Skew left braces and the Yang-Baxter equation

Lindsay N. Childs

Abstract. We give a self-contained, notation-friendly proof that a skew left brace yields a solution of the Yang-Baxter equation.

1. Introduction

A skew left brace is a set $B = (B, \circ, \cdot)$ with two group operations that satisfy the single compatibility condition: for all $x, y, z$ in $B$,

\[(\#) \quad x \circ (y \cdot z) = (x \circ y) \cdot x^{-1} \cdot (x \circ z).\]

The inverse of $x$ in $(B, \circ)$ is denoted $\overline{x}$ and in $(B, \cdot)$ by $x^{-1}$. One easily checks from $(\#)$ that the two groups $(B, \circ)$ and $(B, \cdot)$ share a common identity element, 1. (Let $x = z = 1_\circ$ and $y = 1_\cdot$ in $(\#)$.)

Skew left braces were first defined by Guarneri and Vendramin in [GV17], generalizing the concept of left brace, a concept defined by W. Rump [Ru07] as a generalization of a radical ring.

The primary motivation behind the concept of a brace, and subsequently a skew brace, was to construct algebraic structures that yield set-theoretic solutions of the Yang-Baxter equation. Such a solution is a function $R : B \times B \rightarrow B \times B$ on a set $B$ that satisfies the equation

\[(\ast) \quad (R \times \text{id})(\text{id} \times R)(R \times \text{id})(a, b, c) = (\text{id} \times R)(R \times \text{id})(\text{id} \times R)(a, b, c).\]

for all $a, b, c$ in $B$. This equation has been a question of considerable interest among algebraists since 1990 (motivated by [Dr92]). Solutions of the YBE have been constructed in various settings during the past 25 years (e.g. [LYZ00], [Ru07], [CJO14], [BCJ16]), but the only general descriptions of how a skew left brace yields a solution to the YBE appear in [GV17] and [Ba18].
Beyond their connection to the YBE, skew braces have also been shown in [SV18] to be very closely related to Hopf-Galois structures on Galois extensions of fields—see, for example, [CGK...21] and [ST23].

Skew braces and their role in giving solutions to the YBE were recently introduced to a broad American audience by Vendramin in [Ve24], adapted from a longer survey article [Ve23]. The latter refers only to [GV17] for the proof that a skew brace yields a solution of the YBE. But the proof in [GV17] is not self-contained–it refers to braiding operators, from [LYZ00], and does not explicitly mention Proposition 2.4, below, which is central to the proof.

The referee pointed out that [Ba16], hence [Ba18], gives a self-contained proof of the skew brace-YBE connection that includes Proposition 2.4. But the proofs in [GV17] and [Ba18] involve notation for functions of functions that require multiple layers of subscripts whose complexity obscures what is going on.

This note presents a straightforward, entirely self-contained and notation-friendly proof that a skew left brace yields a solution $R : B \times B \rightarrow B \times B$ of the form

$$R(x, y) = (\sigma_x(y), \tau_y(x))$$

for all $x, y$ in $B$. Here $\sigma_x(y) = x^{-1} \cdot (x \circ y)$ is the well-known $\lambda$-function (or $\gamma$-function, depending on author) associated to a skew brace, and $\tau_y(x)$ is defined by the equation that $\sigma_x(y) \circ \tau_y(x) = x \circ y$. Beyond this equation, the only facts needed for the proof are that $\sigma_x(\sigma_y(z)) = \sigma_x \circ y(z)$ and $\tau_y(\tau_x(z)) = \tau_x \circ y(z)$ (Proposition 2.4), both of which we prove.

The proof of the $\sigma$-result is from [GV17]. The $\tau$-result appears as Lemma 2.4 of [Ba18], but not explicitly in [GV17] and, as will be seen below, is a fundamental contributor to the proof of the main result. There is a proof of the $\tau$ result in [Ba18], but the proof below was obtained independently of [Ba18]. My thanks to the referee for the reference to [Ba16].

