
The (quantum) Yang–Baxter equation
\begin{equation}
R(z)^{12}R(z + w)^{13}R(w)^{23} = R(w)^{23}R(z + w)^{13}R(z)^{12}
\end{equation}

was introduced in the 1960s in the physics literature in the context of statistical mechanics, and it was the starting point of several developments in mathematics, particularly in representation theory and low-dimensional topology, since the 1980s. The Yang–Baxter equation is an equation for a meromorphic function \( z \mapsto R(z) \in \text{End}(V \otimes V) \) (called \( R \)-matrix) of the spectral parameter \( z \in \mathbb{C} \) with values in the endomorphisms of \( V \otimes V \) for some finite-dimensional vector space \( V \). It is an equality in \( \text{End}(V \otimes V \otimes V) \), and the notation with upper indices indicates the factors of a tensor products on which an endomorphism acts. For example,
\[ R(z)^{12} = R(z) \otimes \text{Id}. \]

A basic example of a solution of the Yang–Baxter equation is
\begin{equation}
R(z) = \text{Id} + \frac{1}{z} P, \quad \text{where } P(u \otimes v) = v \otimes u.
\end{equation}

To make immediate contact with a classical topic the reader might be familiar with, let us note that an invertible constant solution \( R \) of (0.1) defines a representation of the braid group \( B_n \) on the \( n \)-fold tensor product \( V^\otimes n \). This representation sends the group generator \( \sigma_i \) that braids the \( i \)th strand with the \((i + 1)\)st strand to \( P^{i,i+1}R^{i,i+1} \). Indeed the Artin relations \( \sigma_i\sigma_j = \sigma_j\sigma_i, |i - j| \geq 2 \) are trivially satisfied and the relation \( \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \) follows from the Yang–Baxter equation.

But before going to the mathematical contexts where this equation appears, let us spend some words on its origin in statistical mechanics and quantum integrable systems. The vector space \( V \) is the space of states of a quantum mechanical system at a site of a one-dimensional crystal. The \( n \)-fold tensor product \( V^\otimes n = V \otimes \cdots \otimes V \) is the state space of the crystal. Out of a solution of the Yang–Baxter equation one constructs \( L(z) = R^{1,n+1}(z) \cdots R(z)^{13}R(z)^{12} \in \text{End}(V \otimes V^\otimes n) \) obeying
\[ R(z)^{12}L(z + w)^{13}L(w)^{23} = L(w)^{23}L(z + w)^{13}R(z)^{12} \in \text{End}(V \otimes V \otimes V^\otimes n). \]

This relation lies at the heart of the quantum inverse scattering method of the Leningrad school (see [26]) that led to the notion of quantum group. One feature is that if \( R(z) \) is invertible for generic \( z \), then the partial traces \( T(z) = \text{Tr}_{V}(L(z)) \in \text{End}(V^\otimes n) \) form a family of commuting endomorphisms. The quantum inverse scattering method is a technique to find simultaneous eigenvalues of this family and is a generalization of the Bethe ansatz, developed by Bethe in the 1930s to determine the spectrum of the Hamiltonian of the Heisenberg spin chain, which appears as the constant term in \( T(z)^{-1}T'(z) \) as \( z \to 0 \) for Yang’s solution (0.2) and two-dimensional
V. In statistical mechanics one considers the trace of $T(z)^m \in \text{End}(V \otimes V)$ which, once written out as a polynomial in the matrix elements of $R$ in some basis, is recognized as a partition function, namely the weighted sum over all configurations of a system on an $n \times m$ rectangular grid on a torus where each link between nearest neighbors can be in $\text{dim}(V)$ different states.

