A standard method in analysis, going back to Hilbert and Schmidt, and much earlier to Green, for studying operators is to find a larger commuting family of operators. A simple example is the Laplacian, an unbounded self-adjoint operator. This commutes with the integral operators $F \mapsto F \ast \phi$, where $\phi \in C_\infty_c(\mathbb{R}^n)$, which are Hilbert-Schmidt operators, hence compact. This method is used extensively in the theory of automorphic forms, for example in the Selberg trace formula.

Baxter’s approach to the six- and eight-vertex models was to embed the row transfer matrix into a larger commuting family of row transfer matrices. Another application of the same idea led to the solution of a problem in one-dimensional quantum mechanics, his analysis of the Heisenberg spin chains, a model of ferromagnetism [5]. Baxter was able to introduce the theory of elliptic functions by finding a commuting family of row transfer matrices from the eight-vertex model.

Baxter [1, 2] knew that two operators arising from physical problems were related to each other. (This fact was also observed by Sutherland [6].) The operators are:

- The row transfer matrices from the field-free 6 or 8 vertex models
- Hamiltonians for Heisenberg spin chains, called the XXZ and XYZ Hamiltonians.

Using the Yang-Baxter equation, the row transfer matrices can be organized into commuting families. This means that the row transfer matrix contains a parameter that can be differentiated, and roughly the Hamiltonian is the logarithmic derivative of the row transfer matrix. Equivalently, the row transfer matrix is an exponentiated Hamiltonian. As a consequence, the Hamiltonian also commutes with this family of row transfer matrices.

The field-free eight vertex model can be solved similarly to the six vertex model, using a parametrized Yang-Baxter equation. Let $a, b, c, d$ be the Boltzmann weights, thus:

<table>
<thead>
<tr>
<th>$a$</th>
<th>$a$</th>
<th>$b$</th>
<th>$b$</th>
<th>$c$</th>
<th>$c$</th>
<th>$d$</th>
<th>$d$</th>
</tr>
</thead>
</table>
| ![Diagram](image.png)

Define

$$\Delta = \frac{a^2 + b^2 - c^2 - d^2}{ab + cd}, \quad \Gamma = \frac{ab - cd}{ab + cd}.$$  \hspace{1cm} (1)

If $a', b', c', d'$ are another set of Boltzmann weights, and $\Delta', \Gamma'$ are defined like $\Delta, \Gamma$, the condition for the Yang-Baxter equation to have a solution is

$$\Delta = \Delta', \quad \Gamma = \Gamma'.$$
Now with $\Gamma$ and $\Delta$ fixed, the solutions $a : b : c : d$ to (1) form an elliptic curve, and indeed, the relevant Yang-Baxter equation is a parametrized Yang-Baxter equation with this curve as its parameter group. The relevant quantum group is an elliptic quantum group ([4]).

Baxter’s work solving the eight-vertex model was carried out on a ship, where he took over the chart room for his calculations. Baxter [2] wrote in Chapter 10:

Sutherland (1970) showed directly that the transfer matrix of any zero-field eight-vertex model commutes with an XYZ operator $X$. They therefore have the same eigenvectors. I was not aware of Sutherland’s result when I solved the eight-vertex model (I did much of the work in the writing room of the P & O liner Arcadia, in the Atlantic and Indian Oceans. This was good for concentration, but not for communication). It should be obvious from Sections 10.4-10.6 that such commutation relations are closely linked with the solution of the problem.

This theory is outside the scope of these lectures, but we will consider the simpler case where $d = 0$, where the field-free six-vertex model is related to the XXZ Hamiltonian. As we know, the relevant quantum group is $U_q(\hat{sl}_2)$.

For the free-fermionic six vertex model (where the quantum group is $U_q(\hat{gl}(1|1))$) a similar result was obtained by Brubaker and Schultz [3]. See also [7].

