

A variety of solutions to the Yang-Baxter equation

Nolan R. Wallach

Introduction.

If V is a finite dimensional vector space of dimension n , at least 2, and if σ is the “flip” on $V \otimes V$ ($x \otimes y \mapsto y \otimes x$) then it has been observed (M. Gerstenhaber, S. Majid and S.D.Schack) that σ is isolated in the set of all invertible solutions to the Yang-Baxter equations. It therefore seems to make little sense to consider solutions of the form $\sigma + P$ (i.e. a deformation of an isolated element). This is precisely what we do for a vary special class of “deformations”: we study the case when P is a rank one transformation. We show that the set of such P with $P \neq 0$ and $R = \sigma + P$ is invertible form an algebraic variety of dimension $n^2 + \lfloor \frac{n}{2} \rfloor$ (see Theorem 2.1).

We began studying such solutions after D. Meyer explained his method [M] of showing the exsistance of the Jones polynomial by looking for solutions in the Brauer algebra (see the examples in section 1). An expanded exposition combining Meyer’s idea and an observation of Turaev [T] can be found in my forthcoming book with Roe Goodman ([G-W]). Using classical invariant theory one sees that if \langle , \rangle is a non-degenerate symmetric bilinear form on V and if $G = O(V)$ is the group of isometries of this form then the space of G -invariant endomorphisms of $V \otimes V$ has basis I, σ, T with T defined as follows: let e_1, \dots, e_n be an orthonormal basis of V then $T(v \otimes w) = \langle v, w \rangle \sum e_i \otimes e_i$. In [G-W] we showed that if R is a G invariant, invertible solution to the Yang-Baxter equation then up to scalar multiple either $R = I$, $R = \sigma$ or $R = \sigma + q(n)T$ with $q(n)^2 + q(n)n + 1 = 0$. Meyer used the last type of solution with $q(n)$ thought of as a trancendental parameter to produce the Jones polynomial. We note that Meyer’s solutions are of the form $R = \sigma + P$ with P of rank 1.

The examples of Meyer led us to ask whether there were other such solutions $R = \sigma + P$ with P of rank 1. As we noted above they exist in profusion. In the case when $n = 2$ the R -matrix of two dimensional irreducible representation of the quantum deformation of the enveloping algebra of SL_2 is (up to scalar multiple) of this form. The family of solutions depending on a parameter q is in a sense the simplest “non-trivial”

family of such solutions (see Example 1 in section 1). In other words these R -matrices appear quite naturally without the introduction of “quantum groups”. In section 2 we show how to construct almost all solutions to the Yang-Baxter equation of our type.

Unfortunately, the solutions of the Yang-Baxter equations of the form studied in this note “only” produce the Jones polynomial (see Section 3). Thus from the perspective of link invariants these solutions are not interesting except possibly as a pedagogically simple way to prove the existence of the Jones polynomial. In the last section of this note we show how an open subset of these solutions give rise to representations of Temperley-Lieb algebras.

1. The main Lemma.

Let V be a free module of rank n over the commutative ring A with unit 1. We choose, $\gamma : V \times V \rightarrow A$, A -bilinear and $z \in V \otimes_A V = V \otimes V$. Let $\sigma(v \otimes w) = w \otimes v$ be the usual flip. Define $P(v \otimes w) = \gamma(v, w)z$. Set $R = \sigma + P$. Recall that if $T \in \text{End}_A(V \otimes V) = \text{End}(V \otimes V)$ then T generates three elements of $\text{End}(V \otimes V \otimes V)$ denoted T_{ij} , $i < j$ with T_{ij} acting as R on the i and j position and the identity in the k position if $\{i, j, k\} = \{1, 2, 3\}$. That is $T_{12}(u \otimes v \otimes w) = T(u \otimes v) \otimes w$, $T_{23}(u \otimes v \otimes w) = u \otimes T(v \otimes w)$ and $T_{13} = \sigma_{23}T_{12}\sigma_{23}$.

