Algebro-geometric aspects of the Bethe equations[†]

Dedicated to the memory of Feza Gürsey

Robert P. Langlands and Yvan Saint-Aubin

Although the Ansatz introduced by Bethe in 1931 ([B]) has been exploited repeatedly by physicists, who have adapted it successfully to a variety of problems, it has never been given a careful mathematical treatment. As a result there is often a disquieting imprecision in its formulation that discourages a resolute pursuit of its analytical consequences; moreover, and more to the point here, its algebraic charm has been little appreciated. Two years ago, the present authors undertook a study of the equations with standard techniques from algebraic geometry. The enterprise, rewarding as it has been, has taken more time and energy than expected. Complete proofs, even adequate understanding, have cost a great deal of effort and patience, and there are still gaps, but the project is nearing completion, and in this paper we describe, albeit in a somewhat provisional form, the principal features of the treatment. Details will appear in [BL].

There is no need here to recall the physical origins of the eigenvalue problem treated by Bethe. The mathematical problem is that of finding the eigenvalues and eigenvectors of an operator on a space of dimension 2^N. This space is

$$\mathfrak{X} = \otimes_1^N \mathbb{C}^2.$$

We take as basis of $\mathbb{C}^2$ two vectors u^+ and u^- and as basis of $\mathfrak{X}$ the vectors

$$u_{m_1,\ldots,m_r} = \otimes u_i, \qquad \{m_1,\ldots,m_r\} \subset \{1,\ldots,N\}$$

where $u_i = u^+$ if $i \in \{m_1,\ldots,m_r\}$ and $u_i = u^-$ otherwise. Thus the index attached to an element of the basis is a sequence of N signs. It is to be thought of as a cyclic sequence, so that each sign has two neighbors, those of the sign at position 1 being those at positions 2 and N. A typical vector will be written

$$\sum_{r=0}^{N} \sum a_{m_1,\ldots,m_r} u_{m_1,\ldots,m_r} = \sum x_r,$$

the inner sum running over all sequences with r positive signs. There is a corresponding decomposition $\mathfrak{X} = \oplus_{r=0}^N \mathfrak{X}_r$.

The operator H whose eigenvalues are to be calculated leaves each of the spaces $\mathfrak{X}_r$ invariant and on $\mathfrak{X}_r$ is given by

$$H_r : \quad x_r \to x'_r$$

with

$$a'_{m_1,\ldots,m_r} = \sum (a_{m_1,\ldots,m_r} - a_{m'_1,\ldots,m'_r}).$$

The sum runs over all sequences $\{m'_1,\ldots,m'_r\}$ that can be obtained from the sequence $\{m_1,\ldots,m_r\}$ by interchanging two adjacent and opposite signs. For example, $- - - - + + - -$ allows just two possibilities: $- - - + - + - -$ and $- - - - + - + -$. Recall that adjacency is to be interpreted in the cyclic sense. For H_r there are $\binom{N}{r}$ eigenvalues and eigenvectors to be found.

[†] First appeared in Strings and Symmetries, Proc. of Gürsey Memorial Conference, Istanbul, Springer-Verlag (1995).

If $z = (z_1, \dots, z_r)$ with z_i complex and $m = (m_1, \dots, m_r)$ we set $z^m = \prod_k z_k^{m_k}$. If P is a permutation of the set $\{1, \dots, r\}$ we set $Pm = m'$, $m'_k = m_{P^{-1}k}$. The Bethe Ansatz is to search for eigenvectors of the form

$$a_{m_1,\dots,m_r} = \sum_P z^{Pm} w_P \tag{1}$$

The sum runs over all permutations of $\{1, \dots, r\}$. The complex number w_P is obtained from a collection of complex numbers $w_{k,l}$, $k \neq l$, with $w_{k,l} = w_{l,k}^{-1}$:

$$w_P = \prod_{\substack{k>l \\ P^{-1}k<P^{-1}l}} w_{k,l}.$$

Bethe, less preoccupied with the algebro-geometric aspects of the equations, chose as variables f_k and $\varphi_{k,l}$ with $z_k = \exp i f_k$, $w_{k,l} = \exp i\varphi_{k,l}$. The formula (1) for the eigenvectors is moreover not quite that of Bethe; we have multiplied his eigenvectors by appropriate constants to obtain more symmetric formulas.

