Journal of Statistical Physics, Vol. 44, Nos. 1/2, 1986

A Comment on Baxter Condition for Commutativity of Transfer Matrices

Eugene Gutkin[1]

Received July 24, 1985; final January 29, 1986

Let T_N and T'_N be the transfer matrices of two vertex models corresponding to two sets of Boltzmann weights. The Baxter condition on Boltzmann weights was known to be sufficient for commutativity of T_N and T'_N for all N. We show that generically it is also necessary.

KEY WORDS: Vertex models; commuting transfer matrices; Baxter condition.

INTRODUCTION

The most important device in the theory of soluble lattice models of statistical mechanics is to imbed the transfer matrix of the model into a family of pairwise commuting transfer matrices.[1] Since the transfer matrix depends on the size N of the lattice, it seems that one needs to verify an infinite number of conditions to check that two transfer matrices T_N and T'_N commute for all N. Baxter introduced a finite number of local conditions which are sufficient for the commutativity of T_N and T'_N for all N (see Ref. 1 and the theorem below). These conditions are written as a matrix equation, and a well-known special case of it is called the Yang–Baxter equation. Similar equations appear in many other situations (see, e.g., Refs. 2, 4, 10).

In this paper we study the commutativity of transfer matrices in a fairly general context in relation with the matrix equation referred to above which we call the Baxter condition. Our main result is that under some

Partially supported by NSF Grant DMS-8403238.

[1] Columbia University and University of Southern California, Department of Mathematics, Los Angeles, California 90089, U.S.A.

0022-4715/86/0700-0193$05.00/0

technical assumptions the Baxter condition is also necessary for the commutativity of transfer matrices.

The material of the paper overlaps with Ref. 9 where the same result is claimed. However, the proof there contains a gap at a crucial point. As the reader will see from the argument below, commutativity of transfer matrices is equivalent to a question in linear algebra which is of its own interest. In the second part of the paper we derive the Yang–Baxter equation from the Baxter condition. (A preliminary version of this paper appeared; Ref. 6. The referee's comments are gratefully acknowledged.)

1. PRELIMINARIES

We start by recalling the basic notions about the lattice models of statistical mechanics, putting them in a convenient form for our presentation. We consider only one class of lattice models, namely the vertex models on rectangular lattices with periodic boundary conditions.

In a vertex model the spins live on the edges. Let $m \geqslant 1$ (resp. $n \geqslant 1$) be the number of spin states living on the horizontal (resp. vertical) edges. We denote the horizontal spin states by i, j $(1 \leqslant i, j \leqslant m)$ and the vertical ones by k, l $(1 \leqslant k, l \leqslant n)$. A vertex model is determined by m^2n^2 numbers $-\infty < e(i, j, k, l) \leqslant \infty$ where $e(i, j, k, l)$ is the energy of an elementary configuration of spins (see Fig. 1).

The other two classes of lattice models which are often considered in the literature are the spin models and the "interactions-round-a-face" models.[1]

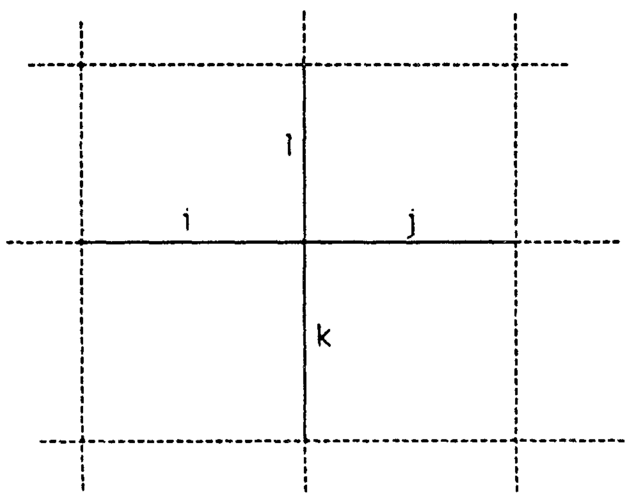

Fig. 1. Elementary configurations in a vertex model.

For any such model we can find a vertex model[5] with the same partition function. In view of this, for the purpose of the present paper it suffices to study the transfer matrices for vertex models.