We are interested in determining when R satisfies the Yang-Baxter equation:

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

We choose a free basis $\{e_1, \dots, e_n\}$ of V and set

$$z = \sum x_{ij}e_i \otimes e_j, \quad y_{ij} = \gamma(e_i, e_j).$$

We are interested in solutions such that R is invertible. We note that

$$R\sigma z = (1 + \text{tr}(xy))z. \tag{1.1}$$

Thus a necessary condition for the invertibility of R is that $1 + \text{tr}(xy)$ is invertible in A if $P \neq 0$. Since $\text{tr}xy = \gamma(\sigma z)$, and if we set $Q = \sigma P$ then $Q^2 = \gamma(\sigma z)Q$ we see that if $P \neq 0$ then R is invertible if and only if $1 + \gamma(\sigma z)$ is a unit in A and

$$(\sigma R)^{-1} = I - \frac{Q}{1 + \gamma(\sigma z)}. \tag{1.2}$$

lem1.1 Assume that $P \neq 0$ and R is invertible and that the rank of V as a free A module is at least 2. Then R satisfies the Yang-Baxter equation if and only if

$$xyx^t y^t + (1 + \text{tr}(xy))I = 0 \quad (1.3)$$

(x^t is the transpose of x).

The proof is elementary but a bit complicated. We note that our hypothesis implies that $1 + \text{tr}(xy)$ is a unit in A . A direct calculation shows that

$$\begin{aligned} & R_{12}R_{13}R_{23}(v_1 \otimes v_2 \otimes v_3) - R_{23}R_{13}R_{12}(v_1 \otimes v_2 \otimes v_3) = \\ & \gamma(v_2, v_3)z \otimes ((1 + \text{tr}(xy))v_1 + \sum_{ijpq} x_{ij}x_{pq}\gamma(v_1, e_j)y_{pi}e_q) - \\ & ((1 + \text{tr}(xy))v_3 + \sum x_{ij}x_{pq}\gamma(e_i, v_3)y_{jq}e_p) \otimes \gamma(v_1, v_2)z. \end{aligned} \quad (1.5)$$

This difference will be 0 if the terms in the braces are both identically 0. In the first replace v_1 by e_k and see that the coefficient of e_q is

$$\delta_{kq}(1 + \text{tr}(xy)) + \sum_{ijp} x_{ij}x_{pq}y_{kj}y_{pi} =$$

$$\delta_{kq}(1 + \text{tr}(xy)) + (x^t y x y^t)_{kq}$$

here the subscript denotes kq matrix coefficient. In the second term we replace v_3 by e_k and have for the coefficient of e_p

$$\delta_{kp}(1 + \text{tr}(xy)) + \sum_{ijq} x_{ij}x_{pq}y_{ik}y_{jq} =$$

$$\delta_{kp}(1 + \text{tr}(xy)) + (y^t x y x^t)_{kp}.$$

This implies that if

$$y^t x y x^t + (1 + \text{tr}(xy))I = 0$$

and if

$$x^t y x y^t + (1 + \text{tr}(xy))I = 0$$

then $\sigma + P$ is a solution to the Yang-Baxter equation.

Since $1 + \text{tr}(xy)$ is a unit, either of these equations implies that x and y are invertible. The first equation is just (1.3) conjugated by y^t and the second is derivable from (1.3) by first conjugating by x^{-1} , then taking the transpose and then conjugating by $(x^t)^{-1}$. This establishes the sufficiency of (1.3).

We now prove the necessity. We first note that if $\sigma + P$ is invertible and $p \neq 0$ then y must be invertible. We now observe that the first part of the proof shows that it is enough to see that for each k there exist $v_2, v_3 \in V$ such that $\gamma(v_2, v_3) \neq 0$ and $\gamma(e_k, v_2) = 0$. But this is an easy consequence of the assumption that the rank of V is at least 2.

In the next section we will do an analysis of the general solution to (1.3). In this section we look at the “simplest” ones.

The first is a class that contains Meyer’s examples. Let $x \in GL(n, \mathbb{C})$ and set $y = qx^{-1}$ then

$$xyx^t y^t + (1 + \text{tr}(xy))I = (q^2 + nq + 1)I.$$

Thus if we take $x = I$ we retrieve Meyer’s examples.