1 Heisenberg Spin Chains

In classical mechanics, observables are functions $A$ on the phase space, which is a parameter space representing the state of a physical system, including the positions and momenta of all particles. Given a state of the system, every observable thus has a definite value.

In quantum mechanics, by contrast, it is possible for the system to be in a state where a given observable does not have a definite value. The state of the system is represented by a vector in a Hilbert space $\mathcal{H}$, and the classical observable $A$ is replaced by a Hermitian (self-adjoint) operator $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ (or an unbounded operator defined on a dense subspace). If $\Psi \in \mathcal{H}$ represents the state of the system, the observable $A$ has a definite value $\lambda$ if $\hat{A}\Psi = \lambda\Psi$.

For simplicity let us assume that $\hat{A}$ has a discrete spectrum. By the spectral theorem, the state $\Psi$ may be expanded as a “Fourier series”

$$\Psi = \sum a_i \Psi_i$$

where $\Psi_i$ are eigenfunctions of $\hat{A}$. If we normalize $\Psi$ so that $|\Psi| = 1$, then the “amplitudes” $a_i$ have a probabilistic interpretation: if the observable $f$ is measured, a definite value $\lambda_i$ is returned, and the “wave function” $\Psi$ collapses to the state $\Psi_i$. The probability of this happening is $|a_i|^2$. By the Plancherel theorem $\sum_i |a_i|^2 = 1$, and so this scheme gives a probability distribution on the spectrum of $\hat{A}$.

In quantum mechanics, two observables $A$ and $B$ can be measured simultaneously if and only if the corresponding operators $\hat{A}$ and $\hat{B}$ commute. In this case, the eigenfunctions $\Psi_i$ can be chosen to be simultaneous eigenfunctions of $\hat{A}$ and $\hat{B}$.
A particular observable is energy, and the corresponding operator is the Hamiltonian. It determines the evolution of the system in time, through Schrödinger’s equation.

An example of a collection of observables that cannot be measured simultaneously are electron spin in different directions. A 2 dimensional Hilbert space $\mathcal{H} = \mathbb{C}^2$ is sufficient to represent a particle such as the electron with spin $\frac{1}{2}$. If the spin is measured along the $z$ axis, it will be found in one of two states, up or down. The spin operator is therefore represented by the matrix

$$\sigma^z = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$  

On the other hand, if the spin is measured along the $x$ or $y$ axes, it will again be found in one of two possible states. The corresponding operators do not commute with $\sigma^z$, and with respect to the same basis, are represented by the matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$ 

The three matrices $\sigma^x$, $\sigma^y$, $\sigma^z$ are called the Pauli spin matrices. They are both Hermitian and unitary. We have an alternative labeling

$$\sigma^1 = \sigma^x, \quad \sigma^2 = \sigma^y, \quad \sigma^3 = \sigma^z, \quad \sigma^4 = I_2.$$  

Heisenberg [5] proposed a quantum mechanical model of ferromagnetism. We consider a sequence of $N$ magnetic atoms such as iron at adjacent sites. We will assume that the sites of the spin chain are arranged in a ring. Consequently the boundary conditions for the six-vertex model will also be periodic, as in Lecture 2.

Each atom is a magnetic dipole whose dipole moment is proportional to the spin. Since the spin module is 2-dimensional, it is represented by a vector in a 2-dimensional space. The Hilbert space of a single magnetic atom is $\mathbb{C}^2$. Therefore the Hilbert space $\mathcal{H}$ of $N$ atoms is $\otimes^N \mathbb{C}^2$. We let $\sigma^x_j$, $\sigma^y_j$ and $\sigma^z_j$ denote the Pauli matrices acting on the $j$-th site, and as the identity operator on all other sites.

