = 0$ on V.*

Proof. We begin by writing the Jordan decomposition of K:

$$V = \bigoplus_{p(x)} V_{(p)}$$

where

- $p(x)$ ranges over irreducible polynomials in $k[x]$.
- $V_{(p)} := \{v \in V : p(K)^n v = 0\}$. Since K is invertible this in particular yields: $V_{(p)} \neq 0 \Rightarrow p(0) \neq 0$.
- $V_{(p)} = V_{(p')}$ if, and only if p and p' are proportional.

Using $KF = q^{-2}FK$ and hence $p(K)F = Fp(q^{-2}K)$, we get:

$$F : V_{(p(x))} \to V_{(p(q^2x))}$$

Moreover, we have $V_{(p(q^r x))} = V_{(p(q^s x))}$ if, and only if $p(q^r x) = p(q^s x)$ (since p has non-zero constant term) which implies $q^{rn} = q^{sn}$ ($n =$ degree of p). This implies $r = s$ since q is not a root of unity. Since V is a finite–dimensional representation we get that F acts nilpotently. The proof for E is absolutely similar. $\square$

4.2.

Proposition. *Assume that $char(k) \neq 2$. If V is a finite–dimensional representation of $\mathcal{U}$ then K acts semisimply on V with eigenvalues $\pm q^a$ ($a \in \mathbb{Z}$).*

Proof. From previous proposition, we know that $F^s = 0$ on V, for some $s \geq 0$.

Claim 1:

$$\prod_{j=1-s}^{s-1} (K - q^j)(K + q^j) = 0$$

Claim 2: Define

$$h_r := \prod_{j=1-r}^{r-1} [K; r - s + j]$$

Then we have $F^{s-r}h_r = 0$ for every $0 \leq r \leq s$.

Note that the second claim for $r = s$ yields the first claim, which directly implies the assertion of the proposition. Thus we are reduced to proving the second claim, for which we shall need the following

Lemma.

$$E^r F^s = \sum_{i=0}^{\min(r,s)} \begin{bmatrix} r \\ i \end{bmatrix}_q \begin{bmatrix} s \\ i \end{bmatrix}_q [i]_q! F^{s-i} \widetilde{h}_i E^{r-i}$$

where we define:

$$\widetilde{h}_i := \prod_{j=1}^{i} [K; i - (r+s) + j]$$

Set $a_i = \begin{bmatrix} r \\ i \end{bmatrix}_q \begin{bmatrix} s \\ i \end{bmatrix}_q [i]_q!$. Assuming the statement of this lemma, we have (for $r \leq s$):

$$\begin{aligned}
0 &= E^r F^s \prod_{j=1}^{r-1} [K; r-s+j] \\
&= \sum_{i=0}^{r} a_i F^{s-i} \widetilde{h}_i E^{r-i} \prod_{j=1}^{r-1} [K; r-s+j] \\
&= a_r F^r \widetilde{h}_r \prod_{j=i}^{r-1} [K; r-s+j] + \sum_{i=0}^{r-1} a_i F^{s-i} \widetilde{h}_i E^{r-i} \prod_{j=1}^{r-1} [K; r-s+j] \\
&= a_r F^r h_r + \sum_{i=0}^{r-1} F^{s-i} h_i \prod_{j=1}^{r-i} [K; i+j-r-s] E^{r-i}
\end{aligned}$$

Now the proof of the second claim follows by induction on r.

Proof of the lemma: Let us begin by writing the straightening relation for unknown elements $h_i(r,s) \in \mathcal{U}^0$:

$$E^r F^s = \sum_{i=0}^{\min(r,s)} F^{s-i} h_i(r,s) E^{r-i} \tag{4.1}$$

Here $\mathcal{U}^0$ is the subalgebra generated by $K^{\pm 1}$. Let σ be the automorphism of $\mathcal{U}^0$ sending K to $q^{-2}K$. Then we have:

- $h_0(0,r) = h_0(r,0) = 1$
- $h_i(r,s) = h_i(s,r)$. To prove this apply the anti–automorphism ω, defined by $\omega(F) = E$, $\omega(E) = F$ and $\omega(K) = K$ to equation (4.1).

Using Lemma 2.1 one can write down the following recursive system, which together with the observations above, determines the elements $h_i(r,s)$:

- For every $0 \leq i \leq r$, $r \leq s$ we have:

$$h_i(r, s+1) = \sigma h_i(r,s) + [r-i+1]_q h_{i-1}(r,s)[K; i-r] \tag{4.2}$$

- For every $r < s$ and $0 \leq i \leq r+1$ we have:

$$h_i(r+1, s) = \sigma h_i(r,s) + [s-i+1]_q h_{i-1}(r,s)[K; i-s] \tag{4.3}$$

  with the convention that $h_{-j}(r,s) = h_{j+\min r,s}(r,s) = 0$ for every $j \geq 1$.

Thus it remains to check that $h_i(r,s)$ given by

$$h_i(r,s) = \begin{bmatrix} r \\ i \end{bmatrix}_q \begin{bmatrix} s \\ i \end{bmatrix}_q [i]_q! \prod_{j=1}^{i} [K; i+j-(r+s)]$$

satisfies this system. The base case $h_0(0,r) = 1$ and symmetry in r,s is clear. We prove (4.2).

$$\sigma h_i(r,s) + [r-i+1]_q [K; i-r] h_{i-1}(r,s) = \frac{[r]_q! [s]_q!}{[r-i]_q! [s-i+1]_q! [i]_q!}$$
$$\left([s-i+1]_q \prod_{j=1}^{i} [K; i+j-r-s-2] + [i]_q [K; i-r] \prod_{j=1}^{i-1} [K; i+j-r-s-1] \right)$$

$$= \frac{[r]_q![s]_q!}{[r-i]_q![s-i+1]_q![i]_q!}\prod_{j=1}^{i-1}[K;i+j-r-s-1]\left([s-i+1]_q[K;i-r-s-1]+[i]_q[K;i-r]\right)$$

One can directly verify that

$$[s-i+1]_K[i-r-s-1]+[i]_q[K;i-r]=[s+1]_q[K;2i-r-s-1]$$

which implies:

$$\begin{aligned}\sigma h_i(r,s)+[r-i+1]_q[K;i-r]h_{i-1}(r,s) &= \frac{[r]_q![s]_q!}{[r-i]_q![s-i+1]_q![i]_q!}[s+1]_q\prod_{j=1}^{i}[K;i+j-r-s-1] \\ &= \begin{bmatrix} r \\ i \end{bmatrix}_q \begin{bmatrix} s+1 \\ i \end{bmatrix}_q [i]_q!\prod_{j=1}^{i}[K;i+j-r-(s+1)] \\ &= h_i(r,s+1)\end{aligned}$$

The proof of (4.3) is similar and hence omitted. □

4.3. From now onwards, we assume that q is not a root of unity and $\mathrm{char}(k)\neq 2$.

Using Proposition 4.2 we have that every finite–dimensional $\mathcal{U}$–module M decomposes as:

$$M=\bigoplus_{\lambda\in k^\times} M_\lambda$$

where $M_\lambda=\{m\in M : Km=\lambda m\}$. Using finite dimensionality of M, we know that there exists $\lambda\in k^\times$ such that $M_\lambda\neq 0$ and $M_{q^2\lambda}=0$. If $v\in M_\lambda$ is any non–zero vector, then $Ev\in M_{q^2\lambda}=0$. Hence v is a highest–weight vector. Thus the submodule M' of M generated by v is a highest–weight module.

Corollary. *Every irreducible finite–dimensional $\mathcal{U}$–module is a highest–weight module for a unique highest–weight λ.*

The uniqueness follows from the fact that q is not a root of unity.

4.4. Highest weight modules. For $\lambda\in k^\times$ there exists a unique Verma module $M(\lambda)$ defined by:

$$M(\lambda):=\mathcal{U}/\mathcal{U}E+\mathcal{U}(K-\lambda)$$

Let us denote by $m_0\in M(\lambda)$ the coset of $1\in\mathcal{U}$. Then a basis of $M(\lambda)$ can be obtained as:

$$\{m_i:=F^i m_0 : i\geq 0\}$$

Moreover the action of $\mathcal{U}$ on $M(\lambda)$ can be written explicitly as:

$$\begin{aligned}Km_i&=\lambda q^{-2i}m_i\\ Fm_i&=m_{i+1}\\ Em_i&=[i]_q\frac{q^{1-i}\lambda-q^{i-1}\lambda^{-1}}{q-q^{-1}}m_{i-1}\end{aligned}$$

Proposition. (1) *$M(\lambda)$ has a unique maximal proper submodule $M(\lambda)'$.*
(2) *If $\lambda\neq\pm q^n$ for any $n\in\mathbb{N}$, then $M(\lambda)$ is irreducible.*
(3) *If $\lambda=\pm q^n$ for some $n\in\mathbb{N}$, then $M(\lambda)'=Span\{m_i : i\geq n+1\}$.*

Proof. (1) is clear. For (2) assume that $M(\lambda)$ is not irreducible, and take M' to be a proper submodule. Then there exists $i > 0$ such that $m_i \in M'$ but $m_{i-1} \notin M'$. Hence $Em_i = 0$, which implies:

$$Em_i = [i]_q \frac{q^{1-i}\lambda - q^{i-1}\lambda^{-1}}{q - q^{-1}} m_{i-1} = 0$$

Therefore, $\lambda^2 = q^{2(i-1)}$ which implies (2). (3) follows from a similar computation. □

Corollary. (1) *For every $\lambda \in k^\times$ there exists a unique irreducible highest–weight module $L(\lambda) := M(\lambda)/M(\lambda)'$ of highest–weight λ.*
(2) *$L(\lambda)$ is finite–dimensional if, and only if $\lambda \in \{\pm q^n : n \in \mathbb{N}\}$.*
(3) *$L(\lambda) \cong L(\mu)$ if, and only if $\lambda = \mu$.*

4.5. Casimir operator. Define:

$$C := FE + \frac{qK + q^{-1}K^{-1}}{(q - q^{-1})^2} \tag{4.4}$$

Using $FE = EF - [K; 0]$ we can rewrite this definition as:

$$C = EF + \frac{q^{-1}K + qK^{-1}}{(q - q^{-1})^2} \tag{4.5}$$

Proposition. (1) *C is a central element of $\mathcal{U}$.*
(2) *If V is a highest–weight module of highest–weight λ then C acts by a scalar on V, given by:*

$$C|_V = \frac{q\lambda + q^{-1}\lambda^{-1}}{(q - q^{-1})^2} Id_V$$

(3) *$C|_{M(\lambda)} = C|_{M(\mu)}$ if, and only if $\lambda = \mu$ or $\lambda = \mu^{-1}q^{-2}$.*

Proof. (1) It is clear that C commutes with K. We check the relation $CE = EC$:

$$\begin{aligned} EC &= EFE + E\frac{qK + q^{-1}K^{-1}}{(q - q^{-1})^2} \\ &= EFE + \frac{q^{-1}K + qK^{-1}}{(q - q^{-1})^2} E \\ &= CE \end{aligned}$$

The proof of $FC = CF$ is same.

(2): Let $v_\lambda \in V$ be a highest–weight vector of highest–weight λ. Then

$$\begin{aligned} Cv_\lambda &= \left(FE + \frac{qK + q^{-1}K^{-1}}{(q - q^{-1})^2} \right) v_\lambda \\ &= \frac{q\lambda + q^{-1}\lambda^{-1}}{(q - q^{-1})^2} v_\lambda \end{aligned}$$

Using the fact that C is central and V is generated by v_λ we obtain the assertion of (2). (3) is clear from (2). □

4.6.

Theorem. *Let M be a finite–dimensional $\mathcal{U}$–module. Then M is completely reducible.*

Proof. We begin by writing the Jordan decomposition of $C|_M$:

$$M = \bigoplus_f M_{(f)}$$

where

- f is an irreducible polynomial in $k[x]$.
- $M_{(f)} := \{m \in M : f(C)^n m = 0 \text{ for } n >> 0\}$

Since C is central, we get that each $M_{(f)}$ is a submodule of M. Thus we may assume that $M = M_{(f)}$ for some $f \in k[x]$ irreducible polynomial.

Now let us consider a composition series of M:

$$0 = M_0 \subset M_1 \subset \cdots \subset M_r = M$$

where each $M_i/M_{i-1} \cong L(\lambda_i)$ is a simple $\mathcal{U}$–module. Hence C acts on the quotient M_i/M_{i-1} by the scalar:

$$C|_{M_i/M_{i-1}} = \frac{q\lambda_i + q^{-1}\lambda_i^{-1}}{(q-q^{-1})^2}\mathrm{Id}_{M_i/M_{i-1}}$$

Let $c_i = \dfrac{q\lambda_i + q^{-1}\lambda_i^{-1}}{(q-q^{-1})^2}$. Then $x - c_i$ divides $f(x)$ for each i. Since f is irreducible, we get that $f(x) = x - c_i$ and $c_i = c_j$ for each i, j.

In particular this implies that $\lambda_i = \lambda_j$ (since the other case: $\lambda_i = q^{-2}\lambda_j$ contradicts the fact that M is finite–dimensional and hence $\lambda \in \pm q^{\mathbb{N}}$).

Thus we have proved that all composition factors of M are isomorphic to $L(\lambda)$ for some λ of the form $\pm q^n$ $(n \in \mathbb{N})$.

Now let $M = \bigoplus_{\mu \in k^\times} M_\mu$ be the weight space decomposition of M. Then we get:

$$\dim(M) = r\dim(L(\lambda)) \qquad \dim(M_\mu) = r\dim(L(\lambda)_\mu)$$

Choose a basis $\{m_1, \cdots, m_r\}$ of M_λ. Let $M' \subset M$ be the submodule generated by $\{m_i : 1 \le i \le r\}$:

$$M' := \sum_{i=1}^{r} \mathcal{U}m_i$$

We claim that $M' = M$ and $M' = \bigoplus \mathcal{U}m_i$. The first assertion follows from the fact that $(M/M')_\lambda = 0$ and the only composition factors of M/M' are isomorphic to $L(\lambda)$. The fact that the sum is direct is an easy consequence of the dimension count. The theorem is proved. $\square$

5. Hopf algebra structure

5.1. Motivation. Let A be a unital associative algebra.

- (Coproduct) For V, W A–modules, we have an A–module structure on $V \otimes W$ if and only if there exists an algebra homomorphism:

$$\Delta : A \to A \otimes A$$

- (Coassociativity) The natural vector space isomorphism $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$ is an A–module isomorphism if and only if

$$(\Delta \otimes 1) \circ \Delta = (1 \otimes \Delta) \circ \Delta$$

- (Counit) The one dimensional vector space k has an A–module structure (trivial A–module) if and only if we have an algebra homomorphism $\varepsilon : A \to k$.
- The natural isomorphisms $k \otimes V \cong V \cong V \otimes k$ are then A–module isomorphisms if and only if

$$(\varepsilon \otimes 1) \circ \Delta = 1 = (1 \otimes \varepsilon) \circ \Delta$$

5.2. Bialgebras. A bialgebra over k is a quintuple $(A, ., 1, \Delta, \varepsilon)$ such that:

(1) $(A, ., 1)$ is a unital associative algebra.
(2) $\Delta : A \to A \otimes A$ is an algebra homomorphism such that

$$(\Delta \otimes 1) \circ \Delta = (1 \otimes \Delta) \circ \Delta$$

(3) $\varepsilon : A \to k$ is an algebra homomorphism such that

$$(\varepsilon \otimes 1) \circ \Delta = 1 = (1 \otimes \varepsilon) \circ \Delta$$

Example. (1) $A = U(\mathfrak{g})$ for a Lie algebra $\mathfrak{g}$ has the following bialgebra structure:

$$\Delta(x) = x \otimes 1 + 1 \otimes x$$
$$\varepsilon(x) = 0$$

for every $x \in \mathfrak{g}[z, z^{-1}]$.

(2) Let Γ be a finite group and $A = k\Gamma$ be the group algebra of Γ. The bialgebra structure on A is given by:

$$\Delta(g) = g \otimes g$$
$$\varepsilon(g) = 1$$

for every $g \in \Gamma$.

(3) Let G be an algebraic group over k and $A = k[G]$ be its coordinate ring. Then A has the following bialgebra structure:

$$\Delta(f)(x, y) = f(xy)$$
$$\varepsilon(f) = f(1)$$

for every $f \in A$, $x, y \in G$.

5.3. Dual modules. If V is an A–module, then in order to have an A–module structure on V^* we need an algebra anti–homomorphism $S : A \to A$ (i.e, $S(ab) = S(b)S(a)$) called antipode. Given such S, we can define an action of A on V^* as follows:

$$(a.\phi)(v) := \phi(S(a)v)$$

for $\phi \in V^*$, $v \in V$ and $a \in A$.

Example.
(1) $A = U(\mathfrak{g})$: the antipode is given by $S(x) = -x$ for every $x \in \mathfrak{g}$.
(2) $A = k\Gamma$: the antipode is given by $S(g) = g^{-1}$ for every $g \in \Gamma$.
(3) $A = k[G]$: the antipode is given by $S(f)(x) = f(x^{-1})$ for every $f \in A$ and $x \in G$.

5.4. Hopf algebra. A Hopf algebra over k is a hextuple $(A, ., 1, \Delta, \varepsilon, S)$ where:

(1) $(A, ., 1, \Delta, \varepsilon)$ is a bialgebra.
(2) $S : A \to A$ is an algebra anti–homomorphism.
(3) The following condition holds:

$$m \circ (S \otimes 1) \circ \Delta = 1\varepsilon = m \circ (1 \otimes S) \circ \Delta$$

Axiom (3) can be interpreted as follows: let V be an A–module. It is natural to require that the following natural homomorphism of vector spaces is an A–module homomorphism:

$$\mathrm{tr} : V^* \otimes V \to k$$

which is equivalent to the assertion that for every $a \in A$, $\phi \in V^*$ and $v \in V$ we have:

$$\mathrm{tr}(a.(\phi \otimes v)) = \varepsilon(a)\phi(v)$$

[Sweedler's notation] For $a \in A$ we write (suppressing the subscript and summation sign): $\Delta(a) = a' \otimes a''$.

$$\begin{aligned}\mathrm{tr}(a(\phi \otimes v)) &= \mathrm{tr}(a'\phi \otimes a''v)\\ &= \phi(S(a')a''v)\\ &= \phi\left((m \circ (S \otimes 1) \circ \Delta)(a)v\right)\\ &= \varepsilon(a)\phi(v)\end{aligned}$$

Similarly we can interpret the second part of axiom (3) as the requirement that the following linear map is an A–module homomorphism:

$$k \to V \otimes V^* \cong End(V)$$

which maps $1 \in k$ to $\mathrm{Id}_V = u_i \otimes u^i$ (here $\{u_i\}$ is a basis of V and $\{u^i\}$ is the dual basis of V^*).

$$\begin{aligned}
a.(u_i \otimes u^i) &= a'u_i \otimes a''u^i \\
&= a'u_i \otimes u^i S(a'') \\
&= a'u_i \otimes u^i (S(a''u_j))u^j \\
&= a'u^i (S(a'')u_j) u_i \otimes u^j \\
&= (a'S(a'')u_j) \otimes u^j = \varepsilon(a) u_j \otimes u^j
\end{aligned}$$

Remark. (1) $S^2 \neq 1$ in general. However in the examples above $S^2 = 1$.
(2) $\varepsilon \circ S = \varepsilon$.
(3) $\Delta \circ S = (S \otimes S) \circ \Delta^{21}$.

5.5. $\mathcal{U}$ as a Hopf algebra.

Theorem. (1) *The following assignment extends to a unique algebra homomorphism $\Delta : \mathcal{U} \to \mathcal{U} \otimes \mathcal{U}$:*

$$\Delta(K) = K \otimes K$$
$$\Delta(E) = E \otimes 1 + K \otimes E$$
$$\Delta(F) = F \otimes K^{-1} + 1 \otimes F$$

(2) *The following assignment extends to a unique algebra homomorphism $\varepsilon : \mathcal{U} \to k$*

$$\varepsilon(K) = 1$$
$$\varepsilon(E) = \varepsilon(F) = 0$$

(3) *The antipode S is the unique algebra anti–homomorphism $S : \mathcal{U} \to \mathcal{U}$ such that $m \circ (S \otimes 1) \circ \Delta = 1\varepsilon = m \circ (1 \otimes S) \circ \Delta$, and is given by:*

$$S(K) = K^{-1}$$
$$S(E) = -K^{-1}E \qquad\qquad S(F) = -FK$$

(4) *We have $S^2(a) = K^{-1}aK$ for every $a \in \mathcal{U}$.*

Proof. (1) We check that $\{\Delta(K), \Delta(E), \Delta(F)\}$ satisfy the defining relations of $\mathcal{U}$:

(QG1) is clear.
(QG2) is proved only for the case of E:

$$\begin{aligned}
\Delta(K)\Delta(E)\Delta(K^{-1}) &= (K \otimes K)(E \otimes 1 + K \otimes E)(K^{-1} \otimes K^{-1}) \\
&= KEK^{-1} \otimes 1 + K \otimes KEK^{-1} \\
&= q^2 \Delta(E)
\end{aligned}$$

(QG3)

$$\begin{aligned}
[\Delta(E), \Delta(F)] &= [E \otimes 1 + K \otimes E, F \otimes K^{-1} + 1 \otimes F] \\
&= [E, F] \otimes K^{-1} + K \otimes [E, F] + KF \otimes EK^{-1} - FK \otimes K^{-1}E \\
&= \frac{1}{q - q^{-1}} \left(K \otimes K^{-1} - K^{-1} \otimes K^{-1} + K \otimes K - K \otimes K^{-1} \right) \\
&= \Delta \left(\frac{K - K^{-1}}{q - q^{-1}} \right)
\end{aligned}$$

(2) We again have to check that the defining relations of $\mathcal{U}$ are satisfied by $\{\varepsilon K = 1, \varepsilon(E) = 0, \varepsilon F = 0\}$ in k, which is clear.

(3) We begin by proving uniqueness. Since $\Delta(K) = K \otimes K$ and $\varepsilon(K) = 1$ we get:

$$1 = \varepsilon(K) = S(K)K \Rightarrow S(K) = K^{-1}$$

Next we use the definition of $\Delta(E)$ and $\varepsilon(E) = 0$:

$$0 = \varepsilon(E) = S(E).1 + K^{-1}E \Rightarrow S(E) = -K^{-1}E$$

The computation of $S(F)$ is similar.

To prove that S extends to an algebra anti–homomorphism, we need to check that $\{S(E), S(F), S(K)\}$ satisfy the relations of $\mathcal{U}$ in $\mathcal{U}^{\mathrm{op}}$:

(QG1) is again clear.

(QG2) is proved for the case of E:

$$\begin{aligned} S(KEK^{-1}) &= S(K^{-1})S(E)S(K) \\ &= K(-K^{-1}E)K^{-1} \\ &= q^2 S(E) \end{aligned}$$

(QG3)

$$\begin{aligned} S([E,F]) &= [S(F), S(E)] \\ &= FKK^{-1}E - K^{-1}EFK \\ &= \frac{K^{-1} - K}{q - q^{-1}} \\ &= S\left(\frac{K - K^{-1}}{q - q^{-1}}\right) \end{aligned}$$

(4) Since S is algebra anti–homomorphism, S^2 is an algebra homomorphism. Thus to prove that $S^2 = Ad(K^{-1})$, we only need to check it on the set of generators $\{K, E, F\}$ of $\mathcal{U}$, which follows directly from (3). □

6. Quasi–triangular structure

6.1. Almost cocommutative bialgebras. We say a bialgebra A is cocommutative if

$$\Delta(a) = \Delta^{21}(a) \text{ for every } a \in A$$

Here $\Delta^{21} = (12) \circ \Delta : A \otimes A \to A \otimes A$. Note that if A is cocommutative, then for any two A–modules, V and W, the natural flip operator $V \otimes W \to W \otimes V$ is an A–module homomorphism.

Definition. A bialgebra A is said to be almost cocommutative, if there exists an invertible element $R \in A^{\otimes 2}$ such that

$$\Delta^{21}(a) = R\Delta(a)R^{-1} \text{ for every } a \in A$$

Lemma. *Let (A, R) be an almost cocommutative bialgebra. Then for any two A–modules, V and W*

$$R^\vee := (12) \circ R : V \otimes W \to W \otimes V$$

is an A–module homomorphism

Proof. We need to prove that $R^\vee(a.(v \otimes w)) = a.(R^\vee(v \otimes w))$ for every $v \in V$, $w \in W$ and $a \in A$. This is equivalent to the following :

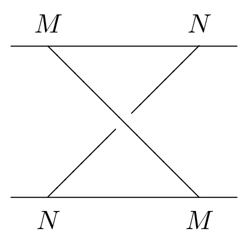
$$\begin{aligned} &R^\vee \circ \Delta(a) = \Delta(a) R^\vee \\ \Longleftrightarrow\ &R\Delta(a) = (12)\Delta(a)(12)R \\ \Longleftrightarrow\ &R\Delta(a) = \Delta^{21}(a)R \end{aligned}$$

Here we have used the fact that $\Delta^{21}(a)(w \otimes v) = (12) \circ \Delta(a) \circ (12)(v \otimes w)$. $\square$

Note that unlike the case of cocommutative bialgebras, the square of "flip" is not necessarily identity:

$$\left(R^\vee\right)^2 = (12)R(12)R = R^{21}R \neq 1$$

6.2. Braid diagrams. To better understand the motivation behind many of the axioms that will follow, it is instructive to represent $R^\vee$ by the following braid:

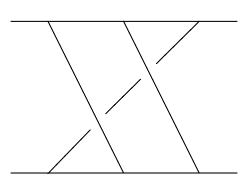


(6.1)

As a general principle the above diagram represents the element $R^\vee$ of $\mathrm{Hom}_A(M \otimes N, N \otimes M)$. We will follow this rule closely, whereby any braid diagram b connecting $(1, 2, \cdots, n)$ to $(\pi(1), \pi(2), \cdots \pi(n))$ (for some $\pi \in \mathfrak{S}_n$) represents an A–module homomorphism:

$$b(M_1, \cdots, M_n) \in \mathrm{Hom}_A(M_1 \otimes \cdots \otimes M_n, M_{\pi(1)} \otimes \cdots \otimes M_{\pi(n)})$$

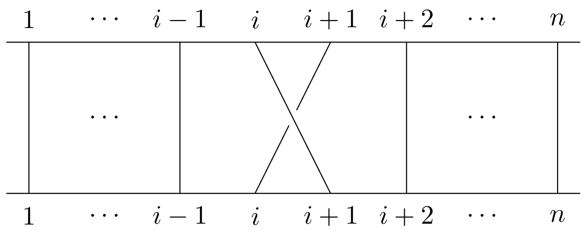
obtained by replacing each subdiagram of the form (6.1) by $R^\vee$, and a subdiagram of the form (6.1) with under–crossing by $(R^\vee)^{-1}$. For instance the following diagram represents an A–module homomorphism $M_1 \otimes M_2 \otimes M_3 \to M_3 \otimes M_1 \otimes M_2$ given by: $R^\vee_{M_1,M_3} \circ R^\vee_{M_2,M_3}$.



Let B_n denote the group of braids on n strands (Artin's braid group). It is well known that B_n is generated by $T_1, \cdots, T_{n-1}$ subject to the following relations:

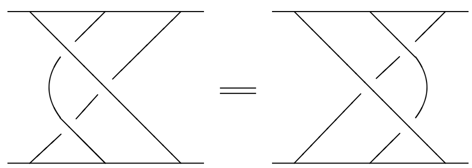
$$\begin{gathered} T_iT_j = T_jT_i \text{ if } |i - j| > 1 \\ T_iT_{i+1}T_i = T_{i+1}T_iT_{i+1} \text{ for } 1 \leq i \leq n-2 \end{gathered}$$

Here T_i is the following braid:



The relation $T_i T_{i+1} T_i = T_{i+1} T_i T_{i+1}$ can be pictorially seen as the following equality of braids:

(6.2)



6.3. Quantum Yang–Baxter equation. Let (A, R) be an almost cocommutative bialgebra and let V be an A–module.

Proposition. *The assignment* $T_i \mapsto R^{\vee}_{i,i+1}$ *extends to a representation of* B_n *on* $V^{\otimes n}$ *if and only if*

$$R_{12} R_{13} R_{23} = R_{23} R_{13} R_{12}$$

holds.

Proof. The relation $T_i T_j = T_j T_i$ for $|i-j| > 1$ is clear. For the braid relations it suffices to consider the case of $n = 3$ and prove

$$R^{\vee}_{12} R^{\vee}_{23} R^{\vee}_{12} = R^{\vee}_{23} R^{\vee}_{12} R^{\vee}_{23}$$

We begin by simplifying the left–hand side :

$$\begin{aligned} \text{L.H.S.} &= (12) R_{12} (23) R_{23} (12) R_{12} \\ &= (12)(23)(12) R_{23} R_{13} R_{12} \end{aligned}$$

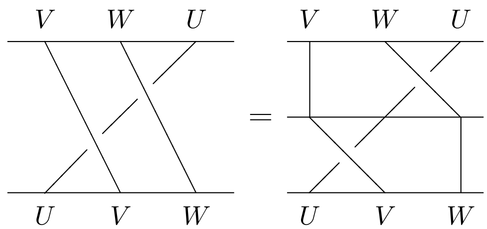
Similarly the right–hand side is same as $(23)(12)(23) R_{12} R_{13} R_{23}$ and we are done. □

6.4. Quasi–triangular Hopf algebras. Let (A, R) be an almost cocommutative Hopf algebra. We say A is quasi–triangular if the following (hexagon) axioms hold:

(QT1) $\Delta \otimes 1(R) = R_{13} R_{23}$
(QT2) $1 \otimes \Delta(R) = R_{13} R_{12}$

The origin of these axioms can be explained using the braid diagrams:

(QT1) Let V, W, U be three A–modules. Consider the following equality of braids:



This implies the following equation for morphisms $V \otimes W \otimes U \to U \otimes V \otimes W$:

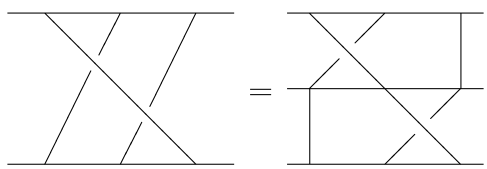
$$R^{\vee}_{V\otimes W,U} = (R^{\vee}_{V,U} \otimes 1_W) \circ (1_V \otimes R^{\vee}_{W,U}$$

which is equivalent to:

$$(123)(\Delta \otimes 1)(R) = (12)R_{12}(23)R_{23}$$

This equation is same as the axiom (QT1).

(QT2) can be similarly explained by the equality of the following braids:



6.5.

Proposition. *If (A, R) is a quasi–triangular Hopf algebra, then the quantum Yang–Baxter equation holds for R.*

Proof.

$$\begin{aligned} R_{12}R_{13}R_{23} &= R_{12}(\Delta \otimes 1)(R) \\ &= (\Delta^{21} \otimes 1)(R)R_{12} \\ &= R_{23}R_{13}R_{12} \end{aligned}$$

□

6.6.

Proposition. *Let (A, R) be a quasi–triangular Hopf algebra. Then we have the following:*

(1)

$$(\varepsilon \otimes 1)(R) = 1 \otimes 1 = (1 \otimes \varepsilon)(R)$$

(2)

$$(S \otimes 1)(R) = R^{-1}$$

(3)

$$(S \otimes S)(R) = R$$

PROOF. (1): apply $(\varepsilon \otimes 1 \otimes 1)$ to both sides of the equation $(\Delta \otimes 1)(R) = R_{13}R_{23}$, and use $(\varepsilon \otimes 1)(\Delta(a)) = 1 \otimes a$ to get:

$$R_{23} = (\varepsilon \otimes 1 \otimes 1)(R_{13})R_{23}$$

which implies that $(\varepsilon \otimes 1)(R) = 1$. The proof of the other equation is similar.

(2): Apply $m_{12} \circ (S \otimes 1 \otimes 1)$ to both sides of the equation $(\Delta \otimes 1)(R) = R_{13}R_{23}$ and use the axiom $(m \circ (S \otimes 1) \circ \Delta)(a) = \varepsilon(a)$ to get:

$$(\varepsilon \otimes 1)(R) = (S \otimes 1)(R).R$$

which together with (1) implies that $(S \otimes 1)(R) = R^{-1}$.