2. The proof

Given a skew brace $B = (B, \circ, \cdot)$, define $\sigma_x : B \rightarrow B$ by

$$\sigma_x(y) = x^{-1} \cdot (x \circ y)$$

for all $x, y$ in $B$. Define

$$\tau_y(x) = \sigma_x(y) \circ x \circ y = x^{-1} \cdot (x \circ y) \circ x \circ y.$$

Then for all $x, y$ in $B$, $\sigma_x$ and $\tau_y$ are one-to-one maps from $B$ to $B$, and by definition of $\tau_y(x)$, $\sigma_x(y) \circ \tau_y(x) = \sigma_x(y) \circ (\sigma_x(y) \circ x \circ y) = x \circ y$. Define

$$R : B \times B \rightarrow B \times B$$

by

$$R(a, b) = (\sigma_a(b), \tau_b(a)) = (\sigma_a(b), \sigma_a(b) \circ a \circ b).$$

for all $a, b$ in $B$. Note that if $R(a, b) = (s, t)$, then $s = \sigma_a(b) \circ \tau_b(a) = a \circ b$.

We will prove:
Theorem 2.1. If \( B \) is a skew left brace and \( R : B \times B \to B \times B \) is defined by \( R(a, b) = (\sigma_a(b), \tau_b(a)) \) for \( a, b \) in \( B \), then \( R \) is a solution of the Yang-Baxter equation: for all \( a, b, c \) in \( B \),

\[
(\ast) \quad (R \times id)(id \times R)(R \times id)(a, b, c) = (id \times R)(R \times id)(id \times R)(a, b, c).
\]

Since \( \sigma_a \) and \( \tau_b \) are one-to-one maps from \( B \) to \( B \) for all \( a, b \) in \( B \), the solution \( R \) of the Yang-Baxter equation is nondegenerate.

Proof. Given a skew brace \( B(\cdot, \cdot) \), for \( x, y \) in \( B \) the maps \( \sigma_x(y) = x^{-1} \cdot (x \circ y) \) and \( \tau_x(x) = \sigma_x(y) \circ x \circ y \) satisfy the following two properties for all \( x, y, z \) in \( B \), as we show below:

(i): \( \sigma \) is a homomorphism from \( (B, \circ) \) to \( \text{Perm}(B) \): ,

\[
\sigma_{x \circ y}(z) = \sigma_x(\sigma_y(z));
\]

(ii): \( \tau \) is an anti-homomorphism from \( (B, \circ) \) to \( \text{Perm}(B) \):

\[
\tau_{x \circ y}(x) = \tau_y(\tau_x(x)).
\]

Beside these two properties, the only other property we need is the property noted above:

(iii) if \( \sigma_u(v) = (\sigma_u(v), \tau_u(u)) = (y, z) \), then \( u \circ v = y \circ z \).

These three properties suffice to show that \( R \) satisfies

\[
(R \times 1)(1 \times R)(R \times 1)(a, b, c) = (1 \times R)(R \times 1)(1 \times R)(a, b, c) \quad (\ast),
\]

for all \( a, b, c \) in \( B \), as follows.

The left side of (\( \ast \)) is:

\[
(R \times 1)(1 \times R)(R \times 1)(a, b, c) = (R \times 1)(1 \times R)(d, e, c) = (R \times 1)(d, f, g) = (h, k, g)
\]

where

\[
d = \sigma_d(b), \quad e = \tau_b(a), \quad \text{so} \quad a \circ b = d \circ e,
\]

\[
f = \sigma_e(c), \quad g = \tau_c(e), \quad \text{so} \quad e \circ c = f \circ g,
\]

and

\[
h = \sigma_f(d), \quad k = \tau_f(d), \quad \text{so} \quad d \circ f = h \circ k.
\]

The right side of (\( \ast \)) is:

\[
(1 \times R)(R \times 1)(1 \times R)(a, b, c) = (1 \times R)(R \times 1)(a, q, r) = (1 \times R)(s, t, r) = (s, v, w),
\]

where

\[
q = \sigma_b(c), \quad r = \tau_r(b), \quad \text{so} \quad b \circ c = q \circ r,
\]

\[
s = \sigma_a(q), \quad t = \tau_q(a), \quad \text{so} \quad a \circ q = s \circ t,
\]

and

\[
v = \sigma_t(r), \quad w = \tau_r(t), \quad \text{so} \quad t \circ r = v \circ w.
\]

We want to show that \( (h, k, g) = (s, v, w) \).