A successful approach to finding solutions of the Yang–Baxter equation is to study the deformation theory of the trivial solution $R = \text{Id}$. Suppose $R_h(z)$ is a solution depending on a parameter $h$ such that $R_h(z) = \text{Id} + h r(z) + O(h^2)$. Then $r(z) \in \text{End}(V \otimes V)$ obeys the classical Yang–Baxter equation

\[ [r(z)^{12}, r(z + w)^{12}] + [r(z)^{12}, r(w)^{23}] + [r(z + w)^{13}, r(w)^{23}] = 0. \]

An important feature of this equation is that it implies that the meromorphic connection on the trivial vector bundle $\mathbb{C}^n \times V \otimes V \to \mathbb{C}^n$ defined by

\[ \nabla = d + k \sum_{i \neq j=1}^n r(z_i - z_j)^{ij}(dz_i - dz_j) \]

is flat for all $k \in \mathbb{C}$. A special case is the Knizhnik–Zamolodchikov \[25\] connection $\nabla_{\text{KZ}}$ of conformal field theory on genus zero curves; see \[27\] for a mathematical introduction. In this case $r(z) = t/z$, where $t \in \mathfrak{g} \otimes \mathfrak{g}$ is an ad-invariant tensor for a semisimple Lie algebra $\mathfrak{g}$ and $V$ is a $\mathfrak{g}$-module.

Let us turn to the dynamical Yang–Baxter equation. Let $V$ be a semisimple finite-dimensional representation of an abelian complex Lie algebra $\mathfrak{h}$. Thus $V = \bigoplus_{\mu \in \mathfrak{h}^*} V[\mu]$ with finitely many nonzero weight spaces $V[\mu]$. The dynamical Yang–Baxter equation is an equation for a meromorphic function $R(z, \lambda) \in \text{End}(V \otimes V)$ of the spectral parameter $z \in \mathbb{C}$ and the dynamical variable $\lambda \in \mathfrak{h}^*$. It is the equation in $\text{End}(V \otimes V)$

\[ R(z, \lambda - h^3)^{12} R(z + w, \lambda)^{13} R(w, \lambda - h^1)^{23} = R(w, \lambda)^{23} R(z + w, \lambda - h^2)^{13} R(z, \lambda)^{12}. \]

The dynamical notation is adopted, meaning that $h^i$ must be replaced by $\mu_i$ when acting on $V[\mu_1] \otimes V[\mu_2] \otimes V[\mu_3]$. This equation appeared in the work of Gervais and Neveu on Liouville theory (without spectral parameter) \[22\]. Its version with spectral parameter was viewed in \[15, 16\] as a genus one theory. The goal was to extend the $R$-matrix formalism to interaction-round-a-face models of statistical mechanics \[2\], which involve elliptic functions. The corresponding dynamical classical Yang–Baxter equation is the flatness condition for Bernard’s extension \[3\] of the Knizhnik–Zamolodchikov connection to conformal field theory on genus one curves. The transfer matrix formalism extends to the dynamical case providing commuting families of difference operators in the dynamical variables; see \[18\]. We discuss these operators below in the setting of dynamical $R$-matrices that are constant as functions of the spectral parameter.

What Etingof and his collaborators recognized is that this theory is interesting and useful in representation theory even without spectral parameter. This book is about this version of the dynamical Yang–Baxter equation and its relation with the representation theory of finite-dimensional Lie algebras and their quantum version.