Let M and N be the number of columns and the number of rows of a lattice and denote by $t > 0$ the normalized temperature. A configuration ω of a vertex model is given by an assignment of spins to all edges of the lattice. The energy $E(\omega)$ of a configuration is given by

$$E(\omega) = \sum e(i, j, k, l) \tag{1}$$

the summation is taken over all elementary configurations (see Fig. 1) contained in ω. The partition function of a model is given by

$$Z(M, N) = \sum_{\omega} \exp[-t^{-1}E(\omega)] \tag{2}$$

and the physical properties of a model are determined by the asymptotics of $Z(M, N)$ when M and N go to infinity.

Now we define the transfer matrix of a vertex model which is useful in studying the partition function. The numbers

$$w(k, l \mid i, j) = \exp[-t^{-1}e(i, j, k, l)]$$

are called the Boltzmann weights of the model. Denote by $E \simeq C^m$ the space of horizontal spin states and by $F \simeq C^n$ the space of vertical spin states of a vertex model. The numbers $w(k, l \mid i, j)$ are the matrix elements of an operator W on the space $F \otimes E$ which we call the Boltzmann matrix of the model. By fixing a pair (k, l) of vertical spins, we determine the

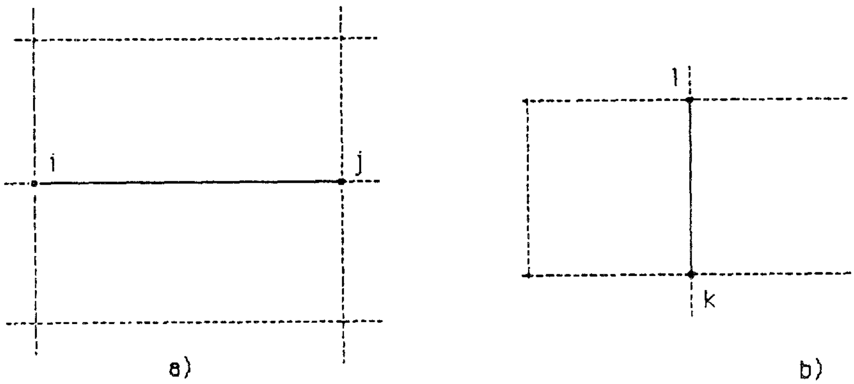

Fig. 2. Elementary configurations in a spin model.

operator W_k^l on E whose matrix elements are $W_k^l(i, j) = w(k, l \mid i, j)$. We call the n^2 operators W_k^l on E the horizontal Boltzmann matrices. In what follows we use the words matrix and operator as synonyms because the spaces we consider always have a canonical basis. For instance, the space $H_M = \otimes^M C^n$ has a canonical basis labeled by M tuples $(k_1,..., k_M)$, $1 \leqslant k_1,..., k_M \leqslant n$. The operator T on H_M whose matrix entries are given by

$$T_{k_1,...,k_M}^{l_1 \ \ l_M} = \mathrm{tr}[W_{k_1}^{l_1},..., W_{k_M}^{l_M}] \tag{3}$$

is called the (row to row) transfer matrix of the model. It is straightforward to check that

$$Z(M, N) = \mathrm{tr}\ T^N \tag{4}$$

Remark. The transfer matrix T defined by (3) is the row to row transfer matrix. Analogously one defines the column to column transfer matrix S on H_N. All considerations of the paper are valid for the transfer matrix S as well.

2. COMMUTING TRANSFER MATRICES

Assume for simplicity of exposition that $M = N$ and denote by T_N the transfer matrix of the lattice of size N for a given model. A model is completely determined by its Boltzmann matrix W acting on $F \otimes E \simeq C^n \otimes C^m$ or, equivalently, by the n^2 horizontal Boltzmann matrices W_k^l, $1 \leqslant k, l \leqslant n$, acting on E. Another model with the same number n, but with m' possibly different from m, is given by its Boltzmann matrix W' acting on $F \otimes E' \simeq C^n \otimes C^{m'}$ and we denote its transfer matrix by T'_N. Both T_N and T'_N act on the same space $H_N \simeq \otimes^N C^n$ and the problem is to find criteria on W and W' under which T_N commutes with T'_N for all N.