The vector (1) will be an eigenvector (or zero) if the points z and $w = (w_{k,l})$ satisfy the equations

$$\begin{aligned} w_{k,l} &= -\frac{z_k z_l - 2z_k + 1}{z_k z_l - 2z_l + 1}, \qquad k \neq l, \\ z_k^N &= \prod_{l \neq k} w_{k,l}. \end{aligned} \tag{2}$$

If it is not zero, the associated eigenvalue is

$$2\epsilon = \sum_1^r 1 - \cos f_k = \sum_1^r 1 - \frac{z_k + z_k^{-1}}{2}.$$

It turns out that these equations evince in algebro-geometrical respects an inconvenient degeneracy. Fortunately this degeneracy is absent for a more general eigenvalue problem, a simple variant of that associated in [TΦ] to the six-vertex model. The equations (2) are replaced by

$$\begin{aligned} w_{k,l} &= -\frac{z_k z_l - 2\Delta z_k + 1}{z_k z_l - 2\Delta z_l + 1}, \qquad k \neq l, \\ R(z_k) &= \prod_{l \neq k} w_{k,l}. \end{aligned} \tag{3}$$

Δ is a parameter whose value will at first be chosen to be generic. The function R is a rational function of degree N with zeros $\alpha_1, \dots, \alpha_N$ and poles $\beta_1, \dots, \beta_N$. The α_i are arbitrary but

$$\beta_i = 2\Delta - 1/\alpha_i = A(\alpha_i). \tag{4}$$

This equation defines the fractional linear transformation A. The equations (3) can be studied for generic values of Δ and generic (with respect to the constraints imposed) R.

We recall that the term generic simply means that the pertinent parameters (at present Δ and the $N+1$ parameters needed to specify R, for example $\alpha_1, \dots, \alpha_N$ and $R(\infty)$, thus $N+2$ in all) are required to lie outside a countable collection of algebraic subvarieties of dimension less than $N+2$. If we can calculate eigenvectors and eigenvalues for generic values of the parameters then, taking limits, we can calculate them for any values. Generic values of the parameters are thus values that do not satisfy any of some countable collection of non-trivial equations, the equations themselves to be determined (explicitly or implicitly) in the course of analyzing the problem. For example, in the present problem the zeros and poles of R are to be distinct and the matrix of A

$$\begin{pmatrix} 2\Delta & -1 \\ 1 & 0 \end{pmatrix} \tag{5}$$

may not have eigenvalues that are roots of unity. Thus Δ may not take any of the values $\cos(a\pi/b)$, a/b rational, although these are far from the sole constraints.

There are several trivial, yet basic, observations to be made about the equations (3). First of all, if Q is a permutation and

$$Q(z)_k = z_{Q^{-1}k}, \qquad Q(w)_{k,l} = w_{Q^{-1}k,Q^{-1}l}$$

then $(Q(z), Q(w))$ is a solution whenever (z, w) is. Since

$$\sum Q(z)^{Pm} Q(w)_P = w_Q^{-1} \sum z^{Pm} w_P,$$

the associated eigenvalue is not changed and the associated eigenvector not changed in any essential way. Moreover if Q is simply an interchange of two integers k and l then $w_Q = w_{k,l}$. Thus whenever two coordinates z_k and z_l are equal $w_Q = -1$; the vector defined by (1) is zero; and the solution of (3) is not admissible because it leads to nothing. Finally, there are solutions of (3) for which $z = \alpha_k$ or $z = \beta_k$. For these solutions, some $w_{k,l}$ is zero or infinity and these solutions are also not admissible because (1) is then not well defined.

A solution will therefore be called *admissible* if:

(1) all the coefficients z_k are different;

(2) no z_k is equal to zero (or infinity) or to a zero or pole of the function R and no $w_{k,l}$ is zero or infinite.

The admissible solutions come in sets with $r!$ elements, any two elements in these subsets differing by a permutation. Since we need $\binom{N}{r}$ vectors we need at least $N(N-1)\dots(N-r+1)$ solutions. The principal theorem of [BL] is easily stated.

Theorem *For generic Δ and R there are exactly $N(N-1)\dots(N-r+1)$ admissible solutions (z, w) of the equations (3). The vector (1) attached to such a solution is not zero and the collection of vectors obtained in this way generate the space $\mathfrak{X}_r$.*

It will become clear that it is easy to find generically at least $N(N-1)\dots(N-r+1)$ admissible solutions whose associated eigenvectors generate the full space $\mathfrak{X}_r$. The difficult assertion is that there are *exactly* this many solutions, and we concentrate on it.

To count the number of admissible solutions we have to count the number of all solutions and then subtract the number of inadmissible solutions. The equations (3) can be regarded as defining the fixed points of an at first imprecisely defined *algebraic correspondence*. Let

$$X = \prod_{k=1}^{r} \mathbb{P}^1 \times \prod_{1 \le k < l \le r} \mathbb{P}^1 = Z \times W.$$

The coordinates are z_k and $w_{k,l}$ with redundant coordinates $w_{l,k} = w_{k,l}^{-1}$. For the moment take $C = X$ and consider the two mappings φ and ψ of C into X defined (inadequately) by

$$\begin{aligned}
&\varphi(z,w) = (z', w'), \quad z'_k = R(z_k), \quad w'_{k,l} = w_{k,l}, \\
&\psi(z,w) = (z'', w''), \ z''_k = \prod_{l \ne k} w_{k,l}, \ w''_{k,l} = Q(z_k, z_l) = -F(z_k, z_l)/F(z_l, z_k),
\end{aligned} \tag{6}$$

with $F(z, z') = F_\Delta(z, z') = zz' - 2\Delta z + 1$. The equations (3) define the points p for which $\varphi(p) = \psi(p)$.