(3): apply $m_{23} \circ (S \otimes S \otimes 1)$ to the equation $(1 \otimes \Delta)(R) = R_{13}R_{12}$ to get:

$$(1 \otimes \varepsilon)(R) = (S \otimes S)(R)(S \otimes 1)(R)$$

which proves (3) using (1) and (2). □

Remark. The category of modules over A, for a quasi–triangular Hopf algebra A is a braided tensor category.

6.7. ***R*–matrix for $\mathcal{U}$.** Recall that the coproduct $\Delta : \mathcal{U} \to \mathcal{U} \otimes \mathcal{U}$ is given by:

$$\Delta(K) = K \otimes K$$
$$\Delta(E) = E \otimes 1 + K \otimes E$$
$$\Delta(F) = F \otimes K^{-1} + 1 \otimes F$$

Let τ be an algebra anti–automorphism of $\mathcal{U}$ defined by $\tau(K) = K^{-1}$ and $\tau(E) = E$, $\tau(F) = F$. We can twist Δ by τ to define another coproduct:

$$\Delta^\tau := (\tau \otimes \tau) \circ \Delta \circ \tau$$

Explicitly, Δ^τ is given by:

$$\Delta^\tau(K) = K \otimes K$$
$$\Delta^\tau(E) = E \otimes 1 + K^{-1} \otimes E$$
$$\Delta^\tau(F) = F \otimes K + 1 \otimes F$$

We aim at constructing Θ such that

$$\Delta(u)\Theta = \Theta\Delta^\tau(u) \text{ for every } u \in \mathcal{U} \tag{6.3}$$

Drinfeld's ansatz: We look for Θ of the following form:

$$\Theta = \sum_{n \geq 0} a_n F^n \otimes E^n \text{ where } a_n \in k$$

Thus we try to solve for $a_n \in k$ so that $\Theta = \sum_{n \geq 0} a_n F^n \otimes E^n$ satisfies (6.3). It is clear that the equation (6.3) holds for $u = K$ without any constraints on $a_n's$. Let us begin by considering (6.3) for $u = E$:

$$(E \otimes 1)\Theta - \Theta(E \otimes 1) = -(K \otimes E)\Theta + \Theta(K^{-1} \otimes E)$$

$$
\begin{aligned}
\text{L.H.S.} &= \sum a_n [E, F^n] \otimes E^n \\
&= \sum a_n [n]_q F^{n-1}[K; 1-n] \otimes E^n \\
\text{R.H.S.} &= -\sum a_n (KF^n - F^n K^{-1}) \otimes E^{n+1} \\
&= -\sum a_n (q-q^{-1}) q^{-n} \frac{Kq^{-n} - Kq^n}{q-q^{-1}} \otimes E^{n+1} \\
&= -\sum a_n q^{-n} (q-q^{-1}) F^n [K; -n] \otimes E^{n+1}
\end{aligned}
$$

Therefore we get the following constraint on the coefficients:

$$a_{n+1} = -a_n \frac{q^{-n}(q-q^{-1})}{[n+1]_q}$$

which has the unique solution given by:

$$a_n = (-1)^n \frac{(q-q^{-1})^n}{[n]_q!} q^{-n(n-1)/2} a_0$$

Next let us consider the equation (6.3) for $u = F$:

$$\Theta(1 \otimes F) - (1 \otimes F)\Theta = (F \otimes K^{-1})\Theta - \Theta(F \otimes K)$$

$$
\begin{aligned}
\text{L.H.S.} &= \sum a_n F^n \otimes [E^n, F] \\
&= \sum [n]_q a_n F^n \otimes E^{n-1}[K; n-1] \\
\text{R.H.S.} &= \sum a_n F^{n+1} \otimes (K^{-1}E^n - E^n K) \\
&= -\sum a_n q^{-n} (q-q^{-1}) F^{n+1} \otimes E^n [K; n]
\end{aligned}
$$

and we obtain the same recurrence relation for $a_n' s$. Hence we have proved:

Proposition.

$$\Theta = \sum_{n \geq 0} (-1)^n q^{-n(n-1)/2} \frac{(q-q^{-1})^n}{[n]_q!} F^n \otimes E^n$$

is the unique solution of (6.3).

Remark. (1) Let us define $\{n\} := \frac{q^{2n}-1}{q^2-1}$ so that we have $[n]_q = q^{-(n-1)}\{n\}$. In this notation Θ can be rewritten as:

$$\Theta = \sum_{n \geq 0} (-1)^n \frac{(q-q^{-1})^n}{\{n\}!} F^n \otimes E^n$$

or more compactly $\Theta = \exp_q(-(q-q^{-1})F \otimes E)$, where we define the q–exponential as:

$$\exp_q(x) := \sum_{n \geq 0} \frac{x^n}{\{n\}!}$$

(2) Θ defined above does not lie in $\mathcal{U} \otimes \mathcal{U}$. However for any two finite–dimensional $\mathcal{U}$–modules M, N, Θ can be evaluated to give a well–defined element $\Theta_{M,N} \in End(M \otimes N)$, which has the following properties:
   (a) $\Theta_{M,N}$ preserves weight spaces of $M \otimes N$.
   (b) $\Theta_{M,N}$ on a given weight space is unipotent (and hence invertible).

6.8. Consider the following subset of $k^\times$:

$$\widetilde{\Lambda} := \{\pm q^n : n \in \mathbb{Z}\} \subset k^\times$$

We have proved that for every finite–dimensional representation of $\mathcal{U}$, say M, we have the weight space decomposition:

$$M = \bigoplus_{\lambda \in \widetilde{\Lambda}} M_\lambda$$

Consider a function $f : \widetilde{\Lambda} \times \widetilde{\Lambda} \to k^\times$ and extend it to an operator $\widetilde{f} : M \otimes N \to M \otimes N$ by:

$$\widetilde{f}|_{M_\lambda \otimes N_\mu} = f(\lambda, \mu).\mathrm{Id}$$

Proposition. *Let* $\Theta^f := \Theta \circ \widetilde{f}$*. Then* Θ^f *satisfies*

$$\Delta(u) \circ \Theta^f = \Theta^f \circ \Delta^{21}(u) \tag{6.4}$$

for every $u \in \mathcal{U}$ *if, and only if*

$$f(q^2\lambda, \mu) = \mu^{-1} f(\lambda, \mu) \qquad f(\lambda, q^2\mu) = \lambda^{-1} f(\lambda, \mu) \tag{6.5}$$

PROOF. Since both Θ and $\widetilde{f}$ preserve the weight space decomposition, the equation (6.4) holds for any f. It remains to check (6.4) for $u = E, F$. Let us begin by rewriting (6.4):

$$\Delta^\tau(u) \circ \widetilde{f} = \widetilde{f} \circ \Delta^{21}(u)$$

For $u = E$, we simplify this equation on $M_\lambda \otimes N_\mu$ as:

$$\begin{aligned}
\text{L.H.S.} &= \Delta^\tau(E) \circ \widetilde{f}|_{M_\lambda \otimes N_\mu} \\
&= f(\lambda, \mu)(E \otimes 1 + K^{-1} \otimes E)|_{M_\lambda \otimes N_\mu} \\
&= f(\lambda, \mu)(E \otimes 1 + \lambda^{-1} 1 \otimes E) \\
\text{R.H.S.} &= \widetilde{f} \circ \Delta^{21}(E)|_{M_\lambda \otimes N_\mu} \\
&= \widetilde{f} \circ (1 \otimes E + E \otimes K)|_{M_\lambda \otimes N_\mu} \\
&= f(\lambda, q^2\mu)(1 \otimes E) + \mu f(q^2\lambda, \mu)(E \otimes 1)
\end{aligned}$$

which finishes the proof. □

Remark. As a corollary we obtain an intertwiner $M \otimes N \to N \otimes M$ for finite–dimensional modules M, N of $\mathcal{U}$:

$$M \otimes N \xrightarrow{\Theta^f \circ (12)} N \otimes M$$

for any solution f of (6.5).

6.9. QYBE. The aim of this section is to prove that the quantum Yang–Baxter equation holds for the intertwiner $(12) \circ \Theta^f$. We will begin by some preparatory results:

Lemma. (1)

$$\Delta(E^n) = \sum_{r=0}^{n} \begin{bmatrix} n \\ r \end{bmatrix}_q q^{r(n-r)} E^{n-r} K^r \otimes E^r$$

(2)

$$\Delta(F^n) = \sum_{r=0}^{n} \left[\begin{array}{c} n \\ r \end{array} \right]_q q^{r(n-r)} F^r \otimes F^{n-r} K^{-r}$$

(3)

$$a_n a_m = q^{nm} \left[\begin{array}{c} n+m \\ n \end{array} \right]_q a_{n+m}$$

Define Θ' *and* Θ'' *as:*

$$\Theta' = \sum_{n \geq 0} a_n F^n \otimes K^n \otimes E^n$$

$$\Theta'' = \sum_{n \geq 0} a_n F^n \otimes K^{-n} \otimes E^n$$

(4)

$$\widetilde{f}_{12} \circ \Theta_{13} = \Theta' \circ \widetilde{f}_{12}$$

(5)

$$\widetilde{f}_{23} \circ \Theta_{13} = \Theta'' \circ \widetilde{f}_{23}$$

(6)

$$\widetilde{f}_{12} \widetilde{f}_{13} \Theta_{23} = \Theta_{23} \widetilde{f}_{12} \widetilde{f}_{13}$$

(7)

$$\widetilde{f}_{23} \widetilde{f}_{13} \Theta_{12} = \Theta_{12} \widetilde{f}_{23} \widetilde{f}_{13}$$

PROOF.

(1) The proof is by induction on n. For $n = 1$ we have by definition:

$$\Delta(E) = E \otimes 1 + K \otimes E$$

Assume the assertion of (1) of Lemma for $n \geq 1$. Then we get:

$$\begin{aligned}
\Delta(E^{n+1}) &= (E \otimes 1 + K \otimes E) \left(\sum_{r=0}^{n} q^{r(n-r)} \left[\begin{array}{c} n \\ r \end{array} \right]_q E^{n-r} K^r \otimes E^r \right) \\
&= \sum_{r=0}^{n+1} \left(q^{r(n-r)} \left[\begin{array}{c} n \\ r \end{array} \right]_q + q^{(r+1)(n+1-r)} \left[\begin{array}{c} n \\ r-1 \end{array} \right]_q \right) E^{n+1-r} K^r \otimes E^r \\
&= \sum_{r=0}^{n+1} q^{r(n+1-r)} \left[\begin{array}{c} n+1 \\ r \end{array} \right]_q E^{n+1-r} K^r \otimes E^r
\end{aligned}$$

(2) the proof is same as that of (1).

(3)

$$\begin{aligned}
a_n a_m &= (-1)^{n+m} (q - q^{-1})^{n+m} q^{-\frac{n(n-1)+m(m-1)}{2}} \frac{1}{[n]_q!} \frac{1}{[m]_q!} \\
&= \left[\begin{array}{c} n+m \\ n \end{array} \right]_q q^{nm} (-1)^{n+m} q^{-\frac{(n+m)(n+m-1)}{2}} (q - q^{-1})^{n+m} \frac{1}{[n+m]_q!} \\
&= q^{nm} \left[\begin{array}{c} n+m \\ n \end{array} \right]_q a_{n+m}
\end{aligned}$$

(4) We compute both sides of (4) on tensor product of weight spaces of weights λ, μ, ν. For each $n \geq 0$ we have:

$$\begin{aligned}\widetilde{f}_{12} \circ (F^n \otimes 1 \otimes E^n)|_{M_\lambda \otimes M'_\mu \otimes M''_\nu} &= f(q^{-2n}\lambda, \mu)(F^n \otimes 1 \otimes E^n) \\ &= \mu^n f(\lambda, \mu)(F^n \otimes 1 \otimes E^n) \\ &= (F^n \otimes K^n \otimes E^n) \circ \widetilde{f}|_{M_\lambda \otimes M'_\mu \otimes M''_\nu}\end{aligned}$$

where we have used the identity (6.5).

(5) the proof is same as that of (4).

(6) Again we compute the operator $\widetilde{f}_{12}\widetilde{f}_{13}\Theta_{23}$ on a tensor product of weight spaces $M_\lambda \otimes M'_\mu \otimes M''_\nu$.

$$\begin{aligned}\widetilde{f}_{12}\widetilde{f}_{13}(1 \otimes F^n \otimes E^n)|_{M_\lambda \otimes M'_\mu \otimes M''_\nu} &= f(\lambda, q^{-2n}\mu) f(\lambda, q^{2n}\nu)(1 \otimes F^n \otimes E^n) \\ &= f(\lambda, \mu) f(\lambda, \nu)(1 \otimes F^n \otimes E^n) \\ &= (1 \otimes F^n \otimes E^n)\widetilde{f}_{12}\widetilde{f}_{13}|_{M_\lambda \otimes M'_\mu \otimes M''_\nu}\end{aligned}$$

(7) the proof is same as that of (6).

$\square$

Proposition. (1)

$$(\Delta \otimes 1)\Theta = (1 \otimes \Theta)\Theta''$$

(2)

$$(1 \otimes \Delta)\Theta = (\Theta \otimes 1)\Theta'$$

PROOF. We prove (1) only.

$$\begin{aligned}(\Delta \otimes 1)\Theta &= \sum a_n \Delta(F^n) \otimes E^n \\ &= \sum_{n \geq 0} a_n \left(\sum_{r=0}^{n} a^{r(n-r)} \begin{bmatrix} n \\ r \end{bmatrix}_q F^r \otimes F^{n-r}K^{-r} \right) \otimes E^n \\ &= \sum_{\substack{n \geq 0 \\ 0 \leq r \leq n}} a_n q^{r(n-r)} \begin{bmatrix} n \\ r \end{bmatrix}_q (1 \otimes F^{n-r} \otimes E^{n-r})(F^r \otimes K^{-r} \otimes E^r) \\ &= \sum_{\substack{n \geq 0 \\ 0 \leq r \leq n}} a_{n-r}a_r (1 \otimes F^{n-r} \otimes E^{n-r})(F^r \otimes K^{-r} \otimes E^r) \\ &= (1 \otimes \Theta)\Theta''\end{aligned}$$

$\square$

Theorem.

$$\Theta^f_{12}\Theta^f_{13}\Theta^f_{23} = \Theta^f_{23}\Theta^f_{13}\Theta^f_{12}$$

PROOF. Using (4)–(7) of Lemma 6.9 the above equation can be shown to be equivalent to:

$$\Theta_{12}\Theta'\Theta_{23} = \Theta_{23}\Theta''\Theta_{12}$$

Using (2) of Proposition 6.9 the left–hand side of this equation is simplified as:

$$\begin{aligned}\Theta_{12}\Theta'\Theta_{23} &= (1 \otimes \Delta)\Theta.\Theta_{23} \\ &= \Theta_{23}(1 \otimes \Delta^\tau)\Theta\end{aligned}$$

Thus we are reduced to proving that $(1 \otimes \Delta^\tau)\Theta = \Theta''\Theta_{12}$ which follows directly from (2) of Proposition 6.9 and definition of τ. $\square$

6.10. Hexagon axioms. Let $\alpha_{M,N} : M \otimes N \to N \otimes M$ be the commutativity constraint given by $\Theta^f \circ (12)$. In order to establish the hexagon axioms for this intertwiner we will need the following constraints on f:

$$f(\lambda\mu, \nu) = f(\lambda, \nu)f(\mu, \nu) \quad f(\lambda, \mu\nu) = f(\lambda, \mu)f(\lambda, \nu) \tag{6.6}$$

Proposition. *The following diagrams commute if and only if* (6.6) *holds for* f.

$$\begin{tikzcd}
M \otimes M' \otimes M'' \arrow[rr, "\alpha_{M\otimes M', M''}"] \arrow[dr, "1\otimes\alpha_{M',M''}"'] & & M'' \otimes M \otimes M' \\
& M \otimes M'' \otimes M \arrow[ur, "\alpha_{M,M''}\otimes 1"'] &
\end{tikzcd}$$

$$\begin{tikzcd}
M \otimes M' \otimes M'' \arrow[rr, "\alpha_{M, M'\otimes M''}"] \arrow[dr, "\alpha_{M,M'}\otimes 1"'] & & M' \otimes M'' \otimes M \\
& M' \otimes M \otimes M'' \arrow[ur, "1\otimes\alpha_{M,M''}"'] &
\end{tikzcd}$$

Proof. We will only focus on the first of the two diagrams. The top arrow is given by:

$$(1 \otimes \Delta)(\Theta \circ \widetilde{f}) \circ (123)$$

which on a tensor product of weight spaces $M_\lambda \otimes M'_\mu \otimes M''_\nu$ is given by:

$$(1 \otimes \Delta)(\Theta)f(\nu, \lambda\mu)(123)$$

Now we compute the other homomorphism:

$$\Theta_{12}\widetilde{f}_{12}(12)\Theta_{23}\widetilde{f}_{23}(23) = \Theta_{12}\widetilde{f}_{12}\Theta_{13}\widetilde{f}_{13}(123)$$

Using the commutation relation (4) of Lemma 6.9 we get

$$= \Theta_{12}\Theta'\widetilde{f}_{12}\widetilde{f}_{13}(123)$$

which evaluated on a tensor product $M_\lambda \otimes M'_\mu \otimes M''_\nu$ is given by $\Theta_{12}\Theta' f(\nu, \lambda)f(\nu, \mu)(123)$. Comparing the two computations the assertion follows by Proposition 6.9. $\square$

6.11. We have proved that the category of finite–dimensional representations of $\mathcal{U}$ has a braided tensor structure with commutativity constraint given by $\Theta^f \circ (12)$ provided we can find f satisfying (6.5) and (6.6). Using (6.5) iteratively, we have the following constraints on f:

$$f(\epsilon_1 q^{2n}, \epsilon_2 q^{2m}) = \epsilon_1^m \epsilon_2^n q^{-2nm} f(\epsilon_1, \epsilon_2)$$
$$f(\epsilon_1 q^{2n+1}, \epsilon_2 q^{2m}) = \epsilon_1^m \epsilon_2^n q^{-(2n+1)m} f(\epsilon_1 q, \epsilon_2)$$
$$f(\epsilon_1 q^{2n}, \epsilon_2 q^{2m+1}) = \epsilon_1^m \epsilon_2^n q^{-n(2m+1)} f(\epsilon_1, \epsilon_2 q)$$
$$f(\epsilon_1 q^{2n+1}, \epsilon_2 q^{2m+1}) = \epsilon_1^m \epsilon_2^n q^{-(2n+1)m-n} f(q\epsilon_1, q\epsilon_2)$$

Thus we can freely choose any values for $f(\epsilon_1 q^a, \epsilon_2 q^b)$ for $\epsilon_1, \epsilon_2 \in \{\pm 1\}$ and $a, b \in \{0, 1\}$ and get a solution of (6.5) using the above relations.

However we can only solve (6.6) if q admits a square–root in k. To see this, we observe that (6.6) implies that $f(q^a, q^b) = f(q,q)^{ab}$. Thus we have

$$\begin{aligned} f(q, q^2) &= q^{-1} \text{ (by (6.5))} \\ &= f(q,q)^2 \text{ (by (6.6))} \end{aligned}$$

Hence $f(q,q)^{-1}$ is a square–root of q.

Even with the assumption that q admits a square–root, a solution f of (6.5) and (6.6) cannot be found for the full $\widetilde{\Lambda}$:

$$\begin{aligned} f(-1,1)^2 &= f(-1,1) = 1 \\ f(-1,q^2) &= f(-1,q)^2 = f(1,q) = 1 \\ &= -f(-1,-1) = -1 \end{aligned}$$

which is a contradiction. Let us define $\Lambda := q^{\mathbb{Z}}$. The previous arguments imply that we can solve for f restricted to Λ, but not for $\widetilde{\Lambda}$. This prompts the following

Definition. A finite–dimensional representation V of $\mathcal{U}$ is said to be of type I (or II respectively) if the set of weights of V, denoted by $P(V)$ is a subset of Λ (or $-\Lambda$ respectively).

The category of type I representations of $\mathcal{U}$ forms a braided tensor category. However the category of type II representations only forms a module category over the category of type I representations.

CHAPTER 3

Quantum affine sl_2: $U_q\hat{\mathfrak{sl}}_2$

1. Two Presentations of $U_q\hat{\mathfrak{sl}}_2$

Recall that $\hat{\mathfrak{sl}}_2$ as a central extension of $\mathfrak{sl}_2[t, t^{-1}]$ was presented on $\{e_i, f_i, h_i\}_{i=0}$ with

$$e_1 = e \otimes t^0 \qquad f_1 = f \otimes t^0 \qquad h_1 = h \otimes t^0 \tag{1.1}$$

$$e_0 = f \otimes t \qquad f_0 = e \otimes t^{-1} \qquad h_1 = -h \otimes t^0 + c \tag{1.2}$$

with the relations

$$[h_i, h_j] = 0 \tag{1.3}$$

$$[h_i, e_j] = a_{ij} e_j \qquad [h_i, f_j] = -a_{ij} f_j \tag{1.4}$$

$$[e_i, f_j] = \delta_{ij} h_i \tag{1.5}$$

$$\mathrm{ad}(e_i)^{1-a_{ij}} e_j = \mathrm{ad}(e_i)^3 e_j = 0 \tag{1.6}$$

$$\mathrm{ad}(f_i)^{1-a_{ij}} f_j = \mathrm{ad}(f_i)^3 f_j = 0 \tag{1.7}$$

where

$$A = (a_{ij}) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

and

$$\mathrm{ad}(x)^n = (\ell(x) - r(x))^n = \sum_{k=0}^{n} \binom{n}{k} \ell(x)^k r(x)^{n-k} (-1)^{-1}$$

1.1. Definition. $U_q\hat{\mathfrak{sl}}_2$ is the associated algebra over k with generators $X_i^{\pm}$, $K_i^{\pm}$ for $i = 0, 1$ and the relations

$$K_i K_i^{-1} = K_i^{-1} K_i = 1 \tag{1.8}$$

$$K_0 K_1 = K_1 K_0 K_i X_j^{\pm} K_i^{-1} = q^{\pm a_{ij}} X_j^{\pm} \tag{1.9}$$

$$[X_i^+, X_j^-] = \delta_{ij} \frac{K_i - K_i^{-1}}{q - q^{-1}} \tag{1.10}$$

$$(X_i^{\pm})^3 X_j^{\pm} - [3](X_i^{\pm})^2 X_j^{\pm} X_i^{\pm} + [3] X_i^{\pm} X_j^{\pm} (X_i^{\pm})^2 - X_j^{\pm} (X_i^{\pm})^3 = 0 \tag{1.11}$$

The above algebra is a Hopf Algebra with coproduct

$$\begin{aligned} \Delta(K_i) &= K_i \otimes K_i \\ \Delta(X_i^+) &= X_i^+ \otimes K_i + 1 \otimes X_i^+ \\ \Delta(X_i^-) &= X_i^- \otimes 1 + K_i^{-1} \otimes X_i^- \end{aligned} \tag{1.12}$$

and antipode

$$S(K_i) = K_i^{-1}$$
$$S(X_i^+) = -X_i^+ K_i^{-1} \tag{1.13}$$
$$S(X_i^-) = K_i \otimes X_i^-$$

In addition we have the following presentation due to Drinfeld:

1.2. Theorem. $U_q\hat{\mathfrak{sl}}_2$ is homomorphic to the associative algebra over $\mathbb{C}$ with generators $x_k^\pm$ for $k \in \mathbb{Z}$, h_k for $k \in \mathbb{Z}^\times$, $K^{\pm 1}$ and central elements $C^{\pm 1}(= q^C)$ such that

(QR 1)
$$CC^{-1} = C^{-1}C = KK^{-1} = K^{-1}K = 1$$

(QR 2)
$$[h_k, h_\ell] = \delta_{k,-\ell}\frac{1}{k}[2K]\frac{C^k - C^{-k}}{q - q^{-1}}$$

(QR 3)
$$Kh_k = h_kK$$

To make sense of the above we can think of $h_k \to h(k) = h \otimes t^k$ as $q \to 1$. To make sense of (QR 2) recall that in $\hat{\mathfrak{sl}}_2$ we have $[h(k), h(\ell)] = [h, h](k + \ell) + k\delta_{k+\ell}0(h, h)c = 2k\delta_{k+\ell,0}c$. Then under the degeneration $q \to 1$
$$\frac{C^k - C^{-k}}{q - q^{-1}} = \frac{q^{kc} - q^{-kc}}{q - q^{-1}} \to kc$$
Note however that $K = q^{h(0)}$ is exponentiated while h_k is not.

(QR 4)
$$Kx_k^\pm K^{-1} = q^{\pm 2}x_K^\pm$$

Where $x_K^+ \underset{q=1}{\longrightarrow} e(k)$ and $x_K^- \underset{q=1}{\longrightarrow} f(k)$.

(QR 4)
$$x_{k+1}^\pm x_\ell^\pm - q^{\pm 2}x_\ell^\pm x_{k+1}^\pm = q^{\pm 2}x_k^\pm x_{\ell+1}^\pm - x_{\ell+1}^\pm x_k^\pm$$

This corresponds to the fact that for $q = 1$, $[x^\pm(k + 1), x^\pm(\ell)] = [x^\pm(k), x^\pm(\ell + 1)]$.

(QR 5)
$$[x_k^\pm, x_\ell^-] = \frac{1}{q - q^{-1}}(C^{k-\ell}\psi_{k+1} - \Phi_{k+1})$$

where

$$\sum_{k\geq 0}\psi_k u^k = K\exp\left((q-q^{-1})\sum_{k\geq 1}h_k u^k\right)$$

$$\sum_{k\geq 0}\Phi_k u^k = K\exp\left(-(q-q^{-1})\sum_{k\geq 1}h_k u^k\right)$$

Remark: We note here that there is another system of generators and relations that can be nested in these relations. Namely,$\psi(u)K^{-1} = \exp((q-q^{-1})h^+(u)$ implies that $h^+(u) = \log(\psi(u)K^{-1})/(q-q^{-1})$.

Proof. We need to show that the algebra we have defined above is indeed homomorphic to $U_q\hat{\mathfrak{sl}}_22$. But we can simply give the homomorphism explicitly as:

$$\begin{array}{lll} K_0 \mapsto CK^{-1}, & K_1 \mapsto K & X_1^{\pm} \mapsto x_0^{\pm} \\ X_0^+ \mapsto x_1^- K^{-1} & X_0^- \mapsto C^{-1}Kx_{-1}^+ & \end{array}$$

The first three maps are obvious, the last two less so but it can be shown that all together they define an isomorphism of algebras. □

Remark: $\forall i\in\mathbb{Z}$ there is a copy of the $U_q\mathfrak{sl}_2$ contained in $U_q\hat{\mathfrak{sl}}_2$ given by

$$E\mapsto x_i^+ \qquad F\mapsto C^{-i}x_{-i}^- \qquad K\mapsto KC^i$$

This can be checked by looking at the classical case: $[e(n), f(-n)] = h(0) + n(e,f)c = h(0) + nc$ and using then checking the relation under the image of the morphism:

$$[x_i^+, C^{-i}x_{-i}^-] = C^{-i}\frac{C^{2i}\psi_0 - \Phi_0}{q-q^{-1}} = \frac{C^iK - C^{-i}K^{-1}}{q-q^{-1}}$$

2. Finite Dimensional Representations of $U_q\hat{\mathfrak{sl}}_2$

Let H be the subalgebra of $U_q\hat{\mathfrak{sl}}_2$ generated by C, K and h_k and let $N_\pm$ be the subalgebra of $U_q\hat{\mathfrak{sl}}_2$ generated by C, K and $X_k^\pm$

Proposition. $U_q\hat{\mathfrak{sl}}_2 \cong N_-\otimes H\otimes N_+$ *(PBW)*

Now, recall that when we spoke of catagory $\mathcal{O}$ of $\hat{\mathfrak{g}}$ we used the triangulation given by

$$\hat{\mathfrak{g}} = \hat{\mathfrak{g}}_-\oplus\hat{\mathfrak{h}}\oplus\hat{g}_+ = \left(t^{-1}\mathfrak{g}[t^{-1}]\oplus n_-\right)\oplus\hat{\mathfrak{h}}\oplus(t\mathfrak{g}[t]\oplus n_+)$$

Definition. A vector Ω in a U-module V is *highest weight* if it is both an eigenvector of $K^{\pm 1}$, $C^{\pm 1}$ and h_k and if $x_k^+\Omega = 0$ $\forall k\in\mathbb{Z}$.

We say V is a *highest weight* module if it is generated by a highest weight vector.

Note, unlike category $\mathcal{O}$ we do not require the q-version of $\hat{\mathfrak{g}}_+$ (ie X_0^+) to annihilate Ω for it to be of highest weight.

Theorem. *Let V be a finite dimensional U-module. Then*

1. $C = C^{-1}$ *on* V
2. $H = \{h_k, K+^{\pm 1}, C^{\pm 1}\}$ *acts by commuting operators on* V

(3) *There exists a highest weight vector Ω in V.*

Proof. (3) We will prove this by contradiction. Suppose there does not exists $v \in V\backslash\{0\}$ annihilated by all x_k^+. Let $v \in V$ be an eigenvector of K, ie $Kv = \lambda K$. Then there exists a never zero sequence $n, x_k^+ v, x_k^+ x_k^+ v, \ldots$ in V. The then these have eigenvalues $\lambda, q^2\lambda, q^4\lambda, \ldots$ contradicting the assumption that dim $V < \infty$. Therefore, $V_0 = \{v \in V | x_k^+ v = 0 \forall k \in \mathbb{Z}\} \neq 0$ and is acted on by K_0 and K_1.

Consider $\Omega \in V_0$ a simultaneous eigenvector of K_0 and K_1. Then the action of $v_1 =< X_1^+, X_1^-, K_1 >$ on Ω is given by $K_1\Omega = \epsilon_1 q^{n_1}\Omega$ for $\epsilon \in \{\pm 1\}$ and $n_1 \geq 0$. Similarly the action of $v_0 =< X_1^+, X_1^-, K_1 >$ is given by $K_0\Omega = \epsilon_0 q^{n_0}\Omega$ where $\epsilon_0 \in \{\pm 1\}$ and $n_0 \leq 0$ since Ω is lowest weight for v_0.

But now, $C\Omega = K_0K_1\Omega = (\epsilon_0\epsilon_1)q^{n_0+n_1}\Omega$ and so $K_1C^i\Omega = \epsilon_1(\epsilon_0\epsilon_1)^i q^{i(n_0+n_1)+n_1}\Omega$. Since Ω is a highest weight vector for $(U_q\mathfrak{sl}_2)^i$ for all $i \in \mathbb{Z}$, $i(n_0+n_1)+n_1 \geq 0$ for all $i \in \mathbb{Z}$. But this is only possible if $n_0 + n_1 = 0$.

From the argument in the preceding paragraph, $C\Omega = \pm\Omega$ for any $\Omega \in V_0$, so $C = C^{-1}$ on V_0. Then, since H preserves V_0, H acts on V_0 by a commuting operator.

(1) Now, since C is central we can decompose V into $V = \bigoplus_\mu V^\mu$ where C acts by μ on V^μ and V^μ is a submodule. But each V^μ had a nonzero V_0^μ and C acts by ± 1 on it so $\mu = \mu^{-1}$ and $\mu = \pm 1$ only. Therefore $C = C^{-1}$ on V whether V is irreducible or not.

(2) Property (1) clearly implies property (2).