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ with Cartan subalgebra $\mathfrak{h}$ and an invariant bilinear form $\langle \,,\, \rangle$ normalized so that short roots have squared length 2. Let us fix a system of simple roots $\alpha_1, \ldots, \alpha_r$, and introduce the root lattice $Q = \bigoplus_{i=1}^r \mathbb{Z} \alpha_i$ and its positive cone $Q_+ = \bigoplus_{i=1}^r \mathbb{Z}_{\geq 0} \alpha_i$. The universal enveloping algebra $U\mathfrak{g} = U_1\mathfrak{g}$ of $\mathfrak{g}$ belongs to a one-parameter family of Hopf algebras, the
Drinfeld–Jimbo quantum universal enveloping algebras $U_q\mathfrak{g}$ \[6\] \[24\], defined for $q = e^{\hbar/2} \in \mathbb{C}^\times$. The representation theory of $U_q\mathfrak{g}$ is very similar to the representation theory of $U\mathfrak{g}$ if $q$ is not a root of unity. In particular the adjoint action of $\mathfrak{h}$ on $U\mathfrak{g}$ extends to a weight decomposition $U_q\mathfrak{g} = \bigoplus_{\alpha \in Q} U_q\mathfrak{g}[\alpha]$ by the root lattice. To the Borel subalgebras $\mathfrak{b}_\pm$ there correspond Hopf subalgebras $U_q\mathfrak{b}_\pm$ whose gradings lie in $\pm Q_+$. There is a category $\mathcal{O}$ of finitely generated $U_q\mathfrak{g}$-modules with a weight decomposition $V = \bigoplus_{\lambda \in \mathfrak{h}^*} V[\lambda]$ and such that $U_q\mathfrak{b}_+V$ is finite dimensional for each $v \in V$. In particular we can view $V \in \mathcal{O}$ as a semisimple $\mathfrak{h}$-module. To each $\lambda \in \mathfrak{h}^*$ there corresponds a one-dimensional representation $\mathbb{C}_\lambda$ of $U_q\mathfrak{b}_+$ and a Verma module $M_\lambda = U_q\mathfrak{g} \otimes_{U_q\mathfrak{b}_+} \mathbb{C}_\lambda \in \mathcal{O}$ with highest weight $\lambda$ generated by a highest weight vector $u_\lambda$, the image of $1 \otimes 1$. Simple modules in $\mathcal{O}$ are isomorphic to quotients $L_\lambda$ of Verma modules $M_\lambda$ by their maximal proper submodules. If $q$ is not a root of unity, all finite-dimensional modules are isomorphic to $L_\lambda$ for some dominant integral weight $\lambda$. Since $U_q\mathfrak{g}$ is a Hopf algebra, tensor products and duals of modules are defined. The $R$-matrices appear when comparing tensor products $V \otimes W$ and $W \otimes V$ of representations. The following results are due to Drinfeld; see \[7\] \[8\].

**Theorem 0.1.** Let $q = e^{\hbar/2} \in \mathbb{C} \setminus \{0\}$ be generic. There exist invertible linear endomorphisms $R_{VW} \in \text{End}_\mathbb{C}(V \otimes W)$ for each pair of modules of the category $\mathcal{O}$ such that $P \cdot R_{VW} : V \otimes W \to W \otimes V$ is an isomorphism of $U_q\mathfrak{g}$ modules obeying the quasi-triangularity property

$$R_{U,V \otimes W} = R_{UW}^{13} R_{UV}^{12}, \quad R_{U \otimes V,W} = R_{UW}^{13} R_{VW}^{23},$$

and the Yang–Baxter equation

$$(0.3) \quad R_{UV}^{12} R_{UW}^{13} R_{VW}^{23} = R_{VW}^{23} R_{UW}^{13} R_{UV}^{12},$$

for any triple of modules in the category $\mathcal{O}$. Moreover, $R_{VW}$ is the image in $\text{End}(V \otimes W)$ of a universal $R$-matrix $\mathcal{R}$ of the form $\mathcal{R} = \mathcal{R}_0 e^{\hbar \sum x_i \otimes x_i/2}$ with

$$\mathcal{R}_0 = 1 + \sum_{\alpha \in Q_+ \setminus \{0\}} \mathcal{R}_\alpha \in U_q\mathfrak{b}_+ \otimes U_q\mathfrak{b}_-.$$

Here $(x_i)$ is any orthonormal basis of $\mathfrak{h}$, and $\mathcal{R}_\alpha$ is of weight $\alpha$ in the first factor. Only finitely many summands in this infinite sum contribute nontrivially when acting on an element of $V \otimes W$.

This result provides constant solutions $R_{VV}$ of the Yang–Baxter equation \[0.3\], and transfer matrices are partial traces $T_V = \text{tr}_V R_{VW} \in \text{End}(W)$ with $W = V \otimes \cdots \otimes V$. The commutativity of the transfer matrices