To give the simplest formulation, we will assume that the chain is periodic, so $\sigma^x_{N+1} = \sigma^x_1$ etc. Adjacent dipoles tend to align in the same direction, which partly explains the form of the Hamiltonian

$$H = \frac{1}{2} \sum_{j=1}^{N} (J_x \sigma^x_j \otimes \sigma^x_{j+1} + J_y \sigma^y_j \otimes \sigma^y_{j+1} + J_z \sigma^z_j \otimes \sigma^z_{j+1}),$$  

for suitable positive constants $J_x, J_y, J_z$. Due to the assumed periodicity, $\sigma_{N+1} = \sigma_N$. If $J_x = J_y$, this is called the XXZ Hamiltonian. It is an endomorphism of $\otimes^N \mathbb{C}^2$.

On the other hand, we can fix Boltzmann weights $a, b, c, d$ and consider the row transfer matrix $T_{a,b,c,d}(\alpha, \beta)$ as follows. We will denote the standard basis of $\mathbb{C}^2$ as

$$v_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$ 

A basis of $\otimes^N \mathbb{C}^2$ consists of vectors $v_\alpha$ where $\alpha = (\alpha_1, \cdots, \alpha_N) \in \{+, -\}^N$, where

$$v_\alpha = v_{\alpha_1} \otimes \cdots \otimes v_{\alpha_N}.$$
We may thus regard both the Hamiltonian $H$ and the row transfer matrices for the 8 vertex model as endomorphisms of the same Hilbert space, $\otimes^N \mathbb{C}^2$. Baxter proved:

- The XXZ Hamiltonian commutes with a family of 6-vertex model row transfer matrices.
- The XYZ Hamiltonian commutes with a family of 8-vertex model row transfer matrices.

We will review the relationship of the XXZ Hamiltonian with the row transfer matrices for the field-free 6-vertex model. We will partly investigate the 8-vertex model, but we will specialize to the 6-vertex model before long, and prove the second statement. See Baxter [1] for the 8-vertex model case, which requires some elliptic and theta functions.

2 Preliminaries

We will make use of the Pauli spin matrices with respect to this basis, and if $\alpha, \beta \in \{+, -\}$, and if $\sigma$ is one of the Paul spin matrices, we will denote by $\sigma_{\alpha, \beta}$ the corresponding matrix entry. Thus $\sigma^y_{+ -} = i$ and $\sigma^y_{+-} = -i$. Let

$$p_1 = \frac{1}{2}(b + d), \quad p_2 = \frac{1}{2}(b - d), \quad p_3 = \frac{1}{2}(a - c), \quad p_4 = \frac{1}{2}(a + c).$$

(4)

Let $v$ be a vertex type. We will denote by $R^{\gamma \delta}_{\alpha \beta}(v)$ the Boltzmann weight

![Diagram](image_url)

Lemma 2.1. Let $\alpha, \beta, \gamma, \delta \in \{+, -\}$. Then

$$R^{\gamma \delta}_{\alpha \beta} = \sum_{k=1}^{4} p_k \sigma_{\beta \gamma}^k \sigma_{\alpha \delta}^k.$$  

(5)

Proof. This can be checked by case-by-case consideration. There are 16 choices for $\alpha, \beta, \gamma, \delta$, but only eight give a nonzero result. Let us consider for example $(\alpha, \beta, \gamma, \delta) = (+, -, +, -)$. Since $\sigma^k_{+ -}$ and $\sigma^k_{- +}$ are nonzero only for $k = 1, 2$, there are two terms:

$$\frac{1}{2}(b + d)\sigma^1_{\beta \gamma} \sigma^1_{\alpha \delta} + \frac{1}{2}(b - d)\sigma^2_{\beta \gamma} \sigma^2_{\alpha \delta} = b.$$  

The remaining cases are similar. \qed

There is a similar identity

$$R^{\gamma \delta}_{\alpha \beta} = \sum_{k=1}^{4} w_k \sigma_{\beta \delta}^k \sigma_{\gamma \alpha}^k$$

where

$$w_1 = \frac{1}{2}(c + d), \quad w_2 = \frac{1}{2}(-c + d), \quad w_3 = \frac{1}{2}(a - b), \quad w_4 = \frac{1}{2}(a + b).$$