$\square$

Corollary. *Every finite dimensional representation of $U_q\hat{\mathfrak{sl}}_2$ is highest weight.*

3. The Drinfeld Polynomial

For a finite dimensional representation there exists $\Omega \in V$ such that $x_k^\pm\Omega = 0$ for all $k \in \mathbb{Z}$. Then

$$\Phi_r\Omega = d_r^+\Omega \tag{3.1}$$

$$\psi_r\Omega = d_r^-\Omega \tag{3.2}$$

for some pair $\underline{d} = \{d_k^+, d_k^-\}_{k\geq 0} \subset \mathbb{C}$. Conversely, we can construct an irreducible highest weight representation such that (3.1) and (3.2) hold: Let

$$V(\underline{d}) = U_q\hat{\mathfrak{sl}}_2/(N_+ + (\Phi_r - d_r^-) + (\psi_r - d_r^+))$$

where

$$d^+(u) = \sum_{r\geq 0} d_r^+ u^r \in \mathbb{C}[[u]]$$

$$d^-(u) = \sum_{r\geq 0} d_r^- u^{-r} \in \mathbb{C}[[u]]$$

Theorem. *(Drinfeld, Chao-Pressley) $V(\underline{d})$ is finite dimensional if and only if there exits a polynomial $P \in \mathbb{C}[u]$ with $P(0) \neq 0$ such that*

$$\psi(u)\Omega = q^{deg\ P}\frac{P(q^{-2}u)}{P(u)}\Omega = \Phi(u)\Omega$$

Where the equality is understood to be as formal power series, ie where $\frac{1}{1+uQ} = 1-uQ+u^2Q^2-\dots$. *Then* P *is called the Drinfeld Polynomial of the representation.*

We will need the following result:

Proposition. *There exists a sequence* $\{P_r\}_{r\geq 0} \subset H =< K, h_K > /(c-1)$ *such that*

(1) $P(u) = \sum_{r\geq 0} u^r P_r \in H[[u]]$ *where* $\psi(u) = K\frac{P(q^{-2}u)}{P(u)}$
(2) $P_r \equiv (-1)^r q^{r^2}(x_0^+)^{(r)}(x_1^-)^{(r)}$ *mod* N_+
(3) $(-1)^r q^{r(r-1)}(x_0^+)^{(r-1)}(x_1)^{(r)} \equiv -\sum_{s=1}^r x_s^- P_{r-s}K^{r-1}$ *mod* N_+ *where* $X^{(r)} = \frac{X^r}{[r]!}$

Proof. (Of Theorem) $\Rightarrow$

First, rephrasing slightly

$$\psi(u)\Omega = q^d \frac{P(q^{-2}u)\Omega}{P(u)}$$

evaluating at $u=0$ we get that $\psi_0 = K\Omega = q^d\Omega$ for some $d \in \mathbb{Z}$ since $V(\underline{d})$ is finite dimensional. So deg $P = d$.

Lets apply (1) to Ω: since $\psi(u)\Omega = d^+(u)\Omega$

$$P(u)\Omega = \sum_{r\geq 0} P_r\Omega = \sum_{r\geq 0}^{d} P_r\Omega = \sum_{r\geq 0}^{d} p_r\Omega$$

where p_r is the eigenvalue of P_r on Ω. Note that the second equality comes from the fact that by (2), $P_r\Omega = 0$ for $r > d$. Then, if $p(u) = \sum_{r\geq 0}^d p_r$,

$$d^+(u)\Omega = q^d \frac{p(q^{-2}u)}{p(u)}\Omega$$

We now need to show that the same $P(u)$ works for $d^-(u)$. To do this we apply (3) to Ω with $r = d+1$:

$$\sum_{s=0}^{d} x_{d+1-s}^- P_s^+ K^d = 0$$

Now, $K\Omega = q^d\Omega$ so $P_s K^d\Omega = (q^{d^2})P_{d+1-s}\Omega$. Apply x_{n-d-1}^+ to both side of the above and use

$$[x_{n-d-1}^+, x_{d+1-s}^-] = \frac{\psi_{n-s} - \Phi_{n-s}}{q-q^{-1}}$$

to get

$$\frac{1}{q-q^{-1}}\sum_{s=0}^{d}(\psi_{n-s} - \Phi_{n-s})P_s\Omega = \sum_{s=0}^{d}(d_{n-s}^+ - d_{n-s}^-)P_s = 0$$

If we take n large enough, the smallest $n-s$ can be is when $s = d$, so for $n \geq d+1$ since d^- has no positive coefficient

(1) for $n \geq d+1$: $\sum_{s=0}^d d_{n-s}^+ P_s^+ = 0$
(2) for $0 \leq n \leq d$: $\sum_{s=0}^n d_{n-s}^+ P_s^+ = \sum_{s=n}^d d_{n-s}^- P_s^+$
(3) for $n \leq -1$: $\sum_{s=0}^d d_{n-s}^- P_s^+ = 0$

Now, from $\psi(u)\Sigma = q^d \frac{P^+(q^{-2}u)}{P^+(u)}\Omega$ we get

$$\sum_{s=0}^{r} P_s^+ d_{r-s}^+ = q^d q^{-2r} p_r^+ \tag{3.3}$$

Since $p_n = 0$ for $n < 0$ we can rewrite (2) as

$$\sum_{s=n}^{d} d_{n-s}^- P_s^+ = 0$$

and so $\forall n \leq d$,

$$\sum_{s=n}^{d} d_{n-s}^- P_s^+ = q^{d-2n} P_n^+ \tag{3.4}$$

We claim that these two facts together imply that

$$\Phi(u)\Sigma = q^d \frac{P(q^{-2}u)}{P(u)} \tag{3.5}$$

when we consider the right hand side as expanded in powers of u^{-1}. To see this we note that

$$P(u) = \sum_{r=0}^{d} P_r u^r = u^d \sum_{r=0}^{d} P_r u = u^d \sum_{r=0}^{d} P_{d-r} u^{-r}$$

and

$$P(q^{-2}u) = q^{-2d} u^d \sum_{r=0}^{d} P_{d-r} u^{-r} q^{2r}$$

so the right hand side of (3.5) is

$$\frac{\sum_{r=0}^{d} q^{-d+2r} P_{d-r} u^{-r}}{\sum_{r=0}^{d} P_{d-r} u^{-r}}$$

and so, given that $\Phi(u)\Omega = \sum_{r\geq 0} u^{-r} d_{-r}^- \Omega$, (3.5) reads

$$\sum_{s=0} d_{-s}^- P_{d-(r-s)}^+ = q^{-d+2r} p_{d-r}^+ \tag{3.6}$$

For $n = d - r$, ($n \leq d \equiv r \geq 0$) this is

$$\sum_{s=0}^{d-n} d_{-s}^- P_{n+s}^+ = \sum_{s=0}^{d} d_{n-s}^- p_s^+ = q^{d-2n} p_n^+$$

which is exactly what we want.

Therefore $V(d)$ is finite dimensional only if there exists a polynomial $P \in \mathbb{C}[u]$ such that $P(0) = 1$ and

$$\sum_{k\geq 0} d_k u^k = q^{\deg P} \frac{P(q^{-2}u)}{P(u)} = \sum_{k\leq 0} d_k u^k$$

We digress briefly now to prove the proposition and will then finish the proof of the theorem.

$\square$

PROOF. (Of Proposition) (1): First,

$$\psi(u) = K\frac{P(q^{-2}u)}{P(u)} \equiv \sum_{s=0}^{r} \Psi_s P_{r-s} = KP_r q^{-2r}$$

for all $r \geq 0$. Now, the left hand side of (1) is

$$\psi_0 P_r + \sum_{s=1} \psi_s P_{r-s}$$

So $K(q^{-2r} - 1)P_r = \sum_{s=1}^{r} \psi_s P_{r-s}$ for $r \geq 1$ or

$$P_r = \frac{K^{-1}}{q^{-2r} - 1} \sum_{s=1}^{r} \psi_0 P_{r-s}$$

(2): This follows from (3) and the recursion above. Start from (3) and multiply on the left by x_0^+ to get

$$[r](-1)^r q^{r(r-1)} (x_0^+)^{(r)} (x_0^-)^{(r)} = -\sum_{s=1}^{r} \psi_s (q - q^{-1})^{-1} P_{r-s} K^{r-1} = K^r q^{-r} [r] P_r$$

as claimed.

(3): Recall that we cant multiply on the right since we're working modulo a right ideal so to prove by induction we need away to write $x_1^-(x_0^+)^{(r)}(x_1)^{(r)}$ in terms of $(x_0^+)^{(r)}(x_1)^{(r+1)}$ so we need to perform the following computations:

(QGR 1)

$$\begin{aligned}
[x^-, (x_0^+)^{r+1}] &= \frac{1}{[r+1]!} \sum_{j=0}^{r} (x_0^+)[x_1^-, x_0^+](x_0^+)^{r-j} \\
&= \frac{1}{[r+1]!} \sum_{j=0}^{r} (x_0^+)^j K h_1 (x_0^+)^{r-j} \\
&= \frac{-K}{[r+1]!} \sum_{j=0}^{r} q^{-2j} (x_0^+)^j h_1 (x_0^+)^{r-j}
\end{aligned}$$

(QGR 2)

$$\begin{aligned}
[h_1, (x_0^+)^r] &= \sum_{s=0}^{r-1} (x_0^+)^j [h_1, x_0^+](x_0^+)^{r-1-j} \\
&= [2] \sum_{j=0}^{r-1} (x_0^+)^j X_1^+ (X_0^+)^{r-1-j} \\
&= [2] X_1^+ (X_0^+)^{r-1} \sum_{j=0}^{r-1} q^{-2j} \\
&= [2][r] x_1^+ (x_0)^{r-1} q^{-r+1}
\end{aligned}$$

Where the second equality come from $[h_1, x_k^+] = [2]x_{k+1}$ and the third from

$$\sum_{j=0} r-1 q^{-2j} = \frac{q^{2r}-1}{q^{-2}-1} = q^{-r-1}[r]$$

and

$$x_1^+x_0^+ - q^2x_0^+x_1^+ = q^2x_0^+x_1^+ - x_1^+x_0^+$$

(QGR 1) *Continued.* Now:

$$\begin{aligned}
[x_1^-, (x_0^+)^{(r+1)}] &= -\frac{K}{[r+1]!}(x_0^+)^r h_1 q^{-r}[r+1] \\
&= -\frac{K}{[r+1]!}\sum_{j=0}^{r} q^{-2j}(x_0^+)^j[2][r-j]x_1^+(x_0^+)^{r-j-1}q^{-r+j+1}[r+j] \\
&= -q^{-r}K(x_0^+)^r h_1 - q^{-2(r-1)}Kx_1^+(x_0^+)^{(r-1)}
\end{aligned}$$

(QGR 3) By a similar proof we have

$$[x_0^+, (x_1^{-(r+1)}] = q^rKx_1^{-(r)}h_1 - Kx_1^{-(r-1)}x_2^-$$

Now, $(-1)^r q^{r(r-1)}(x_0^+)^{(r-1)}(x_1^-)^{(r)} = -\sum_{s=1}^r x_s^- P_{r-x}K^{r-1} \mod N_+$ so by multiplying both sides by x_1^- on the left we get

$$\begin{aligned}
&[r+1](-1)^r q^{r(r-1)}(x_0^+)^{(r-1)}(x_1^-)^{(r+1)} \\
-(-1)^r q^{r(r-1)}\left(q^{-r}K(x_0^+)^{(r)}h_1 + q^{-2r+2}Kx_1^+(x_0^+)^{(r-1)}\right)(x_1^-)^{(r)} &= -\sum_{s=1}^{r} x_1^- x_s^- P_{r-s}K^{r-1}
\end{aligned}$$

Multiplying by x_0^+ on the left we get

$$\begin{aligned}
&[r][r+1](-1)^r q^{r(r-1)}(x_0^+)^{(r)}(x_1^-)^{(r+1)} = \\
(-1)^r q^{r(r-1)}x_0^+\left(q^{-r}K(x_0^+)^{(r)}h_1 + q^{-2r+2}Kx_1^+(x_0^+)^{(r-1)} - \sum\right)(x_1^-)^{(r)} &- \sum_{s=1}^{r} x_0^+x_1^-x_s^-P_{r-s}K^{r-1}
\end{aligned}$$

□

PROOF. (Of Theorem) $\Leftarrow$

We want to prove that if there exists a polynomial P such that

$$d_+(u) = q^{\deg P}\frac{P(q^{-2}u)}{P(u)} = d_-(u)$$

then $V(d)$ is finite dimensional. We will do this by constructing $V(d)$ explicitly.

Our main tool will be a map $\mathrm{ev}_a : U_q\hat{\mathfrak{sl}}_2 \to U_q\mathfrak{sl}_2$. Classically, $\hat{\mathfrak{sl}}_2$ is just a loop algebra so for any $a \in \mathbb{C}^\times$ we just evauate at a. Assuming we ca find such a map, we can then pull representations of $U_q\mathfrak{sl}_2$ back to $U_q\hat{\mathfrak{sl}}_2$. Since we have a Hopf algebra structure on $U_q\mathfrak{sl}_2$ we can then take tensors of $U_q\mathfrak{sl}_2$ modules.

Proposition. *For all $a \in \mathbb{C}^\times$ there exists an algebra homomorphism $ev_a : U_q\hat{\mathfrak{sl}}_2 \to U_q\mathfrak{sl}_2$ such that $ev_a(x_k^+) = q^{-ka^kK^kE}$ and $ev_a(x_k^-) = q^{-ka^kK^kE}$.*

Proof. We will construct this map using the quantum KM generators:

$$
\begin{aligned}
\mathrm{ev}_a(K_1) &= K & \mathrm{ev}_a(K_0) &= K^{-1} \\
\mathrm{ev}_a(X_1^+) &= E & \mathrm{ev}_a(X_0^+) &= q^{-1}aF \\
\mathrm{ev}_a(X_1^-) &= E & \mathrm{ev}_a(X_0^-) &= q^{-1}a^{-1}E
\end{aligned}
\tag{3.7}
$$

so that the composition

$$U_q\mathfrak{sl}_2 \hookrightarrow U_q\hat{\mathfrak{sl}}_2 \xrightarrow{\mathrm{ev}_a} U_q\mathfrak{sl}_2$$

of evaluation with the inclusion map of the constant loops is the identity:

$$K,\ E,\ F \mapsto K_1,\ X_1^\pm \mapsto K,\ E,\ F$$

We claim that this extends to a homomorphism. We need to check that

$$\mathrm{ev}_a : (X_i^\pm)^3X_{\bar{i}}^\pm - [3](X_i^\pm)^2X_{\bar{i}}^\pm X_i^\pm + [3]X_1^\pm X_{\bar{i}}(X_i^\pm)^2 - X_{\bar{i}}(X_i^\pm)^3 \mapsto 0$$

□

Let check for the case $i = 1$, $\bar{i} = 0$, $+$:

$$\ldots \mapsto q^{-1}a\left(F^3E - [3]F^2EF + [3]FEF^2 - EF^3\right) = 0$$

we can rewrite this as

$$
\begin{aligned}
[E, F^3] - [3]F[F, E]F &= 0 \\
[3]F^2[K; -2] - [3]\frac{FK - K^{-1}F}{q - q^{-1}} &= 0 \\
[3]F^2[K; -2] - [3]F^2[K; -2] &= 0
\end{aligned}
$$

So the relation check out. In addition the following relations are clear: ev_a :

$$
\begin{aligned}
C = K_0K_1 &\longrightarrow 1 \\
K = K_1 &\longrightarrow K \\
X_1^\pm &\longrightarrow X_0^\pm \\
X_0^+ &\longrightarrow x_1^-K^{-1} \\
X_0^- &\longrightarrow Kx_{-1}^+ \\
x_1^- = X_0^+K &\longrightarrow q^{-1}aFK \\
x_{-1}^+ = K^{-1}X_0^- &\longrightarrow qa^{-1}K^{-1}E
\end{aligned}
\tag{3.8}
$$

We can use then derive the rest of the map from the above:

$$\mathrm{ev}_a(\psi_1) = (q - q^{-1})[x_0^+, x_1^+] = (q - q^{-1})h_1K$$

and

$$\begin{aligned}\mathrm{ev}_a(h_1) =&\mathrm{ev}_a([x_0^+, x_1^-]K^{-1})\\ =&q^{-1}a([E, FK]K^{-1})\\ =&q^{-1}a\frac{K-K^{-1}}{q-q^{-1}} - q^{-1}aF[E,K]K^{-1}\\ =&q^{-1}a\frac{K-K^{-1}}{q-q^{-1}} + a(q-q^{-1})FE\end{aligned}$$

□

Next,

$$[H_1, x_k^{\pm}] = \pm[2]x_{k+1}^{\pm}$$

so we can compute $\mathrm{ev}_a(x_k^{\pm})$ by induction on $k \geq 0$. For the other relations we can use $\Phi_1 = -[x_0^+, x_{-1}^-](q-q^{-1})$ to get $\mathrm{ev}_a(\Phi_1)$, $\Phi_1 = -K^{-1}h_{-1}(q-q^{-1})$ to get $\mathrm{ev}_a(h_{-1})$ and $[h_{-1}, x_k^{+}] = \pm[2]x_{k-1}^{\pm}$ to get $\mathrm{ev}_a(x_k^{\pm})$ for $k < 0$.

Recall. If V_n is the unique irreducible Type I representation of $U_q\mathfrak{sl}_2$ of dim $n+1$ then it has a basis $v_0, \dots v_n$ and $U_q\mathfrak{sl}_2$ acts on this basis by

$$\begin{aligned}Kv_i =&q^{n-2i}v_i\\ Ev_i =&[n-i+1]v_{i-1}\\ Fv_i =&[i+1]v_{i+1}\end{aligned}$$

Definition. Let $V_n(a) := \mathrm{ev}_a^* V_n$ an irrational representation of $U_q\hat{\mathfrak{sl}}_2$ of type (I, I). In particular C acts by $+1$ since $\mathrm{ev}_a C = 1$.

Recall. There are two important things to note here: First ev_a is *not* a Hopf Algebra homomorphism. Second, Although we have shown this evaluation map to exits for $U_q\hat{\mathfrak{sl}}_2$ and indeed it can be shown to exist for $U_q\hat{\mathfrak{sl}}_n$ for a general lie algebra $\mathfrak{g}$ there is no guarantee it exists.

Corollary. *For $n \in \mathbb{N}$ and $a \in \mathbb{C}^*$ the action of the loop generators $x_k^{\pm}$ on $V_n(a)$ is given by*

$$\begin{aligned}ev_a(x_k^+)v_i =&q^{-k}a^kK^kEv_i\\ =&q^{-k}a^k[n-i+1]q^{k(n-2i+2)}v_{i-1}\\ =&a_kq^{k(n-2i+1)}[n-i+1]v_{i-1}\end{aligned}$$

$$\begin{aligned}ev_a(x_k^-)v_i =&q^{-k}a^kFK^kv_i\\ =&a_kq^{k(n-2i-1)}[i+1]v_{i+1}\end{aligned}$$

so $V_n(a)$ is a highest weight representation with highest weight vector v_0.

The Drinfeld Polynomial of $V_n(a)$. For $k \geq 1$

$$(q - q^{-1})\psi_k v_0 = [x_k^+, x_0]v_0 = x_k^+ x_0^- v_0 = x_k^+ v_1 = a^k q^{k(n-1)}[n]v_0 = (aq^{n-1})^k [n]v_0$$

And $\psi_0 v_0 = K v_0 = q^n v_0$. So $(aq^{n-1})^k[n]$ and q^n are the eigenvalues of ψ_k so

$$\begin{aligned} d_+(u) =& q^n + \sum_{k\geq 1}(aq^{(n-1)})^k (q^n - q^{-n})u^k \\ =& q^n + (q^n - q^{-n})\frac{aq^{n-1}u}{1 - aq^{n-1}u} \\ =& q^n \frac{1 - aq^{-n-1}u}{1 - aq^{n-1}u} \\ =& q^n \frac{P(q^{-2}u)}{P(u)} \end{aligned}$$

Where

$$\begin{aligned} P(u) &= (1 - aq^{n-1}u)(1 - aq^{n-3}u)(1 - aq^{n-5})\dots \\ P(q^{-2}u) &= (1 - aq^{n-3}u)(1 - aq^{n-4}u)(1 - aq^{n-7})\dots \end{aligned}$$

Example. For $n = 1$, $V_1 \cong \mathbb{C}^2$ we have $P_{V_1(a)} = (1 - au)$.

So now we're going to try to start from a polynomial of the form $(1-aq^m u)\dots(1-aq^{-m}u)$ and show that it's the Drinfeld polynomial of some representation. Take $P \in \mathbb{C}[u]$ such that $P(0) = 1$. Then

$$P = (1 - a_1 u)\dots(1 - a_m u)$$

and we can construct $V(d)$ from $V_1(a_1) \otimes \dots \otimes V_1(a_m)$ as follows:

Theorem. *$V(d)$ is a subquotient of $V_1(a_1) \otimes \dots \otimes V_1(a_m)$ where $V_1(a_i) \cong \mathbb{C}^2$.*

In order to prove this we need to understand Δ on $\mathcal{U}_q(\mathrm{Lsl}_2)$. Recall that by (1.12)

$$\begin{aligned} \Delta(K_i) &= K_i \otimes K_i \\ \Delta(X_i^+) &= X_i^+ \otimes K_i + 1 \otimes K_i^+ \\ \Delta(X_i^-) &= X_i^- \otimes 1 + \otimes K_i^{-1} \otimes X_i^- \end{aligned}$$

But what we're really interested in of course is $\Delta(\psi_i)$ and $\Delta(\Phi_i)$, however these are not known in full in loop generators.

Proposition. *Let $\Xi_\pm \subset \mathcal{U}_q(Lsl_2)$ be the subspace generated by the $x_k^\pm$, $k \in \mathbb{Z}$. Then*

(QDS 1) *Modulo $\mathcal{U}\Xi_+^2 \otimes \mathcal{U}\Xi_-$*

$$\Delta(x_k^+) = x_k^+ \otimes K + 1 \otimes x_k^+ + \sum_{i=1}^{k} x_{k-i}^+ \otimes \psi_i, \qquad k \geq 0 \tag{3.9}$$

$$\Delta(x_{-k}^+) = x_{-k}^+ \otimes K^{-1} + 1 \otimes x_{-k}^+ + \sum_{i=1}^{k-1} x_{-k-i}^+ \otimes \phi_i, \qquad k > 0 \tag{3.10}$$

(QDS 2) *Modulo* $\mathcal{U}\Xi_+ \otimes \mathcal{U}\Xi_-^2$

$$\Delta(x_k^-) = x_k^- \otimes 1 + K \otimes x_k^- + \sum_{i=1}^{k-1} \psi_i \otimes x_{k-i}^-, \qquad k > 0 \tag{3.11}$$

$$\Delta(x_{-k}^-) = x_{-k}^- \otimes 1 + K^{-1} \otimes x_{-k}^- + \sum_{i=1}^{k} \phi_{-i} \otimes x_{-k-i}^+, \qquad k \geq 0 \tag{3.12}$$

(QDS 3) *Modulo* $\mathcal{U}\Xi_+ \otimes \mathcal{U}\Xi_- + \mathcal{U}\Xi_- \otimes \mathcal{U}\Xi_+$

$$\Delta(\psi_k) = \sum_{i=0}^{k} \psi_i \otimes \psi_{k-i} \qquad (\equiv \Delta(\psi(u)) = \psi(u) \otimes \psi(u)) \tag{3.13}$$

$$\Delta(\Phi_k) = \sum_{i=0}^{k} \Phi_i \otimes \Phi_{-k+i} \qquad (\equiv \Delta(\Phi(u)) = \Phi(u) \otimes \Phi(u)) \tag{3.14}$$

Or, more compactly, for $X_{\geq 0}^{\pm} := \sum_{k\geq 0} x_k^{\pm} u^k$ *and* $X_{<0}^{\pm} := \sum_{k<0} x_k^{\pm} u^k$

(QDS 1) *Modulo* $\mathcal{U}\Xi_+^2 \otimes \mathcal{U}\Xi_-$

$$\Delta X_{\geq 0}^+ = X_{\geq 0}^+ \otimes \psi + 1 \otimes X_{\geq 0}^+ \tag{3.15}$$

$$\Delta X_{<0}^+ = X_{<0}^+ \otimes \Phi + 1 \otimes X_{<0}^+ \tag{3.16}$$

(QDS 2) *Modulo* $\mathcal{U}\Xi_+ \otimes \mathcal{U}\Xi_-^2$

$$\Delta X_{\geq 0}^- = X_{\geq 0}^- \otimes 1 + \psi \otimes X_{\geq 0}^- \tag{3.17}$$

$$\Delta X_{<0}^- = X_{<0}^- \otimes 1 + \Phi \otimes X_{<0}^- \tag{3.18}$$

(QDS 3) *Modulo* $\mathcal{U}\Xi_+ \otimes \mathcal{U}\Xi_- + \mathcal{U}\Xi_- \otimes \mathcal{U}\Xi_+$

$$\Delta\psi = \psi \otimes \psi \tag{3.19}$$

$$\Delta\Phi = \Phi \otimes \Phi \tag{3.20}$$

PROOF. We will use the follwoing scheme for this proof:

(QDP 1) Prove the above for $x_k^{\pm}$ for $k = 0, 1$
(QDP 2) Use $\psi_1 = \{x_1^+, x_0^-\}$ to prove the above for ψ_1. Use $h_1 = (q - q^{-1})K^{-1}\psi_1$ to prove the above for h_1.
(QDP 3) Use $[h_1, x_k^{\pm}] = \pm[2]x_{k+1}^{\pm}$ to prove the above for $x_k^{\pm}$
(QDP 4) Compute $\Delta(\Phi_1)$ then $\Delta(h_{-1})$ then use $[h_1, x_k^{\pm}] = \pm[2]x_{k+1}^{\pm}$.
(QDP 5) Finally, use $\psi_k = \{x_k^+, x_0^-\}$.

3.1. (QDP 1). By Drinefelds Realization (3.8) and the Coproduct Structure (1.12) we have:

$$\begin{aligned}\Delta x_0^+ =& x_0^+ \otimes K + 1 \otimes x_0^+\\ \Delta x_0^- =& x_0^- \otimes 1 + K^{-1} \otimes x_0^-\end{aligned}$$

and

$$\begin{aligned}\Delta x_1^- &= \Delta(X_0^+)\Delta(K)\\ &= (X_0^+ \otimes K_0 + 1 \otimes X_0^+)K_1 \otimes K_1\\ &= (x_1^- K_1^{-1} \otimes K_0 + 1 \otimes x_1^- K_1^{-1})K_1 \otimes K_1\\ &= x_1^- \otimes K_0 K_1 + K_1 \otimes x_1^-\\ &= x_1^- \otimes C + K \otimes x_1^-\end{aligned}$$

Similarly,

$$\begin{aligned}\Delta x_{-1}^+ &= CK^{-1} \otimes CK^{-1}(X_0^- \otimes 1 + K_0^{-1} \otimes X_0^-)\\ &= CK^{-1} \otimes CK^{-1}(C^{-1}Kx_{-1}^+ \otimes 1 + K_0^{-1} \otimes C^{-1}Kx_{-1}^+)\\ &= x_{-1}^+ \otimes CK^{-1} + 1 \otimes x_{-1}^+\end{aligned}$$

3.2. (QDP 2). Now, since $\psi_1 = C(q-q^{-1})[x_0^+, x_1^-]$ we have

$$\begin{aligned}\Delta\psi_1 =&(q-q^{-1})C \otimes C[x_0^+ \otimes K + 1 \otimes x_0^+, x_1^- \otimes C + K \otimes x_1^-]\\ =&\psi_1 \otimes KC^2 + KC \otimes \psi_1 + (q-q^{-1})C \otimes C[x_0^+ \otimes K, K \otimes x_1^-]\\ =&\psi_1 \otimes KC^2 + KC \otimes \psi_1 + (q-q^{-1})^2 KC \otimes KC[2]x_0^+ \otimes x_1^-\end{aligned}$$

where the last equality follows from

$$x_0^+ K \otimes K x_1^- - K x_0^+ \otimes x_1^- K = K \otimes K(q^{-2}x_0^+ \otimes x_1^- - q^2 x_0^+ \otimes x_1^-)$$

Now: $\psi_1 = (q-q^{-1})Kh_1 \equiv h_1 = (q-q^{-1})K^{-1}\psi_1$ implies that

$$\begin{aligned}\Delta h_1 =&(q-q^{-1})^{-1}K^{-1} \otimes K^{-1}\Delta\psi_1\\ =&h_1 \otimes C^2 + C \otimes h_1 - (q-q^{-1})[2]C \otimes C \cdot x_0^+ \otimes x_1^-\end{aligned}$$

3.3. (QDP 3). Since $[h_1, x_0^+] = [2]x_1^+$ we have

$$\begin{aligned}
\Delta_{x_1^+} &= [2]^{-1}\left[h_1 \otimes C^2 + C \otimes h_1 - (q-q^{-1})[2]C \otimes C \cdot x_0^+ \otimes x_1^-, x_0^+ \otimes K + 1 \otimes x_0^+\right] \\
&= x_1^+ \otimes C^2K + C \otimes x_1^+ - (q-q^{-1})C \otimes C[x_0^+ \otimes x_1^-, x_0^+ \otimes K + 1 \otimes x_0^+] \\
&= x_1^+ \otimes C^2K + C \otimes x_1^+ - (q-q^{-1})C \otimes C\left\{(q^2-1)(x_0^+)^2 \otimes Kx_1^- - (q-q^{-1})x_0^+ \otimes C^{-1}\psi_1\right\} \\
&= x_1^+ \otimes C^2K + C \otimes x_1^+ - (q-q^{-1})(q^2-1)C(x_0^+)^2 \otimes CKx_1^- + Cx_0^+ \otimes \psi_1 \\
&= x_1^+ \otimes C^2K + Cx_0^+ \otimes \psi_1 + C \otimes x_1^+ - (q-q^{-1})(q^2-1)C(x_0^+)^2 \otimes CKx_1^-
\end{aligned}$$

By induction this yields the coproduct structure $\Delta(x_k^+)$ for all $K \geq 2$ (although in practice this might be tedious to compute). (QDP 4) and (QDP 5) follow from the induction above.

□

4. Tensor Products of Irreducible $U_q\hat{\mathfrak{sl}}_2$-modules

Recall the definition of ev_a the evaluation map (3.7) and S the antipode map (1.13):

$$\begin{aligned}
\mathrm{ev}_a(K_1) &= K & \qquad \mathrm{ev}_a(K_0) &= K^{-1} \\
\mathrm{ev}_a(X_1^+) &= E & \qquad \mathrm{ev}_a(X_0^+) &= q^{-1}aF \\
\mathrm{ev}_a(X_1^-) &= E & \qquad \mathrm{ev}_a(X_0^-) &= q^{-1}a^{-1}E
\end{aligned}$$

$$\begin{aligned}
S(K_i) &= K_i^{-1} \\
S(X_i^+) &= -X_i^+K_i^{-1} \\
S(X_i^-) &= K_i \otimes X_i^-
\end{aligned}$$

Proposition. *$ev_a \circ S = S \circ ev_{q^2a}$ where the S on the right hand side is that of $U_q\mathfrak{sl}_2$.*

Proof. We will check that these agree on the KM generators K_i, $X_i^{\pm}$. This is clear for $i = 1$ since ev_a on these is just "the identity," and for K_0. For $X_0^{\pm}$ we have the following:

$$\mathrm{ev}_a \circ S(X_0^+) = \mathrm{ev}_a(-X_0^+K_0^{-1}) = -q^{-1}aFK = -qaKF$$

while

$$S \circ \mathrm{ev}_b(X_0^+) = S(-q^{-1}bF) = -q^{-1}bKF$$

so these match provided $b = q^2a$. A similar proof holds for X_0^-.

□

Corollary. $V_n(a)^* \cong V_n(q^2a)$

PROOF. The action of $X \in U_q\hat{\mathfrak{sl}}_2$ on the left hand side is given by

$$Xg = g \circ \mathrm{ev}_a(S(X)) = g \circ S(\mathrm{ev}_{q^2a}(X))$$

But as a $U_q\mathfrak{sl}_2$-mod, $V_n^* \cong V_n$ so the result follows.

□

We want to determine when the tensor product of two irreducible representations of is again irreducible.

Definition. Let $S \subset \mathbb{C}^\times$ be a finite subset. We call S a q-string if it is of the form $\{\zeta, q^{-2}\zeta, q^{-4}\zeta, \ldots, q^{-2r}\zeta\}$. eg

$$P(u)_{V_n(s)} = (1 - aq^{n-1}u)\ldots(1 - aq^{-n-1}u)$$

has roots $S = \{a^{-1}q^{n+1}, \ldots, a^{-1}q^{-n+1}\}$ which are a q-string of length $n+1$.