To show that \( h = s \) uses property (i): \( \sigma_{y \circ z}(x) = \sigma_{x}(\sigma_{z}(x)) \), as follows:

\[
s = \sigma_d(q) = \sigma_d(\sigma_b(c)) = \sigma_{a \circ b}(c);
\]

\[
h = \sigma_d(f) = \sigma_d(\sigma_e(c)) = \sigma_{d \circ e}(c);
\]
and
\[ d \circ e = \sigma_d(b) \circ \tau_b(a) = a \circ b. \]
So
\[ h = \sigma_{d \circ e}(c) = \sigma_{a \circ b}(c) = s. \]
To show that \( w = g \) uses property (ii): \( \tau_{z \circ y}(x) = \tau_y(\tau_z(x)) \), as follows:
\[ g = \tau_c(e) = \tau_c(\tau_b(a)) = \tau_{b \circ c}(a); \]
\[ w = \tau_r(t) = \tau_r(\tau_q(a)) = \tau_{q \circ r}(a) \]
and
\[ q \circ r = \sigma_b(c) \circ \tau_c(b) = b \circ c. \]
So
\[ w = \tau_{q \circ r}(a) = \tau_{b \circ c}(a) = g. \]
Finally, to show that \( k = v \) we just use property (iii) many times, that for any \( u, v \), if \( R(u, v) = (m, n) \), then \( m \circ n = u \circ v \):
The left side of equation (*) is \( (h, k, g) \); the right side is \( (s, v, w) \), and using all of the equalities above, we have that
\[ s \circ v \circ w = a \circ b \circ c = h \circ k \circ g: \]
For
\[
\begin{align*}
\circ(v \circ w) &= \circ(\sigma_r(t) \circ \tau_r(t)) = \circ(t \circ r) \\
&= (s \circ t) \circ r = \sigma_a(q) \circ \tau_a(q) = \sigma_b(c) \circ \tau_b(c) = \sigma_a(q) \circ \tau_a(q) \\
&= a \circ (q \circ r) = a \circ (\sigma_b(c) \circ \tau_c(b)) = a \circ (b \circ c);
\end{align*}
\]
while
\[
\begin{align*}
(a \circ b) \circ c &= (\sigma_a(b) \circ \tau_b(a)) \circ c = (d \circ e) \circ c \\
&= d \circ (e \circ c) = d \circ (\sigma_b(c) \circ \tau_c(e)) = d \circ (f \circ g) \\
&= (d \circ f) \circ g = \sigma_d(f) \circ \tau_d(f) \circ \tau_d(g) = (h \circ k) \circ g.
\end{align*}
\]
So \( s \circ v \circ w = h \circ k \circ g \). Since \( w = g \), and \( h = s \) in the group \( B, \circ \), it follows that \( k = v \). Given properties (i) and (ii), that completes the proof. \( \square \)

To prove properties (i) and (ii) we need the following consequence of the compatibility condition (#) for a skew brace (c.f. [GV17], Lemma 1.7(2)):

**Lemma 2.2.** For all \( a, b \) in \( B \), \( a^{-1} \cdot (a \circ b^{-1}) \cdot a^{-1} = (a \circ b)^{-1} \).

**Proof.** The compatibility condition(#) for a skew brace is that for all \( x, y, z \) in \( B \),
\[ x \circ (y \cdot z) = (x \circ y) \cdot x^{-1} \cdot (x \circ z), \]
hence
\[ x \cdot (x \circ y)^{-1} \cdot (x \circ (y \cdot z)) = x \circ z \]
or
\[ x \circ z = x \cdot (x \circ y)^{-1} \cdot (x \circ (y \cdot z)). \]
Set \( x = a, y = b, z = b^{-1} \) to get
\[ a \circ b^{-1} = a \cdot (a \circ b)^{-1} \cdot a, \]
or 
\[ a^{-1} \cdot (a \circ b^{-1}) \cdot a^{-1} = (a \circ b)^{-1}. \]

Here is property (i): it is Proposition 1.9 (2) of [GV17].