Before discussing the failings of the definition of the correspondence (φ, ψ) we observe that the inadmissible solutions can be defined as the fixed points of similar correspondences. Suppose that $\{A_1, \dots, A_s, B'_1, \dots, B'_t, B''_1, \dots, B''_t\}$ (denoted more compactly (A, B)) is a disjoint decomposition of $\{1, \dots, r\}$ into non-empty subsets and that to each l, $1 \le l \le t$, there is associated a zero $\alpha(l)$ of R. Set $i \equiv j$ if i and j belong to the same A_l or if $i \in B'_l$, $j \in B'_m$, and $\alpha(l) = \alpha(m)$, or finally if $i \in B''_l$, $j \in B''_m$, and $\alpha(l) = \alpha(m)$. Define a sub-variety $X^{A,B}$ of X by the conditions

(1) $z_i = z_j$ and $w_{i,m} = w_{j,m}$ if $i \equiv j$;

(2) if $i \equiv j$ then $w_{i,j} = -1$;
(3) if $i \in B'_l$ then $z_i = 0$;
(4) if $i \in B''_l$ then $z_i = \infty$;
(5) $i \in B'_{l'}$ and $j \in B''_{l''}$ with $\alpha(l') = \alpha(l'')$ then $w_{i,j} = 0$ and $w_{j,i} = \infty$;
(6) if $\alpha' = \alpha(l') \neq \alpha(l'') = \alpha''$ and $\beta' = A(\alpha')$, $\beta'' = A(\alpha'')$ then
  (a) $w_{i,j} = Q(\alpha', \alpha'')$ if $i \in B'_{l'}$ and $j \in B'_{l''}$,
  (b) $w_{i,j} = Q(\alpha', \beta'')$ if $i \in B'_{l'}$ and $j \in B''_{l''}$,
  (c) $w_{i,j} = Q(\beta', \alpha'')$ if $i \in B''_{l'}$ and $j \in B'_{l''}$,
  (d) $w_{i,j} = Q(\beta', \beta'')$ if $i \in B''_{l'}$ and $j \in B''_{l''}$.

If $t = 0$ and $s = r$ then $X^{A,B}$ is X itself. It is convenient to set

$$\mathbb{B}'_\alpha = \cup_{\{l|\alpha=\alpha(l)\}} B'_l, \qquad \mathbb{B}''_\alpha = \cup_{\{l|\alpha=\alpha(l)\}} B''_l,$$

and $\mathbb{B}_\alpha = \mathbb{B}'_\alpha \cup \mathbb{B}''_\alpha$. It is not excluded that one or the other of $\mathbb{B}'_\alpha$ and $\mathbb{B}''_\alpha$ is empty but then both are.

The variety C is replaced, again for a provisional definition of the correspondence, by a sub-variety $C^{A,B}$ defined by the conditions (1) (2) (5) and (6) together with the following modifications of (3) and (4):

(3) if $i \in B'_l$ and $\alpha = \alpha(l)$ then $z_i = \alpha$;
(4) if $i \in B''_l$ and $\alpha = \alpha(l)$ then $z_i = \beta = A(\alpha)$. It is clear how to restrict (φ, ψ) to $C^{A,B}$ to obtain $(\varphi^{A,B}, \psi^{A,B})$.

There is an obvious order on the decompositions $\{A, B\}$. The decomposition $\{\tilde{A}, \tilde{B}\}$ is deeper than $\{A, B\}$ if it imposes more conditions. We write $\{\tilde{A}, \tilde{B}\} \prec \{A, B\}$. We attach to a fixed point $p = (z, w)$ the decomposition $\{A(p), B(p)\}$ in which each B_l is of the form $B'_l = \{i|z_i = \alpha\}$, $B''_l = \{i|z_i = \beta = A(\alpha)\}$, $\alpha(l) = \alpha$, and for which i and j are in a common A_l if and only if $z_i = z_j$ and z_i is neither a zero nor a pole of R. It is pretty clear, apart from the unresolved ambiguities in the correspondences, that a fixed point in X lies in $X^{A,B}$ if and only if $\{A(p), B(p)\}$ is deeper than $\{A, B\}$.

To each set A_l or $B_l = B'_l \cup B''_l$ of a decomposition we attach the weight $(-1)^{n-1}(n-1)!$, n being the number of integers in the set. To the decomposition itself we attach the weight $\omega(A, B)$ equal to the product of the weights of all terms in the decomposition. Thus the weight of the decomposition $\{1, 2\}$ into $\{\{1\}, \{2\}\}$ is 1 and of that into $\{\{1, 2\}\}$ is -1. That of $\{1, 2, 3, 4, 5, 6, 7\}$ into $\{\{1, 2, 3\}, \{4, 5, 6, 7\}\}$ is -12. The following lemma is easily verified.