Definition. Two q-strings S_1 and S_2 are in general position if either

(1) $S_1 \cup S_2$ is not a q-string
(2) $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$

Lemma. *Any finite subset of $\mathbb{C}^\times$ with multiplicities can be written uniquely as a union of q-strings in general position.*

Theorem. *A representation $V_{n_1}(a_1) \otimes \ldots \otimes V_{n_r}(a_r)$ is irreducible if and only if the strings $s_{n_1}(a_i)$ are in general positions.*

Example. The representation $V_1(a_1) \otimes \ldots \otimes V_r(a_r) = \mathbb{C}^2(a_1) \otimes \ldots \otimes \mathbb{C}^2(a_r)$ is irreducible if and only if for all $i \neq j$ either $a_i = a_j$ or $a_i \neq q^{\pm 2}a_j$. In particular $V_1(a)^{\otimes n}$ is irreducible. Note that this implies that ev_a is not a Hopf algebra morphism since $V_1(a)$ is reducible as a $U_q\mathfrak{sl}_2$-module.

PROOF. (Of Theorem) First, consider the case $r = 2$, ie for $V_m(a) \otimes V_n(b)$. We may assume that $m \geq n$ since $V \otimes W$ is irreducible iff $W \otimes V$ is.

Recall the Clebsch-Gordan rules:

$$V_m \otimes V_n \cong V_{m+n} \oplus V_{m+n-2} \oplus \ldots \oplus V_{m-n}$$

as a $U_q\mathfrak{sl}_2$-module. In fact, the highest weight vector Ω_p corresponding to $V_{m+n-2p} \subseteq V_m \otimes V_n$ is given by

$$\Omega_p = \sum_{i=0}^{p} (-1)^i q^{i(n-i+1)}[m-p+i]![n-i]! v_{p-i}^{(n)} \otimes v_i^{(m)}$$

Take $p \geq 0$. We will compute the action of x_-^+ on Ω_p. First,

$$\Delta(x_{-1}^+) = x_{-1}^+ \otimes K^{-1} + 1 \otimes x_{-1}^+$$

and

$$\mathrm{ev}_a(x_k^+)v_i = a^k q^{k(n-2i+1)}[n-i+1]v_{i-1}$$

so

$$x^+_{-1}\Omega_p = \sum_{i=0}^{p}(-1)^i q^{i(n-i+1)}[m-p+i]![n-i]!$$
$$\times\Big(q^{-(n-2i)}a^{-1}q^{-(m-2p+2i+1)}[m-p+i+1]v^{(m)}_{p-i-1}\otimes v^{(n)}_i\delta_{i<p}$$
$$+ a^{-1}q^{-(n-2i+1)}[n-i+1]v^{(m)}_{p-i}\otimes v_{i+1}\delta_{i>0}\Big)$$

$$= a^{-1}\left\{\sum_{i=0}^{p-1}(-1)^i q^{i(n-i)+i-n-m-2p-1}[m-p+i+1]![n-i]!v^{(m)}_{p-i-1}\otimes v^{(n)}_i\right\}$$
$$-b^{-1}\left\{\sum_{i=0}^{p-1}(-1)^i q^{i(n-i)+i+1}[m-p+i+1]![n-i]!v^{(m)}_{p-i-1}\otimes v^{(n)}_i\right\}$$

$$= \sum_{i=0}^{p-1}(-1)^i[m-p+i+1]![n-i]!q^{i(n-i)+i+1}v^{(m)}_{p-i}\otimes v^{(n)}_i(a^{-1}q^{-n-m+2p-2}-b^{-1})$$

So $x^+_{-1}\Omega_p = 0$ if and only if $b/a = q^{n+m-2p+2}$.

Now, this then implies that

$$0 = [2]x^+_{-1}\Omega_p = [h_{-1}, x^+_0]\Omega_p = x^+_0 h_{-1}\Omega_p$$

so $x^+_0 h_{-1}\Omega_p = 0$. But $h_{-1}\Omega_p$ has the same weight as Ω_p so $h_{-1}\Omega_p = \lambda\Omega_p$ and

$$[h_{-1}, x^-_{-k+1}][2]^{-1}\Omega_p = x^+_{-k}\Omega_p = 0$$

for all $k \geq 0$. Similarly, $x^+_k\Omega_p = 0$ For all $k \geq 0$.

In summery, if $\frac{b}{a} = q^{n+m-2p+2}$ then $V_m(a)\otimes V_n(b)$ contains a subrepresentation not containing it's highest weight component (ie the one generated by Ω_0). But similarly, any such subrepresentation is also a subrepresentation for $U_q\mathfrak{sl}_2 \subset U_q\hat{\mathfrak{sl}}_2$ and so a direct sum of V_{m+n-2p}'s for some $p's \geq 1$.

Next we will show that $V_m(a)\otimes V_n(b)$ has a proper subrepresentation containing Ω_0 iff $(V_m(a)\otimes V_n(b))^* \cong V_n(q^2b)\otimes V_m(q^2a)$. This in turn holds iff $\frac{b}{a} = q^{-n+m-2p'+2}$ where $p' = 1\ldots n$. Thus $V_m(a)\otimes V_n(b)$ is an irreducible representation iff $\frac{b}{a} \notin \{q^{\pm(n+m-2p+2)}\}^n_{p=1}$.

Exercise. This is equivalent to $S_m(a)$ and $S_n(b)$ begin in general position.

General Case. First, assume that $\bigotimes V_{n_i}(a_i)$ is irreducible but $S_{n_i}(a_i)$ and $S_{n_j}(a_j)$ are not in general position. Then $\bigotimes V_{n_i}(a_i)$ is homomorphic to $\bigotimes V_{n_{\sigma(1)}}(a_{\sigma(1)})$ for all $\sigma \in S_r$ so we may assume that $i = 1$ and $j = 2$, But then by the previous analysis $V_{n_1}(a_1) \otimes V_{n_2}(a_2)$ are reducible and so $\bigotimes V_{n_i}(a_i)$ is as well.

It remains to show that $\bigotimes V_{n_i}(a_i)$ is reducible only if the $S_{n_i}(a_i)$ are in general position. We will need the following:

Lemma. *Suppose the $S_{n_i}(a_i)$ are in general position and that $n_1 \leq \ldots \leq n_r$. Then $V_{n_1}(a_1) \otimes \ldots \otimes V_{n_r}(a_r)$ is generated by the tensor product of the highest weight vectors $v_0^{(n_1)} \otimes \ldots \otimes v_0^{(n_r)}$.*

Assuming the lemma holds, then starting from $V_{n_1}(a_1) \otimes \ldots \otimes V_{n_r}(a_r)$ with $S_{n_i}(a_i)$ in general position. By our analysis of $r = 2$ and the fact that S_n is generated by transposes $(i \quad i+1)$ the above is homomorphic to $V_{n_{\sigma(1)}}(a_{\sigma(1)}) \otimes \ldots \otimes V_{n_{\sigma(r)}}(a_{\sigma(r)})$ for all $\sigma \in S_n$ so we may assume as above that $n_1 \leq \ldots \leq n_r$.

Now, by the lemma there exists no proper subrepresentation containing the highest weight component V, ie the $U_q\hat{\mathfrak{sl}}_2$-mod generated by $v_0^{(n_1)} \otimes \ldots \otimes v_0^{(n_r)}$ since if $W \subset \bigotimes V_{n_i}(a_i)$ is a subrepresentation containing V then $W^\perp \subset (\bigotimes V_{n_i}(a_i))^* \cong \bigotimes V_{n_i}(q^2 a_i)$ contains the highest weight component. Then then we have a contradiction since the $S_{n_i}(q^2 a_i)$ are in general position.

$\square$

Proof. (Of Lemma) We will prove the lemma by induction on r. First, for $r = 2$ we are done by the analysis in the proof of the theorem.

Let $V_n(\underline{a}) := \bigotimes V_{n_i}(a_i)$. We claim that $V_n(\underline{a})$ is generated by

$$\Omega' \otimes v_{n_r}^{(n_r)} = v_0^{(n_1)} \otimes \ldots \otimes v_0^{n_{r-1}} \otimes v_{n_r}^{(n_r)}$$

by applying the lowering operators x_k^- to $\Omega' \otimes v_{n_r}^{(n_r)}$ to get, by the induction hypothesis, $\bigotimes_{i<r} V_{n_i}(a_i) \otimes v_{n_r}^{(n_r)}$. Then we can get the rest of V by applying x_0^+.

Now, we claim that $\Omega' \otimes v_i^{(n_r)} \in V_N$ where V_N is the highest weight component of $V_n(\underline{a})$. First, this is true for $i = 0$ by definition so if we can use induction on i to prove it for $i = n_r$ this will prove that $V_N = V$.

Assume that $\Omega' \otimes v_i \in V_N := \mathcal{U}\Omega$. Let $k > 0$. Then $x_k^-(\Omega' \otimes v_i) \in V_n$ and

$$\begin{aligned} x_k^-(\Omega' \otimes v_i) &= x_k^-\Omega' \otimes v_i + \sum_{j=0}^{k-1} \psi_j \Omega' \otimes x_{k-j}^- v_i \\ &= x_k^-\Omega' \otimes v_i + \sum_{j=0}^{k-1} d_{j,r-1} a_r^{k-j} \Omega' \otimes F v_i q^{(k-j)(n_r-2j)} \\ &= x_k^-\Omega' \otimes v_i + \sum_{j=0}^{k-1} d_{j,r-1} b_r^{k-j} \Omega' \otimes F v_i \end{aligned}$$

where $\psi_j\Omega' = d_{j,r-1}\Omega'$ and $b_r = a_r q^{n_r-2i+1}$. If we let $A_{k,r-1} := \sum_{j=0}^{k-1} d_{j,r-1}b_r^{k-j}$ we get the far more readable formula

$$x_k^-(\Omega'\otimes v_i) = x_k^-\Omega'\otimes v_i + A_{k,r-1}\Omega'\otimes Fv_i$$

More generally we get

$$x_k^-\Omega'\otimes v_i = \sum_{j=0}^{r-1} A_{k,j}v_0\otimes\ldots\otimes Fv_0\otimes\ldots\otimes v_i$$

where F is in the $j+1$ position. The matrix $A_{k,j}$ is given by $A_{k,j} = \sum_{p=0}^{k-1} d_{p,j}b_{j+1}^{k-p}$ for

$$\psi_p v_0^{(n_1)}\otimes\ldots\otimes v_0^{(n_j)} = d_{p,j}v_0^{(n_1)}\otimes v_0^{(n_j)}$$

and $b_j = q_j q^{n_j-1}$ when $j\le r-1$ and $b_j = a_j q^{n_j-2i+1}$ when $j=r$.

Now, if $(A_{k,j})_{j=0\ldots r-1,k=1\ldots r}$ is invertible then $\Omega'\otimes v_i$ is a linear combination of $(x_k^-\Omega'\otimes v_i$ where $k=1\ldots r$ so we're done. That A is invertible follows from the fact that

$$\det A = q^{\sum_{j=1}^{r-1} n_j}\prod_j b_j\prod_{j<k}(b_k - q^{2n_j}b_j)$$

so $\det A = 0$ implies (since $b_j\neq 0$) that $b_k = q^{2n_j}b_j$ for some $k<j$ and so, if $k<r$, $a_k = a_j q^{n_j+ni}$ which violates our assumptions.

□

Corollary. *Any finite dimensional representation of $U_q\hat{\mathfrak{sl}}_2$ of type (I,I) is a tensor product of irreducible representations. Two such products are isomorphic if and only if they are obtained by the same tensor factors.*

Proof. If V is finite dimensional and irreducible then $P = P_V = (1-a_1u)\ldots(1-a_mu)$. Write $S = \{a_i^{-1}\}$ as a union of q-strings in general position $S = \bigcup_{j=1}^r S_{n_j}(a_j)$. Then $\bigotimes_{j=1}^r V_{n_j}(a_j)$ is irreducible with Drinfeld Polynomial P and so is V. □

CHAPTER 4

Introduction to Statistical Mechanics

References

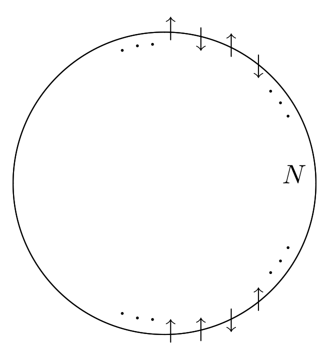
- Reif. "Fundamentals of Statistical and Thermal Physics"
- Pathria. "Statistical Mechanics"
- Tolman. "The Principles of Statistical Mechanics"

1. Motivation

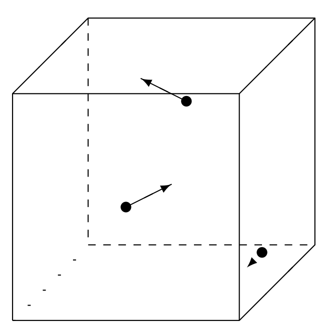
Statistical Mechanics is the branch of physics that studies the behavior of a system with a large number of degrees to freedom $N >> 1$. For example, if we wanted to study all of a particles of air in a room then $N \sim 10^{23}$ (Avagadro's number); other examples include say electrons in a piece of coper or magnetic spins in a bar of iron.

The basic question is how can go from the micro (eq Maxwell's Equations) to the macro (eg the Ideal Gas Law $PV = nRT$)?

1.1. Example: Ising Spins. A lattice (say 1-dim) with N sites and each site labeled by $+ \equiv\uparrow$ or $- \equiv\downarrow$ on a circle. We call configurations of up and down arrows Microstates (eg. $(+, +, -, \ldots)$). In the model here there are 2^N possible Microstates.



1.2. Example: Continuous Systems. Consider a volume V containing N particles. We impose periodic boundary conditions on V so we can actually think of the volume as a 3-Torus. Then each particle has position $\vec{q}$ and momentum $\vec{p}$ and is endowed with continuous degrees of freedom:



- Position $\vec{\varphi}_i = (x, y, z) \in T^3$
- Momentum $\vec{p}_i = (p_x, p_y, p_z) \in \mathbb{R}^3$

So this gives use a phase space $\mathcal{M} = T^*T^3 \times \ldots \times T^*T^3$ which is N copies of the tangent bundle.

In both of the above systems we can constrain the microstates of the system by introducing external conditions, for example conservation of energy for the continuous system leads to

$$E = \sum_{i=1}^{N} \frac{\vec{p_i}^2}{2m}$$

and conservation of energy for the Ising model give us, if the magnetic moment M = number of up spins - number of down spins, ie

$$M = \mu(\mu_+ - \mu_-)$$

the conservation law $E = -MH$ where H is come constant.

We note that in the following it will often be useful to us to use a "discreteized" phase space, given by dividing the phase space into cells of a given symplectic volume (eq. $\partial p \partial q = h$ ($= \hbar$ in quantum mechanics))

2. The Microcanonical Ensemble

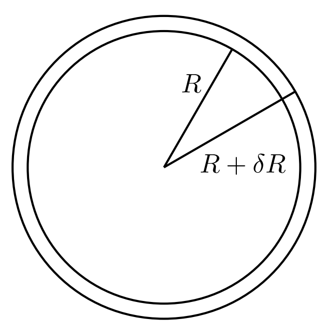
The Microcanonical Ensemble (MCE) is the probability distribution specified by the hypothesis of equal a priori probabilities, ie all microstates with (N,V,E) are all equally likely.

Set $\Omega(N, V, E) = \#$ of microstates given (N, V, E). For example, in the Ising spin model we have

$$E = -\mu H(N_+ - N_-) = \mu H(N - 2N_+)$$

so

$$\Omega(N, E) = \Omega(N, N_+) = \binom{N}{N_+} = \frac{N!}{N_+!(N - N_+)!}$$



2.1. Example. Lets consider an ideal gas over a spherical hypersurface in momentum space $\mathbb{R}^{3n}$. Then

$$E = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m}$$

and $R = \sqrt{2mE}$. Then

$$\Omega(N, V, E) = \omega(N, V, E) \sim V^N \frac{R^{3N-1}\delta R}{h^{3N}} = V^N E^{\frac{3N}{2}-1} \frac{\delta E}{h^{3N}}$$

where V^N is the position factor and $\delta R / h^{3N}$ is the size of the unit cell.

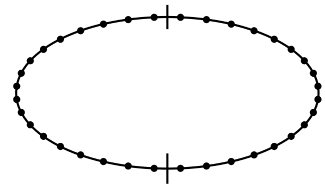
Let ρ be any probability distribution on the (classical) phase space M. We say that we are studying "Equilibrium Statistical Mechanics" if p is invariant under time, ie if $\frac{dp}{dt} = 0$. So, is the uniform distribution at equilibrium? Recall that under Hamiltonian time evolution $\mathcal{H}(p,q) = E$. In general, by Liouville's Theorem

$$\frac{dp}{dt} = \{p, \mathcal{H}\} = \frac{\partial p}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} - \frac{\partial p}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i}$$

Now, the uniform distribution is independent of both p_i and q_i and so the above derivative is 0. Therefore it is time independent. It's importnat to note that there are of course many more invariant distributions; indeed $\frac{dp}{dt} = 0$ if $p = p(E)$ since $\{\mathcal{H}, \mathcal{H}\} = 0$.

3. Temperature

For the Ising model, assume that we've divided our space into two systems. The total energy is given by $E_{tot} = E_A + E_B + E_{interaction}$ where in our example we can assume $E_{interaction} = 0$. Let $\Omega(E_{tot}, E_A)$ be # of states in the combined system $A \oplus B$ with combined energy (E_A, E_B). So



System A System B

$\Omega(E_{tot}, E_A) = \Omega_A(E_A)\Omega_B(E_B) = \Omega_A(E_A)\Omega_B(E_{tot} - E_A)$

Since all the states with a fixed E_{tot} are equally likely the probability that A has energy E_A is maximized when

$$0 = \frac{d}{dE_A}\ln\Omega(E_{tot}, E_A) = \frac{d}{dE_A}\ln\Omega_A(E_A) - \frac{d}{dE_B}\ln\Omega_B(E_B) = \beta_A - \beta_B$$

where

$$\beta = \frac{\partial}{\partial E}\ln\Omega = \frac{1}{k_B T}$$

is the definition of temperature and

$$S = k_B \ln\Omega$$

is entropy. Note that

$$\frac{\partial S}{\partial E} = \frac{1}{T}$$

4. The Canonical Ensemble

In the previous discussion, if $B >> A$ we say that B is a "heat bath." In general, a particular microstate of A is energy E_A occurs with probability $p_j \propto \Omega(E_{tot} - E_A)$ by assumption where $E_A << E_{tot}$. So we can use a Taylor expansion to get

$$\ln\Omega(E_{tot} - E_A) \cong \ln\Omega(E_{tot}) - \left.\frac{\partial \ln\Omega}{\partial E}\right|_{E_{tot}} + \ldots$$

or, putting $\beta = \frac{\partial \ln\Omega}{\partial E}$

$$\Omega(E_{tot} - E_A) \approx \Omega(E_{tot})e^{-\beta E_A}$$

then $p_j = ce^{-\beta E_A}$ where c is constant and $e^{-\beta E_A}$ is called the Boltzman weight.

Now, in the MCE we fixed the energy, in the Canonical Ensemble we will instead weight by the above. We normalize the Boltzman distribution by the requirement that

$$\sum_j p_j = \sum_j ce^{-\beta E_j} = 1$$

so $c^{-1} = \sum_j e^{-\beta E_j} = \mathcal{Z}$ where $\mathcal{Z}$ is the partition function. So once we know $\mathcal{Z}$ we can easily find

$$p_j = \frac{e^{-\beta E_j}}{\mathcal{Z}}$$

5. An Alternative Distribution

Consider $\mathcal{N}$-systems, all identical (eg a Ising chain), with total energy $\mathcal{E}$. Let j label the systems with the same energy and let n_j be the number of systems with energy E_j. Then

$$\sum_j n_j = \mathcal{N}$$
$$\sum_j n_j E_j = \mathcal{E}$$

As for Ising spins,

$$\Gamma(\{n_j\}) = \frac{\mathcal{N}!}{\prod_j n_j!} \qquad \text{(Multinomial Coefficients)}$$

is the number of configurations with occupation numbers in n_j.

Now, we want to look for a distribution of $\{n_j\}$ which maximizes Γ subject to the constraint that for

$$\begin{aligned}
\ln\Gamma = &\ln\mathcal{N}! - \sum_j \ln n_j! \\
=&(\mathcal{N}\ln\mathcal{N} - \mathcal{N}) - \sum_j n_j \ln n_j - n_j \\
=&\mathcal{N}\ln\mathcal{N} - \sum_j n_j \ln n_j
\end{aligned}$$

(where the second equality is by the Sterling Approximation and the second is by canceling the $\mathcal{N}$ with the n_j) we have

$$\delta \ln\Gamma = -\sum_j (n_j \ln n_j + 1)\delta n_j = 0$$

where $\sum_j \delta n_j = \sum_j E_j \delta n_j = 0$.

We now introduce Lagrange multipliers (α, β) such that

$$\sum_j (\ln n_j + \alpha + \beta E_j)\partial n_j = 0$$

So now $\ln \widehat{n}_j + \alpha + \beta E_j = 0$ which implies that $\widehat{n}_j = e^{-\alpha-\beta E_j}$ is the most likely distribution. So, choose α such that

$$\sum_j \widehat{n}_j = e^{-\alpha} \sum_j e^{-\beta E_j} = \mathcal{N}$$

Then

$$e^{-\alpha} = \frac{\mathcal{N}}{\sum_j e^{-\beta E_j}}$$

Next, we choose β such that

$$\frac{1}{\mathcal{N}} \sum_j \widehat{n}_j E_j = \frac{\sum_j \widehat{n}_j E_j e^{-\beta E_j}}{\mathcal{N} \sum_j e^{-\beta E_j}} = \frac{\mathcal{E}}{\mathcal{N}}$$

Finally, we have $p_j = \frac{n_j}{\mathcal{N}}$ and so

$$\begin{aligned}
\ln \Gamma =& \mathcal{N} \ln \mathcal{N} - \sum_j (\mathcal{N} p_j) \ln(\mathcal{N} p_j) \\
=& \mathcal{N} \ln \mathcal{N} - \mathcal{N} \sum_j p_j (\ln \mathcal{N} + \ln p_j) \\
=& -\mathcal{N} \sum_j p_j \ln p_j
\end{aligned}$$

CHAPTER 5

Quantum Groups and Statistical Mechanics

1. One-dimensional Ising model

1.1. Introduction. *Phase transitions* = a discontinuity of a miscroscopic observable M in terms of one of the parameters of the
system (E, T).

Examples:

(a) melting of ice (drastic change of density)
(b) boiling of water
(c) magnetization - here we have a discontinuity at low (room) temperature; at high temperature there is no discontinuity.

The phase transition can happen only in the thermodynamic over infinite volume limit ($N = \#\text{particles} \to \infty$)

We will consider the microscopic (rather than measure) observable

$$M : \Sigma \to \mathbb{R} \qquad \text{with} \qquad \langle M \rangle = \sum_{\sigma \in \Sigma} \frac{M(\sigma) \exp(-E(\sigma)/kT)}{Z}$$

where Σ is the state space, $\Sigma = \text{Maps}\{1, 2, \dots N\} \to \{\pm 1\} = \{\pm 1\}^N$. The visualization of this simple model is a line with N particles on it.

$$\text{———}\ \overset{1}{\bullet}\ \text{———}\ \overset{2}{\bullet}\ \text{———}\ \overset{3}{\bullet}\ \text{———}\ \cdots\ \text{———}\ \overset{N}{\bullet}\ \text{———}$$

The energy is:

$$E(\sigma) = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} - H \sum_{i=1}^{N} \sigma_i$$

where the first sum of the RHS expresses the internal couplings, and the second one the external couplings. In this formula we have the following:

- k is the Boltzmann constant;
- $H \in \mathbb{R}$ is the magnetic field strength;
- $J \in \mathbb{R}$ is the coupling constraint;
- the periodic boundary conditions hold: $\sigma_{N+1} = \sigma_1$.

The Boltzmann weight is

$$p(\sigma) = \exp(-E(\sigma)/kT)Z^{-1} = \exp(K\sum_{i=1}^{N}\sigma_i\sigma_{i+1} + h\sum_{i=1}^{N}\sigma_i)Z^{-1}$$

where we used the notation $K = \dfrac{J}{kT}$, $h = \dfrac{H}{kT}$. Z is called the partition function as is equal to

$$Z = \sum_{\sigma}\exp(K\sum_{i}\sigma_i\sigma_{i+1} + h\sum_{i}\sigma_i).$$

Using the periodic boundary conditions we infer that the system is **translation invariant**:

Take $\sigma_i : \Sigma \to \mathbb{R}$ and $F(\sigma_{i_1}, \sigma_{i_2}, \dots \sigma_{i_k})$ (for example, $M = \dfrac{1}{N}(\sigma_1 + \sigma_2 + \dots + \sigma_N)$). Then

$$\langle F(\sigma_{i_1}, \dots \sigma_{i_k})\rangle = \sum_{\sigma\in\Sigma} F(\sigma_{i_1}, \dots, \sigma_{i_k})\exp(-E(\sigma)/KT) = \langle F(\sigma_{i_1+1}, \dots \sigma_{i_k+1})\rangle$$

In particular, we have $M = \dfrac{1}{N}\langle\sigma_1+\sigma_2+\dots+\sigma_N\rangle = \langle\sigma_1\rangle = \langle\sigma_2\rangle = \dots = \langle\sigma_N\rangle$.

1.2. The partition function and the transfer matrix. We can express the partition function Z_N in a more preferable way using the transfer matrix.

$$Z_N = \sum_{\sigma}\exp(K\sum_{i}\sigma_i\sigma_{i+1}+h\sum_{i}\sigma_i) = \sum_{\sigma}\exp(K\sigma_1\sigma_2+\frac{h}{2}(\sigma_1+\sigma_2))\dots\exp(K\sigma_N\sigma_1+\frac{h}{2}(\sigma_N+\sigma_1))$$

Analyzing the possible values of a term $\exp(K\sigma_i\sigma_{i+1}+\frac{h}{2}(\sigma_i+\sigma_{i+1}))$, we consider a 2×2 matrix T indexed by ± 1:

$$(T_{\epsilon\epsilon'})_{\epsilon,\epsilon'\in\{\pm 1\}} = \begin{pmatrix}\exp(K+h) & \exp(-K)\\ \exp(-K) & \exp(K-h)\end{pmatrix}$$

T is called the transfer matrix. Computing the trace of the Nth power of T leads precisely to the partition function:

$$Z_N = \mathrm{Tr}(T^N) = \lambda_1^N + \lambda_2^N$$

where λ_1, λ_2 are the eigenvalues of T so that $\lambda_1 > |\lambda_2|$ (by Perron-Frobenius, for example). From here we get

$$\frac{\ln Z_n}{N} = \ln\lambda_1 + \frac{1}{N}\ln\left(1+\left(\frac{\lambda_2}{\lambda_1}\right)^N\right) \stackrel{N\to\infty}{\longrightarrow} \ln\lambda_1$$

Consider the free energy, $F_N = -KT \ln Z$, then the free energy per unit site is:

$$f = \lim_{N\to\infty} \frac{F_N}{N} = -KT \ln \lambda_1$$

Also consider the internal energy of the system

$$\langle E \rangle = Z^{-1} \sum_{\sigma\in\Sigma} E(\sigma) \exp(-E(\sigma)/kT) = kT^2 \frac{\partial \ln Z}{\partial T}$$

If we compute λ_1 explicitly, we get

$$\lambda_1 = \exp(K)\cosh(h)\sqrt{\exp(2K)\sinh^2 h + \exp(-2K)}$$

,

so plugging this in

$$f(H,T) = -KT \ln\left(\exp(K)\cosh(h)\sqrt{\exp(2K)\sinh^2 h + \exp(-2K)}\right)$$

We used that $\mathrm{Tr}(T) = 2\exp(K)\cosh(h)$ and $\det T = \exp(2K) - \exp(-2K) = 2\sinh(2K)$

1.3. Magnetization. Lemma: $M(H,T) = -\dfrac{\partial f(H,T)}{\partial H}$.

Proof: $M(H,T) = \dfrac{1}{N} Z_N^{-1} \sum_\sigma (\sigma_1 + \cdots + \sigma_N) \cdot \exp(-\dfrac{1}{kT}(E_0(\sigma) - H\sum_i \sigma_i))$, where $E_0 = -\sum_i J\sigma_i\sigma_{i+1}$

On the other hand, we have $Z = \sum_\sigma \exp(-\dfrac{1}{kT}(E_0(\sigma) - H\sum \sigma_i))$. Differentiating with respect to H we get

$$\frac{\partial Z}{\partial H} = \frac{Z}{kT}\sum_\sigma\left(\sum_i \sigma_i\right)\exp(\sum_i \sigma_i/kT) = \frac{Z}{kT}\sum_i \langle\sigma_i\rangle = \frac{NZ}{KT}M(H,T)$$

From $f_N = \dfrac{F_N}{N} = -KT/N \ln Z$ we get the result by differentiating with respect to H. ■

Hence we may express

$$M(H,T) = \frac{\exp(K)\sinh(h)}{\sqrt{\exp(2K)\sinh^2(h) + \exp(-2K)}}$$

Hence there is no phase transition with respect to H, which is "bad" in some sense.

1.4. Correlation functions. Recall the entries of the transfer matrix T: $T_{\sigma_i\sigma_j} = \exp(K\sigma_i\sigma_j + \frac{h}{2}(\sigma_i\sigma_j))$.

To understand the correlation we consider the following

$$\begin{aligned}\langle\sigma_1 \cdot \sigma_3\rangle = Z_N^{-1}\sum_\sigma \sigma_1 T_{\sigma_1\sigma_2}T_{\sigma_2\sigma_3}\sigma_3 \ldots T_{\sigma_n\sigma_1} = \\ = Z_N^{-1}\operatorname{Tr}(ST^2ST^{N-2})\end{aligned}$$

where $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

From here we easily deduce the general formulas:

$$\langle\sigma_i\rangle = Z_N^{-1}\operatorname{Tr}(ST^N)$$

$$\langle\sigma_i\sigma_j\rangle = Z_N^{-1}\operatorname{Tr}(ST^{j-i}T^{N-j+i})$$

Define the eigenmatrix $P = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ where $0 < \phi < \pi/2$ such that $\cot 2\phi = \exp(2K)\sinh(h)$

A straightforward verification shows that P satisfies the following properties:

- It diagonalizes T, i.e. $P^{-1}TP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, where $\lambda_1 > |\lambda_2|$;
- $P^{-1}SP = \begin{pmatrix} \cos 2\phi & -\sin 2\phi \\ -\sin 2\phi & -\cos 2\phi \end{pmatrix}$

Using this eigenmatrix, we get

$$\begin{aligned}\langle\sigma_i\sigma_j\rangle = Z_N^{-1}\operatorname{Tr}\left(P^{-1}SP\begin{pmatrix} \lambda_1^{j-i} & 0 \\ 0 & \lambda_2^{j-i} \end{pmatrix}P^{-1}SP\begin{pmatrix} \lambda_1^{N-j+i} & 0 \\ 0 & \lambda_2^{N-j+i} \end{pmatrix}\right) \overset{N\to\infty}{=} \\ = \cos^2(2\phi) + \left(\frac{\lambda_2}{\lambda_1}\right)^{j-i}\sin^2(2\phi)\end{aligned}$$

Similarly, we get $\langle\sigma_i\rangle = 2\cos 2\phi$.

So we get the correlation functions

$$g_{ij} = \langle\sigma_i\sigma_j\rangle - \langle\sigma_i\rangle\langle\sigma_j\rangle = \left(\frac{\lambda_2}{\lambda_1}\right)^{j-i}\sin^2(2\phi)$$

We use the notation $\left(\frac{\lambda_2}{\lambda_1}\right)^{j-i} = \exp\left(-\frac{j-i}{\xi}\right)$, where

$$\xi = \left[\ln\left(\frac{\lambda_1}{\lambda_2}\right)\right]^{-1} = \left[\ln\left(\frac{\exp(K)\cosh(h) + \sqrt{\exp(2K)\sinh^2(h) + \exp -2K}}{\exp(K)\cosh(h) - \sqrt{\exp(2K)\sinh^2(h) + \exp -2K}}\right)\right]^{-1}$$

When $H = 0 \Rightarrow h = 0$, we get $\xi = [\ln(\coth K)]^{-1} \to \infty$, as $T \to 0$.

So we get a criticality in the origin.

2. Two-dimensional Ising model on a square lattice

This model will be a more satisfactory one since it has a criticality (1944 Onsager), hence it is more realistic than the previous one. However, it has been solved only for 0 magnetic field.