Our second general type of example is a bit more subtle. As before take $x \in GL(n, \mathbb{C})$ but take $y = q(x^t)^{-1}$. This time we have

$$xyx^t y^t + (1 + \text{tr}(xy))I = (q^2 + \text{tr}(x(x^t)^{-1})q + 1)I.$$

Thus q must be a solution of the quadratic equation in the parenthesis. The simplest case of this form that is not of the previous form involves taking

$$x = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}, y = q(x^t)^{-1}.$$

If we set $\lambda = a^{-1}b$ then the equation above for q becomes

$$q^2 + (\lambda + \lambda^{-1})q + 1 = 0.$$

That is, $q = -\lambda^{\pm 1}$. After a bit of high school algebra this yields the following example.

Example 1. Let $A = \mathbb{Z}[q, q^{-1}]$ and let $V = A \times A$ with standard basis e_1, e_2 . Set

$$x = \begin{bmatrix} 0 & q^{-1} \\ -1 & 0 \end{bmatrix}, y = \begin{bmatrix} 0 & 1 \\ -q^{-1} & 0 \end{bmatrix}.$$

Then

$$xy = \begin{bmatrix} -q^{-2} & 0 \\ 0 & -1 \end{bmatrix}, x^t y^t = \begin{bmatrix} -1 & 0 \\ 0 & -q^{-2} \end{bmatrix}.$$

Thus (1.3) is clearly satisfied. If we take $e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2$ as an ordered free basis of $V \otimes V$ then if R is the corresponding element of $\text{End}(V \otimes V)$ then R has the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & q^{-1} & 1 - q^{-2} & 0 \\ 0 & 0 & q^{-1} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This can be recognized as q^{-1} times the R -matrix of the basic representation of $U_q(\mathfrak{sl}(2))$.

We give one final “obvious” class of solutions. Let $V = \mathbb{C}^n$ and let $x \in O(n, \mathbb{C})$ (i.e. $xx^t = I$) and $y = qI$. Then x, y produces a solution to (1.3) if and only if

$$q^2 + \text{tr}(xy)q + 1 = 0.$$

This leads to the following one parameter family of solutions.

Example 2. Let $A = \mathbb{C}$ and $V = \mathbb{C}^2$ with standard basis e_1, e_2 . Let

$$x = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}, y = \lambda I.$$

We solve for the λ such that (5) is satisfied. This time $\text{tr}xy = 2\lambda \cos \theta$, $xyx^t y^t = \lambda^2 I$. This λ must satisfy

$$\lambda^2 + 2\lambda \cos \theta + 1 = 0.$$

One checks easily that $\lambda = -e^{\pm i\theta}$ are the solutions. Take $\lambda = -e^{i\theta}$ then the corresponding R has matrix (as in example 1)

$$\begin{bmatrix} 1 - e^{i\theta} \cos \theta & 0 & 0 & -e^{i\theta} \cos \theta \\ -e^{i\theta} \sin \theta & 0 & 1 & -e^{i\theta} \sin \theta \\ e^{i\theta} \sin \theta & 1 & 0 & e^{i\theta} \sin \theta \\ -e^{i\theta} \cos \theta & 0 & 0 & 1 - e^{i\theta} \cos \theta \end{bmatrix}.$$

If we write $q = e^{i\theta}$ then we have in $\mathbb{Q}[i][q, q^{-1}]$

$$\begin{bmatrix} \frac{1-q^2}{2} & 0 & 0 & -\frac{1+q^2}{2} \\ \frac{1-q^2}{2i} & 0 & 1 & \frac{1-q^2}{2i} \\ \frac{q^2-1}{2i} & 1 & 0 & \frac{q^2-1}{2i} \\ -\frac{1+q^2}{2} & 0 & 0 & \frac{1-q^2}{2} \end{bmatrix}$$

2. The variety of solutions to equation (1.3).

Let $X = \{(x, y) \in GL(n, \mathbb{C}) \mid (x, y) \text{ satisfies (1.3)}\}$. We note that if $(x, y) \in X$ and if $g \in GL(n, \mathbb{C})$ then $(gxg^t, (g^t)^{-1}yg^{-1}) \in X$. We therefore have an action of $G = GL(n, \mathbb{C})$ on X .