For $1 \leqslant k, l \leqslant n$ we define the operators U_k^l and V_k^l on $E \otimes E' = C^m \otimes C^{m'}$ by setting their matrix entries $U_k^l(i, i'; j, j')$ and $V_k^l(i, i'; j, j')$, respectively, to be

$$U_k^l(i, i'; j, j') = \sum_{p=1}^{n} w(p, l \mid i, j)\, w'(k, p \mid i', j') \tag{5}$$

and

$$V_k^l(i, i'; j, j') = \sum_{p=1}^{n} w'(p, l \mid i', j')\, w(k, p \mid i, j) \tag{6}$$

Theorem 1. Let W and W' be the Boltzmann matrices of two lattice models and let U_k^l and V_k^l, $1 \leqslant k, l \leqslant n$, be the two sets of operators on $C^m \otimes C^{m'}$ defined by them.

1. Assume that the operators U_k^l or V_k^l, $1 \leqslant k, l \leqslant n$, have no nontrivial common invariant subspace. Then T_N and T'_N commute for all N if and only if there is an invertible operator R on $C^m \otimes C^{m'}$ such that for $1 \leqslant k, l \leqslant n$
$$RU_k^l = V_k^l R \tag{7}$$
2. If (7) is satisfied then T_N and T'_N commute for all N without any other assumptions.

To prove the Theorem we need the following Lemma, which explains the role of operators U_k^l and V_k^l.

Lemma. The matrix elements

$$(T_N T'_N)_{k_1,\dots,k_N}^{l_1,\dots,l_N} \qquad \text{and} \qquad (T'_N T_N)_{k_1,\dots,k_N}^{l_1,\dots,l_N}$$

of $T_N T'_N$ and $T'_N T_N$ are given by

$$(T_N T'_N)_{k_1,\dots,k_N}^{l_1 \quad l_N} = \mathrm{tr}[U_{k_1}^{l_1},\dots, U_{k_N}^{l_N}] \tag{8}$$

$$(T'_N T_N)_{k_1,\dots,k_N}^{l_1 \quad l_N} = \mathrm{tr}[V_{k_1}^{l_1},\dots, V_{k_N}^{l_N}] \tag{9}$$

respectively.

Postponing the proof of the Lemma, we now prove the Theorem.

Proof of the Theorem. By the Lemma, T_N and T'_N commute for all N if and only if for any N-tuples $(k_1,\dots, k_N)$ and $(l_1,\dots, l_N)$

$$\mathrm{tr}[U_{k_1}^{l_1},\dots, U_{k_N}^{l_N}] = \mathrm{tr}[V_{k_1}^{l_1},\dots, V_{k_N}^{l_N}] \tag{10}$$

If (7) is satisfied then

$$V_{k_1}^{l_1},\dots, V_{k_N}^{l_N} = R[U_{k_1}^{l_1},\dots, U_{k_N}^{l_N}]\, R^{-1} \tag{11}$$

implying (10) and providing the second assertion of the Theorem.

Assume for concreteness that the operators U_k^l, $1 \leqslant k, l \leqslant n$, have no common invariant subspace. Consider the free algebras $P(U)$ and $P(V)$ over C generated by the symbols U_k^l and V_k^l, respectively. Elements of $P(U)$ (resp., $P(V)$) are the noncommutative polynomials $p(U_{k_i}^{l_i})$ (resp., $p(V_{k_i}^{l_i})$) $i = 1, 2,\dots$, which, for brevity, we denote by $p(U_k^l)$ and $p(V_k^l)$, respectively. The correspondence

$$V_k^l \to U_k^l, \qquad 1 \leqslant k, l \leqslant n \tag{12}$$

obviously extends to the isomorphism of algebras

$$\psi: P(V) \to P(U) \tag{13}$$

given by $\psi[p(V_k^l)] = p(U_k^l)$. Denote by M the full matrix algebra on $C^m \otimes C^{m'} \simeq C^{mm'}$ and let $\mathbf{U}$ and $\mathbf{V}$ be the subalgebras generated by U_k^l and V_k^l, respectively. Matrix algebras $\mathbf{U}$ and $\mathbf{V}$ are the quotients of $P(U)$ and $P(V)$, respectively, and we want to show that ψ descends to an algebra homomorphism