We gather our notations: $N = \#$ sites, $\Sigma = \{\pm 1\}^N$

$$Z_N = \sum_{\sigma \in \Sigma} \exp(K \sum_{i-j} \sigma_i \sigma_j + L \sum_{|^i_j} \sigma_i \sigma_j)$$

is the partition function, where $K = J/kT$, $L = J'/kt$ with k the Boltzmann constant, $i - j$ are the horizontal bonds (edges) and $|^i_j$ are the vertical ones.

We impose free boundary conditions (so there are no torus-like interactions at the boundaries).

In the following, we represent the partition function in two ways, depending on the temperature.

2.1. Low temperature representation of Z_N.

$$\begin{array}{cccc} \text{Lattice } L & +\bullet & -\bullet & +\bullet \\ & & & \\ & +\bullet & +\bullet & -\bullet \\ & & & \\ & +\bullet & -\bullet & -\bullet \\ & & & \\ & +\bullet & -\bullet & +\bullet \end{array}$$

Given a lattice L, we define its dual lattice $L^\vee$, that is:

- vertices of $L^\vee$ correspond to faces of L.
- edges of $L^\vee$ correspond to pairs of adjacent faces of L.

Given $\sigma \in \Sigma$, we can isolate the $+$ vertices of L from the $-$ vertices of L, forming so-called "islands". To give a configuration $\sigma \in \Sigma$ is the same to give spins of the faces of $L^\vee$. Dividing the islands, we get a **polygon configuration** $\mathcal{P}(\sigma)$ on $L^\vee$.

Remarks:

1) $\mathcal{P}(\sigma) = -\mathcal{P}(\sigma') \Leftrightarrow \sigma = \pm\sigma'$
2) $E(\sigma) = E(-\sigma)$, because $H = 0$.
3) The horizontal edges in L that have opposite signs on their ends correspond precisely to the vertical edges of the polygon configuration $\mathcal{P}$.
4) As above, the vertical edges in L that have opposite signs on their ends correspond to the horizontal edges in $\mathcal{P}$.

Let $r = |P_v| = \#$ vertical edges in $\mathcal{P}$, $s = |P_h| = \#$ horizontal edges in $\mathcal{P}$, and $M = \#$ horizontal edges in $L = \#$ vertical edges in L.

Using remarks 3) - 4) we get

$$-\frac{E(\mathcal{P})}{kT} = K\sum_{i-j}\sigma_i\sigma_j + L\sum_{|^i_j}\sigma_i\sigma_j = K(M-r)-Kr+L(M-s)-Ls = K(M-2r)+L(M-2s)$$

So the partition function is

$$Z_N = \sum_\sigma \exp(K(M-2r)+L(M-2s)) = 2\exp(M(K+L))\sum_{\mathcal{P}}\exp(-2Kr-2Ls)$$

The last equality holds since polygonial configurations in the dual lattice correspond to states up to signs (remark 1), which also accounts for the factor 2.

Denote by $v^* = exp(-2K)$ and $w^* = exp{-2L}$ (it will become clear in the next section why we use these notations)

$$Z_N = 2\exp(M(K+L))\sum_{\mathcal{P}} v^{*|P_v|}w^{*|P_h|}$$

2.2. High temperature representation of Z_N. We define a similar formula for the partition function for high temperature. We start with the simple observation:

$$\exp(K\sigma_i\sigma_j) = \cosh K + \sinh \sigma_i\sigma_j$$

Using this, and denoting $v = \tanh K$, $w = \tanh L$, we have

$$Z_N = \sum_\sigma \prod_{i-j}(\cosh K + \sinh K\sigma_i\sigma_j)\prod_{|^i_j}(\cosh L + \sinh L\sigma_i\sigma_j) =$$

$$= (\cosh K\cosh L)^M \sum_\sigma \prod_{i-j}(1+v\sigma_i\sigma_j)\prod_{|^i_j}(1+w\sigma_i\sigma_j) =$$

$$= (\cosh K\cosh L)^M \sum_\sigma \sum_{P\subset \text{edges of } L} v^{|P_h|}w^{|P_v|}\prod_{i-j\in P}\sigma_i\sigma_j\prod_{|^k_l\in P}\sigma_k\sigma_l =$$

$$== (\cosh K\cosh L)^M \sum_{P\subset \text{edges of } L} v^{|P_h|}w^{|P_v|}\sum_\sigma\prod_{i-j\in P}\sigma_i\sigma_j\prod_{|^k_l\in P}\sigma_k\sigma_l$$

Now it's not hard to see that

$$\sum_\sigma\prod_{i-j\in P}\sigma_i\sigma_j\prod_{|^k_l\in P}\sigma_k\sigma_l = 0$$

except when every vertex has an even number of incoming edges which happens precisely when P is a polygonial configuration on L. In this case, each term in the sum is 1, hence the sum will be 2^N. Hence we got

$$Z_N = 2^N(\cosh K\cosh L)^M \sum_{\mathcal{P}\text{ polygonial config. on } L} v^{|P_h|}w^{|P_v|}$$

Define K^*, L^* by

- $\tanh K^* = v^* = \exp(-2L)$

- $\tanh L^* = w^* = \exp(-2K)$

Substituting these into the formula for low temperature representation we get a similar one to the high temperature representation.

It is an easy exercise to show that the preceding definitions are equivalent to

- $\sinh(2L)\sinh(2K^*) = 1$
- $\sinh(2L^*)\sinh(2K) = 1$

From this we deduce that the relation is involutive, i.e.

$$(K, L) \to (L^*, K^*)$$

exchanges horizontal with vertical (dual lattice), and exchanges high temperature representation with the low temperature representation.

Taking $N, M \to \infty \quad (N/M \to 1)$ we have

$$-\psi = \lim_{N\to\infty} \frac{1}{N} \ln Z_N \qquad (= f/kT, \text{ where f is the free energy per unit site})$$

For low temperature we get

$$-\psi_L = K + L + \Phi_{L^\vee}(v^*, w^*)$$

where $\Phi_{L^\vee}(v^*, w^*) = \lim_{N\to\infty} \frac{1}{N} \ln \sum_{\mathcal{P}(L^\vee)} v^{*|P_h|} w^{*|P_v|}$, and we assume this limit exists.

For high temperature we get

$$-\psi_L = \ln(2\cosh K \cosh L) + \Phi_L(v, w)$$

Now we apply these formulas to the dual lattice:

$$-\psi_{L^\vee}(K^*, L^*) \stackrel{\text{high temp.}}{=} K^* + L^* + \Phi_L(v, w) \stackrel{\text{low temp.}}{=} \ln(2\cosh K^* \cosh L^*) + \Phi_{L^\vee}(v^*, w^*)$$

After eliminating $\Phi_{L^\vee}(v^*, w^*)$, we get

$$-\psi_L(K, L) = K + L - \ln(2\cosh K^* \cosh L^*) - \psi_{L^\vee}(K^*, L^*)$$

An easy calculation shows that $\ln(2\cosh K^* \cosh L^*) = \frac{1}{2}\ln(\sinh 2K^* \sinh 2L^*)$, and using that at infinity $L \approx L^\vee$, we arrive to

$$\psi_L(K^*, L^*) = \psi_L(K, L) - \frac{1}{2}\ln(\sinh 2K^* \sinh 2L^*)$$

2.3. Criticality. First consider the isotropic case, i.e. $K = L$

If there is $K = K_C$ a critical value, then ψ_L is not analytic in K_C. From the formula above we conclude that it is not analytic at K_C^* either.

If this critical value is unique, then $K_C = K_C^*$, and computation shows that the only value can be $K_C = 0.4406\ldots$ (Kramers-Wannier duality)

Peierls showed that there is at least one critical point K_C. Onsager proved that this is indeed unique, so the preceding discussion is valid.

In the anisotropic case, if there exists a unique critical curve, it is at criticality $\sinh(2L_C)\sinh(2K_C) = 1$.

CHAPTER 6

Ising Model on the Honeycomb Lattice

1. Low and High Temperature Duality

The partition function for the hexagonal Ising lattice (the black part of Figure 1) will given by

$$Z_N^{\mathcal{H}}(L) = \sum_{\sigma \in \{\pm 1\}^N} \exp(L_1 \sum_{i \diagup j} \sigma_i \sigma_j + L_2 \sum_{j \diagdown i} \sigma_i \sigma_j + L_3 \sum_{i \updownarrow j} \sigma_i \sigma_j$$

Where the magnetic field $H = 0$, N denote the number of states and L denotes the interaction coefficients $L_i = J_i/k_g T$ where J_i's are the coupling constants. Now, we have a dual triangular lattice given by the red part of Figure 1. The partition function for this lattice is

$$Z_N^{\mathcal{H}}(L) = \sum_{\sigma \in \{\pm 1\}^N} \exp(K_1 \sum_{j \diagdown i} \sigma_i \sigma_j + K_2 \sum_{i \diagup j} \sigma_i \sigma_j + K_3 \sum_{j - i} \sigma_i \sigma_j$$

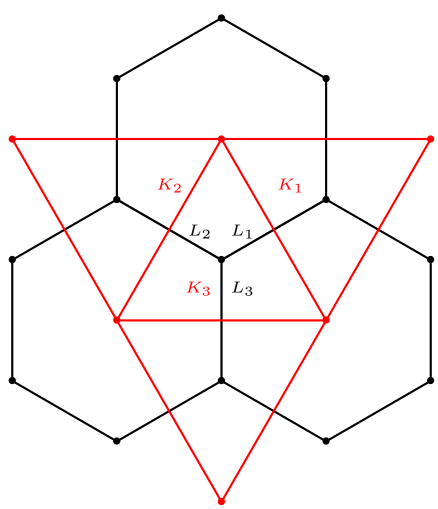


FIGURE 1. The Ising Model on a Low Temperature Hexagonal Lattice representation $\mathcal{H}$ overlaid with it's dual High Temperature Triangular Lattice $\mathcal{C}$.

In the previous section, we observed that for a spin configuration on vertices of the square lattice

$$\frac{\sigma \in \{\pm 1\}^{V(\mathcal{L})}}{\mathbb{Z}_2} \longleftrightarrow \mathcal{P}(\sigma)$$

where $\mathcal{P}(\sigma)$ is a polygon configuration on the dual lattice $\mathcal{L}_{\mathcal{H}}^{\vee}$ and the $\mathbb{Z}_2$ factor denotes that this correspondence is up to an overall sign. Since $H = 0$,

$$E(\sigma) = E(-\sigma) = E(\mathcal{P}(\sigma)) = -M(L_1 + L_2 + L_3) + 2L_1 r_1 + 2L_2 r_2 + 2L_3 r_3$$

where r_i is the number of edges in $\mathcal{P}(\sigma)$ "perpendicular" to and edge in $\mathcal{L}$ of type i and M is the number of edges in $\mathcal{L}^{\vee}$ parallel to a given direction. Note that M is only the same for each direction in the limit. Then

$$Z_{2N}^{H}(L) = 2\exp(M(L_1 + L_2 + L_3)) \sum_{\mathcal{P} \subset \mathrm{edges}(\mathcal{L}^{\vee}} \exp\left((-2L_1 r_1 - 2L_2 r_2 - 2L_3 r_3\right)$$

is the Low Temperature Representation. At low temperature, $L_i = \frac{1}{T}$ so the sum of exponent above is small. Take $2N$ to be the number of vertices in $\mathcal{L}_{\mathcal{H}}$ so that $2N$ is the number of faces in $\mathcal{L}_{\mathcal{H}}^{\vee}$. Then, in the limit

$$\begin{aligned}
&\# \text{ of edges in } \mathcal{L}_{\mathcal{H}}^{\vee} \approx \tfrac{6N}{2} = 3N \\
&\# \text{ of verticies in } \mathcal{L}_{\mathcal{H}} \approx \tfrac{6N}{2} = 3N
\end{aligned}$$

so the partition function becomes

$$Z_{2N}^{H}(L) = 2\exp(N(L_1 + L_2 + L_3)) \sum_{\mathcal{P} \subset \mathrm{edges}(\mathcal{L}^{\vee})} \exp\left((-2L_1 r_1 - 2L_2 r_2 - 2L_3 r_3\right)$$

2. High Temperature Representation for $\mathcal{L}_H^{\vee} = \mathcal{L}_{\mathcal{C}}$ on the triangle lattice

Recall that for $\epsilon^2 = 1$

$$\begin{aligned}
\exp(\epsilon K) &= \cosh(K) + \epsilon \sinh(K) \\
&= \cosh(K)(1 + \epsilon \tanh(K)) \\
&=: \cosh(K)(1 + \epsilon v)
\end{aligned}$$

where $v = \tanh(K)$. Then for $K := (K_1, K_2, K_3)$, the partition function for the triangular lattice $\mathcal{C}$ becomes

$$\begin{aligned}
Z_N^{\mathcal{C}}(K) = {} & [\cosh(K_1)\cosh(K_2)\cosh(K_3)]^N \times \\
& \sum_{\sigma \in \{\pm 1\}}^{V(\mathcal{C})} \prod_{j\,i} (1 + v_1 \sigma_i \sigma_j) \prod_{j\,i} 1 + v_2 \sigma_i \sigma_j) \prod_{j\,i} (1 + v_3 \sigma_i \sigma_j) \\
& = [2\cosh(K_1)\cosh(K_2)\cosh(K_3)]^N \sum_{\mathcal{P} \subset \mathrm{edges}(\mathcal{L}_{\mathcal{C}})} v_1^{r_1} v_2^{r_2} v_3^{r_3}
\end{aligned}$$

where r_i is the number of edges in $\mathcal{P}//K_i$ and $v_i = \tanh(K_i)$. Now we want to do a matching: let $v_i = \exp(-2L_i)$. Then it can be shown that this imply that

$$\cosh(K_i) = e^{L_i}(2\sinh(2L_i))^{1/2}$$

which in turn implies that

$$\sinh(2L_i)\sinh(2K_i) = 1 \tag{2.1}$$

We now have the following duality:

$$Z_N^{\mathcal{C}}(K) = \frac{1}{2}(2\sinh(2L_1)\sinh(2L_2)\sinh(2L_3))^{\frac{N}{2}} Z_{2N}^{H}(L)$$

provided (2.1) holds.

3. The Star-Triangle Duality

Observe that the vertices of the honeycomb lattice can be partitions into two distinct sets: $V(\mathcal{L}_{\mathcal{H}}) = A \sqcup B$ such that for any edge connecting i and j we have $i \in A$ and $j \in B$ or $j \in A$ and $i \in B$. Then

(3.1)

$$\begin{aligned}
\mathcal{L}_{\mathcal{H}} &= \sum_{\sigma\in\{\pm1\}^{A\sqcup B}} \exp\left(\sum L_1\sigma_i\sigma_j + \sum L_2\sigma_i\sigma_j + \sum L_3\sigma_i\sigma_j \right) \\
&= \sum_{\sigma\in\{\pm1\}^{A}} \sum_{\sigma\in\{\pm1\}^{B}} \prod_{b\in B} \exp\sigma_b \left(\sum L_1\sigma_i\sigma_j + \sum L_2\sigma_i\sigma_j + \sum L_3\sigma_i\sigma_j \right) \\
&= \sum_{\sigma\in\{\pm1\}^{A}} \prod_{b\in B} 2\cosh\left(\sum L_1\sigma_i\sigma_j + \sum L_2\sigma_i\sigma_j + \sum L_3\sigma_i\sigma_j \right) \\
&= \sum_{\sigma\in\{\pm1\}^{A}} \prod_{a \,\underset{b}{\nabla}\, a'} \cosh\left(\sum L_1\sigma_i\sigma_j + \sum L_2\sigma_i\sigma_j + \sum L_3\sigma_i\sigma_j \right)
\end{aligned}$$

The idea here is that we want to write the above as a partition function on the triangle lattice with vertices A and edges

$a \longrightarrow a'$ in A if and only if there is a vertex b such that $a \,\underset{b}{\nabla}\, a'$

Now we will start parsing Eq. (3.1). We start by writing

(3.2)

$$2\cosh(L_1\sigma_{a_1} + L_2\sigma_{a_2} + L_3\sigma_{a_3}) = R\exp(K_1\sigma_{a_2}\sigma_{a_3} + K_2\sigma_{a_1}\sigma_{a_3} + K_3\sigma_{a_1}\sigma_{a_2})$$

The left hand side takes 8 values but by the evenness of cosh we can reduce this to 4. The right hand side on the other hand only takes on 3 values so we must add the R above. If R and K satisfy Eq. (3.2) then $Z_{2N}^{\mathcal{H}}(L) = R^N Z_N^{\mathcal{C}}(K)$. This is the so called Star-Triangle relation since we've replaced

$$\underset{a_3}{a_1 \overset{b}{\curlyvee} a_2} \qquad \text{with} \qquad \underset{a_3}{a_1 \,\nabla\, a_2}$$

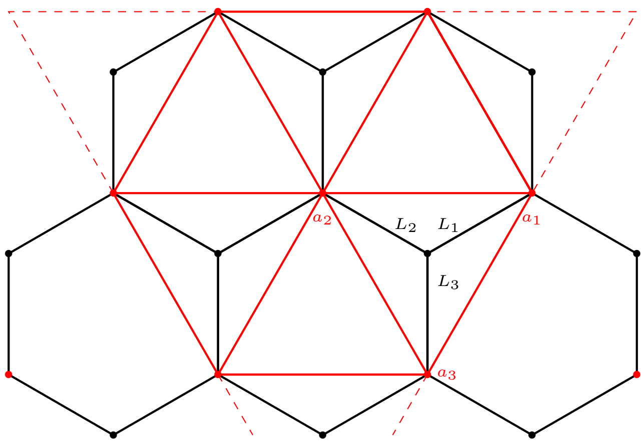


FIGURE 2. The Star-Triangle Duality.

Let solve Eq. (3.2). By evenness, assume that the number of -'s is less than the number of +'s. Then

σ_1	σ_2	σ_3	
1	1	1	$2\cosh(L_1+L_2+L_3) = R\exp(K_1+K_2+K_3)$
-1	1	1	$2\cosh(-L_1+L_2+L_3) = R\exp(K_1-K_2-K_3)$
1	-1	1	$2\cosh(L_1-L_2+L_3) = R\exp(-K_1+K_2-K_3)$
1	1	-1	$2\cosh(L_1+L_2-L_3) = R\exp(-K_1-K_2+K_3)$

Set $c = \cosh(L_1+L_2+L_3)$, $c_i = \cosh(-L_i+L_j+L_k)$ then the above give us $\exp(4K_1) = cc_1/c_2c_3$. Recall the basic hyperbolic trig identities:

$$\cosh(a)\cosh(b) = \frac{\cosh(a+b)+\cosh(a-b)}{2}$$
$$\sinh(a)\sinh(b) = \frac{\cosh(a+b)-\cosh(a-b)}{2}$$

Then

$$\begin{aligned}
\exp(4K_1) - 1 &= \frac{cc_1 - c_2c_3}{c_2c_3} \\
&= \frac{\cosh(L_1+L_2+L_3)\cosh(-L_1+L_2+L_3) - \cosh(L_1-L_2+L_3)\cosh(L_1+L_2-L_3)}{c_2c_3} \\
&= \frac{\cosh(2L_2+2L_3) - \cosh(2L_2-2L_3)}{2c_2c_3} \\
&= \frac{\sinh(2L_2)\sinh(2L_3)}{c_2c_3} \\
&= \frac{e^{-2K_1}\left(e^{4K_1}-1\right)\sinh(2L_1)}{2c_2c_3} \\
&= \left(\frac{c_2c_3}{cc_1}\right)^{\frac{1}{2}} \frac{\sinh(2L_1)\sinh(2L_2)\sinh(2L_3)}{2c_2c_3} \\
&= \frac{\sinh(2L_1)\sinh(2L_2)\sinh(2L_3)}{2(cc_1c_2c_3)^{-1/2}} \\
&=: \bar{K}^{-1}
\end{aligned}$$

Now by symmetry $\sinh(K_i)\sinh(2L_i) = \bar{K}^{-1}$ for i =1, 2, 3. Multiplying all 4 initial relations, we get

$$R^4 = 16cc_1c_2c_3$$

or

$$\begin{aligned}
R^2 &= 4(cc_1c_2c_3)^{1/2} \\
&= 2K\sinh(L_1)\sinh(L_2)\sinh(L_3) \\
&= \frac{2}{K^2}(\sinh(2K_1)\sinh(2K_2)\sinh(2K_3))^{-1}
\end{aligned}$$

Essentially what we're doing here is averaging over the b points in Figure 2. This is the temperature preserving since if $T >> 0$, L_i and K_i are both small so $\sinh(2L_2) \approx \sinh(2L_3) \approx 0$ and $c_2 \approx c_3 \approx 1$. Then

$$e^{4K-1} = \frac{sh(2L_2)\sinh(2L_3)}{c_2c_3} \approx 0$$

We will now use both dualities:

$$\begin{aligned}
Z_N^{\mathcal{C}}(K) &\overset{\text{H-L Temp Dual}}{=} \frac{1}{2}(2\sinh(2L_1)\sinh(2L_2\sinh(2L_3)^{-\frac{N}{2}} Z_{2N}^{\mathcal{H}}(L) \\
&\overset{\text{Star-Tri Dual}}{=} \frac{1}{2}(2\sinh(2L_1)\sinh(2L_2\sinh(2L_3)^{-\frac{N}{2}} (R^2)^{\frac{N}{2}} Z_N^{\mathcal{C}}(K^*) \\
&= \frac{1}{2}K^{-\frac{N}{2}} Z_N^{\mathcal{C}}(K^*)
\end{aligned}$$

where $\sinh(2K_i^*)(= (\sinh 2L_i)^{-1}) = \bar{K}\sinh(2K_1)$ where the first equality is by the High-Low Temp. Duality and the second is by the Star-Tri Duality.

Exercise. Show that the map $K \mapsto K^*$ is an involution.

Now assuming that the free energy for $\mathcal{L}_3$ converges as $N \to \infty$ and there exists a unique critical hypersurface in K space then by the exercise we will obtain a critical point when $K_i = K_i^*$. But this implies that $\bar{K} = 1$ and so $K_i = K_i^C = 0.27465$. Similarly for the honeycomb model we have $L_i = L_i^C = 0.6584$

4. Renormalization

Consider the $4N$ honeycomb lattice given by L_1, L_2 and L_3 which is dual to the $2N$ triangle lattice given by R, K_1, K_2 and K_3 by

$$\sinh(2K_i)\sinh(2L_i) = \bar{K}^{-1}$$

By taking the duality again we get that this is dual to the N honeycomb lattice $\mathcal{H}'$:

$$Z_{2N}^{\mathcal{C}} = \sum_{\sigma\in\{\pm 1\}} \prod_{\substack{\hexagon \\ a_6 \quad a_1 \quad a_2}} 2\cosh\left(\sum_{i=1}^{6} K_i \sigma_{a_i}\right)$$

Then by requiring

$$2\cosh\left(\sum_{i=1}^{6} K_i \sigma_{a_i}\right) = \exp(L_\bullet^* \sigma_i \sigma_j)$$

you should get a new lattice with parameters L_i^{RG}. It can be shown that at criticality $L = L^{RG}$.

CHAPTER 7

Ising Model on the Square Lattice

1. Commuting transfer matrices

Given a lattice as above we slice the lattice along the diagonal and turn it to yield a "stack" of one dim spin chains. Now, assume that $H = 0$, we have periodic boundry conditions and that there are m rows, where m is even. Then the partition function has the following form:

$$Z_N = \Sigma_{\phi_1 \dots \phi_n \in \{\pm 1\}^n} V_{\phi_1 \phi_2} W_{\phi_2 \phi_3} V_{\phi_3 \phi_4} \dots V_{\phi_{m-1} \phi_m} W_{\phi_m \phi_1}$$

where:

$$V_{\phi \phi'} = exp(\Sigma_{i=1}^{n} (L \sigma_i \sigma_i' + K \sigma_{i+1} \sigma_i'))$$

and

$$W_{\phi \phi'} = exp(\Sigma_{i=1}^{n} (L \sigma_i \sigma_{i+1}' + K \sigma_i \sigma_i'))$$

The partition function then becomes:

$$Z_N = tr(VWVW \dots VW) = tr((VW)^{\frac{m}{2}})\ \lambda_1^{m/2}$$

Where $\rightarrow \lambda_1$ is the largest eigen value.

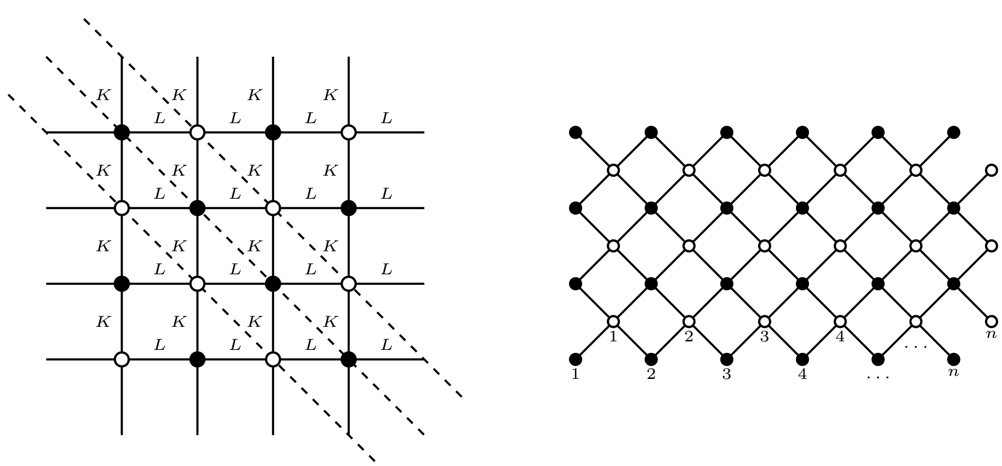
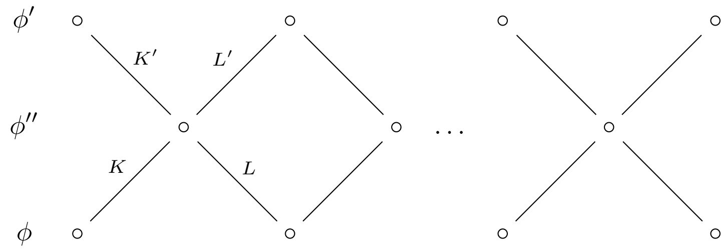


Figure 1. The Ising Model on a Square Lattice

Our aim is to study products of the from: $V(K, L)W(K', L')$. This corresponds to this part of the lattice:



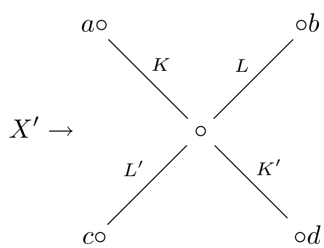
So

$$\begin{aligned}(VW)_{\phi\phi'} =&\Sigma_{\phi''}V_{\phi\phi''}W_{\phi''\phi'}\\ =&\Sigma_{\sigma_1'',\dots,\sigma_n''}\Pi_{j=1}^{n}exp[\sigma_j''(L\sigma_j + K\sigma_{j+1} + K'\sigma_j' + L'\sigma_{j+1}')\\ =&\Pi_{j=1}^{n}X_j(\sigma_j, \sigma_{j+1}, \sigma_j', \sigma_{j+1}')\end{aligned}$$

where

$$X_j(a, b, c, d) = \Sigma_{f\in\{\pm1\}}exp(f(La + Kb + K'c + L'd)$$

Graphically we have



2. Commutation

We are looking for constants K', L' such that the following matrices equation is satified:

$$V(K, L)W(K', L') = V(K', L')W(K, L) \tag{2.1}$$

Note first that $\Pi_{j=1}^{n}X_j(\sigma_j, \sigma j + 1, \sigma_j', \sigma_{j+1})$ is invariant if

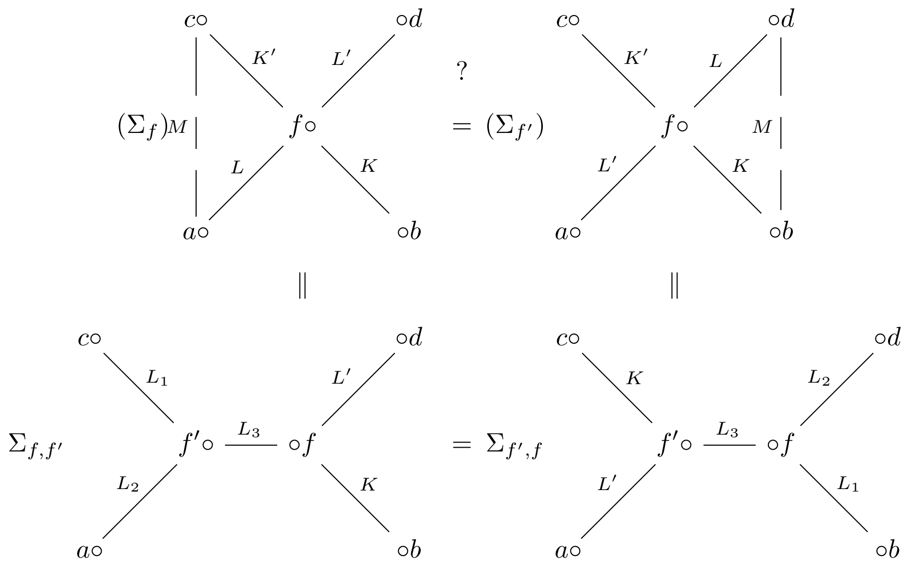
$$X(a, b, c, d) \to exp(Mac)X(a, b, c, d)exp(-Mbd)$$

Then, if there exists $M \in \mathbf{C}$ such that

$$exp(Mac)X(a, b, c, d) = X'(a, b, c, d)exp(Mbd)$$

where $X' = X$ but with L and K exchanged with L' and K' respectively.Then condition (2.1) holds.

Schematically this is



Where $(M, K', L) \mapsto (L_3, L_2, L_1)$ and where the bottom two graphs are the Star-Triangle Duality. The bottom equality holds for $L_2 = L', L_1 = K$ and implies that

$$\sinh(2L)\sinh(2K) = \sinh(2L')\sinh(2K')$$

3. Inversion

We want the product: $V(K,L)W(K',L')$ to be "near diagonal" i.e we want

(SL1) $X(a,b,c,d) = 0$ if $a \neq c$ and $b = d$

3.1. Corollary. If the (SL1) holds then

$$(VW')_{\phi\phi'} = \prod_j X_j(\sigma_j, \sigma_{j+1}, \sigma'_j, \sigma'_{j+1}) = 0$$

unless $\phi = \pm\phi'$.

In addition, the (SL1) is equivalent to the following:

$$\cosh(L - K' + (K + L')) = 0 \tag{3.1}$$

$$\cosh(L - K' - (K + L')) = 0 \tag{3.2}$$

3.2. Exercise 1. (SL1)$\Leftrightarrow$ $K' = L + (\pi_1/2)m_1$ and $L' = -K + (\pi_1/2)m_2$ where $m_1, m_2 \in \mathbf{Z}$ of opposite parity

3.3. Exercise 2. Exercise 1 $\Rightarrow \sinh(2L)\sinh(2K) = \sinh(2L')\sinh(2K')$

From now on we will restrict to the case $m_1 = 1; m_2 = 0$.