Theorem 2.1. *The dimension of X as an algebraic variety is $n^2 + \lfloor \frac{n}{2} \rfloor$ ($\lfloor x \rfloor$ is the largest integer less than or equal to x).*

Let $\Phi : G \times G \rightarrow G \times G$ be defined by $\Phi(x, y) = (xy, x^t y^t)$. Let $Int(g)x = gxg^{-1}$ for $x, y \in G$. In the proof of Theorem 2.1 we will use the following Lemma.

Lemma 2.2. $\Phi(G \times G) = \{(u, v) \in G \times G \mid v^t \in Int(G)u\}$. If $\Phi(x, y) = (u, v)$ then $\Phi^{-1}(u, v) = \{(gx, yg^{-1}) \mid g \in G, gu = ug\}$.

Proof. If $u = xy$, $v = x^t y^t$ then

$$v^t = yx = Int(x^{-1})(xy) = Int(x^{-1})u.$$

Conversely, if $(u, v) \in G \times G$ and $v^t = x^{-1}ux$ then set $y = x^{-1}u$. Then $xy = u$ and $x^t y^t = x^t u^t (x^t)^{-1} = (v^t)^t$. This proves the first assertion. If $xy = x_1 y_1$ and $x^t y^t = x_1^t y_1^t$ then set $x_1 = gx$ we then have $x^t y^t = x^t g^t y_1^t$. Hence, $y = y_1 g$. Thus $x_1 = gx$ and $y_1 = yg^{-1}$. Now $xy = gx y g^{-1}$ implies that the condition in the second assertion is necessary. The sufficiency is even easier.

We now continue the proof of Theorem 2.1. We set $\tilde{X} = \{(x, y, z) \in G \times G \times \mathbb{C} \mid xyx^t y^t + z^2 I = 0\}$. We define $\Psi : X \times \mathbb{C}^\times \rightarrow \tilde{X}$ by $\Psi(x, y, \lambda) = (x, \lambda y)$. Then we note that Ψ is surjective and the inverse image of a point has one or two elements. This implies that

$$\dim \tilde{X} = \dim X + 1.$$

Let $\tilde{Y} = \{(u, v, z) \in G \times G \times \mathbb{C} \mid uv + z^2I = 0, v^t \in \text{Int}(G)u\}$. In light of Lemma 2.2,

$$\tilde{Y} = \{\Phi(x, y) \mid (x, y) \in \tilde{X}\}.$$

Finally, set $X_1 = \{(x, y) \in G \times g \mid xyx^ty^t = I\}$ and $Y_1 = \{u \in G \mid u^{-1} \in \text{Int}(G)u^t\}$. Let $\mu(x, y) = xy$. Then $\mu : X_1 \rightarrow Y_1$ is surjective. Also, if $g \in G$, $u \in Y_1$ then $gug^{-1} \in Y_1$. Furthermore, $\Psi : X_1 \times \mathbb{C}^\times \rightarrow \tilde{X}$ is surjective and the inverse image of a point has cardinality at most 2. Thus $\dim X_1 = \dim X$. If $u \in M_n(\mathbb{C})$ then set $\det(u + qI) = \sum_{i=0}^n q^i d_{n-i}(u)$. Define $\Lambda : M_n(\mathbb{C}) \rightarrow \mathbb{C}^{n+1}$ by $\Lambda(u) = (d_0(u), \dots, d_n(u))$. Clearly, $d_0(u) = 1$. Also, if $u \in Y_1$ then

$$\det(u + qI) = \det(u^t + qI) = \det(u^{-1} + qI) =$$

$$q^n (\det u)^{-1} \det(u + q^{-1}).$$

Since $d_n(u) = \det(u)$ we see that

$$\sum q^i d_n(u) d_{n-i}(u) = \sum q^{n-i} d_{n-i}(u).$$

This implies that

$$\Lambda(Y_1) = \{(z_0, \dots, z_n) \in \mathbb{C}^{n+1} \mid z_0 = 1, z_n z_{n-i} = z_i, i = 0, \dots, n\}.$$

If $z \in \Lambda(Y_1)$ then $\dim \mu^{-1} \Lambda^{-1}(z) = n^2$. Indeed, if the centralizer of u is of dimension r then $\mu^{-1}(u)$ has dimension r . Thus $\dim \mu^{-1}(\text{Int}(G)u)$ has dimension n^2 . Since $\Lambda^{-1}(z)$ is a union of a finite number of G orbits (the possible Jordan canonical forms for a given characteristic polynomial) the observation follows. Thus $\dim X_1 = n^2 + \dim \Lambda(Y_1)$. Since it is clear that $\dim \Lambda(Y_1) = \lfloor \frac{n}{2} \rfloor$, the proof is complete.