$$f: \mathbf{V} \to \mathbf{U} \tag{14}$$

This is true if and only if any relation $p(V_k^l) = 0$ as an operator on $C^{mm'}$ implies $p(U_k^l) = 0$ in the same sense. Assume the opposite, i.e., let p be a polynomial such that $p(U_k^l) \neq 0$ but $p(V_k^l) = 0$. Notice that (10) implies that for any polynomial q we have

$$\operatorname{tr} q(U_k^l) = \operatorname{tr} q(V_k^l) \tag{15}$$

A matrix A is equal to zero if and only if $\operatorname{tr}(AB) = 0$ for any matrix B. Since $p(U_k^l) \neq 0$ there exists a matrix B such that $\operatorname{tr}[U_k^l B] \neq 0$. Since the operators U_k^l, $1 \leqslant k, l \leqslant n$, have no common invariant subspace, the matrix algebra $\mathbf{U}$ is irreducible; hence, by Burnside's Theorem (Ref. 3, p. 182), $\mathbf{U} = M$. Thus, $B = q(U_k^l)$ for some polynomial q. Hence, $\operatorname{tr}[p(U_k^l)\, q(U_k^l)] \neq 0$, therefore, by (15), $\operatorname{tr}[p(V_k^l)\, q(V_k^l)] \neq 0$ contrary to the assumption that $p(V_k^l) = 0$.

We have shown that the correspondence (12) uniquely extends to the homomorphism f of matrix algebras. Since $\mathbf{U} = M$ and f, by definition, is onto, $\mathbf{V} = M$ and f is an automorphism of the full matrix algebra. By Skolem's Theorem (cf. Ref. 7, p. 99), any automorphism of the full matrix algebra is an inner automorphism, i.e., there exists an invertible matrix $R \in M$ such that for all $X \in M$

$$f(X) = RXR^{-1} \tag{16}$$

which implies (7) and finishes the proof.

Corollary 1. Let assumption 1 of the Theorem be satisfied. If the operator R satisfying (7) exists, then it is unique up to a scalar factor.

Proof. Let R_1 and R_2 be two operators satisfying (7). They define the automorphisms $f_1: X \to R_1 X R_1^{-1}$ and $f_2: X \to R_2 X R_2^{-1}$ of the full matrix algebra M. The automorphism $f_1 f_2^{-1}$ defined by $R_1 R_2^{-1}$ does not change the operators U_k^l, $1 \leqslant k, l \leqslant n$. Since these operators generate M, $f_1 f_2^{-1} = Id$. Thus, $R_1 R_2^{-1}$ commutes with any $X \in M$, that is, $R_1 R_2^{-1}$ is a scalar matrix.

Remarks. 1. We call (7) the Baxter condition. When $n=m=m'=2$ it becomes (9.6.7) in Ref. 1. 2. Assertion 2 of Theorem 1 is of course well-known and is included only for completeness.

Proof of Lemma. Using (3) and the definition of Boltzmann matrices we have for any two N tuples $(k_1,..., k_N)$ and $(l_1,..., l_N)$ (we omit the subscript N in T_N and T'_N).

$$\begin{aligned}(T'T)^{l_1 \quad l_N}_{k_1,...,k_N} &= \sum_{p_1,...,p_N} T'^{l_1 \quad l_N}_{p_1,...,p_N} T^{p_1 \quad p_N}_{k_1,...,k_N} \\ &= \sum_{p_1,...,p_N} \operatorname{tr}[W'^{l_1}_{p_1} ... W'^{l_N}_{p_N}] \operatorname{tr}[W^{p_1}_{k_1} ... W^{p_N}_{k_N}] \\ &= \sum_p \left[\sum_{i'} W'^{l_1}_{p_1}(i'_1, i'_2) ... W'^{l_N}_{p_N}(i'_N, i'_1) \right] \\ &\quad \times \left[\sum_i W^{p_1}_{k_1}(i_1, i_2) ... W^{p_N}_{k_N}(i_N, i_1) \right]\end{aligned}$$