Lemma *For each fixed point p in X the sum*

$$\sum_{(A(p),B(p))\prec(A,B)} \omega(A, B)$$

is 0 unless p is admissible, but then it is 1.

Thus if $n(A, B)$ is the number of fixed points on $X^{A,B}$ it suffices to verify that

$$N(N-1)\dots(N-r+1) = \sum_{(A,B)} \omega(A, B)n(A, B). \tag{7}$$

Long ago, in the twenties, Lefschetz ([L]) introduced a topological formula not for the number of fixed points of a correspondence but for the number of fixed points counted with multiplicities. If all multiplicities are one, or even if a given fixed point p has the same multiplicity for all correspondences $(\varphi^{A,B}, \psi^{A,B})$ of which it is a fixed point and this multiplicity is one if the point is admissible, then it is legitimate to substitute for $n(A, B)$ in (7) the number $\lambda(A, B)$ of fixed points counted with multiplicity. Since part of our strategy is to show that all these multiplicities are one for generic values of the parameters, we shall use this new form of (7).

There is a disagreeable complication. The Lefschetz formula expresses the number of fixed points in terms of topological data associated to the correspondence. These data are relatively easy to determine, at least in a combinatorial form, from the inadequately defined correspondences with which we have dealt so far because the underlying topological spaces are products of the Riemann sphere with itself, so that their homology has a simple structure. The true correspondences are obtained by blow-ups that are yet to be described. One of our principles has been to make the count, in a manner that cannot be completely correct, with the ill-defined correspondences, leaving the complete justification, which should pose no problems, until the more serious algebraic difficulties are out of the way.

The homology of Z is spanned by products of an arbitrary number of the factors $\mathbb{P}^1$ appearing in Z. An element of this basis is therefore given by a subset of $\{1,\ldots,r\}$. An element of the analogous basis of the homology of W is given by a collection of unordered pairs $\{k,l\}$, $k\neq l$. A basis for the homology of X is obtained by taking all products of a basis element for Z with one for W. In particular, the homology groups of X are non-zero only in even dimensions. There are maps $\varphi_i : H_i(C)\to H_i(X)$, and by Poincaré duality an associated map $\varphi^i : H_i(X)\to H_i(C)$, as well as a map $\psi_i : H_i(C)\to H_i(X)$. The composition $\psi_i\varphi^i$ takes $H_i(X)$ to itself and the Lefschetz formula is

$$\lambda(X)=\sum_{i=0}^{dim\,X}\mathrm{Tr}\,(\psi_{2i}\varphi^{2i}). \tag{8}$$

There is of course a similar formula for $\lambda(X^{A,B})=\lambda(A,B)$.

The map φ is so simple that it is easy to determine φ_{2i}. It multiplies a basis element whose Z component is the product of s factors $\mathbb{P}^1$ by N^s. From this it is easy to determine φ^{2i}. The map ψ on the other hand interchanges the two factors Z and W. According to the formulas (6), $w_{k,l}$ is for fixed z_l a fractional linear function of z_k and for fixed z_k a fractional linear function of z_l. As a consequence the class in $H_2(Z)$ associated to the index k is mapped to the sum of the classes in $H_2(W)$ associated to those indices $\{l,m\}$ for which $m=k$ or $l=k$. In the same way z_k is a linear function of $w_{k,l}$ for each l. As only the classes of the product $X=Z\times W$ that appear in the representation of their own image contribute to the trace, we obtain a combinatorial expression for this trace. It is the sum over a collection of oriented graphs G on the set $\{1,\ldots,r\}$ of $N^{\pi(G)}$, $\pi(G)$ being the number of connected components of G.

The graphs are subject to two constraints:

(1) they are trees;

(2) there is at most one bond issuing from each point.

The meaning of these conditions is easily explained. Suppose S is a subset of $R=\{1,\ldots,r\}$ and T a subset of unordered pairs $\{k,l\}$. The basis element

$$\eta=\prod_{k\in S}\mathbb{P}^1\times\prod_{\{l,m\}\in T}\mathbb{P}^1$$

will contribute either 0 or 1 to the trace. To each basis element contributing 1 we associate a graph G. For such a basis element the cardinality of S must be that of T. Moreover it must be possible to assign to each $\{l,m\}\in T$ a k in $\{l,m\}$ and in S so that as z_k moves so does $w_{l,m}$. Thus $k\in\{l,m\}$ and $\{l,m\}$ is the bond issuing from k. Since a given z_k can not be responsible for the movement of two factors in W, we attach different vertices to different bonds, and the equality $|S|=|T|$ assures us that there is exactly one bond issuing from a vertex in S. The vertices of $\{1,\ldots,r\}$ not in S are the final points of G and some may be isolated.