**Proposition 2.3.** For all \( x, y, z \) in \( B \),
\[ \sigma_{x \circ y}(z) = \sigma_x(\sigma_y(z)). \]

**Proof.** (from [GV17]) The right side of
\[ \sigma_{x \circ y}(z) = \sigma_x(\sigma_y(z)) \]
is
\[ \sigma_x(\sigma_y(z)) = x^{-1} \cdot (x \circ \sigma_y(z)) = x^{-1} \cdot (x \circ (y^{-1} \cdot (y \circ z))) = x^{-1} \cdot (x \circ y^{-1}) \cdot x^{-1} \cdot (x \circ y \circ z) \text{ (by (#))}. \]

By Lemma 2.2, this is
\[ = (x \circ y)^{-1} \cdot (x \circ y \circ z) = \sigma_{x \circ y}(z). \]

(We note that [GV17] proves that given a set \( B \) with two group operations, \( \cdot \) and \( \circ \), and \( \sigma_x(y) = x^{-1} \cdot (x \circ y) \), then for all \( x, y, z \) in \( B \),
\[ \sigma_x(\sigma_y(z)) = \sigma_{x \circ y}(z) \]
if and only if the compatibility condition (\#) holds, if and only if \( B \) is a skew left brace: see Proposition 1.9 of [GV17].)

Finally, we prove property (ii):

**Proposition 2.4.** \( \tau \) is an anti-homomorphism from \( (B, \circ) \) to \( \text{Perm}(B) \): for all \( x, y, z \) in \( B \),
\[ \tau_{y \circ z}(x) = \tau_z(\tau_y(x)). \]

**Proof.** We begin with the definition of \( \sigma_x(q) \):
\[ x^{-1} \cdot (x \circ y) = \sigma_x(y) \]
Rearrange the equation and use that \( x \circ y = \sigma_x(y) \circ \tau_y(x) \), to get:
\[ \sigma_x(y)^{-1} \cdot x^{-1} = (\sigma_x(y) \circ \tau_y(x))^{-1} \]
Apply the Lemma 2.2 formula, \( (a \circ b)^{-1} = a^{-1} \cdot (a \circ b^{-1}) \cdot a^{-1} \) to the right side, to get:
\[ \sigma_x(y)^{-1} \cdot x^{-1} = \sigma_x(y)^{-1} \cdot (\sigma_x(y) \circ \tau_y(x))^{-1} \cdot \sigma_x(y)^{-1} \]
Cancel \( \sigma_x(y)^{-1} \) on the left and multiply both sides by \( (x \circ y \circ z) \) on the right:
\[ x^{-1} \cdot (x \circ y \circ z) = (\sigma_x(y) \circ \tau_y(x))^{-1} \cdot \sigma_x(y)^{-1} \cdot (x \circ y \circ z) \]
Apply the definition of $\sigma$ to the left side and use that $x \circ y = \sigma_x(y) \circ_{\tau_y}(x)$ on the right side:
\[
\sigma_x(y \circ z) = (\sigma_x(y) \circ_{\tau_y}(x)^{-1}) \cdot \sigma_x(x)^{-1} \cdot (\sigma_x(y) \circ (\tau_y(x) \circ z))
\]

Apply the skew brace formula (#) to the right side:
\[
\sigma_x(y \circ z) = \sigma_x(y) \circ (\tau_y(x)^{-1} \cdot (\tau_y(x) \circ z))
\]

Use the definition of $\sigma$ on the far right side:
\[
\sigma_x(y \circ z) = \sigma_x(y) \circ \sigma_{\tau_y(x)}(z)
\]

Take the $\circ$-inverse of both sides, and multiply both sides by $\sigma_x(y \circ z)$:
\[
\overline{\sigma_x(y \circ z)} \circ x \circ y \circ z = \overline{\sigma_{\tau_y(x)}(z)} \circ (\sigma_x(y) \circ x \circ y) \circ z
\]

Use the definition of $\tau$: $\tau_y(a) = \sigma_a(b) \circ a \circ b$ on the right side:
\[
\overline{\sigma_x(y \circ z)} \circ x \circ (y \circ z) = \overline{\sigma_{\tau_y(x)}(z)} \circ \sigma_{\tau_y(x)}(x) \circ z,
\]
then on both sides:
\[
\tau_{y \circ z}(x) = \tau_x(\tau_y(x))
\]

So $\tau$ is an anti-homomorphism on $(B, \circ)$. □

References


(Lindsay N. Childs) Department of Mathematics and Statistics, University at Albany, Albany, NY 12222, USA
lchilds@albany.edu

This paper is available via [http://nyjm.albany.edu/j/2024/30-29.html](http://nyjm.albany.edu/j/2024/30-29.html).