$$T_UT_V = T_V T_U$$

is an easy consequence of \[0.3\]. Solutions of the Yang–Baxter equation with spectral parameter are obtained from $U_q\mathfrak{g}$, where $\mathfrak{g}$ is an affine Kac–Moody algebra, to which the Drinfeld–Jimbo theory also applies. The technical difficulty is that finite-dimensional representations of $U_q\mathfrak{g}$ do not belong to the category $\mathcal{O}$ in general. Still a well-developed theory exists; see \[21\] \[1\] \[5\] \[20\]. In particular there are finite-dimensional evaluation modules $V(z)$ associated to finite-dimensional representations $V$ of $\mathfrak{g}$, and $R$-matrices $R_{VV}(z-w) : V(z) \otimes V(w) \to V(z) \otimes V(w)$ defined for generic $z-w$, so that $PR_{VV}(z-w)$ is a morphism of $U_q\mathfrak{g}$-modules.
Let us go back to finite-dimensional Lie algebras and explain the appearance of the dynamical Yang–Baxter equation in representation theory. Let $V$ be a finite-dimensional $U_q\mathfrak{g}$-module and $v \in V[\mu]$ a vector of weight $\mu$. Then for each generic $\lambda \in \mathfrak{h}^*$ there is a unique morphism of modules

$$\Phi^\mu_\lambda: M_\lambda \to M_{\lambda-\mu} \otimes V$$

sending the highest weight vector $u_\lambda \in M_\lambda$ to $u_{\lambda-\mu} \otimes v$. The fusion operator $J_{VW}(\lambda): V \otimes W \to V \otimes W$ for finite-dimensional $U_q(\mathfrak{g})$-modules $V, W$ is induced by the composition of intertwining operators: if $v \in V[\mu]$ and $w \in W[\nu]$, then

$$(\Phi^\nu_{\lambda-\mu} \otimes \operatorname{Id})\Phi^\mu_\lambda = \Phi^{\mu+\nu}_\lambda: M_\lambda \to M_{\lambda-\mu-\nu} \otimes V \otimes W$$

for some uniquely determined $u = J(\lambda)(v \otimes w) \in V \otimes W$.

**Definition 0.2.** Let $V, W$ be $U_q(\mathfrak{g})$-modules of the category $O$. The exchange operator is

$$R_{VW}(\lambda) = (J_{VW}(\lambda))^{-1} R^{21}_{WV}J_{WV}(\lambda)^{21}.$$

The exchange operators relate the composition of intertwining operators to the composition taken in the opposite order.

**Theorem 0.3.** Let $\rho \in \mathfrak{h}^*$ be the half-sum of positive roots and denote by $\bar{\lambda}$ the image of $\lambda \in \mathfrak{h}^*$ by the isomorphism $\mathfrak{h}^* \to \mathfrak{h}$ defined by the bilinear form. Let $\theta(\lambda) = \bar{\lambda} + \bar{\rho} - \frac{1}{2} \sum x_i^2$ where $x_i$ is any orthonormal basis of $\mathfrak{h}$. Exchange and fusion operators obey the ABRR equation

$$J_{VW}(\lambda)(\operatorname{Id} \otimes q^{2\theta(\lambda)}) = \mathcal{R}_0(\operatorname{Id} \otimes q^{2\theta(\lambda)})J_{VW}(\lambda)$$

and the dynamical Yang–Baxter equation

$$R^{12}_{UV}(\lambda - h^3) R^{13}_{UV}(\lambda) R^{23}_{UV}(\lambda - h^3) = R^{23}_{UV}(\lambda) R^{13}_{UV}(\lambda - h^3) R^{12}_{UV}(\lambda).$$

Moreover, $J_{VW}(\lambda)$ is the image in $\text{End}(V \otimes W)$ of a universal dynamical twist

$$J(\lambda) \in U_q\mathfrak{b}_- \hat{\otimes} U_q\mathfrak{b}_+.$$

**Remark 0.4.** The ABRR equation is due to Arnaudon, Buffenoir, Ragoucy, and Roche [7] who noticed that if $R$ is a universal $R$-matrix in the sense of Drinfeld [7], then $J(\lambda)^{-1} R^{21}J(\lambda)^{21}$ obeys the dynamical Yang–Baxter equation provided $J(\lambda)$ obeys a cocycle condition, which is a dynamical version of Drinfeld’s twist equation. They showed that the dynamical twist equation is implied by the ABRR equation that being linear can be solved by an explicit recursive procedure. Explicit expressions are given in [1] and [23]. The interpretation of the image of $J(\lambda)$ in $\text{End}(V \otimes W)$ as the fusion operator is due to Etingof and Varchenko [13, 9].