Suppose there were a cycle in G with bonds $\{k_1,k_2\},\ldots,\{k_p,k_1\}$. Then the number of remaining bonds is equal to the number of remaining vertices in S and each of the remaining bonds is responsible for the movement of one of the remaining vertices. Thus the effective movement in $(z_{k_1},\ldots,z_{k_p})$ is achieved by the bonds in the cycle. However if $P=\{k_1,\ldots,k_p\}$ and Q is its complement in R

$$\prod_{k\in P}z_k''=\prod_{k\in P}\prod_{l\in Q}w_{k,l}. \tag{9}$$

Thus the product is independent of the variables $w_{k_i,k_{i+1}}$ and the image of the class attached to η is zero. Equation (9), valid for any set P, is basic and will be used again, but without comment.

The Lefschetz numbers attached to $(\varphi^{A,B}, \psi^{A,B})$ are calculated in a similar way with a similar result. The vertices of the graphs are now $A_1, \ldots, A_s$ and $B_1, \ldots, B_t$. Each vertex carries a weight, the cardinality of the respective set; so does each graph, the product over the directed bonds of the weights of the vertices at which the bond ends. Moreover no bond may issue from a vertex B_i. Thus B_i is *forbidden* as the *source* of a bond.

Although it is possible to carry out a good part, but not all, of the combinatorial discussion without them, it turns out that these numbers defined combinatorially can be expressed by simple algebraic formulas that, together with their proofs, are due to Fan Chung. It is sufficient to give a formula for the number of connected graphs of the described type. The number of forbidden vertices is then at most one. Suppose there are n vertices with weights $\omega_1, \ldots, \omega_n$.

Lemma(F. Chung)

(1) If there are no forbidden vertices the number of connected graphs counted with weight is $(\omega_1 + \ldots + \omega_n)^{n-1}$.

(2) If there is one forbidden vertex of weight ω *the number of connected graphs counted with weight is* $\omega(\omega_1 + \ldots + \omega_n)^{n-2}$.

This lemma simplifies, for example, the proof of formula (7). The remaining obstacle is the proof that the multiplicites of the fixed point are generically one, and it is surprisingly difficult to overcome. The bulk of our effort has had to be devoted to this.

A closer examination of the description of the effect of ψ, or more generally $\psi^{A,B}$, on cycles reveals the inadequacy of the formulas (6). It is possible that $F(z_k, z_l)$ and $F(z_l, z_k)$ vanish simultaneously so that $w''_{k,l}$ is not well defined, not even as infinity, or that $w_{k,l} = 0$ while $w_{k,m} = \infty$ so that z''_k is not well defined. The second possibility has to be dealt with, but its consequences are largely technical, while an adequate understanding of the consequences of the first is central to the argument.

It is necessary to blow up C by introducing new coordinates $(u_{k,l}, u_{l,k})$ and $\lambda_{k,l}$. The first of these are projective coordinates, so that only their ratio matters, and to fix $\lambda_{k,l}$ it can be supposed that one or the other is 1. They are subject to one basic relation and to all relations that can be deduced from it: $u_{k,l} F(z_l, z_k) = u_{l,k} F(z_k, z_l)$. Of course these new coordinates are redundant except where $F(z_l, z_k) = F(z_k, z_l) = 0$. Additional blow-ups are necessary at $\Delta = 0$ when one of the z_k is a zero or a pole of the function $R(z)$. The blow-ups are carried out in a systematic way, so that the result is a family of correspondences parametrized by Δ with generically smooth C_Δ, or more generally $C_\Delta^{A,B}$.

The fiber is not smooth at $\Delta = 0$. We proceed, however, by first analyzing the fixed points of the correspondence at this fiber, and then studying its deformations. The first point, less evident than we at first thought, is that in a neigborhood of $\Delta = 0$ and thus generically the fixed points of the modified, well-defined correspondences $(\varphi_\Delta^{A,B}, \psi_\Delta^{A,B})$ are isolated, although perhaps with multiplicity greater than one. This is not true at $\Delta = 0$ where the variety of fixed points has a large number of components of positive dimension.

The next point is to show that appropriate fixed points at $\Delta = 0$ deform to give at least $\lambda(A, B)$ distinct fixed points in its neighborhood. It then follows immediately from the Lefschetz formula for $(\varphi_\Delta^{A,B}, \psi_\Delta^{A,B})$ that these are all the fixed points and that they have multiplicity one. The constraints imposed by (A, B) do not affect the form of the equations. They need only be taken into account when counting the number of solutions.

At $\Delta = 0$

$$F(z_k, z_l) = F(z_l, z_k) = z_k z_l + 1$$

so that unless $z_k z_l = -1$ at a fixed point the value of $w_{k,l}$ there is -1. As a consequence at a fixed point p the collection $\{1, \ldots, r\}$ breaks up naturally into subsets $E_1, F_1, \ldots, E_m, F_m$, two indices k and l lying in the same subset if and only if $z_k = z_l$ and lying in *opposed* subsets E_i and F_i if and only if $z_k z_l = -1$. (For a generic choice of the function R the values of the z_k at a fixed point are finite and not zero, and $z_k^2 \neq -1$.) Moreover when k and l do not lie in opposed subsets then $w_{k,l} = -1$.