We won’t need this but mention it for completeness.
3 Six-vertex model and the XXZ Hamiltonian

Now we specialize to the six-vertex model, referring to [1] for the general case. Thus now \( d = 0 \), and as a consequence of this simplification we will have \( J_x = J_y \) in the Hamiltonian.

\[
\begin{align*}
\begin{array}{ccc}
 a & a & b \\
+ & - & + \\
+ & - & - \\
+ & + & - \\
+ & + & +
\end{array}
\begin{array}{ccc}
 b & b & c \\
+ & - & + \\
+ & - & - \\
+ & + & - \\
+ & + & +
\end{array}
\begin{array}{ccc}
 c & c & c \\
+ & - & + \\
+ & - & - \\
+ & + & - \\
+ & + & +
\end{array}
\end{align*}
\]

Let \( T_{a,b,c}(\alpha, \beta) \) be the corresponding row transfer matrix.

We saw in Lecture 4 that if \( \Delta \) is fixed, then we have a parametrized Yang-Baxter equation involving \( a, b, c \) such that
\[
a^2 + b^2 - c^2 = \frac{\Delta}{2ab}.
\]

Let \( q \) be such that \( \Delta = \frac{1}{2}(q + q^{-1}) \). We may parametrize the solutions by a map
\[
R_\Delta : \mathbb{C}^\times \rightarrow \{\text{field free Boltzmann weights } a, b, c\}
\]
given by
\[
R_\Delta(x) = (a,b,c) = \left(\frac{xq - (xq)^{-1}}{q - q^{-1}}, \frac{x - x^{-1}}{q - q^{-1}}, 1\right).
\]

These are the Boltzmann weights from Theorem 5.2 in Lecture 4, divided by the constant \( \frac{1}{2}(q - q^{-1}) \). We saw that this gives a parametrized Yang-Baxter equation. (Dividing by a constant does not affect this since both sides of the Yang-Baxter equation are divided by the same constant.)

Then by Theorem 1.1 of Lecture 2, the row transfer matrices \( T_{a,b,c}(\alpha, \beta) \) form a commuting family. We choose fixed \( \chi \) so that \( e^{i\chi} = q \). Then we choose variable \( \theta \) so that \( e^{i\theta} = xq \). We slightly modify the notation, omitting \( \Delta \) from the notation \( R_\Delta \) because it is fixed, and regarding \( R \) as a function of \( \theta \) instead of \( x \). We will use the notation \( R(\theta)_{\gamma\delta}^{\alpha\beta} \) as explained in Section 2.

Lemma 3.1. When \( \theta = \chi \), we have
\[
R(\chi)_{\gamma\delta}^{\alpha\beta} = \begin{cases} 
1 & \text{if } \alpha = \delta \text{ and } \beta = \gamma; \\
0 & \text{otherwise}.
\end{cases}
\]

Proof. Note that when \( \theta = \chi \) we have \((a,b,c) = (1,0,1)\). So this follows from the definition of the Boltzmann weights. \( \square \)

Let \( T_\theta(\alpha, \beta) \) be the row transfer matrix \( T_{a,b,c}(\alpha, \beta) \) with this parametrization.

Remark 1. At the special point \( \theta = \chi \), since \( b = 0 \), we are in a 5-vertex model case in which the particles are allowed to move to the right but not straight down. In fact \( T_\chi(\alpha, \beta) \) is the right shift operator, moving each particle one step to the right. Obviously \( T_\chi \) is invertible, the inverse being the left shift operator.
We may differentiate the operator $T_{\theta}$ with respect to $\theta$. The derivative $T'_{\theta}$ commutes with $T_{\theta}$, and we may consider the logarithmic derivative at $\theta = \chi$.

$$L = \frac{1}{2} T_{\chi}^{-1} T'_{\chi}.$$  

It is at this point $\theta = \chi$ that there is a relationship between the XXZ Hamiltonian and the six-vertex model. Regard the $p_i$ in (4) as functions of $\theta$.

$$J_x = \frac{1}{2} p'_1(\chi), \quad J_y = \frac{1}{2} p'_2(\chi), \quad J_z = \frac{1}{2} p'_3(\chi). \quad (6)$$

Since $d = 0$, we have $J_x = J_y$. Let $H$ be the XXZ Hamiltonian (3).