3.4. Exercise 3. Prove that both

$$X(a,b,c,d) = 2i * sinh(2L)$$

and

$$X(a,b,-a,-b) = -2iab * sinh(2K)$$

Now,

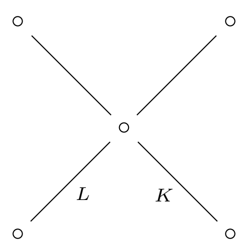
$$(VW)_{\phi,\phi'} = \delta_{\phi,\phi'}(2i * sinh(2L))^n + \delta_{\phi,-\phi'}(-2i * sinh(2K))^n$$

and

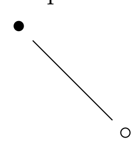
$$V(K,L)W(L+(i\pi/2);-K) = (2\pi i * sinh(2L))^n I + (-2i * sinh(2K))^n R$$

where $R_{\phi,\phi'} = \delta_{\phi,-\phi'}$. Now, $R^2 = 1$ so the right hand side of the above is easy to invert. Therefore (VW^*) is easy to invert and so V is easy to invert. In addition we have the relations:

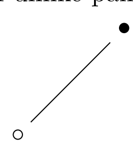
(1) $(K \longleftrightarrow L; \phi \longleftrightarrow \phi') \Rightarrow V \longleftrightarrow W$
(2) $V(-K,-L) = RV(K,L) = V(K,L)R$ for some W
(3) Fix ϕ and ϕ' and consider the following



Let $r =$ the number of unlike pairs



and let $s =$ the number of unlike pairs



Then $r+s =$ the number of sign changes in sequence $\sigma_1', \sigma_2, \sigma_2', \ldots, \sigma_1'$ and

$$V_{\phi\phi'} = exp[(m-2r)K + (m-2s)L]$$

Assume now that $m = 2p$ is even. Then for $r' \in [0,p]$

$$m - 2r = 2(p-r) \triangleq \pm 2r'$$

so

$$V_{\phi\phi'} = exp[\pm 2r'K \pm 2s'L]$$

for r',$s' \in [0,p]$

and

$$V_{\phi\phi'}(K \pm \frac{\pi i}{2}, L \pm \frac{\pi i}{2}) = V_{\phi\phi'}(K, L)$$

3.5. Some more relations.

- $C_{\phi\phi'} = \delta_{\sigma_1\sigma_2'}\delta_{\sigma_2\sigma_3'} \dots \delta_{\sigma_n\sigma_1'}$ is called the "Coxeter element"
- Translation invariance in the direction $\rightarrow XXXXXXXXX$ implies that

$$\begin{aligned} V(K, L) =& C^{-1}V(K, L)C \\ W(K, L) =& C^{-1}W(K, L)C \end{aligned}$$

  Moreover we have the following properties: $W(K, L) = V(K, L)C$ by inspection.
- Since $V(K, L)W(K', L') = V(K', L')W(K, L)$ we use the above to get $V(K, L)V(K', L')C = V(K', L')V(K, L)C$ and so

$$[V(K, L), V(K', L')] = 0$$

  Or, to summarize, $V(K, L), V(K', L'), C, R$ all commute.
- **Inversion identity.**

$$V(K, L)V(L + \frac{\pi i}{2}, -K)C = (2ish(2L))^n I + (-2ish(2K))^n R$$

Consider joint eigenvector x of above matrices, ie

$$\begin{aligned} V(K, L)x =& v(K, L)x \\ Cx =& cx \\ Rx =& rx \\ C^n = R^2 =& 1 = c^n = r^n \end{aligned} \tag{3.3}$$

Then the inversion relation becomes:

$$v(K, L)v(L + \frac{\pi i}{2}, -K)c = (2i\sinh(2L))^n + (-2i\sinh(2K))^n r$$

We are interested in the eigenvalues of $V \bullet W = V^2C$. Define $\Lambda(K, L) \triangleq v(K, L)c^{1/2}$ so that evaluation of VW are the square of the $\Lambda(K, L)$. The the inversion formula takes the form

$$\Lambda(L, K)\Lambda(L + \frac{\pi i}{2}, -K) = (2ish(2L))^n + (-2ish(2K))^n r$$

Now, we need to parametrize all K and L such that

$$\sinh(2K)\sinh(2L) = k^{-1}$$

To do this, we will restrict to the case $k = 1$ (if you don't the proof still goes through but it requires elliptic, not hyperbolic, functions). One such parametrization is given by the following. Set

$$\sinh(2K) = \tan(u)$$
$$\sinh(2L) = \cot(u)$$

For $u \in (0, \frac{\pi}{2})$. Note, this makes sense since we are assuming $K, L \in \mathbb{R}^+$ and for u in the interval, $\tan u$ is a strictly increasing function covering $\mathbb{R}^+$. Complexifying we get

$$\begin{Bmatrix} \sinh(2K) = \tan u \\ \sinh(2L) = \cot u \end{Bmatrix} \subset \mathbf{C}^3 \ni \{u, K, L\}$$

The functional relations then become

$$\Lambda(u)\Lambda(u + \pi/2) = (2icot(u))^n + (2itan(u))^n r$$

Where

$$K \mapsto L + 2\pi \Rightarrow \sinh(K) \to -\sinh(2L)$$
$$L \mapsto -K \Rightarrow \sinh(L) \to -\sinh(2K)$$

Next, we can derive the following:

$$\begin{aligned} e^{2K} &= \cosh(2K) + \sinh(2K) \\ &= \sqrt{1 + \sinh^2(2K)} + \sinh(2K) \\ &= \sqrt{1 + \tan^2(u)} + \tan(u) \\ &= \frac{1}{\sqrt{\cos^2(u)}} + \frac{\sin(u)}{\cos(u)} \\ &= \frac{1 + \sin(u)}{\cos(u)} \end{aligned}$$

We get a similar result for e^{-2K} and $e^{\pm 2L}$:

$$e^{\pm 2K} = \frac{1 \pm \sin(u)}{\cos(u)}$$
$$e^{\pm 2L} = \frac{1 \pm \cos(u)}{\sin(u)}$$

Where as always $u \in (0, \frac{\pi}{2})$.

Now, recall that

$$V_{\phi,\phi'} = exp(\pm 2r'K \pm 2s'L) = (\frac{1 \pm \sin(u)}{\cos(u)})^{r'} (\frac{1 \pm \cos(u)}{\sin(u)})^{s'}$$

where $r', s' \in (1, 2, \dots, \frac{n}{2})$ and r', s' are even, and so

$$V_{\phi,\phi'} = \frac{t(u)}{(\sin(u)\cos(u))^p} \tag{3.4}$$

where $t(u)$ is a polynomial in sin(u), cos(u) of total degree p. Now the eigenvalue equation$V(K,L)x = v(K,L)x$ implies that $v(K,L)$ is of the form given in (3.4) and so $\Lambda(K,L)$ are of the form given in (3.4).

We will now check the periodicity of $\Lambda(u)$ and notice that under the $u \to u+\pi$ we have:

$$e^{\pm 2K} = \frac{1 \pm sin(u)}{cos(u)} \to -\frac{1 \pm sin(u)}{cos(u)} = -e^{\pm 2K}$$

so $K \longrightarrow -K \pm i\frac{\pi}{2}$ and similarly, $L \longrightarrow L \pm i\frac{\pi}{2}$. Now,

$$V(-K \pm i\frac{\pi}{2}; -L \pm i\frac{\pi}{2}) = V(-K,L) = V(K,L)R$$

so we must have

$$\begin{aligned} v(u+\pi) =& v(u)r \\ \Lambda(u+\pi) =& \Lambda(u)r \end{aligned} \tag{3.5}$$

and

$$\Lambda(u) = \frac{e^{-2ipu}(c_0 + c_1 e^{iu} + \cdots + c_{2n}e^{nipu})}{(sin(u)cos(u))^p} =$$

Now, (3.5) implies that if $r = 1$ then $c_{2k+1} = 0$ and if $r = -1$ then $c_{2k} = 0$. We now state the following claim:

3.6. Claim. For some $p, u_j \in \mathbf{C}$,

$$\Lambda(u) = P(sin(u)cos(u))^{-p}\Pi_{j=0}^{l} sin(u - u_j)$$

where $l = 2p$ if $r = 1$ and $l = 2p - 1$ if $r = -1$.

Proof. Let

$$p(u) = c_0 + c_1 e^{iu} + \cdots + c_{2n}e^{nipu} = (c_0 + c_1\omega_2 + \ldots + c_{2n}\omega_{2n})$$

for $\omega = e^{iu}$. For $r = 1$ above we have

$$\begin{aligned} q(\omega) &= a\prod_{j=0}^{n}(\omega - \omega_j)(\omega - \omega_j) \\ &= a\prod_{j=0}^{n}(\omega^2 - \omega_j^2) \\ &= a\prod_{j=0}^{n}(e^{2iu} - e^{2iu_j} \\ &= \tilde{a}\prod_{j=0}^{n} e^{iu}\frac{(e^{i(u-u_j)} - e^{-i(u-u_j)})}{2i}e^{iu_j} \end{aligned}$$

for some u_j.

Thus we have

$$\Lambda(u)\Lambda(u+\frac{\pi}{2}) = (2i\cot(u))^n + (-2i\tan(u))^n r$$

We now plug the above into the claim. The left and side becomes

$$\rho^2(\sin(u)\cos(u))^{-n}\prod_{j=0}^{l}(-1)^p \sin(u-u-j)\cos(u-u_j)$$

and the right hand side becomes

$$(-1)^p 2^{2p}[\frac{\cos(u)^n}{\sin(u)^n} + \frac{\sin(u)^n}{\cos(u)^n}]$$

Canceling between both sides we get

$$\rho^2\prod_{j=0}^{l}(-1)^p \sin(u-u-j)\cos(u-u_j) = 2^{2p}[\cos(u)^{2n} + \sin(u)^{2n}p] \tag{3.6}$$

Now, set $z = e^{2iu}$ and $z_j = e^{2iu_j}$ so that

$$\sin(u-u_j)\cos(u-u_j) = \frac{z^{-1}(z^2-z_j^2)z_j^{-1}}{4i}$$

Then the Equation (3.6) becomes

$$z^{-l}\rho^2\prod_{j=0}^{p}\frac{z^2-z_j^2}{z_j}(iu)^l = \frac{z^{l-2p}}{2^{2p}}[(z+1)^{4p} - (z-1)^{4p}]$$

This determines z_j^2 and ρ^2 exactly. So we get

$$z_j^2 = -\tan^2\theta_j/2$$

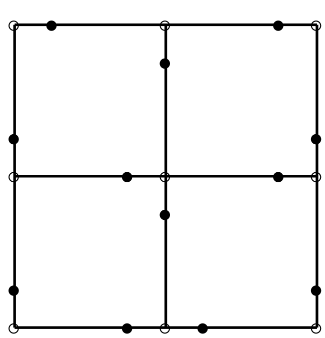
and

$$\theta_j = \begin{cases} \frac{\pi(j-1/2)}{2p} & \text{r=1} \\ j\frac{\pi}{2p} & \text{r=-1} \end{cases}$$

This then implies that $\phi_j = \frac{1}{2}\ln(\tan(\frac{\theta_j}{2}))$ and so that $\mu_j = \pm\frac{\pi}{4} - i\phi_j$. $\square$

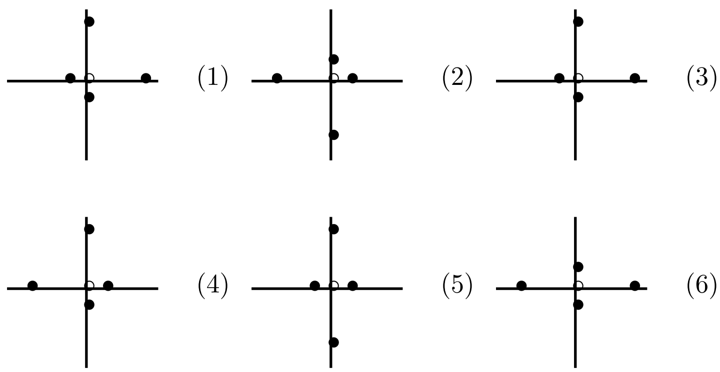
4. Ice Type Models

We will base our 2-dim Ice Type Model on a square lattice of the type pictured.



The physical model that we have in mind for this mathematical models is the structure of H_2O ice. Each vertex is occupied by (fixed) oxygen atoms and each edge is an ion of hydrogen (H^+) which is closer to one of the two endpoints. The "Ice rule" (by Slater) states that each for each of the four ions surrounding each atom, two of them are closer that atom and two are closer to the respective neighboring atoms. This is called the Electric Neutrality.

We have the following six configurations of H^+ ions:



The partition function for this model is

$$Z = \sum_{\text{orientation on edges}} exp(\frac{-\epsilon}{K_\beta T})$$

for $\epsilon = n_1\epsilon_1 + \cdots + n_6\epsilon_6$ where

$$\begin{aligned} n_i &= \text{number of atoms in configuration } i \\ \epsilon_i &= \text{energy of configuration } i \end{aligned}$$

Another name for the above model is the Six Vertex Model.

4.1. Choices of Configurations. There are two main choices of configuration we will consider:

1) **The** $2d$ **Ice Model** is given by the assumption $\epsilon_1 = \cdots = \epsilon_6$ which implies $\epsilon_i = 0$. Then all states have the same 0 energy and Z is simply the number of states.

2) **The Ferroelectric Model.** For T small enough all the dipoles point in the same direction(Slater) and we have $\epsilon_1 = \epsilon_2 = 0$ and $\epsilon_3 = \ldots = \epsilon_6 > 0$. In this case, the ground state consists of either all of the configurations of type (2) or type (3) above.

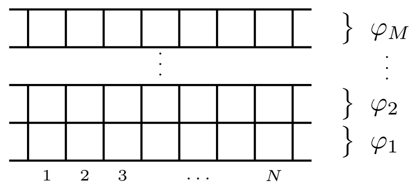


FIGURE 2. Row Model for Ice Type Models

3) **The Antiferroelectric Model.** Similar to the Ferroelectric Model, we assume that $\epsilon_1 = \cdots = \epsilon_4 > 0$ and that $\epsilon_5 = \epsilon_6 = 0$. In the ground states only (5) and (6) occur. i.e there exists only two such states.

Assumption: $\epsilon_1 = \epsilon_2$; $\epsilon_3 = \epsilon_4$; $\epsilon_5 = \epsilon_6$;
Comment: ϵ is invariant if we reverse all the configurations in zero external field.
Remark: Under periodic boundary conditions we can always assume $\epsilon_5 = \epsilon_6$

5. Transfer Matrix

In Figure (2) let φ_r be the state of row $r \in \{\text{up}, \text{down}\}^N$. Then

$$V(\varphi, \varphi') = \sum \exp(-(m_1\epsilon_1 + \ldots + m_6\epsilon_6)/k_\beta T$$

where m_i is the number of atoms in state i and the sum is taken over configurations of horizontal lines. In addition we have periodic boundry conditions:

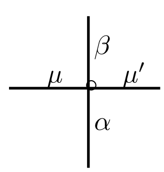
$$Z = \sum_{\varphi_1 \cdots \varphi_M} V(\varphi_1, \varphi_2)V(\varphi_2, \varphi_3)\ldots V(\varphi_M, \varphi_1) = \mathrm{Tr} V^M \sim \Lambda_{\max}^M$$

Let $\alpha \in \{\pm 1\}$ be the vector of spins associated to φ such that $\alpha_i = +1$ if the H^+ ion is close to the top atom and $\alpha_i = -1$ if the ion is close the bottom atom. Then

$$V_{\alpha\beta} := V(\varphi, \varphi') = \sum_{\mu_1, \ldots, \mu_N} \omega(\mu_1, \alpha_1|\beta_1, \mu_2)\ldots\omega(\mu_M, \alpha_M|\beta_M, \mu_1)$$

where $\omega(\mu, \alpha|\beta, \mu') = \exp(-\epsilon_i/k_\beta T)$ is the Boltzman weight of

and

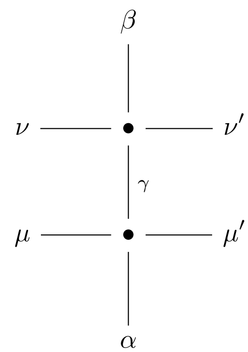


$$\begin{aligned} a :=& \omega(++\,|\,++) = \omega(--\,|\,--) \\ b :=& \omega(+-\,|\,-+) = \omega(-+\,|\,+-) \\ c :=& \omega(+-\,|\,+-) = \omega(-+\,|\,-+) \end{aligned}$$

Let $V = V(a, b, c)$. We are going to look for another transfer matrix $V' = V(a', b', c')$ such that V and V' commute, ie such that $VV' = V'V$. Now,

$$\begin{aligned}(VV')_{\alpha\beta} &= \Sigma_{\mu_1\ldots\mu_N}\Sigma_{\gamma_1\ldots\gamma_N}\Sigma_\gamma\omega(\mu_1, \alpha_1|\gamma_1, \mu_2)\omega(\mu_2, \alpha_2|\gamma_2, \mu_3)\ldots\\ &\quad \omega(\mu_{N-1}, \alpha_N|\gamma_N, \mu_1)\omega'(\nu_1, \gamma_1|\beta_1, \gamma_2)\omega'(\nu_2, \gamma_2|\beta_2, \nu_3)\\ &\quad \ldots\omega(\nu_N, \gamma_N|\beta_N, \nu_1)\\ &= \Sigma_{\mu,\nu\in\mathbf{Z}_2^N}\prod_{i=1}^N s(\mu_i, \nu_i|\mu_{i+1}, \nu_{i+1}|\alpha_i, \beta_i)\end{aligned}$$

where $s(\mu, \nu|\mu', \nu'|\alpha, \beta) = \Sigma_\gamma\omega(\mu\alpha|\gamma\mu')\omega'(\nu\gamma|\beta\nu)$ corresponds to the diagram



Let $s(\alpha, \beta)$ be the matrix with entries $s(\mu\nu|\mu'\nu'|\alpha\beta)$, (ie $\in$End$(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Then

$$(VV')_{\alpha\beta} = \mathrm{Tr}_{\mathbb{C}^2\otimes\mathbb{C}^2}\left(s(\alpha_1, \beta_1)\ldots s(\alpha_N, \beta_N)\right)$$

and similarly, if $s' = s$ with $a' \mapsto a$, $b' \mapsto b$, $c' \mapsto c$

$$(VV')_{\alpha\beta} = \mathrm{Tr}_{\mathbb{C}^2\otimes\mathbb{C}^2}\left(s'(\alpha_1, \beta_1)\ldots s'(\alpha_N, \beta_N)\right)$$

5.1. Ansatz. If there exists $M \in \mathbb{C}^2 \otimes \mathbb{C}^2$ such that $s(\alpha, \beta) = Ms'(\alpha, \beta)M^{-1}$ then secretly $VV' = V'V$.

6. Star-Triangle Relation

Let $M := \omega''(\nu, \mu|\nu', \mu'')$, then the Ansatz reads, in the form $s(\alpha, \beta)M = Ms'(\alpha, \beta)$:

$$\begin{aligned}&\sum_{\gamma,\mu'',\nu''}\omega(\mu, \alpha|\gamma, \mu'')\omega'(\nu, \gamma|\beta, \nu'')\omega''(\nu'', \mu''|\nu', \mu')\\ =&\sum_{\gamma,\mu'',\nu''}\omega''(\nu, \mu|\nu'', \mu'')\omega(\mu'', \alpha|\gamma, \mu')\omega'(\nu'', \gamma|\beta, \nu')\end{aligned}$$

For all $\mu, \nu, \mu'\nu', \alpha, \beta$. See Figure (3):

The unknowns in this equation are $a', b', c'; a'', b'', c''$, $3 - 1 + 3 - 1 = 4$. There will be one equation for each of the $\alpha, \beta, \mu, \mu', \nu, \nu'$ so there would seem to be 2^6 equations. But the Ice Rule gives tell us that $\omega(\mu, \alpha|\beta, \nu) = 0$ unless $\mu + \alpha = \beta + \nu$ so the left hand side of the Star-Triangle relation is 0 unless

$$\mu + \alpha + \nu = \beta + \nu' + \mu'$$

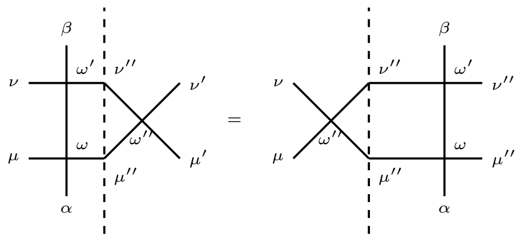


FIGURE 3. Star-Triangle Relation

or the following all hold.

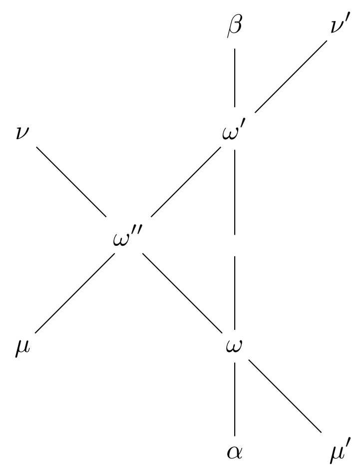
$$\mu + \alpha = \gamma + \mu''$$
$$\gamma + \nu = \beta + \nu''$$
$$\nu'' + \mu'' = \nu' + \mu'$$

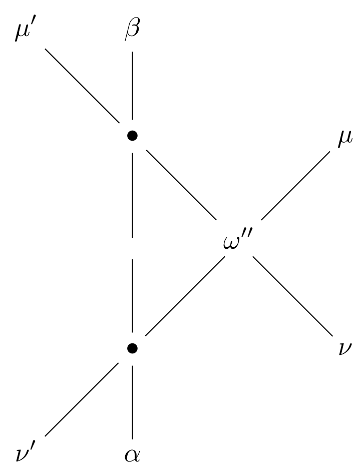
$\Rightarrow \mu + \alpha + \nu = \beta + \nu' + \mu'$

We can brutally enumerate the possibilities to get the following list:

μ	α	ν	β	ν'	μ'
1	1	1	1	1	1
1	1	-1	1	1	-1
			1	-1	1
			-1	1	1
1	-1	1			
-1	1	1			
-1	-1	1			
-1	1	-1			
1	-1	-1			
-1	-1	-1	-1	-1	-1

So there are $20 = 1 + 9 + 9 + 1$ possibilities. There is also symmetry under flipping in any direction so that cuts the number of equations in half to 10. By interchanging $\alpha \leftrightarrow \beta$, $\mu \leftrightarrow \nu'$ and $\nu \leftrightarrow \mu'$ we see that $2^3/2$ equations are trivially satisfied:





The remaining six come in three pairs:

μ	α	ν	μ'	β	ν'
1	1	-1	1	-1	1
1	1	-1	-1	1	1
1	-1	1	-1	1	1

The actual equations are:

(1)

$$\begin{tikzcd} ac'a'' \arrow[r, equal] & bc'b'' + ca'c'' \\ \omega'(-11|-11) \arrow[u] & \omega''(11|11) \arrow[ul] \end{tikzcd}$$

(2) $ab'c'' = ba'c'' + cc'b''$

(3) $cb'a'' = ca'b'' + bc'c''$

Now, we can eliminate c'', b'', a'' out of the above equations to get

$$\frac{a^2+b^2-c^2}{ab} = \frac{(a')^2+(b')^2-(c')^2}{a'b'}$$

and so we see that V and V' commute if

$$2\Delta = \frac{a^2+b^2-c^2}{ab} = \frac{(a')^2+(b')^2-(c')^2}{a'b'} = 2\Delta' \tag{6.1}$$

7. Parameterizing Solutions of Eq (6.1)

To parametrize solutions of $\frac{a^2+b^2-c^2}{ab} = 2\Delta$ we will assume that Δ and a are constants. Then

$$\frac{1+(b/a)^2+(c/a)^2}{b/a} = 2\Delta$$

Set $x = b/c$ and $y = c/a$. Then we ca rewrite the above as $1 + x^2 + y^2 = 2x\Delta$ so

$$y = (1 + x^2 - 2x\Delta)^{1/2}$$

so $c = a(\dots)^{1/2}$. Now, we want to parameterize expansions of the form $f(x) = [(x-x_1)(x-x_2)]^{1/2}$ so we'll substitute

$$t = \left(\frac{x-x_1}{x-x_2}\right)^2 \qquad \Rightarrow \qquad t^2 = \frac{x-x_1}{x-x_2}$$

and therefore

$$x = \frac{x_1 - t^2 x_2}{1-t^2}$$

So we can write

$$F(x) = t(x-x_2) = \frac{t(x_1-x_2)}{1-t^2} = F(t)$$

In our case, $x_1x_2 = 1$ and $\Delta = \frac{1}{2}(x_1+x_2) = \frac{x_1+x_1^{-1}}{2}$ so we have the equations

$$a = \quad a \tag{7.1}$$

$$b = \quad ax = \tfrac{a(x_1-t^2x_1^{-1})}{1-t^2} \tag{7.2}$$

$$c = \quad a\tfrac{(x_1-x_1^{-1})t}{1-t^2} \tag{7.3}$$

By removing denominators we get

$$a = \quad \rho'(1-t^2)x_1 \tag{7.4}$$

$$b = \quad \rho'(x_1^2-t^2) \tag{7.5}$$

$$c = \quad \rho'(x_1^2-1)t \tag{7.6}$$

And finally by making one last change of variables we get:

$$x_1 = \quad -\exp(-\lambda) \tag{7.7}$$

$$t = \quad \exp(\tfrac{1}{2}(v-\lambda)) \tag{7.8}$$

$$\rho' = \quad \tfrac{1}{2}\rho t^{-1}x_1^{-1} \tag{7.9}$$

These finally imply that

$$a = \rho\sinh(\tfrac{\lambda - v}{2}) \tag{7.10}$$
$$b = \rho\sinh(\tfrac{\lambda + v}{2}) \tag{7.11}$$
$$c = \rho\sinh(\lambda) \tag{7.12}$$

In summary we've made the following changes of variables: $V = V(a, b, c) = V(\rho', t, x) = V(\rho, v, \lambda) \triangleq V(v)$. Now, $V(v)$ is holomorphic/entire function of v and $V(v)$ and $V(u)$ commute for any $u, v \in \mathbf{C}$. This implies that simultaneus eigenvalues of these matrices will be holomorphic functions in v.

Now $\Delta'' = \Delta(a'', b'', c'') = \Delta = \Delta'$ so we can can parametrize a'', b'', c'' as above for the same λ for some v'' and ρ''. Plug these into equations you get the following constraints:

$$\sinh\left(\frac{\lambda + v - v' + v''}{4}\right) = 0$$

This implies that $v' = \lambda + v + v''$. Lets write $u = \frac{\lambda + v}{2}$ (and similarly for u' and u'') then

$$a = \rho\sinh(\lambda - u) \tag{7.13}$$
$$b = \rho\sinh(u) \tag{7.14}$$
$$c = \rho\sinh(\lambda) \tag{7.15}$$

So $v' = \lambda + v + v'$ becomes $u' = u + u''$ and we have the following relations

$$R(u)R(u')R(u'') = R(u'')R(u')R(u)$$

which becomes

$$R(u)R(u + u'')R(u'') = R(u'')R(u + u'')R(u) \tag{7.16}$$

These are the Yang Baxter Equations written as braid relations on R.

8. Yang-Baxter on R Implies Commutativity of V

Now, recall that

$$V = \sum_{\mu_1, \ldots, \mu_N} \omega(\mu_1, \alpha_1|\beta_1, \mu_2) \ldots \omega(\mu_N, \alpha_N|\beta_N, \mu_1)$$

where $V \in$End$((\mathbb{C}^2)^{\otimes N})$, $V = \mathrm{Tr}_{\mathbb{C}_0^2}(R_{01} \ldots R_{0N})$ where $R_{0i} \in$End$(\mathbb{C}_0^2 \otimes C_i^2)$ and have matrix entries

$$R_{\mu\alpha}^{\mu\beta} = \omega(\mu\alpha|\mu'\beta)$$

Then the partition function over M rows is given by

$$Z_M = \mathrm{Tr}_{\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}(V^M)$$

We want to show that the Yang-Baxter Equation for R implies the commutativity of V. We will write (7.16) in a more standard form:

$$R_{12}(\lambda)R_{13}(\mu)R_{23}(\nu) = R_{23}(\nu)R_{13}(\mu)R_{12}(\lambda) \tag{8.1}$$

Claim. If (8.1) holds then the RTT-Relasions

$$R_{0\bar{0}}(\lambda)T_0(\mu)T_{\bar{0}}(\nu) = T_{\bar{0}}(\nu)T_0(\mu)R_{0\bar{0}}(\lambda)$$

also hold, where

- The identity occurs in $\mathrm{End}(\mathbb{C}_0^2 \otimes \mathbb{C}_{\bar{0}}^2 \otimes (\mathbb{C}^2)^{\otimes N})$
- T_0 stands for $R_{01}R_{02} \dots R_{0N}$ is the monodromy matrix.
- $T_{\bar{0}}$ similarly stands for $R_{\bar{0}1}R_{\bar{0}2} \dots R_{\bar{0}N}$

8.1. Corollary. The RTT-Relasion implies the commutation relation $V(\mu)V(\nu) = V(\nu)V(\mu)$.

Proof. First, we'll rewrite the RTT as

$$R_{0\bar{0}}(\lambda)T_0(\mu)T_{\bar{0}}(\nu)R_{0\bar{0}}^{-1}(\lambda) = T_{\bar{0}}(\nu)T_0(\mu)$$

so that both sides act on the tensor product of two spaces. Then taking the trace we see that the $R_{0\bar{0}}(\lambda)$ and the $R_{0\bar{0}}^{-1}(\lambda)$ terms cancel and we get

$$V(\mu)V(\nu) = V(\nu)V(\mu)$$

□

Proof. (of Claim) We will write out the left hand side of (8.1) and use the fact that $R_{0i}(\mu)$ and $R_{\bar{0}i}(\nu)$ commute to move each element $R_{0i}(\mu)$ next to its pair $R_{\bar{0}i}(\nu)$. We then use the Yang-Baxter Equation to commute $R_{\bar{0}0}(\lambda)$ past the pair and finally use the commutation to reorder the equation into the proper ordering. The LHS is:

$$\begin{aligned}
&R_{0\bar{0}}(\lambda)R_{01}(\mu)\dots R_{0N}(\mu)R_{\bar{0}1}(\nu)\dots R_{0\bar{N}}(\nu)\\
&\qquad = R_{0\bar{0}}(\lambda)R_{01}(\mu)R_{\bar{0}1}(\nu)\dots R_{0N}(\mu)R_{\bar{0}2}(\nu)\dots R_{\bar{0}N}(\nu)\\
&\qquad = R_{\bar{0}1}(\nu)R_{01}(\mu)R_{\bar{0}0}(\lambda)\dots R_{0N}(\mu)R_{\bar{0}2}(\nu)\dots R_{\bar{0}N}(\nu)\\
&\qquad = \dots = R_{\bar{0}1}(\nu)R_{01}(\mu)R_{\bar{0}2}(\nu)R_{02}(\mu)\dots R_{\bar{0}N}(\nu)R_{0N}(\mu)R_{\bar{0}0}(\lambda)\\
&\qquad = R_{\bar{0}1}(\nu)R_{\bar{0}2}(\nu)\dots R_{\bar{0}N}(\nu)R_{01}(\mu)R_{02}(\mu)\dots R_{0N}(\mu)R_{\bar{0}0}(\lambda)
\end{aligned}$$

□

Recall from the square Ising Model that we could find eigenvalues with functional relations between the eigenvalues the commuting transfer matrices. We now need functional relations of the eigenvalues of $V(u)$. Instead of postulating functional relations between eigenvalues of $V(u)$ we will postulate functional relations of V and see how it descends to the eigenvalues:

8.2. Baxter Ansatz. There exists commuting matrices $Q(v)$ depending holomorphically on v such that

(1) $Q(v)$ is invertible for at least one $v_0 \in \mathbb{C}$.
(2) $V(u)Q(v) = Q(v)V(u)$ for all v and u.
(3) $Q(v)$ commutes with the spin operator

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

(4) The entries of $Q(v)$ are of the form

$$\sum_{-\mu<r<\mu} t' \exp(rV/2)$$

(5) Setting $\phi(v) = \rho^N \sinh^N(V/2)$, where ρ is the parameter from before, and $\lambda' = \lambda - i\pi$, we have

$$V(v)Q(v) = \phi(\lambda - v)Q(u + 2\lambda') + \phi(\lambda + v)Q(v - 2\lambda')$$

Recall that

$$\Delta = -\cosh(x) = \frac{a^2 + b^2 - c^2}{2ab} = \frac{x_1 + x_1^{-1}}{2}$$

for $x_1 = e^{-\lambda}$ and λ fixed. So then VQ is a linear combination of two shifted versions of Q. This then is a second order differential equation in Q and comes from the observation that in the model the eigenvalues actually do satisfy these requirements.

Now, to go from (5) to the eigenvalues of V we diagonalize $V(v)$ and $Q(v)$ simultaneously, ie see if there exists $P \in$GL$((\mathbb{C}^2)^{\otimes N})$ such that $P^{-1}V(v))P$ and $P^{-1}Q(u)P$ are diagonal. Now we can check that

$$\begin{aligned} V(v)^T =& V(-v) \\ Q(u)^T =& Q(-v) \end{aligned}$$

so these are unitary automorphisms and so are simultaneously diagonalizable.