Thus the equations decouple and as a first approximation it is convenient to deform into a neighborhood of $\Delta = 0$ as though $w_{k,l}$ continued to be -1 when k and l are not opposed.

Let γ_i and δ_i be the values of the coordinates in E_i and F_i, of which one can be empty but not both. There are two types to distinguish: either $\{\gamma_i, \delta_i\}$ is not a pair $\{\alpha, \beta\}$ for some zero and matching pole of R or it is. In the second case neither set is empty.

The first case is, however, the easiest to treat. If one of the two sets is empty then for all k in the other

$$R(z_k) = (-1)^{r-1}.$$

For generic R this equation has N distinct solutions (of which no two are opposed) and they deform uniquely to a whole neighborhood of $\Delta = 0$. Choosing distinct $z_1, \ldots, z_r$ among these solutions we obtain $N(N-1)\ldots(N-r+1)$ admissible fixed points. (In each pair (E_i, F_i) of the associated configuration one set is then empty and the other contains a single element.) Thus it is to be expected that all other fixed points will deform to inadmissible fixed points. The existence of these solutions means that the existence generically of admissible solutions sufficient to generate all eigenvectors is rather simple. The significant mathematical assertion is that there is no redundancy. It is unlikely that this is without consequences for the set of eigenvalues, but we have not had an opportunity to pursue the matter very far.

Continuing to suppose that γ_i is neither a zero nor a pole of R we suppose that neither E_i nor F_i is empty. Let their cardinalities be μ and ν. The number $\rho = R(z)R(-1/z)$ is independent of z at $\Delta = 0$ and may be supposed generic. Multiplying the equations $R(z_k) = \prod_l w_{k,l}$, $k \in E$ and $k \in F$, together we find that

$$(-1)^{\mu(r-\nu)} R(\gamma_i)^{\mu} \rho^{\nu} = (-1)^{\nu(r-\mu)} R(\gamma_i)^{\nu}. \tag{10}$$

Since ρ is generic we conclude that $\mu \neq \nu$ and that there are $|\mu - \nu| N$ possibilities for γ_i.

At $\Delta = 0$ and for $k \in E_i$, $l \in F_i$ it is best not to use the projective coordinates $(u_{k,l}, u_{l,k})$ but new coordinates $(t_{k,l}, v_{k,l})$ such that

$$\frac{u_{k,l}}{u_{l,k}} = \frac{v_{k,l} - 2a_{k,l}t_{k,l}}{v_{k,l} - 2t_{k,l}}, \qquad a_{k,l} = z_k/z_l.$$

When $z_k = \gamma_i$ and $z_l = \beta_i = -1/\gamma_i$ then $a_{k,l} = a_i = -\gamma_i^2$, and for generic parameters this is never 1 at a fixed point. Disregarding the possibility that there may be solutions with some $t_{k,l} = 0$, we take all $t_{k,l}$, $k \in E_i$, $l \in F_i$ equal to 1. Then the $v_{k,l}$ are not independent but may be written as $v_{k,l} = r_k + s_l$. There is an indeterminacy in these new parameters; all the r_k may be replaced by $r_k + t$ and the s_l by $s_l - t$.

Upon decoupling, the equations that refer to the coordinates in E_i and F_i are at $\Delta = 0$

$$\begin{aligned} c_i &= \prod_{l \in F_i} \frac{r_k + s_l - 2a_i}{r_k + s_l - 2}, \qquad & d_i &= \prod_{k \in E_i} \frac{r_k + s_l - 2a_i}{r_k + s_l - 2}, \\ c_i &= (-1)^{r-\mu} R(\gamma_i), & d_i &= (-1)^{r-\nu} R^{-1}(\delta_i). \end{aligned} \tag{11}$$

The values of c_i and d_i in (11) are generic, and in contrast to the Bethe equations themselves it is possible to show directly that each solution of these equations is of multiplicity one, an assertion that remains meaningful even when the equations are recoupled in a neighborhood of $\Delta = 0$.

The solutions of (10) are counted differently for the different correspondences $(\varphi^{A,B}, \psi^{A,B})$ because there are different constraints imposed. For given (A, B) the set E_i and the set F_i must each be the union of the A_l it contains, for at present γ_i is neither an α_k nor a β_k. If E_i contains e elements of weights $\mu_1, \ldots, \mu_e$ and F_i contains f elements of weights $\nu_1, \ldots, \nu_f$ then $\mu = \mu(E_i) = \sum \mu_i$, $\nu = \mu(F_i) = \sum \nu_j$ and the number of solutions of (11), counted by determining the degrees of the maps involved – not forgetting that (11) implies (10) – is $|\mu - \nu| \mu^{f-1} \nu^{e-1} N$.