With this result one can apply the machinery of transfer matrices and construct families of commuting operators. Let $V$ be a finite-dimensional $U_q\mathfrak{g}$-module with zero weight space $V[0]$. Then each finite-dimensional module $W$ gives rise to a difference operator acting on $V[0]$-valued functions on $\mathfrak{h}^*$:

$$D_W f(\lambda) = \sum_{\nu \in \mathfrak{h}^*} \operatorname{Tr}|_{W[\nu]} R_{WV}(-\lambda - \rho)f(\lambda + \nu).$$

The dynamical Yang–Baxter equation implies that $D_W D_U = D_U D_W$ for any $W, U$. Common eigenvectors of these difference operators are obtained from traces of intertwining operators: the trace function $\psi_V: \mathfrak{h}^* \times \mathfrak{h}^* \to \text{End}(V[0])$ is

$$\psi_V(\lambda, \mu): v \mapsto \operatorname{Tr}|_{M_\mu}(\Phi^\mu_\lambda e^{h\lambda}).$$
Etingof and Varchenko showed that a suitably normalized version of the trace function is a common eigenfunction of the commuting difference operators and has several additional properties.

**Theorem 0.5.** Let \( \delta_q(\lambda) = \prod_{\alpha > 0} (q^{\langle \lambda, \alpha \rangle} - q^{-\langle \lambda, \alpha \rangle}) \) and \( Q(\mu) = \sum_i S(a_i) b_i \), where \( J(\mu) = \sum_i a_i \otimes b_i \) and \( S \) is the antipode of the Hopf algebra \( U_q g \). The normalized trace function \( F_V(\lambda, \mu) = \delta_q(\lambda) \psi_V(\lambda, -\mu - \rho) Q(-\mu - \rho)^{-1} \), regarded as a function of \( \lambda \), is an eigenvector of \( D_W \) with eigenvalue \( \chi_W(q^{-2\mu}) = \sum_\nu q^{-2\langle \mu, \nu \rangle} \dim V[\nu] \). Moreover, if \( V^* \) denotes the representation dual to \( V \),

\[
F_V(\mu, \lambda) = F_V(\lambda, \mu)^*.
\]

A consequence is that \( F_V(\lambda, \mu) \) also obeys a dual difference equation with respect to \( \mu \).

In the special case of \( g = sl_n \) with \( V \) a symmetric power of the vector representation, the difference operators reduce to the Macdonald difference operators, and the Macdonald polynomials for \( sl_n \) can be constructed out of trace functions \([14]\).

Moreover, as shown in \([11]\), the normalized trace function obeys several further identities, including the invariance under the dynamical Weyl group \([10]\), orthogonality relations, and the qKZB heat equation. These identities (and the terminology) are degenerate versions of identities proposed in an attempt \([19]\) to develop a quantum group version of conformal field theory on elliptic curves. Conjecturally, trace functions \( \psi_V(\lambda, \mu) \) are degenerate limits of trace functions for quantum universal enveloping algebras of affine Lie algebras. The classical \((q = 1)\) version of the latter trace functions appears in conformal field theory. They are solutions of the Knizhnik–Zamolodchikov–Bernard equations \([3]\) and have integral representations of hypergeometric type \([17]\). The corresponding statements for \( q \neq 1 \) are known only in very special cases, cf. \([19]\).

This book originates from lecture notes of a course given by the first author at MIT. It contains a hands-on introduction to the representation theory of quantum groups and develops the theory of intertwining operators and the dynamical Yang–Baxter equation from scratch, starting from the classical case \( q = 1 \). The book is an excellent and accessible introduction to the subject. It is a good complement to \([12]\), based on another course of Etingof, where the quantization problem, namely the construction of \( R \)-matrices from solutions of the classical Yang–Baxter equation, and the geometric interpretations in terms of Poisson–Lie group(oid)s are discussed.

**References**


7. ibid., *On quasitriangular quasi-Hopf algebras and on a group that is closely connected with Gal(Q/Q)*, Algebra i Analiz 2 (1990), no. 4, 149–181. MR1080203 (92f:16047)


Giovanni Felder
Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland

E-mail address: felder@math.ethz.ch