**Theorem 3.2.** With these notations, we have

$$L = H + c I \otimes C_2,$$

where $c$ is an explicit constant. The operator $H$ commutes with the 6-vertex row transfer matrices $T_{\theta}$.

**Proof.** We will label the interior edges of the single-layer grid whose partition function is $T_{\theta}$ by $\lambda_1, \cdots, \lambda_N$, thus:

\[
\begin{array}{cccccccc}
\lambda_1 & \alpha_1 & & & \alpha_2 & & & \alpha_3 & & \alpha_{N-1} & & \alpha_N \\
\beta_1 & & \lambda_2 & & \lambda_3 & & \ldots & & \lambda_N & & \beta_1 \\
& \beta_2 & & \beta_3 & & \beta_{N-1} & & \beta_N & & \\
\end{array}
\]

Due to the periodic boundary conditions, $\lambda_{N+1} = \lambda_1$. Thus

$$T_{\theta}(\alpha, \beta) = \sum_{\lambda} \prod_{i=1}^{N} R(\theta)^{\lambda_{i+1} \lambda_i}.$$  

Differentiating with respect to $\theta$ and setting $\theta = \chi$,

$$T'_{\chi}(\alpha, \beta) = \sum_{j=1}^{N} \sum_{\lambda} \left[ \frac{d}{d\theta} R(\theta)^{\lambda_{j+1} \lambda_j} \right]_{\theta = \chi} \prod_{i \neq j} R(\chi)^{\lambda_{i+1} \lambda_i}.$$  

By Lemma 3.1 if $i \neq j$ then $R(\chi)^{\lambda_{i+1} \lambda_i} = 1$ provided $\lambda_i = \beta_i$ and $\alpha_i = \lambda_{i+1}$, and is zero otherwise. Therefore the $j$-th term only contributes if $\beta_i = \alpha_{i-1}$ when $i \neq j, j + 1$. Assuming this, since we are summing over $\lambda$, we may omit these factors and take $\lambda_j = \alpha_{j-1}$, $\lambda_{j+1} = \beta_{j+1}$ to obtain

$$T'_{\chi}(\alpha, \beta) = \sum_{j=1}^{N} \frac{d}{d\theta} R(\theta)^{\beta_{j+1} \alpha_j} \big|_{\theta = \chi}.$$
Now we substitute (5) to obtain

\[ T'_\chi(\alpha, \beta) = \sum_{k=1}^{4} p'_k(\chi) \sum_{j=1}^{N} \sigma^k_{\alpha_j \beta_{j+1}} \sigma^k_{\alpha_{j-1} \beta_j}. \]

But now we remember that it is not \( T'_\chi \) that we are trying to compute, but \( \frac{1}{2} T^{-1}_\chi T'_\chi \), and \( T^{-1}_\chi \) is the left-shift operator by Remark 1. Thus

\[ \mathcal{L} = \frac{1}{2} \sum_{k=1}^{4} p'_k(\chi) \sum_{j=1}^{N} \sigma^k_{\alpha_j \beta_j} \sigma^k_{\alpha_{j-1} \beta_{j-1}}. \]

The first three terms produce the XXZ Hamiltonian, with \( J_x = J_y = p'_1(\chi) \) and \( J_z = p'_3(\chi) \). The last term produces \( cI_{2^N \mathbb{C}} \) with the constant \( c = \frac{1}{2} N p'_1(\chi) \).

References