Let $\Lambda(v)$ be an eigenvalue of $V(v)$ and $q(v)$ is the corresponding eigenvalue of $Q(v)$. Now,

$$\Lambda(v)q(v) = \phi(\lambda - v)q(v + 2\lambda') + \phi(\lambda + v)q(v - 2\lambda')$$

by the Ansatz (called Baxters TQ Relation). The assumptions on the entries of $Q(v)$ give us

$$q(v) = c\prod_{\ell=1}^{n} \sinh\left(\frac{v - v_\ell}{2}\right)$$

For some $n \leq N$ and some $v_1, \ldots, v_\ell$. Now,

$$\Lambda(v) = \frac{\phi(\lambda - v)q(v + 2\lambda') + \phi(\lambda + v)q(v - 2\lambda')}{q(v)}$$

and since $\Lambda(v)$ is entire the numerator must vanish on any $v_1, \ldots, v_N$. This gives us the relation that for $j = 1, \ldots, N$

$$\frac{\phi(\lambda - v_j)}{\phi(\lambda + v_j)} = \frac{q(v_j - 2\lambda')}{q(v_j + 2\lambda')} \tag{8.2}$$

We will call this equation the Ansatz Equation.

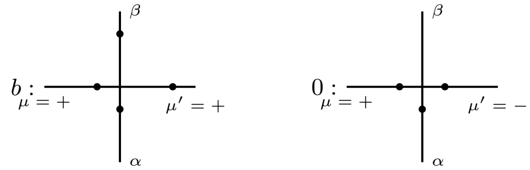


FIGURE 4. Diagrams associated to the weight in the matrices $G_i(+)$

9. The Construction of Baxters Q-Matrix

Let g be a column of Q. We will assume that g is a pure tensor:

$$g = g_1 \otimes g_2 \otimes \ldots \otimes g_N$$

where $g_i \in \mathbb{C}^2$. Then

$$Vg(v) = \mathrm{Tr}_{\mathbb{C}_0^2}(R_{01}g_1, \ldots, R_{0N}g_n)$$

where $R_{0i}g_i \in$End$(\mathbb{C}_0^2) \otimes \mathbb{C}^2$. Note that the factor of $\mathbb{C}^2$ corresponds to one complex dimension of spin up, on of spin down. Define $G_i := R_{0i}g_i$ and let $\alpha = \pm 1$ so that $G_i(\alpha) \in$End$(\mathbb{C}_0^2)$. Then

$$G_i(\alpha)_{\mu\mu'} = \sum_\beta \omega(\mu\alpha|\beta\mu')g_i(\beta)$$

so

$$G_i(+) = \begin{array}{cc} & \begin{array}{cc} \mu'{=}+ & \mu'{=}- \end{array} \\ \begin{array}{c} \mu{=}+ \\ \mu{=}- \end{array} & \begin{pmatrix} ag_i(+) & 0 \\ cg_i(-) & bg_i(+) \end{pmatrix} \end{array}$$

and

$$G_i(-) = \begin{array}{cc} & \begin{array}{cc} \mu'{=}+ & \mu'{=}- \end{array} \\ \begin{array}{c} \mu{=}+ \\ \mu{=}- \end{array} & \begin{pmatrix} bg_i(-) & cg_i(+) \\ 0 & ag_i(+) \end{pmatrix} \end{array}$$

where, we have drawn the diagrams corresponding to the weights a and to the 0 in Figure 4. Note that the 0's occurs because the configuration isn't allowable.

Now,

$$[Vg]_\alpha = \mathrm{Tr}_{\mathbb{C}_0^2}(G_1(\alpha_1) \ldots G_N(\alpha_n)) \tag{9.1}$$

Ansatz. Assume now that we can, by a suitable change of basis, make all the G_i upper or lower triangular. Ie there exists P_i for $i = 1, \ldots, N$ such that $G_i(\alpha_i) = P_i H_i(\alpha) P_{i+1}^{-1}$ where the $H_i(\alpha)$ are upper triangular. If so, then

$$[Vg]_\alpha = \mathrm{Tr}_{\mathbb{C}_0^2}(H_1(\alpha_1) \ldots H_N(\alpha_n)) = g_1'(\alpha_1) \ldots g_N'(\alpha_N) + g_1''(\alpha_1) \ldots g_N''(\alpha_N)$$

Now, to solve Equation (9.1). By the Ansatz

$$G_i(\alpha_i)P_{i+1} = P_i \begin{pmatrix} g_i'(\alpha_i) & * \\ 0 & g_i'(\alpha_i) \end{pmatrix}$$

where $P_{i+1} = [p_{i+1}|q_{i+1}]$. Then

$$G_i P_i = g_i'(\alpha_i) \begin{pmatrix} p_i^1 \\ p_i^2 \end{pmatrix}$$

and we can write $G_i(\alpha_i)P_{i+1} = g_i'(\alpha_i)P_i$, where $\alpha = \pm 1$. In coordinates this reads

$$\sum_{\beta,\mu'} \omega(\mu\alpha|\beta\mu')g_i(\beta)P_{i+1}(\mu') = g_i'(\alpha)P_i(\mu)$$

for all $\alpha, \mu \in \{\pm 1\}$. There are $2^2 = 4$ possibilities and we will simply list them:

μ	α	
+	+	$ag_i(+)P_{i+1}(+) = g_i'(+)P_i(+)$
+	-	$bg_i(-)P_{i+1}(+) + cg_i(+)P_{i+1}(-) = g_i'(-)P_i(+)$
-	+	$cg_i(-)P_{i+1}(+) + bg_i(+)P_{i+1}(-) = g_i'(+)P_i(-)$
-	-	$ag_i(-)P_{i+1}(-) = g_i'(-)P_i(-)$

The unknowns in the above equations are the (g_i)'s, (g_i')'s, and the p_i's. Now, we can use the equations corresponding to $(++)$ and $(--)$ to eliminate $g_i'(\pm)$. The second and third equations then become

$$\begin{aligned} bg_i(-)P_{i+1}(+)P_i(-) + cg_i(+)P_{i+1}(-)P_i(-) =& ag_i(-)P_{i+1}(-)P_i(+) \\ cg_i(-)P_{i+1}(+)P_i(+) + bg_i(+)P_{i+1}(-)P_i(+) =& ag_i(+)P_{i+1}(+)P_i(-) \end{aligned}$$

Set $r_i = P_i(-)/P_i(+)$, then we have

$$\begin{aligned} bg_i(-)r_i + cg_i(+)r_{i+1}r_i =& ag_i(-)r_{i+1} \\ cg_i(-) + bg_i(+)r_{i+1} =& ag_i(+)r_i \end{aligned}$$

We can now eliminate $g_i(-)$:

$$\begin{aligned} g_i(-) &= \frac{ar_i - br_{i+1}}{c} \cdot g_i(+) \\ &= \frac{cr_ir_{i+1}}{ar_{i+1} - br_i} \cdot g_i(+) \end{aligned}$$

This implies that

$$(ar_i - br_{i+1})(ar_{i+1} - br_i) = c^2 r_i r_{i+1}$$

Dividing both sides by r_ir_{i+1} we have

$$e^{-\lambda} + e^{\lambda} = 2\Delta = \frac{a^2 + b^2 - c^2}{ab} = \frac{r_i}{r_{i+1}} + \frac{r_{i+1}}{r_i}$$

so

$$r_{i+1} = -r_i e^{\sigma_i \lambda}$$

for $\sigma_i \in \{\pm 1\}$ and so finally

$$r_i = (-1)^i r \exp((\sigma_1 + \ldots + \sigma_{i-1})\lambda)$$

where r is arbitrary and $\sum \sigma_i = 0$ so as to have periodicity, note that this implies that N is even.

Now, fix $P_i(+) := 1$ so that $P_i(-) = r_i = (-1)^i \exp[\lambda\sigma_1, \ldots, \sigma_{i-1}]$. We also have the freedom now to define $g_i(+) := 1$ which in turn implies that $g_i(-) =$

$r_i \exp((\lambda+\nu)\sigma_i/2)$. Recall as well that

$$\begin{aligned} a =& \rho \sinh \frac{\lambda-\nu}{2} \\ b =& \rho \sinh \frac{\lambda+\nu}{2} \\ c =& \rho \sinh \lambda \\ g_i'(+) =& a \\ g_i'(-) =& -a r_i \exp \frac{3\lambda+\nu}{2}\sigma_i \end{aligned}$$

We can compute $g_i''(+)$ by taking the determinant of $G_i P_{i+1} = \dots$ from before and setting $\det(P_{i+1}) = 1$. The result is

$$\begin{aligned} g_i''(+) &= b \\ g_i''(-) &= -b r_i \exp \frac{\lambda-\nu}{2}\sigma_i \end{aligned}$$

So

$$g_i = \begin{pmatrix} 1 \\ r_i \exp(\frac{\lambda\nu}{2}\sigma_i \end{pmatrix} =: h_i(\nu)$$

and

$$g_i' = a \begin{pmatrix} 1 \\ -r_i \exp(\frac{3\lambda+\nu}{2}\sigma_i \end{pmatrix} = a$$

We claim that g_i' is the as before but shifted so

$$g_i' = a h_i(\nu_1 + 2\lambda')$$

Similarly,

$$g_i'' = h \begin{pmatrix} 1 \\ r_i \exp(\frac{\lambda\nu}{2}\sigma_i) \end{pmatrix} = h h_i(\nu - 2\lambda')$$

So

$$[Vg]_\alpha = g_1'(\alpha) \dots g_N'(\alpha) + g_1''(\alpha) \dots g_N''(\alpha)$$

and so for $y(v) = h_1(v) \otimes h_N(v) \in (\mathbb{C}^2)^{\otimes N}$

$$V_{(v)} y(v) = a^n y(v + 2\lambda') + b^N y(v - 2\lambda') = \phi(\lambda - v) + \phi(\lambda + v) \tag{9.2}$$

So this is a vector solution to the Baxters TQ relation.

Now, recall that $a = \rho \sinh((\lambda - v)/2$, $b = \rho \sinh((\lambda + v)/2$ and $c = \rho \sinh \lambda$. Thus, $v \to -v$ exchanges the elements a and b. Now, $V(b, a) = V(a, b)^T$, in fact recall that V has entries

$$V_{\alpha\beta} = \sum_{\mu_1 \dots \mu_N} \omega(\mu_1\alpha_1|\beta_1\mu_2) \dots \omega(\mu_N\alpha_N|\beta_N\mu_1)$$

so

$$V^T_{\alpha\beta} = \sum_{\mu_1 \dots \mu_N} \omega(\mu_1\beta_1|\alpha_1\mu_2) \dots \omega(\mu_N\beta_N|\alpha_N\mu_1)$$

9.1. Remark. The basic fact that we're using here is that $\omega(\mu\beta|\alpha\mu') = \omega(-\mu\alpha|\beta-\mu')$ where the negation is required to preserve

$$\omega(+-\,|\,+-) = \omega(-+\,|\,-+) = c$$
$$\omega(++\,|\,--) = \omega(--\,|\,++) = 0$$

but then negating μ and μ' takes $a = \omega(++\,|\,--)$ to $\omega(-+\,|\,+-) = b$ as required. Thus, indeed $V(b,a) = V(a,b)^T$ and so $V(-v) = V(v)^T$.

Let Q_R be a 2^N by 2^N matrix whose whose columns are linear combinations of the vector $y(v)$ above and let $Q_L(v) = Q_R(-v)^T$. Then it follows from (9.2) that

$$V(v)Q_R(v) = \phi(\lambda - v)Q_R(v + 2\lambda') + \phi(\lambda.+)Q_R(v - 2\lambda') \tag{9.3}$$

By switching the sign of u, v this gives us.

$$V(-v)Q_R(-v) = \phi(\lambda + v)Q_R(-v + 2\lambda') + \phi(\lambda - v)Q_R(-v - 2\lambda')$$
$$\Downarrow$$
$$Q_L(v)V(v) = \phi(\lambda.v)Q_L(v + 2\lambda') + \phi(\lambda.c)Q_L(v - 2\lambda')$$

Next, we want to claim that for any u, v we have

$$Q_L(u)Q_R(v) = Q_L(v)Q_R(u)$$

We will call a typical column of $Q_r(u)$ $y(u|r,\sigma)$ where $\sigma = (\sigma_1, \ldots, \sigma_N)$, $\sigma_i \in \{\pm 1\}$ subject to the condition that $\sigma_1 + \ldots + \sigma_N = 0$. We then need to show that

$$y(-u|r',\sigma')^T y(v|r,\sigma) = y(-v|r',\sigma')^T y(u|r,\sigma)$$

Now, the left hand side is

$$\prod_{i=1}^{m}(1 + r_i r_i' \exp[\frac{1}{2}(\lambda - u)\sigma_i' + \frac{1}{2}(\lambda + v)\sigma_i]$$
$$= \prod_{i=1}^{m}(1 + r_i r_i' \exp[\lambda(\sigma_1 + \ldots + \sigma_{i-1} + \sigma_1' + \ldots \sigma_{i-1}') + \frac{1}{2}(\lambda - u)\sigma_i' + \frac{1}{2}(\lambda + v)\sigma_i]$$

so following Baxter we will call the above $J(u,v|\sigma_1,\ldots,\sigma_N)$ and look at

$$J(u,v|\ldots,\sigma_{j+1},\sigma_j,\ldots)/J(u,v|\ldots,\sigma_j,\sigma_{j+1},\ldots)$$
$$= \frac{1 + r_i r_i' \exp[\lambda(\sigma_1 + \ldots + \sigma_{j-1} + \sigma_1' + \ldots \sigma_{j-1}') + \frac{1}{2}(\lambda - u)\sigma_j' + \frac{1}{2}(\lambda + v)\sigma_j]}{1 + r_i r_i' \exp[\lambda(\sigma_1 + \ldots + \sigma_{j-1} + \sigma_1' + \ldots \sigma_{j-1}') + \frac{1}{2}(\lambda - u)\sigma_j' + \frac{1}{2}(\lambda + v)\sigma_{j+1}]}$$
$$\times \frac{1 + r_i r_i' \exp[\lambda(\sigma_1 + \ldots + \sigma_j + \sigma_1' + \ldots \sigma_j') + \frac{1}{2}(\lambda - u)\sigma_{j+1}' + \frac{1}{2}(\lambda + v)\sigma_{j+1}]}{1 + r_i r_i' \exp[\lambda(\sigma_1 + \ldots + \sigma_j + \sigma_1' + \ldots \sigma_j') + \frac{1}{2}(\lambda - u)\sigma_{j+1}' + \frac{1}{2}(\lambda + v)\sigma_j]}$$

We want to show that this is symmetric in u and v. First, note that we can assume $\sigma_{j+1} = -\sigma_j$ since otherwise the result is obvious. Then there are only $2^3 = 8$ cases to check which are left as an exercise. Thus $Q_L(u)Q_R(v) = Q_L(v)Q_R(u)$ for any u, v.

9.2. $Q(u)Q(v) = Q(v)Q(u)$**.** To show that $Q(u)$ and $Q(v)$ commute for any u and v we will use the above and the fact that $Q_R(v_0)$ is invertible for some (and hence a generator) v_0. Set

$$Q(v) := Q_R(u)Q_R^{-1}(v_0) = Q_L^{-1}(v_0)Q_L(v)$$

Now recall that from Equation (9.3) we have

$$V(v)Q_R(v) = \phi(\lambda - v)Q_R(v + 2\lambda') + \phi(\lambda + v)Q_R(v2\lambda')$$

Multiplying on the right by $Q_R^{-1}(v_0)$ we have

$$V(v)Q(v) = \phi(\lambda - v)Q(v + 2\lambda') + \phi(\lambda + v)Q(v2\lambda')$$

but we also have, by transposing Equation (9.3) and letting $v \mapsto -v$ we get

$$Q_L(v)V(v) = \phi(\lambda - v)Q_L(v + 2\lambda') + \phi(\lambda + v)Q_L(v2\lambda')$$

so

$$V(v)Q(v) = Q(v)V(v) = \phi(\lambda - v)Q(v + 2\lambda') + \phi(\lambda + v)Q(v2\lambda')$$

And so V and Q commute. Finally,

$$\begin{aligned} Q(u)Q(v) =& Q_L(v_0)^{-1}Q_L(u)Q_R(v)Q_R(v_0)^{-1} \\ =& Q_L(v_0)^{-1}Q_L(v)Q_R(u)Q_R(v_0)^{-1} \\ =& Q(v)Q(u) \end{aligned}$$

Thus $Q(u)$ and $Q(v)$ commute and $V(v)$ and $Q(v)$ commute.

10. Spin Operators

We just have to check now that Q commutes with the spin operators. Let

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{\otimes N}$$

Then we can check that

$$Sy(v) = y(v + 2\pi i)$$

which implies that $SQ_R(v) = Q_r(v + 2\pi i)$ so

$$Q_L(v)S = Q_R^T(-v)S = Q_r^T(-v + 2\pi i) = Q_r^T(-v - 2\pi i) = Q_L(v + 2\pi i)$$

Where the 4'th equality is by the $4\pi i$ periodicity. Thus

$$SQ(v) = Q(v + 2\pi i) = Q(v)S$$

Also, $SV(v) = V(v)S$ since

$$\begin{aligned} (SV(v)S^{-1})_{\alpha\beta} =& \sum_{\mu_1\dots\mu_N} \alpha_1\omega(\mu_1\alpha_1|\beta_1\mu_2)\beta_1\alpha_2\dots\alpha_N\omega(\mu_N\alpha_N|\beta_N\mu_1) \\ =& \sum_{\mu_1\dots\mu_N} \mu_1\omega(\mu_1\alpha_1|\beta_1\mu_2)\mu_2\mu_2\dots\mu_N\omega(\mu_N\alpha_N|\beta_N\mu_1)\mu_1 \\ =& \sum_{\mu_1\dots\mu_N} \omega(\mu_1\alpha_1|\beta_1\mu_2)\dots\omega(\mu_N\alpha_N|\beta_N\mu_1) \\ =& (V(v))_{\alpha\beta} \end{aligned}$$

since that fact that $\mu_i\alpha_i\beta_i\mu_{i+1} = 1$ means that $\mu_i\mu_{i+1} = \alpha_i\beta_i$ and so we can make the change in the third line.

11. Consequences of Baxters Ansatz

It follows from the construction of Q that

$$SQ(v) = Q(v)S = Q(v + 2\pi i)$$

so $S^2Q(v) = Q(v)$ by the $4\pi i$ periodicity of $Q(v)$, so Q is a function of $e^{v/2}$. Moreover, if $Q_\pm$ are the diagonal blocks of Q corresponding to the ± 1 eigenvalues of S then

$$Q_\pm(v + 2\pi i) = \pm Q_\pm(v)$$

So Q_+ is actually a function of e^v and Q_- is a a function of $e^{v/2}$ not involving even powers. From the construction of Q is also follows that Q grows at most as fast as $\exp(Nv/2)$. So Property 6 can be replaced by these conditions.

Further consiquences:

$$q(v) = \sum_{|r|<N} d_r \exp((r)v/2) = e^{-\frac{Nv}{2}} \sum_{|r|<N} d_r \exp((r + N)v/2)$$

will be a typical entry of $Q(v)$. We can factor it as

$$c \cdot e^{-\frac{Nv}{2}}(e^{\frac{v}{2}} - z_1)\dots(e^{\frac{v}{2}} - z_n)$$

so that

$$q(v) = \sum_{|r|<N-1} d_r x^r$$

is a Laurent Polynomial in $x = e^{\frac{v}{2}}$ which is either even (when r's have even powers) or odd (when r's have even powers). We can factor it as above to get

$$q(v) = e^{-(N-1)} \sum_{|r|<N-1} d_r x^{r+N-1} = e^{-(N-1)}(x-x_1)\dots(x-x_{n'}) = cx^{-n}(x-x_1)\dots(x-x_{2n})$$

for some $n \leq 2(N-1)$ and some constant c. Thus,

$$\begin{aligned}
q(v) =& cx^{-n}(x - x_1)(x + x_1)\dots(x - x_n)(x + x_n) \\
=& c\prod_{i=1}^{n} \frac{x^2 - x_i^2}{x} \\
=& c\prod_{i=1}^{n}(x_i) \prod_{i=1}^{n} \frac{x}{x_i} - \frac{x_i}{x} \\
=& c\prod_{i=1}^{n}(x_i) 2^n \prod_{i=1}^{n} \sinh\left(\frac{v - v_i}{2}\right)
\end{aligned}$$

where $x = \exp\left(\frac{v}{2}\right)$ and $x_i =: \exp\left(\frac{v_i}{2}\right)$ so indeed q is proportional to

$$\prod_{i=1}^{n} \sinh\left(\frac{v - v_i}{2}\right)$$

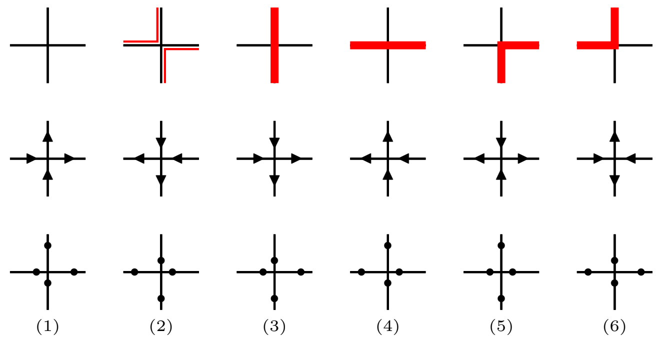


FIGURE 5. The six dot configurations at a vertex and the corresponding arrow and line configurations.

for some $v_i \in \mathbb{C}$ defined up to $4\pi i$. Thus if $\Lambda(v)$ and $q(v)$ are a pair of joint eigenvalues of $T(v)$ and $Q(v)$ respectively then to summarize

$$q(v) = \prod_{i=1}^{n} \sinh\left(\frac{v - v_i}{2}\right) \tag{11.1}$$

$$\Lambda(v) = \frac{\phi(\lambda - v)q(v + 2\lambda') + \phi(\lambda + v)q(v - 2\lambda')}{q(v)} \tag{11.2}$$

12. The Six Vertex Model Continued

13. Line Configurations

We will now replace the dot configurations above with arrow on line configurations as shown in Figure (1). Then the lines move upward or to the right (if one starts moving up along a line) and never move down or to the left. Now, a line might be able to not move upwards if a row is purely of type (4) but due to the periodic boundary conditions if a line has started moving up it will continue to do so. Moreover, because of the vertical boundary it will continue moving upward.

In conclusion, each non-horizontal line meats every row of vertical edges one because it has to continue moving up and only one because it cannot go down. In addition the number of lines in a given row of vertical edges is constant in any state.

13.1. Corollary. $V_{\varphi\varphi'} = 0$ unless the number of down arrows of φ is the same as the number of down arrows of φ'.

Call the number of down arrows n. Our aim now is to find eigenvectors of V among the states with "quantum number" n. We will specify such states by $1 \leq x_1 < \ldots < x_n \leq N$ where x_i is the position of the i'th down arrow and let $X = \{x_1, \ldots, x_n\}$. Then

$$g(X) = \uparrow \otimes \ldots \otimes \uparrow \otimes \underset{x_i}{\downarrow} \otimes \uparrow \otimes \ldots \in (\mathbb{C}^2)^{\otimes N}$$

Note, n corresponds to the weight of $g(X)$ under the ${}_2$ action on $(\mathbb{C}^2)^{\otimes N}$ so the horizontal line configuration corresponds to the the fact that R is of 0 weight. Now,

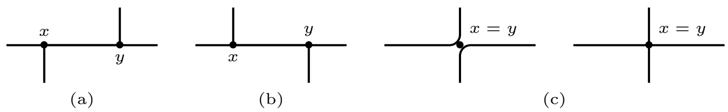


FIGURE 6. Diagrams of line configurations in the case $n = 1$

the eigenvalues $Vg = \Lambda g$, where $g = \{g(X)\}$ are of the form

$$\Lambda g(X) = \sum_Y V(X, Y) g(X)$$

where

$$V(X, Y) = \sum a^{m_1+m_2} b^{m_3+m_4} c^{m_5+m_6}$$

where X is a configuration of lines on one row of vertical edges, Y is the configuration of the row above and the sum is taken over all allowed configurations of lines on the intervening horizontal edges.

For example, note that the horizontal lines are the intervening configurations of $X = Y = \emptyset$ so

$$V(\emptyset, \emptyset) = a^N + b^N = \Lambda$$

13.2. Case $n = 1$. If we have one and only one down arrow then $g(X) =: g(x)$ where $X = \{x\}$ and we want to solve $\Lambda g(x) = \sum_y V(x, y) g(y)$.

- $x < y$ We only have one possible configuration, shown in Figure (2.a). This gives us

$$V(x, y) = cb^{y-x-1} ca^{N-(y-x+1)}$$

where the number of vertices between x and y is denoted $y - x + 1$.
- $x > y$ Again, we only have one possible configuration, shown in Figure (2.b). This gives us

$$V(x, y) = cb^{N+y-x-1} ca^{x-y-1)}$$

where the number of vertices between x and y is denoted $N - (x - y + 1)$.
- $x = y$ We have two possibilities here, shown in Figure (2.c). This gives us

$$V(x, y) = b^{N-1} a + a^{N-1} b$$

Counting for all the various cases, the eigenvalue equation is

$$g(x) = a^{N-1} b g(x) + \sum_{y=x+1}^{N} c^2 b^{N+y-x-1} a^{x-y-1)} g(y) + ab^{N-1} g(x) + \sum_{y=1}^{x-1} c^2 b^{y-x-1} a^{N-(y-x+1)} g(y) \tag{13.1}$$

Bethe Ansatz. Assume that $g(x) = z^x$ for some z as of yet undetermined. Then

$$\begin{aligned}\sum_{y=x+1}^{N} c^2 b^{N+y-x-1} a^{x-y-1)} z^y =& c^2 a^{N-2} z^{x+1)} \sum_{y=x+1}^{N} \left(\frac{bx}{a}\right)^{y-x-1} \\ =& c^2 a^{N-2} z^{x+1)} \frac{1-\left(\frac{bx}{a}\right)^{N-x}}{1-\frac{bx}{a}} \\ =& \frac{c^2 a^{N-2} z^{x+1)}}{a^2 - abz} - \frac{a^{x-1} b^{N-x} c^2 z^{N+1}}{a - bz}\end{aligned}$$

Similarly,

$$sum_{y=1}^{x-1} c^2 b^{y-x-1} a^{N-(y-x+1)} z^y = -\frac{c^2 a b^{N-1} z^{x)}}{a^2 - abz} + \frac{a^{x-1} b^{N-x} c^2 z}{a - bz}$$

so adding these we get for the right hand side of (1.1):

$$\begin{aligned}& a^{N-1} b z^x + \frac{c^2 a^{N-2} z^{x+1)}}{a^2 - abz} - \frac{a^{x-1} b^{N-x} c^2 z}{a - bz} (z^N - 1) + a b^{N-1} z^x - \frac{c^2 a b^{N-1} z^{x)}}{a^2 - abz} \\ =& a^N \left(\frac{ab + (c^2 - b^2) z}{a^2 - abz}\right) z^x - \frac{a^{x-1} b^{N-x} c^2 z}{a - bz} (z^N - 1) + b^N \left(\frac{a^2 - c^2 - abz}{ba - b^2 z}\right) z^x\end{aligned}$$

we can eliminate the unwanted term (the coefficient of $(z^N - 1)$ by simply setting $z^N = 1$. Then

$$\Lambda = a^N \left(\frac{ab + (c^2 - b^2) z}{a^2 - abz}\right) z^x + b^N \left(\frac{a^2 - c^2 - abz}{ba - b^2 z}\right) z^x$$

All the eigenvalues corresponding to the roots of $z^N = 1$. We would need to do some additional work in the general case.

13.3. Proposition. The preferred eigenvector of V for $n = 1$ is the vector $g(x) = 1$ for all $x = 1, \ldots, N$ and have a corresponding preferred eigenvalues

$$\Lambda = a^N \left(\frac{ab + (c^2 - b^2)}{a^2 - ab}\right) z^x + b^N \left(\frac{a^2 - c^2 - ab}{ba - b^2}\right)$$

13.4. Case $n = 2$**.** Now, $g(X) = g(x_1, x_2)$ for $1 \leq x_1 < x_2 \leq N$ and $V(X, Y) = 0$ unless either $x_1 \leq y_1 \leq x_2 \leq y_2$ or $y_1 \leq x_1 \leq y_2 \leq x_2$.

The contribution for the first type is

$$a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) E(x_2, y_2) e^{N - y_2}$$

where

$$E(x, y) = \begin{cases} cb^{y-x-1} & \text{if } x < y \\ b/c & \text{if } y = x \end{cases}$$

and

$$D(y, x) = \begin{cases} ca^{x-y-1} & \text{if } y < x \\ a/c & \text{if } y = x \end{cases}$$

Now, consider the following case:

- $x_1 < y_1$: In this case we have one vertex of type (5) for x_1 and $y_1 - x_1 - 1$ vertices of type (4); for y_1 we could have a vertex of type (2) or (6). Then $E(x_1, y_1) = cb^{y_1 - x_1 - 1}$.
- $x = y$: In this case we have one vertex of type (1) so we have $E(x_1, y_1) = b/c$.
- $y_1 < x_2$: Here, if $x_1 = x_2$ we've already counted y_1. If on the other hand if $x_1 < y_1$ then we add these contributions $x_1 < y_1$: $D(y_1, x_1) = ca^{x_2 - y_1 - 1}$.
- $y_1 = x_2$: In this case we have the partition $x_1 < y_1 = x_2 < y_2$ which corresponds to a vertex of type (2). Then $D(x, y) = a/c$
- $x_2 < y_2$:Similar to the above, we have $E(x_2, y_2) = cb^{y_2 - x_2 - 1}$
- $x_2 = y_2$:Similar to previous two, $E(x_2, y_2) = b/c$.

Thus the eigenvalue equation is

$$\begin{aligned}\Lambda g(x_1, x_2) = &\sum_{x_1 \le y_1 \le x_2} \sum_{x_2 \le y_2 \le N}^{*} a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) E(x_2, y_2) ca^{N - y_2} g(y_1, y_2) \\ &+ \sum_{1 \le y_1 \le x_1} \sum_{x_1 \le y_2 \le x_2}^{*} b^{y_1 - 1} E(y_1, x_1) D(x_1, y_2) E(y_2, x_2) cb^{N - x_2} g(y_1, y_2)\end{aligned}$$

where the star indicates that we're adding the condition that $y_1 < y_2$.

13.5. Ansatz for the Eigenvector. We will assume g takes the form $g(x_1, x_2) = A_{12} z_1^{x_1} z_2^{x_2} A_{12}$ where A_{12} is constant.