Before commenting on the difficult case that $\{\gamma_i, \delta_i\}$ is a pair $\{\alpha, \beta\}$, we examine the nature of the combinatorial problem that remains. The Lefschetz formula yields the number of fixed points with multiplicities as a polynomial in N. The coefficient of a given power N^s is the number of graphs of a certain type counted with weights but with s connected components. This number is thus the sum over all decompositions of $\{1, \ldots, r\}$ into s subsets of the product from 1 to s of the number, counted with weights, of *connected* graphs on the elements in the corresponding subset. This structure is shared by the other number to which it is to be compared. The set $\{1, \ldots, r\}$ is decomposed into the sets $D_i = E_i \cup F_i$ and the equations, to a first approximation that then has to be improved, decoupled. For each element of the decomposition there are equations to be solved, and the total number of solutions is obtained by summing over all partitions of the product of the number of solutions attached to the elements of the partition.

For each collection of vertices D and all possible decompositions $D = E \cup F$ (the elements defining the decomposition of $\{1, \ldots, r\}$ are now themselves decomposed) we have to count the solutions of each type observing that if D contains forbidden vertices only solutions of the second type are possible. In all cases, it is essential that the result is a combinatorial factor, independent of N, times N. This is so for solutions of the first type. In all there are

$$N \sum_{E \cup F = D} |\mu(E) - \nu(F)| \mu(E)^{f-1} \mu(F)^{e-1} \tag{12}$$

where e and f are, as before, the numbers of elements in E and F. The sum runs over disjoint partitions of D without regard to order. It is not difficult to show that

$$\begin{gathered} \sum_{E \cup F = D} (\mu(E) + \nu(F)) \mu(E)^{f-1} \mu(F)^{e-1} = \mu(D)^{d-1}, \\ d = |D|, \quad \mu(D) = \sum_{A_l \in D} \omega_l \end{gathered} \tag{13}$$

so that, if D contains no forbidden vertices, there are N times

$$2 \sum_{E \cup F = D} min\{\mu(E), \nu(F)\} \mu(E)^{f-1} \mu(F)^{e-1} \tag{14}$$

deformations of solutions of the second type to be found. If D contains forbidden vertices there are only solutions of the second type and the pertinent analogue of (14) is

$$\omega \mu(D)^{d-2} = \sum \{ (\mu(E) + \omega')^f (\mu(F) + \omega'')^e - \mu(E)^f \mu(F)^e \} \tag{15}$$

if $\omega = \omega' + \omega''$ is the weight of the forbidden vertex of D and the sum is over decompositions of the set $\{A_1, \ldots, A_s\}$ of the set of remaining vertices into two subsets E and F. Recall that the weight $\omega = |B|$ of a forbidden vertex is the sum of $\omega' = |B'|$ and $\omega'' = |B''|$.

For pairs (E_i, F_i) of the second type the source of the factor N is clear for there are N possibilities for α. The combinatorics are, however, less transparent because a crude count of possibilities does not distinguish between two disjoint pairs (E_i, F_i) and (E_j, F_j) associated to the same α_k and the pair $(E_i \cup E_j, F_i \cup F_j)$. Although we have yet to deal with this problem in general, it cannot be major. A second difficulty that we also expect to overcome is more serious. The decoupled equations of the first type define at $\Delta = 0$ a variety of positive dimension. Moreover, although only a finite number of points on this variety admit a deformation to a neighborhood of $\Delta = 0$, there may be several such deformations, so that we have to count not the number of such points but the number of curves (parametrized by Δ) passing through them.

We do not attempt to describe what we know at present, for our knowledge is incomplete. Yet a few words to indicate the source of the difficulties are appropriate. At $\Delta = 0$ and when γ_i is a zero α of R and $\delta_i = A(\alpha)$ it is necessary, in order to define the correspondences completely, to blow C up further by the introduction of projective coordinates (p_k, q_k), $k \in E_i \cup F_i$, such that

$$\begin{aligned} z_k - \alpha = \xi_k p_k, \quad \Delta = \xi_k q_k, \quad k \in E_i, \\ \tilde{z}_k - \tilde{\beta} = \xi_k p_k, \quad \Delta = \xi_k q_k, \quad k \in F_i. \end{aligned}$$

The coordinates ξ_k have to be introduced at the same time as the supplementary projective coordinates, and (for convenience) in equations containing $\tilde{z}_k = 1/z_k$ rather than z_k.

In a form that should be taken as a first symbolic approximation the equations for the variables attached to $E_i \cup F_i$ are

$$
\begin{aligned}
A\Delta p_k &= (-1)^{r-\mu} \prod_{l \in F_i} \frac{p_k + p_l}{p_k + p_l - 2b}, \\
B\Delta p_l &= (-1)^{r-\nu} \prod_{k \in E_i} \frac{p_k + p_l}{p_k + p_l - 2b}.
\end{aligned} \tag{16}
$$

The number $b = 1 + \alpha^2$ is not 0. Moreover A is the derivative of $R(z)$ at α and $B = -A/\rho$. As a result of the presence of Δ on the left side these equations can be solved at $\Delta = 0$ simply be demanding that for each k in E_i or F_i there be an l in the other set such that $p_k + p_l = 0$. To find the solutions that deform and their deformations one looks not for numerical solutions in p_m but for solutions that are Puiseux expansions in Δ, not all constant.