The proof of the theorem above suggests a method of finding an open subvariety of solutions to (1.3). We will look at the case $n = 2k$ and leave the case when n is odd to the reader. Let D_p be the space of all invertible $p \times p$ matrices. Let U_1 be the set of diagonal $k \times k$ matrices with diagonal entries $\lambda_1, \dots, \lambda_k$ such that $\lambda_i^2 \neq \lambda_j^{\pm 2}$ for $i \neq j$ and $\lambda_i^4 \neq 1$. If $u \in D_k$ set

$$\tilde{u} = \begin{bmatrix} u & 0 \\ 0 & u^{-1} \end{bmatrix}.$$

We set

$$x_o = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

with I the $k \times k$ identity matrix. If $(g, u_1, u_2) \in GL(n, \mathbb{C}) \times D_k \times D_n$ then we set

$$\Phi(g, u_1, u_2) = (gu_2x_og^t, (g^t)^{-1}x_o\tilde{u}_1(u_2)^{-1}g^{-1}) = (x, y).$$

Then $xy = g\tilde{u}_1g^{-1}$ and $x^ty^t = gx_o\tilde{u}_1x_og^{-1}$ thus $(x, y) \in X_1$. One checks that Φ has rank $n^2 + k$ on $GL(n, \mathbb{C}) \times U_1 \times D_n$. If $(g, u_1, u_2) \in GL(n, \mathbb{C}) \times D_k \times D_n$ and if q satisfies

$$q^2 + (\text{tr}u_1 + \text{tr}(u_1^{-1}))q + 1$$

then $(gu_2x_og^t, q(g^t)^{-1}x_o\tilde{u}_1(u_2)^{-1}g^{-1}) \in X$.

3. The “enhanced R -matrices”.

Let A and V be as in section 1. If $T \in \text{End}(V \otimes V)$ then we define $\text{tr}_1(T)(v) = \sum_i (e_i^* \otimes I)(T(e_i \otimes v))$ for e_1, \dots, e_n a free basis of V over A and e_1^*, \dots, e_n^* the dual basis. Similarly $\text{tr}_2(T)(v) = \sum_i (I \otimes e_i^*)(T(v \otimes e_i))$. Let $P(v) = \gamma(v)z$ for $v \in V \otimes V$ as in the previous section. We set $t = \gamma(\sigma z)$ and assume that $1 + t$ is a unit in A . We note that if $Q = \sigma P$ then

1. $\text{tr}_1(Q)$ has matrix xy and $\text{tr}_2(Q)$ has matrix x^ty^t .

Thus (1.3) becomes

2. R satisfies (1.3) if and only if

$$\text{tr}_1(\sigma P)\text{tr}_2(\sigma P) + (1 + t)I = 0.$$

Turaev has introduced the notion of enhanced Yang-Baxter operators. We recall his definition. Let R be an inverible solution to (1) in section 1. Then the triple (R, f, λ, μ) with $f \in GL(V)$, λ, μ units in A is an enhanced Yang-Baxter operator if

i) $f \otimes f$ commutes with R .

ii) $\text{tr}_2(\sigma R(f \otimes f)) = \lambda \mu f$ and $\text{tr}_2((\sigma R)^{-1}(f \otimes f)) = \lambda^{-1} \mu f$.

We note that under our conditions on $R = \sigma + P$ that the first condition in (ii) forces $f = ux^ty^t = u\text{tr}_2(\sigma P)$ with u a unit in A .