Use the ansatz, the first double sum becomes

$$\begin{aligned}&\sum_{x_1 \le y_1 \le x_2} \sum_{x_2 \le y_2 \le N}^{*} a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) E(x_2, y_2) ca^{N - y_2} z_1^{y_1} z_2^{y_2} \\ = &\sum_{x_1 \le y_1 \le x_2} a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) cz_1^{y_1} \sum_{x_2 \le y_2 \le N} E(x_2, y_2) a^{N - y_2} z_2^{y_2}\end{aligned}$$

The second summand here can be rewritten as follows:

$$\begin{aligned}&= \frac{b}{c} a^{N - x_2} z_2^{x_z} + \sum_{y_2 = x_2 + 1}^{N} cb^{y_2 - x_2 - 1} a^{N - y_2} z_2^{y_2} \\ &= \frac{b}{c} a^{N - x_2} z_2^{x_z} + ca^{N - x_2 - 1} z_2^{x_2 + 1} \sum_{y_2 = x_2 + 1}^{N} \left(\frac{bz_2}{a} \right)^{y_2 - x_2 - 1} \\ &= \frac{b}{c} a^{N - x_2} z_2^{x_z} + ca^{N - x_2 - 1} z_2^{x_2 + 1} \frac{1 - \left(\frac{bz_2}{a} \right)^{N - x_2}}{1 - \frac{bz_2}{a}}\end{aligned}$$

The first sum then is the sum of the $y_1 = x_1$ part, the $y_1 = x_2$ part and standard part.

$$\begin{aligned}&a^{x_1 - 1} ba^{x_2 - x_1 - 1} cz_1^{x_1} + a^{x_1 - 1} cb^{x_2 - x_1 - 1} z_2^{x_2} \\ &+ \sum_{y_1 = x_1 + 1}^{x_2 - 1} \left(\frac{bz}{a} \right)^{y_1 - x_1 - 1} = \qquad a^{x_2 - 3c^3}\end{aligned}$$

By Baxter, the result will eventually be

$$\begin{aligned}
&A_{12}\{a^{x_1}(L_1 a^{x_2-x_1} z_1^{x_1} + M_1 b^{x_2-x_1} z_1^{x_2})\\
&\times(L_2 a^{N-x_2} z_2^{x_2} + P_2 b^{N-x_2} z_2^{N})\\
&-a^{N+x_1-x_2-1} b^{x_2-x_1}(z_1 z_2)^{x_2}\}
\end{aligned}$$

Where

$$L_i = L(z_i) := \frac{ab + (c^2 - b^2) z_i}{a^2 - abz_i}$$

$$M_i = M(z_i) := \frac{ab + c^2 - abz_i}{a^2 - abz_i}$$

$$P_i = P(z_i) := \frac{c^2 z_i}{a^2 - abz_i}$$

Similarly, the second sum is

$$\begin{aligned}
A_{12}\{(a^{x_1} P_1 + M_1 b^{x_1} z_1^{x_1}) \times (L_2 a^{x_2-x_1} z_2^{x_1} + M_2 b^{x_2-x_1} z_2^{x_2}) b^{N-x_2}\\
-a^{x_2-x_1} b^{N+x_1-x_2}(z_1 z_2)^{x_1}\}
\end{aligned}$$

CHAPTER 8

The Six Vertex Model Continued

1. Line Configurations

We will now replace the dot configurations above with arrow on line configurations as shown in Figure (1). Then the lines move upward or to the right (if one starts moving up along a line) and never move down or to the left. Now, a line might be able to not move upwards if a row is purely of type (4) but due to the periodic boundary conditions if a line has started moving up it will continue to do so. Moreover, because of the vertical boundary it will continue moving upward.

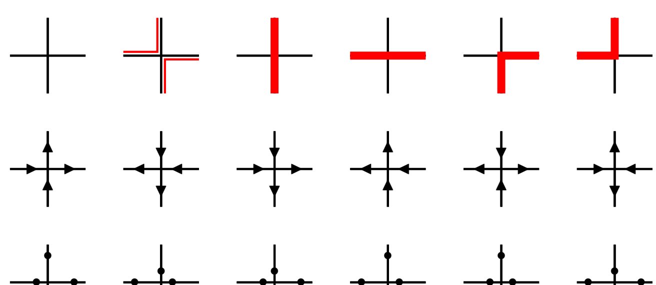
In conclusion, each non-horizontal line meats every row of vertical edges one because it has to continue moving up and only one because it cannot go down. In addition the number of lines in a given row of vertical edges is constant in any state.

1.1. Corollary. $V_{\varphi\varphi'} = 0$ unless the number of down arrows of φ is the same as the number of down arrows of φ'.

Call the number of down arrows n. Our aim now is to find eigenvectors of V among the states with "quantum number" n. We will specify such states by $1 \leq x_1 < \ldots < x_n \leq N$ where x_i is the position of the i'th down arrow and let $X = \{x_1, \ldots, x_n\}$. Then

$$g(X) = \uparrow \otimes \ldots \otimes \uparrow \otimes \underset{x_i}{\downarrow} \otimes \uparrow \otimes \ldots \in (\mathbb{C}^2)^{\otimes N}$$

Note, n corresponds to the weight of $g(X)$ under the $_2$ action on $(\mathbb{C}^2)^{\otimes N}$ so the horizontal line configuration corresponds to the the fact that R is of 0 weight. Now,



(1) (2) (3) (4) (5) (6)

Figure 1. The six dot configurations at a vertex and the corresponding arrow and line configurations.

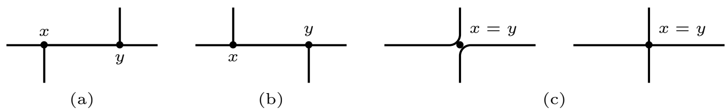


FIGURE 2. Diagrams of line configurations in the case $n = 1$

the eigenvalues $Vg = \Lambda g$, where $g = \{g(X)\}$ are of the form

$$\Lambda g(X) = \sum_{Y} V(X,Y)g(X)$$

where

$$V(X,Y) = \sum a^{m_1+m_2} b^{m_3+m_4} c^{m_5+m_6}$$

where X is a configuration of lines on one row of vertical edges, Y is the configuration of the row above and the sum is taken over all allowed configurations of lines on the intervening horizontal edges.

For example, note that the horizontal lines are the intervening configurations of $X = Y = \emptyset$ so

$$V(\emptyset, \emptyset) = a^N + b^N = \Lambda$$

1.2. Case $n = 1$**.** If we have one and only one down arrow then $g(X) =: g(x)$ where $X = \{x\}$ and we want to solve $\Lambda g(x) = \sum_y V(x,y)g(y)$.

- $x < y$ We only have one possible configuration, shown in Figure (2.a). This gives us

$$V(x,y) = cb^{y-x-1}ca^{N-(y-x+1)}$$

  where the number of vertices between x and y is denoted $y - x + 1$.
- $x > y$ Again, we only have one possible configuration, shown in Figure (2.b). This gives us

$$V(x,y) = cb^{N+y-x-1}ca^{x-y-1)}$$

  where the number of vertices between x and y is denoted $N - (x - y + 1)$.
- $x = y$ We have two possibilities here, shown in Figure (2.c). This gives us

$$V(x,y) = b^{N-1}a + a^{N-1}b$$

Counting for all the various cases, the eigenvalue equation is

(1.1)

$$g(x) = a^{N-1}bg(x) + \sum_{y=x+1}^{N} c^2 b^{N+y-x-1} a^{x-y-1)} g(y) + ab^{N-1}g(x) + \sum_{y=1}^{x-1} c^2 b^{y-x-1} a^{N-(y-x+1)} g(y)$$

Bethe Ansatz. Assume that $g(x) = z^x$ for some z as of yet undetermined. Then

$$\begin{aligned}\sum_{y=x+1}^{N} c^2 b^{N+y-x-1} a^{x-y-1)} z^y =& c^2 a^{N-2} z^{x+1)} \sum_{y=x+1}^{N} \left(\frac{bx}{a}\right)^{y-x-1} \\ =& c^2 a^{N-2} z^{x+1)} \frac{1-\left(\frac{bx}{a}\right)^{N-x}}{1-\frac{bx}{a}} \\ =& \frac{c^2 a^{N-2} z^{x+1)}}{a^2-abz} - \frac{a^{x-1} b^{N-x} c^2 z^{N+1}}{a-bz}\end{aligned}$$

Similarly,

$$sum_{y=1}^{x-1} c^2 b^{y-x-1} a^{N-(y-x+1)} z^y = -\frac{c^2 ab^{N-1} z^{x)}}{a^2-abz} + \frac{a^{x-1} b^{N-x} c^2 z}{a-bz}$$

so adding these we get for the right hand side of (1.1):

$$\begin{aligned}&a^{N-1} bz^x + \frac{c^2 a^{N-2} z^{x+1)}}{a^2-abz} - \frac{a^{x-1} b^{N-x} c^2 z}{a-bz}(z^N-1) + ab^{N-1} z^x - \frac{c^2 ab^{N-1} z^{x)}}{a^2-abz} \\ =& a^N \left(\frac{ab+(c^2-b^2)z}{a^2-abz}\right) z^x - \frac{a^{x-1} b^{N-x} c^2 z}{a-bz}(z^N-1) + b^N \left(\frac{a^2-c^2-abz}{ba-b^2 z}\right) z^x\end{aligned}$$

we can eliminate the unwanted term (the coefficient of (z^N-1) by simply setting $z^N = 1$. Then

$$\Lambda = a^N \left(\frac{ab+(c^2-b^2)z}{a^2-abz}\right) z^x + b^N \left(\frac{a^2-c^2-abz}{ba-b^2 z}\right) z^x$$

All the eigenvalues corresponding to the roots of $z^N = 1$. We would need to do some additional work in the general case.

1.3. Proposition. The preferred eigenvector of V for $n = 1$ is the vector $g(x) = 1$ for all $x = 1, \ldots, N$ and have a corresponding preferred eigenvalues

$$\Lambda = a^N \left(\frac{ab+(c^2-b^2)}{a^2-ab}\right) z^x + b^N \left(\frac{a^2-c^2-ab}{ba-b^2}\right)$$

2. Case $n = 2$

Now, $g(X) = g(x_1, x_2)$ for $1 \leq x_1 < x_2 \leq N$ and $V(X, Y) = 0$ unless either $x_1 \leq y_1 \leq x_2 \leq y_2$ or $y_1 \leq x_1 \leq y_2 \leq x_2$.

The contribution for the first type is

$$a^{x_1-1} E(x_1, y_1) D(y_1, x_2) E(x_2, y_2) e^{N-y_2}$$

where

$$E(x,y) = \begin{cases} cb^{y-x-1} & \text{if } x < y \\ b/c & \text{if } y = x \end{cases}$$

and

$$D(y,x) = \begin{cases} ca^{x-y-1} & \text{if } y < x \\ a/c & \text{if } y = x \end{cases}$$

Now, consider the following case:

- $x_1 < y_1$: In this case we have one vertex of type (5) for x_1 and $y_1 - x_1 - 1$ vertices of type (4); for y_1 we could have a vertex of type (2) or (6). Then $E(x_1, y_1) = cb^{y_1 - x_1 - 1}$.
- $x = y$: In this case we have one vertex of type (1) so we have $E(x_1, y_1) = b/c$.
- $y_1 < x_2$: Here, if $x_1 = x_2$ we've already counted y_1. If on the other hand if $x_1 < y_1$ then we add these contributions $x_1 < y_1$: $D(y_1, x_1) = ca^{x_2 - y_1 - 1}$.
- $y_1 = x_2$: In this case we have the partition $x_1 < y_1 = x_2 < y_2$ which corresponds to a vertex of type (2). Then $D(x, y) = a/c$
- $x_2 < y_2$:Similar to the above, we have $E(x_2, y_2) = cb^{y_2 - x_2 - 1}$
- $x_2 = y_2$:Similar to previous two, $E(x_2, y_2) = b/c$.

Thus the eigenvalue equation is

$$\begin{aligned}\Lambda g(x_1, x_2) = &\sum_{x_1 \le y_1 \le x_2} \sum^{*}_{x_2 \le y_2 \le N} a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) E(x_2, y_2) c a^{N - y_2} g(y_1, y_2) \\ &+ \sum_{1 \le y_1 \le x_1} \sum^{*}_{x_1 \le y_2 \le x_2} b^{y_1 - 1} E(y_1, x_1) D(x_1, y_2) E(y_2, x_2) c b^{N - x_2} g(y_1, y_2)\end{aligned}$$

where the star indicates that we're adding the condition that $y_1 < y_2$.

2.1. Ansatz for the Eigenvector. We will assume g takes the form $g(x_1, x_2) = A_{12} z_1^{x_1} z_2^{x_2} A_{12}$ where A_{12} is constant.

Use the ansatz, the first double sum becomes

$$\begin{aligned}&\sum_{x_1 \le y_1 \le x_2} \sum^{*}_{x_2 \le y_2 \le N} a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) E(x_2, y_2) c a^{N - y_2} z_1^{y_1} z_2^{y_2} \\ = &\sum_{x_1 \le y_1 \le x_2} a^{x_1 - 1} E(x_1, y_1) D(y_1, x_2) c z_1^{y_1} \sum_{x_2 \le y_2 \le N} E(x_2, y_2) a^{N - y_2} z_2^{y_2}\end{aligned}$$

The second summand here can be rewritten as follows:

$$\begin{aligned}&= \frac{b}{c} a^{N - x_2} z_2^{x_z} + \sum_{y_2 = x_2 + 1}^{N} c b^{y_2 - x_2 - 1} a^{N - y_2} z_2^{y_2} \\ &= \frac{b}{c} a^{N - x_2} z_2^{x_z} + c a^{N - x_2 - 1} z_2^{x_2 + 1} \sum_{y_2 = x_2 + 1}^{N} \left(\frac{b z_2}{a} \right)^{y_2 - x_2 - 1} \\ &= \frac{b}{c} a^{N - x_2} z_2^{x_z} + c a^{N - x_2 - 1} z_2^{x_2 + 1} \frac{1 - \left(\frac{b z_2}{a} \right)^{N - x_2}}{1 - \frac{b z_2}{a}}\end{aligned}$$

The first sum then is the sum of the $y_1 = x_1$ part, the $y_1 = x_2$ part and standard part.

$$a^{x_1-1}ba^{x_2-x_1-1}cz_1^{x_1} + a^{x_1-1}cb^{x_2-x_1-1}z_2^{x_2} + \underbrace{\sum_{y_1=x_1+1}^{x_2-1}\left(\frac{bz}{a}\right)^{y_1-x_1-1}}_{=a^{x_2-3c^3}}$$

By Baxter, the result will eventually be

$$A_{12}\{a^{x_1}(L_1a^{x_2-x_1}z_1^{x_1} + M_1b^{x_2-x_1}z_1^{x_2}) \times (L_2a^{N-x_2}z_2^{x_2} + P_2b^{N-x_2}z_2^{N}) \\ -a^{N+x_1-x_2-1}b^{x_2-x_1}(z_1z_2)^{x_2}\}$$

Where

$$L_i = L(z_i) := \frac{ab + (c^2 - b^2)z_i}{a^2 - abz_i}$$

$$M_i = M(z_i) := \frac{ab + c^2 - abz_i}{a^2 - abz_i}$$

$$P_i = P(z_i) := \frac{c^2z_i}{a^2 - abz_i}$$

Similarly, the second sum is

$$`A_{12}\{(a^{x_1}P_1 + M_1b^{x_1}z_1^{x_1}) \times (L_2a^{x_2-x_1}z_2^{x_1} + M_2b^{x_2-x_1}z_2^{x_2})b^{N-x_2} \\ -a^{x_2-x_1}b^{N+x_1-x_2}(z_1z_2)^{x_1}\}$$

Now, opening these sums we get three types of terms: wanted terms, unwanted internal terms and unwanted boundary terms. We will deal with them in order.

- Wanted Terms The wanted terms are those proportional to $g(x_1, x_2)$ ie the terms

$$A_{12}\left(a^NL_1L_2 + b^NM_1M_2\right)z_1^{x_1}z_2^{x_2}$$

  Assuming we can cancel out the remaining terms, these will give rise to an eigenvector with eigenvalue

$$\Lambda = \Lambda(z_1, z_2) = a^NL_1L_2 + b^NM_1M_2$$

  where z_1 and z_2 have yet to be determined.

- Unwanted Internal Terms These are terms that contain $(z_1z_2)^{x_1}$ or $(z_1z_2)^{x_2}$ and include in particular (but not exclusively) those coming from the subtracted terms above, ie

$$A_{12}\left(a^{N+x_1-x_2}b^{x_2-x_1}(M_1L_2-1)(z_1z_2)^{x_2}\right) + A_{12}\left(a^{x_2-x_1}b^{N+x_1-x_2}(M_1L_2-1)(z_1z_2)^{x_1}\right)$$

Exercise. Show that $M_1L_2 - 1 = -c^2s_{12}/[(a - bz_1)(a - bz_2)]$ where

$$s_{12} = 1 - 2\Delta z_2 + z_1z_2$$

$$\Delta = \frac{a^2 + b^2 - c^2}{2ab}$$

- Unwanted Boundary Terms Let $R_j(x,x') := L_j a^{x'-x} z_j^x + M_j b^{x'-x} z_j^{x'}$. These terms are then
$$-A_{12}a^{x_1}b^{N-x_2}R_1(x_1,x_2)P_2z_2$$
and
$$-A_{12}c^{x_1}b^{N-x_2}R_2(x_1,x_2)P_2$$

3. Elimination of Unwanted Terms

3.1. Elimination of Unwanted Internal Terms. We will try to eliminate the internal terms of the form $\sum_r A_{12}^{(r)} z_{1,r}^{x_1} z_{2,r}^{x_2}$. We will first require that Λ defined as above is independent of r. Our second requirement will be that we can fins z_1' and z_2' such that $z_1z_2 = z_1'z_2'$ and $\Lambda = \Lambda'$.

By eliminating z_2' we are left with a quadratic equation so the solution must be either $z_i = z_i'$ or $z_i' = z_{3-i}$. Recall now the ansatz:
$$g(x_1,x_2) = A_{12}z_1^{x_1}z_2^{x_2} + A_{21}z_2^{x_1}z_1^{x_2}$$
Then the unwanted terms will cancel if
$$(M_1L_2-1)A_{12} + (M_2L_1-1)A_{21} = 0$$
But using the formula from the exercise this simplifies to
$$s_{12}A_{12} + s_{21}A_{21} = 0$$
but these are equations in A_{12} and A_{21} in terms of z_1 and z_2 so the solution is
$$\frac{A_{12}}{A_{21}} = -\frac{s_{21}}{s_{12}}$$

3.2. Elimination of Unwanted Boundary Terms. We can rewrite the contribution from the boundary terms as
$$-a^{x_1}b^{N-x_2}\left\{P_2R_1(x_1,x_2)(z_2^NA_{12}-A_{21}) - P_1R_2(x_1,x_2)(z_1^NA_{21}-A_{12})\right\}$$
So these terms vanish if
$$z_1^N = \frac{A_{12}}{A_{21}} = -\frac{s_{21}}{s_{12}}$$
$$z_2^N = \frac{A_{21}}{A_{12}} = -\frac{s_{12}}{s_{21}}$$

Or, explicitly
$$z_1^N = \frac{1-2\Delta z_1 + z_1z_2}{1-2\Delta z_2 + z_1z_2}$$
$$z_2^N = \frac{1-2\Delta z_2 + z_1z_2}{1-2\Delta z_1 + z_1z_2}$$

Before moving on, its interesting to observe that if we multiply these two equations together we get $(z_1z_2)^N = 1$ so $z_1z_2 = \kappa$ is an N'th root of unity of 1. This then implies that
$$g(x_1+1,x_2+1) = z_1z_2g(x_1,x_2) = \kappa g(x_1,x_2)$$
Now, by the PF $\kappa = 1$ since otherwise the entries (?) could be all of $\mathbb{R}_+^*$. Thus $z_1z_2 = 1$ and

$$\begin{aligned} g(x_1, x_2) =& A_{12} z_1^{x_1} z_2^{x_2} + A_{21} z_2^{x_1} z_1^{x_2} \\ =& A_{21} \left(\frac{A_{12}}{A_{21}} z_1^{x_1} z_2^{x_2} + z_2^{x_1} z_1^{x_2} \right) \\ =& A_{21} \left(z_1^{N+\lambda_1} z_2^{x_2} + z_2^{x_1} z_1^{x_2} \right) \end{aligned}$$

Set $z_1 = \exp(ik)$. Then

$$\begin{aligned} g(x_1, x_2) &= A_{21} \exp(iK(N + x_1 - x_2)) + \exp(-iK(x_1 - x_2)) \\ &= A_{21} \exp\left(\frac{iNk}{2}\right) 2 \cos\left(K\left(x_1 - x_2 + \frac{N}{2}\right)\right) \\ &\propto \cos\left(K\left(x_1 - x_2 + \frac{N}{2}\right)\right) \end{aligned}$$

Now, A_{21} only depends on z_1 and z_2 not on x_1 or x_2 and so is constant and can be dropped from the eigenvector. As $1 \leq x_1 \leq x_2 \leq N$ the quantity $x_1 - x_2 + N/2$ ranges over $[-1 + N/2, 1 - N/2] = [-r, r]$ where $r = N/2 - 1$. To make sure that $g(x_1, x_2)$ is real and positive it is sufficient to check that either $k \in [-\pi/2r, \pi/2r]$ or $k \in i\mathbb{R}^*$ is purely imaginary.

Now,

$$z_1^N = -\frac{1 - 2\Delta z_1 + 1}{1 - 2\Delta z^{-1} + 1} = -\frac{1 - \Delta z_1}{1 - \Delta z_1^{-1}}$$

But then $z_1^N - \Delta z_1^{N-1} = -1 + \Delta z_1$ and so

$$\begin{aligned} \Delta &= \frac{z_1^N + 1}{z_1^{N-1} + z_1} \\ &= \frac{e^{iNK} + 1}{e^{i(N-1)K} + e^{iK}} \\ &= \frac{\cos((r+1)K)}{\cos(rk)} \end{aligned}$$

This implies two things:

(1) If $\Delta < 1$ then the above equation has a unique real solution in $[0, \pi/2r]$ and no purely imaginary solutions.
(2) If $\Delta > 1$ then the above equation has no solutions in $[0, \pi/2r]$ and a single purely imaginary solution.

This solution (which depends on Δ) is the PF solution.

4. Case: General $n \geq 3$, $x \leq N$

Let $X = \{x_i\}$ be the vector of increasing points $0 \leq x_1 < \ldots < x_n \leq N$ and let $Y = \{y_1\}$ be the vector $0 \leq y_1 < \ldots < y_n \leq N$. Again, $V(X, Y) = 0$ unless X and Y interlace, ie

$$x_1 \leq y_1 \leq \ldots \leq x_n \leq y_n \tag{4.1}$$

or

$$y_1 \leq x_1 \leq \ldots \leq y_n \leq x_n \tag{4.2}$$

The corresponding configurations are then given in Figure 3

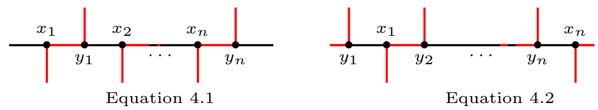


FIGURE 3. Diagrams of line configurations in the case of general n

And the eigenvector reads

$$\Lambda g(X) = \sum_{X \le Y}^{*} a^{x_1 - 1} E_{11} D_{12} E_{22} D_{23} \dots E_{nn} c a^{N - y_n} g(Y)$$
$$+ \sum_{Y \le X}^{*} b^{y_1 - 1} D_{11} E_{12} D_{22} E_{23} \dots D_{nn} c b^{N - x_n} g(Y)$$

where the $*$ indicates that $y_i = y_{i+1}$ is not allowed and where $D_{ij} = D(y_i, x_j)$ and $E_{ij} = E(x_i, y_j)$.

Now, the first sum gives

$A_{1\dots n}\{a^{x_1} R_1(x_1, x_2) \dots R_{n-1}(x_{n-1}, x_n) \times (L_n a^{N-x_n} z_n^{x_n} - P_n b^{N-x_n} z_n^N) - \text{terms for } y_1 = y_{i+1}\}$

and the second sum gives

$A_{1\dots n}\{(P_1 a^{x_1} + M_1 b^{x_1} z_1^{x_1}) R_2(x_1, x_2) \dots R_n(x_{n-1}, x_n) b^{N-x_2} - \text{terms for } y_1 = y_{i+1}\}$

Each R is the sum of 2 terms, so the above are really 2^{n-1} terms. Only one of these terms is "wanted":

$$\Lambda = a^N L_1 \dots L_n + b^N M_1 \dots M_n$$

aside from the boundary terms, other terms contain at least one of the following factors: $(z_j z_{j+1})^{x_{j+1}}$ (corresponding to $y_j = x_{j+1} = y_{j+1}$) or $(z_j z_{j+1})^{x_j}$). This leads to that following.

Bethe Ansatz.

$$g(x_1, \dots, x_n) = \sum_{\sigma \in S_n} A_\sigma z_{\sigma(1)}^{x_1} \dots z_{\sigma(n)}^{x_n}$$

Now, the unwanted terms cancel if

$$S_{\sigma(j)\sigma(j+1)} A_\sigma + S_{\sigma(j+1)\sigma(j)} A_{(jj+1)\sigma} = 0$$

The boundary terms contain one factor of P_j for some value of j so by replacing $z_1 \dots z_n$ by $z_2 \dots z_n, z_)1$ we can show that the boundary terms vanish if

$$-z_1^N A_{2\dots n1} + A_{1\dots n} = 0$$

and more generally that they vanish if

$$z_{\sigma(i)}^N = \frac{A_{\sigma(1)\dots\sigma(n)}}{A_{\sigma(2)\dots\sigma(n)\sigma(1)}} \tag{4.3}$$

The above are sufficient to get an eigenvector. Now let consider the implications.

$$S_{\sigma(j)\sigma(j+1)} A_\sigma + S_{\sigma(j+1)\sigma(j)} A_{(jj+1)\sigma} = 0$$

has the solution

$$A_\sigma = (-1)^\sigma \prod_{i<j} s_{\sigma(j)\sigma(i)}$$

so substituting into Eq. (4.3) we get

$$z^N_{\sigma(i)} = (-1)^{n-1} \prod_{\ell=2}^{n} \frac{s_{\sigma(\ell)\sigma(1)}}{s_{\sigma(1)\sigma(\ell)}}$$

which give the **Bethe Equations**

$$z^n_j = (-1)^{n-1} \prod_{\ell=2}^{n} \frac{s_{\ell,1}}{s_{1,\ell}}$$

5. The Maximum Eigenvalue

Consider the "trivial" case $\Delta > 1$. Then let Λ_n denote the maximum eigenvalue of V on the subspace with quantum number n. It was proved that $\Lambda_0 > \Lambda_i$ for $1 \le i < n$ and that $\Lambda_0 \ge \Lambda_n$. But this implies that

$$\Lambda_{\max} = \Lambda_0 = a^N + b^N$$

so $f = \min(\epsilon_1, \epsilon_2)$ and the most likely state of the system is one where all arrows point up ($n = 0$) or down ($n = N$) and is unique.

6. Thermodynamic Limit: $\Delta < 1$

We will proceed as we did in the case $n = 2$. If $z_j = \exp(iK_j)$ then

$$\frac{s_{\ell,j}}{s_{j,\ell}} = \frac{1 - 2\Delta e^{iK_j} + e^{i(K_j+K_\ell)}}{1 - 2\Delta e^{iK_\ell} + e^{i(K_j+K_\ell)}} =: \exp(-i\theta(K_j, K_\ell))$$

Note that if $z_j \neq 0$ then $s_{\ell,j}, s_{j,\ell} \neq 0$ and so their $s_{\ell,j}/s_{j,\ell} \in \mathbb{C}^\times$. Now, assuming that z_j lies on the unit circle (so K_j are real) we have

$$\begin{aligned}
(6.1)\qquad \exp(-i\theta(p,q)) &= \frac{1 - 2\Delta e^{ip} + e^{i(p+q)}}{1 - 2\Delta e^{iq} + e^{i(p+q)}} \\
&= \frac{(1 - 2\Delta e^{ip} + e^{i(p+q)})(1 - 2\Delta e^{-iq} + e^{-i(p+q)})}{(1 - 2\Delta e^{iq} + e^{i(p+q)})(1 - 2\Delta e^{-iq} + e^{-i(p+q)})} \\
&= \frac{2 + 2\cos(p+q) - 4\Delta(e^{ip} + e^{-iq}) + 4\Delta^2 e^{i(p-q)}}{2 + 4\Delta^2 + e^{-i(p+q)} - 2\Delta(e^{-ip} + e^{ip}) - 2\Delta(e^{-iq} + e^{iq})} \\
&= \frac{2 + 2\cos(p+q) - 2\Delta(e^{ip} + e^{-iq}) + 4\Delta^2 e^{i(p-q)}}{2 + 4\Delta^2 + 2\cos(p+q) - 4\Delta\cos(p) - 4\Delta\cos(q)}
\end{aligned}$$

Note, it's clear that this last sum is symmetric under the change $p \leftrightarrow q$ so x_j and $s_{\ell,j}$ are indeed on the unit circle.

Now, consider $\frac{z}{w}$ where $|z| = |w|$. Then

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} = \frac{a + ib}{w\bar{w}} = \rho e^{i\theta}$$

where $\rho = 1$ and $\theta = \tan^{-1} b/a$. Now, we will need the following relations:

$$\begin{aligned} e^{ip} + e^{-iq} &= e^{\frac{i}{2}(p-q)} \left(e^{\frac{i}{2}(p+q)} + e^{-\frac{i}{2}(p+q)} \right) \\ &= 2e^{\frac{i}{2}(p-q)} \cos\left(\frac{i}{2}(p+q)\right) \end{aligned}$$

and

$$2\Delta e^{i(p-q)} - (e^{ip} + e^{-iq}) = 2\left(\Delta e^{i(p-q)} - e^{\frac{i}{2}(p-q)} \cos\left(\frac{p+q}{2}\right)\right)$$

Now, it's clear that the denominator of the last line of Eq. (6.1) is actually the norm squared of a diagonal fraction of modulus 1 (since it is symmetric in p and q). The numerator is

$$\begin{aligned} &2 + 2\cos(p+q) - 2\Delta(e^{ip} + e^{-iq}) + 4\Delta^2 e^{i(p-q)} \\ =&2(1 + \cos(p+q) - 4\Delta e^{\frac{i}{2}(p+q)} \cos\left(\frac{p+q}{2}\right) + 2\Delta^2 e^{i(p-q)} \end{aligned}$$

After further simplification we find that this is

$$= e^{i\Theta(p,q)}$$

where

$$\Theta(p,q) = 2\tan^{-1}\left\{ \frac{\Delta \sin \frac{p-q}{2}}{\cos \frac{p+q}{2} - \Delta \cos \frac{p-q}{2}} \right\}$$

7. Summery

Give that $\Theta(p,p) = 0$,

$$\begin{aligned} \exp(iNK_j) &= (-1)^{n-1} \prod_{\ell=1,\ell\neq j}^{n} \exp[i\Theta(K_j, K_\ell)] \\ &= (-1)^{n-1} \prod_{\ell=1}^{n} \exp[i\Theta(K_j, K_\ell)] \end{aligned}$$

Taking log of both sides we get

$$NK_j = 2\pi I_j - \sum_{\ell=1}^{n} \Theta(K_j, K_\ell) \tag{7.1}$$

for some $I_j \in \mathbb{Z}$ if n is odd and $I_j \in \mathbb{Z} + \frac{1}{2}$ if n is even. What do we want to take for our I_j:

- K_j should be distinct but packed as closely as possible.
- Since $K_j \approx 0$ and $\Theta(p,q) \approx 0$ for $p,q \approx 0$ we should have $K_j \approx I_j 2\pi/N_j$
- To have K_j distinct and as dense as possible we must make sure that $I_{j+1} - I_j = 1$.
- Symmetry about 0. Ie, $I_j + I_{n-j+1} = 0$, or equivalently that $I_1 = -\frac{n-1}{2}$. Then

$$I_j = j - \frac{n+1}{2}$$

Now, Yang and Yang showed that there exists a unique real solution of (7.1) with I_j given by the above conditions. Then in the Thermodynamic Limit these stay solutions of (7.1) as $N \to \infty$ and $\frac{n}{N}$ fixed since $\frac{n}{N}$ is the probability of a given arrow to be an up arrow in the configuration. Thus $n \to \infty$ and $K_1, \ldots, K_n$ (or rather $\partial K_1 + \ldots + \partial K_n$) tends to a distribution on $\mathbb{R}$.

Define a function

$$\rho(K)dK = \frac{\text{\# of zero's between } K \text{ and } K + dK}{N}$$

Then

$$NK_j = 2\pi\left(j - \frac{n+1}{2}\right) - \sum_{\ell=1}^{n} \Theta(K_j, K_\ell)$$

Taking $n, N \to \infty$ we have

- $$\frac{N}{n}K_j \to \frac{N}{n}K$$
- $$\sum_{\ell=1}^{n} \frac{\Theta(K_j, K_\ell)}{n} \to \int \Theta(k, k')\rho(k')dk'$$
- $$\frac{j}{n} \to \int_{-\infty}^{K} \rho(K')dK'$$

  since there are $j(-1)$ K_ℓ's less than K_j.

So in the limit we get

$$\frac{N}{n}K = 2\pi \int_{-\infty}^{K} \rho(K)dK' - \pi - \int \Theta(K, K')\rho(K')dK'$$

which implies that

$$\frac{N}{n} = 2\pi\rho(K) - \int \frac{\partial\Theta}{\partial K}(K, K')\rho(K')dK'$$

so in the end we get an integral equation

$$2\pi\rho(K) = \frac{N}{n} + \int \partial_1\Theta(K, K')\rho(K')dK'$$

This can be solved by Fourier Integrals.