The strategy is to find as many solutions as demanded by the combinatorics and the Lefschetz formula, and therefore to deduce indirectly that we have all. We have no idea whether this might be done directly; it certainly is not promising. We also observe that it is by no means obvious that the deformations of solutions inadmissible at $\Delta = 0$ should remain inadmissible. Yet this is so – or rather our analysis confirms it so far! There are now two combinatorial elements that arise in the algebra and that must be taken into account in the combinatorial comparison with (14) and (15). Not only must the number of solutions of various equations be found, but also, as the first step, the possible leading exponents of the Puiseux expansions calculated. Even for very low values of e and f the analysis is far from transparent.

The ultimate purpose of the Bethe Ansatz is useful insight into the behavior of the eigenvalues of the operator H. To this end Bethe introduced the notion of *Wellenkomplex* for which the customary terminology is now the less formal *string*. At $\Delta = 1$, the value of the parameter that Bethe treated, this notion has limitations, so that clarifications are necessary, but near $\Delta = \infty$ there is a precise, general definition. Since admissible solutions remain admissible as the parameters are deformed, the strings can be transported to 1, but their relation with those of Bethe will not be examined here.

For large $|\Delta|$, set $\lambda_\pm = 1 \pm \sqrt{1 - 1/\Delta^2}$. The matrix

$$
V = \begin{pmatrix} \kappa\lambda_+ & \lambda_- \\ \kappa/\Delta & 1/\Delta \end{pmatrix}
$$

diagonalizes the matrix (5),

$$
V^{-1}AV = \Delta \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}
$$

As a consequence the Bethe equations take a simpler form when the unknowns z_i are replaced by $u_i = V^{-1}(z_i)$. Define κ so that $\kappa^2 = \lambda_-/\lambda_+$ and so that κ behaves like $1/2\Delta$ when $\Delta \to \infty$.

Adapting the ideas of [B] to our purposes we fix a decomposition of $\{1, \dots, r\}$ into ordered subsets $S_1 = \{u_{1,1}, \dots, u_{1,L_1}\}, S_2, \dots$ of lengths $L_1, L_2, \dots$, and look for solutions that are Laurent series in κ,

$$
u_{l,i} = \kappa^{L_l + 1 - 2i} v_{l,i} + \dots, 1 \le i \le L_l.
$$

The zeros α_i have to be so chosen that $\alpha_i = \epsilon_i \kappa$, where ϵ_i is itself a power series in κ with generic initial term, and

$$
R(z) = c \prod \frac{z - \alpha_i}{z/\beta_i - 1},
$$

where c is also a power series in κ with generic initial term. Once again, if we find enough solutions of this type generically admissible in a neighborhood of $\Delta = \infty$ then we have them all. It turns out that the correct number of solutions is obtained by taking all possible decompositions, and by choosing the leading coefficients $v_{l,i} = v_l$

to be independent of i, to be different if the associated strings have the same length ($v_l \neq v_{l'}$ if $L_l = L_{l'}$), and to satisfy the system of equations

$$d_l v_l^N = \prod_{k \neq l} \left(\frac{v_l}{v_k} \right)^{X(L_l, L_k)} \tag{17}$$

with

$$X(l,m) = \begin{cases} 2l - 1, & \text{if } l = m \\ 2\min(l,m), & \text{if } l \neq m. \end{cases}$$

The number d_l is determined by the leading coefficients ϵ_i^0 and c_0 of the series for ϵ_i and c,

$$d_l = c_0^{L_l} (-1)^{N + L_l(r - L_l)} \prod_{i=1}^{N} (\epsilon_i^0)^{L_l - 1}.$$

That there are enough solutions of (17) is a simple combinatorial problem; the implicit function theorem can then be used to establish that the desired solutions in series exist. We observe that the notion of strings is only useful when $2r \leq N$ and that there is a duality in the equations that permits the replacement of r by $N - r$, adding that our argument does not yet deal with the case $2r = N$.

We have explicit and complete sets of admissible solutions near 0 and near ∞ that can be analytically continued along paths, either closed, starting near 0 or ∞ and ending where they began, or passing from 0 to ∞. It may well be of considerable mathematical interest to examine the effect on the solutions and the partitions into strings of different choices of path for there is very likely ramification at some non-generic points. We have, however, only begun to experiment.

Bibliographie

[B] H. Bethe, *Zur Theorie der Metalle,* Zeit. f. Physik, **71**, 205-226 (1931)

[BL] R. P. Langlands and Y. Saint-Aubin, *Les sommes de Bethe et la formule de Lefschetz,* in preparation.

[L] S. Lefschetz, *Intersections and transformations of complexes and manifolds,* Trans. A.M.S., **28**, 1-49 (1926)

[TΦ] Л. А. Тахтаджян и Л. Д. Фаддеев, Квантовой метод обратной задачи у XYZ модель Гейзенберга, Успехи Мат. Наух, **34**, вып. 5, 13-63 (1979)