Lemma 3.1. *Let $R = \sigma + P$ satisfy 2. above. Set $f = x^t y^t = \text{tr}_2(\sigma P)$. If $\mu = \pm\sqrt{-(1+t)}$, $B = A[\mu, \mu^{-1}]$ then $(R, f, -\frac{1}{\mu}, \mu)$ is an enhanced Yang-Baxter operator over B .*

One checks (i) directly and observes that that $\text{tr}_2(\sigma R(f \otimes f)) = -f$ and $\text{tr}_2((\sigma R)^{-1}(f \otimes f)) = (1+t)f$.

The reason Turaev introduced the above notion was to give a simple proof of the existence of link invariants. Under the condition of Lemma 3.1 one can produce an invariant of links in \mathbb{R}^3 relative to the action of orientation preserving diffeomorphisms of \mathbb{R}^3 (ambient isotopy) with values in the base ring. A discussion of this can be found in Chari-Pressley section 15.2. Using Proposition 15.3.5 in [C-P] each $(x, y) \in X$ gives a link invariant, $p(x, y)(l)$ (l is the link). The method involves the notion of the completion of a braid. A discussion can be found in [C-P] (or [G-W]). If (R, f, λ, μ) is an enhanced Yang-Baxter operator and if b is a braid on m strands then set

$$p(b) = \lambda^{-\alpha(b)} \mu^{-m+1} \text{tr}(\rho_m(b) \circ (\otimes^m f))$$

where α is the map from the braid group of m strands to \mathbb{Z} that assigns to a braid the signed length of the word defining it (using the standard generators S_i , $i = 1, \dots, m-1$ of the braid group on m strands) and ρ_m is the representation of the braid group corresponding to $S_i \mapsto (\sigma R)_{ii+1}$. The point is that the above expression depends only on the class of the completion of b under ambient isotopy (i.e. is invariant under both Markov moves). If we argue as in [G-W, 10.4.4] we find that the invariant satisfies the skein relation of the Jones polynomial in μ^2 . A calculation of the value on the ‘‘unknot’’ (that is $\text{tr}(f \otimes f)$) we find that the invariant is $J(l)(\mu^2)\mu$ with $\mu = \pm\sqrt{-(1 + \text{tr}(xy))}$ and $J(l)$ is the Jones polynomial of the link.

Since, the Jones polynomial has an intimate relationship with the Temperley-Lieb algebra, it seems reasonable to guess that the points in X might correspond to representations of this algebra. That is the content of the next section.

4. Representations of the Temperley-Lieb algebra.

If $\lambda \in \mathbb{C}$ then the Temperley-Lieb (cf. [C-P], p.333) algebra $\mathcal{TL}_n(\lambda)$ is the universal associative algebra with unit generated by $n-1$ elements, t_1, \dots, t_{n-1} with the defining relations

$$t_i^2 = t_i, \quad t_i t_{i\pm 1} t_i = \lambda t_i, \quad t_i t_j = t_j t_i, \quad |i - j| > 1.$$

Lemma 4.1. *Let $(x, y) \in X \subset GL(n, \mathbb{C}) \times GL(n, \mathbb{C})$ (as in section 2). Let $V = \mathbb{C}^n$. If $\text{tr}(xy) \neq 0$ set $\lambda = -\frac{(1+\text{tr}(xy))}{\text{tr}(xy)^2}$ and $T = \frac{1}{\text{tr}(xy)}\sigma P$ (P corresponds to (x, y) as in section 1) then $t_i \mapsto T_{ii+1}$ on $\otimes^n V$ defines a representation of $\mathcal{TL}_n(\lambda)$.*

The proof of this result is a direct calculation using the observation that $P^2 = \text{tr}(xy)P$ and is left to the reader.

References

- [C-P] V. Chari and A. Pressley, *A Guide to Quantum Groups*, Cambridge University Press, Cambridge, 1994.
- [G-W] R. Goodman and N. Wallach, *Representations and Invariants of the Classical Groups*, Cambridge University Press, Cambridge, 1997.
- [M] D. Meyer, State models for link invariants from classical Lie groups, in *Knots 90*, Walter de Gruyter & Co., Berlin, 559-592.
- [T] V. G. Turaev, The Yang-Baxter equation and invariants of links, *Invent. math.*, 92(1988), 527-553.