Interchange the order of summation in the last expression, and do summation over p first (p, i, and i' denote the multiindices). We have

$$\begin{aligned}(T'T)^{l_1 \quad l_N}_{k_1,...,k_N} &= \sum_{i'} \sum_i \sum_p [W'^{l_1}_{p_1}(i'_1, i'_2)\, W^{p_1}_{k_1}(i_1, i_2)] ... \\ &\quad \times [W'^{l_N}_{p_N}(i'_N, i_1)\, W^{p_N}_{k_N}(i_N, i_1)]\end{aligned} \tag{17}$$

Using (6) we rewrite (17) as

$$\begin{aligned}(T'T)^{l_1 \quad l_N}_{k_1,...,k_N} &= \sum_{i,i'} V^{l_1}_{k_1}(i_1, i'_1; i_2, i'_2) ... V^{l_N}_{k_N}(i_N, i'_N; i_1, i'_1) \\ &= \operatorname{tr}[V^{l_1}_{k_1} ... V^{l_N}_{k_N}]\end{aligned}$$

which proves (9). Switching the order of T and T' we obtain (8).

3. BAXTER CONDITION AND YANG–BAXTER EQUATION

We start by introducing new notation. The space of linear operators on a vector space E is denoted by $M(E)$. Let E, E', and F be vector spaces and let $W \in M(F \otimes E)$, $W' \in M(F \otimes E')$. There is a bilinear operation (pairing)

$$(W, W') \to W \underset{F}{*} W' \in M(F \otimes E \otimes E') \tag{18}$$

To define the pairing (18) we choose a basis $\{f_k, 1 \leqslant k \leqslant n\}$ in F. Then W

and W' correspond to the n^2 operators $W_k^l \in M(E)$ (resp. $W'^l_k \in M(E')$), $1 \leqslant k, l \leqslant n$, by the formulas

$$W(f_k \otimes e) = \sum_{l=1}^{n} f_l \otimes W_k^l e$$

and

$$W'(f_k \otimes e') = \sum_{l=1}^{n} f_l \otimes W'^l_k e'$$

respectively. We set

$$(W \underset{F}{*} W')^l_k = \sum_{p=1}^{n} W_p^l \otimes W'^p_k \tag{19}$$

The n^2 operators $(W *_F W')^l_k \in M(E \otimes E')$ determine the operator $W *_F W' \in M(F \otimes E \otimes E')$. It is not hard to check that $W *_F W'$ does not depend on the choice of basis in F. One can think of the pairing (18) as the convolution with respect to F indices.

Another way to describe this pairing is to extend W and W' to operators on $F \otimes E \otimes E'$ by identity on the third, respectively, second, factor. Denote the extended operators by W_{12}, $W'_{13} \in M(F \otimes E \otimes E')$, respectively. It is straightforward to see that our pairing satisfies

$$W \underset{F}{*} W' = W_{12} W'_{13} \tag{20}$$

The same way we define $W' *_F W \in M(F \otimes E' \otimes E)$ which, after the natural identification of $E' \otimes E$ with $E \otimes E'$, becomes an element of $M(F \otimes E \otimes E')$. Then we have

$$W' \underset{F}{*} W = W'_{13} W_{12} \tag{21}$$

Now let $R \in M(E \otimes E')$. We extend R to an element of $M(F \otimes E \otimes E')$ by identity on the first factor and denote the result by R_{23}.

Proposition 1. The Baxter condition (7) on the Boltzmann matrices $W \in M(F \otimes E)$, $W' \in M(F \otimes E')$ is equivalent to

$$R_{23} W'_{13} W_{12} = W_{12} W'_{13} R_{23} \tag{22}$$

Proof. It suffices to notice that the operators U_k^l, $V_k^l \in M(E \otimes E')$ defined by (5) and (6) are equal to $(W *_F W')^l_k$ and $(W' *_F W)^l_k$, respectively.

Let F be as above and let G be an arbitrary vector space. The algebra $M(F\otimes G)$ naturally contains $M(G)$ as the subalgebra of operators on $F\otimes G$, which are identity on the first factor. Choosing a basis in F we identify $M(F\otimes G)$ with the algebra $M[n, M(G)]$ of $n\times n$ matrices with entries in $M(G)$. The subalgebra $M(G)$ corresponds in this representation to the scalar $n\times n$ matrices. We call such matrices F-scalar. Thus, $M(G)$ imbeds into $M(F\otimes G)$ as the subalgebra of F-scalar matrices. The definition does not depend on the choice of basis in F; thus we can talk about F-scalar operators on $F\otimes G$.

Definition. An operator $W\in M(F\otimes G)$ is said to be in general position (with respect to F) if the only F-scalar operators commuting with W are scalar operators. In this language, we have the following result.

Theorem 2. Let $W\in M(F\otimes E)$ and $W'\in M(F\otimes E')$ be the Boltzmann matrices of two vertex models, where $F\simeq C^n$, $E\simeq C^m$, $E'\simeq C^{m'}$. Consider the operators $W'_{13}W_{12}$ and $W_{12}W'_{13}$ belonging to $M(F\otimes E\otimes E')$. If either $W'_{13}W_{12}$ or $W_{12}W'_{13}$ are in general position then the Baxter condition

$$R_{23}W'_{13}W_{12}=W_{12}W'_{13}R_{23}$$

is necessary and sufficient for the commutativity of the transfer matrices T_N and T'_N for all N.

Proof. Let G be any vector space, let $A\in M(F\otimes G)$, and let A^l_k, $1\leqslant k, l\leqslant n$, be defined as above. By Schur's Lemma, A is in general position if and only if the operators A^l_k have no nontrivial common invariant subspace. It remains to set $G=E\otimes E'$ and notice that $(W'_{13}W_{12})^l_k=U^l_k$ and $(W_{12}W'_{13})^l_k=V^l_k$. The assertion now follows from Theorem 1.

Assume now that $F=E=E'\simeq C^n$ and that we have a one-parameter family $A(z)\in M(C^n\otimes C^n)$ of Boltzmann matrices. The Yang–Baxter equation (see, e.g., Refs. 2, 8, 11) is

$$A_{23}(u-v)\,A_{13}(u)\,A_{12}(v)=A_{12}(v)\,A_{13}(u)\,A_{23}(u-v) \tag{23}$$

Proposition 2. If (23) is satisfied for some u and v, then the transfer matrices $T_N(u)$ and $T_N(v)$ corresponding to $A(u)$ and $A(v)$, respectively, commute.

Proof. Equation (23) implies (22) with $W'=A(u)$, $W=A(v)$, and $R=A(u-v)$.

CONCLUSION

The Baxter condition (22) which is sufficient for the commutativity of the transfer matrices T_N and T'_N (for all N) of two vertex models with the Boltzmann matrices W and W', respectively, is also necessary under a certain technical assumption on W and W'. The Yang–Baxter equation (23) is much stronger than (22) and therefore it should not, in general, be necessary for the commutativity of T_N and T'_N.

REFERENCES

1. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, New York, 1982).
2. A. Bovier, *J. Math. Phys.* **23**:631–641 (1983).
3. C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Wiley, New York, 1962).
4. E. Gutkin, Bethe Ansatz and functional equations, preprint.
5. E. Gutkin, Equivalence of lattice models of statistical mechanics, in preparation.
6. E. Gutkin, A Comment on Yang–Baxter Relations, preprint.
7. I. N. Herstein, *Noncommutative Rings*, Carus Mathematics Monographs (Wiley, New York, 1968).
8. P. P. Kulish and E. K. Sklyanin, *J. Sov. Math.* **19**:1596–1620 (1982).
9. P. Lochak and J. M. Maillard, Necessary Versus Sufficient Conditions for Exact Solubility, preprint.
10. L. A. Takhtadzhan and L. D. Faddeev, *Russ. Math. Surv.* **34**:5 (1979), p. 11–68.
11. C. A. Tracy, *Physica* **16D**:203–220 (1985).
12. C. N. Yang, *Phys. Rev. Lett.* **19**:1312–1315 (1967).