Operated groups, differential groups and Rota-Baxter groups with an emphasis on the free objects

Xing Gao\textsuperscript{a,b}, Li Guo\textsuperscript{c}, Yanjun Liu\textsuperscript{d}, and Zhi-Cheng Zhu\textsuperscript{a}

\textsuperscript{a}School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu, China; \textsuperscript{b}School of Mathematics and Statistics, Qinghai Nationalities University, Xining, China; \textsuperscript{c}Department of Mathematics and Computer Science, Rutgers University, Newark, New Jersey, USA; \textsuperscript{d}School of Mathematics and Statistics, Jiangxi Normal University, Nanchang, China

ABSTRACT

Groups with various types of operators, in particular the recently introduced Rota-Baxter groups, have generated renowned interest with close connections to numerical integrals, Yang-Baxter equation, integrable systems and post-Hopf algebras. This paper gives the general notion of operated groups and provides explicit constructions of free operated groups, free differential groups and free Rota-Baxter groups.

1. Introduction

Linear operators are ubiquitous in mathematics. Notable examples are the endomorphisms and automorphisms in Galois theory, derivations and integral operators in analysis and related areas. Even though groups do not carry a linear structure, the study of operations on groups has a long history. Groups with endomorphisms were extensively studied by Emmy Noether and her school in the 1920s. In fact, it is this concept that she employed in her original formulation of the three Noetherian isomorphism theorems \[42, 52\]. In the classical treatment of groups by Bourbaki \[10, \text{pp. 30–31}\], the notion of groups with endomorphisms was introduced at the very beginning and was then specialized to define modules (in analogy to rings with operators used to define algebras). See also MacLane \[38, \text{p. 41}\]. In addition to the endomorphisms, operators on groups with other properties had appeared, such as crossed homomorphisms in group cohomology \[50\].

Thanks to the flexibility afforded by the linear structure, linear operators on various algebras have flourished. Along with the ones mentioned above, there were also the difference operators, differential operators with weight, Rota-Baxter operators, Reynolds operators, averaging operators and Nijenhuis operators, as well as their multi-operator analogs, abstracted from geometry, probability, fluid mechanics, analysis, combinatorics, differential equations and mathematical physics, and with applications in broad areas from quantum field renormalization to mechanical proofs of geometric theorems \[2, 7, 11, 15, 16, 21, 23–25, 30, 32, 33, 41, 43, 46, 47, 49, 53–56\].

Putting the corresponding operator identities in a unified framework, Rota raised the question on a classification of operator identities that can be defined on algebras \[48\]. Attempts to address Rota’s question have led to a deeper understanding of operator identities in terms of operated
algebras, rewriting systems and Gröbner-Shirshov bases, and to the discovery of new operator identi-
cities [18, 19, 22, 27, 57].

Recently, such enduring enthusiasm on linear operators on algebras has led to renewed interest in
operators on groups, beginning with [26] motivated by the classical Yang-Baxter equation and Poisson
geometry [37, 49]. More precisely, the notion of Rota-Baxter operators on (Lie) groups was defined
in [26] for which the differentiation gives Rota-Baxter operators on the corresponding Lie algebras.
As an application, the fundamental factorization of a Lie group in [44, 45, 49] originally obtained
indirectly from locally integrating a factorization of a Lie algebra now comes directly from a global
factorization of the Lie group, equipped with the Rota-Baxter operator. Incidentally, the formal inverse
of a Rota-Baxter operator on a Lie group is none other than the crossed homomorphism on the Lie
group, and differentiates to a differential operator on the Lie algebra of left-translation-invariant vector
fields.

After this work, there has been a boom of studies on the theory and applications of differential and
Rota-Baxter operators on groups.

In [5], general properties of Rota-Baxter groups, especially extensions of Rota-Baxter groups and
Rota-Baxter operators on sporadic simple groups, were studied. In [4], it was shown that Rota-Baxter
groups give rise to braces and the more general skew left braces in quantum Yang-Baxter equation.
In [13], a characterization of the gamma functions on a group that come from Rota-Baxter operators
as in [4] was given in terms of the vanishing of a certain element in a suitable second cohomology group.
In [12], Rota-Baxter operators on Clifford semigroups were introduced as a useful tool for obtaining
dual weak braces. In [6, 51], Rota-Baxter groups and skew braces were further studied.

In [29], a cohomological theory of Rota-Baxter Lie groups was developed which differentiates to
the existing cohomological theory of Rota-Baxter Lie algebras. In [34], a one-to-one correspondence
was established between factorizable Poisson Lie groups and quadratic Rota-Baxter Lie groups. [20]
introduced and studied a Rota-Baxter operator on a cocommutative Hopf algebra that generalizes the
notions of a Rota-Baxter operator on a group and a Rota-Baxter operator of weight 1 on a Lie algebra.
Finally, in [3], post-groups and pre-groups were introduced that can be derived from Rota-Baxter groups
and capture the extra structures on (Lie-)Butcher groups in numerical integration, braces, Yang-Baxter
equation and post-Hopf algebras [1, 4, 14, 17, 36, 39, 40].

With all these activities on Rota-Baxter groups, differential groups and related structures, it is time
to give a systematic study of Rota-Baxter groups and the more general operated groups. As with any
algebraic structure, the free objects should play a fundamental role in the study of such group structures.
For instance, being able to give an explicit construction of the free objects affords a clear picture
on how the operations behave, and allows every other object to be expressed as a quotient of a free
object.

Our main goal of this paper is the explicit constructions of the free objects, specifically free operated
groups, free differential groups and free Rota-Baxter groups. In Section 2, after some general discussion
of operated groups, free operated groups are constructed. Then properties and free objects of differential
groups are obtained in Section 3. Finally in Section 4, free Rota-Baxter groups are constructed.

2. Operated groups
In this section we give the notion of operated groups and recall the special cases of Rota-Baxter groups
and differential groups. We then construct free operated groups.

2.1. Notations and examples
We first give the notion of operated groups. We then provide some examples in both concrete and general
terms, some of which will be studied further in later sections.
Definition 2.1. An **operated group** is a pair \((G, P)\) consisting of a group \(G\) and a map \(P : G \rightarrow G\). A **homomorphism** from an operated group \((G, P)\) to another one \((G', P')\) is a group homomorphism \(f : G \rightarrow G'\) such that \(P'f = fP\).

Well-known examples of operated groups are groups equipped with group endomorphisms as mentioned in the introduction. In the classical literature such as [10], the term groups with operators means that the operators are group endomorphisms; while for us, the operators do not need to satisfy any condition. For distinction, a group equipped with an group endomorphism will be called an **endo group**. Further examples of operated groups are differential groups coming from crossed homomorphisms [26], and Rota-Baxter groups arising recently from studies of integrable systems and Hopf algebras [26].

Recall that a **differential operator (of weight \(\lambda\))** [25] is an associative algebra \(A\) with a linear operator \(d : A \rightarrow A\) such that

\[
d(uv) = d(u)v + ud(v) + \lambda d(u)d(v) \quad \text{for all } u, v \in A.
\]

The term comes from its classical example as the operator of differential quotient:

\[
d(f)(x) := (f(x + \lambda) - f(x))/\lambda
\]

in analysis. Its study in the weight zero case originated from the work of Ritt [46] on algebraic study of differential equations. Then through the work of Kolchin [32] and many others (see [43] and the references therein), the study has expended into an area with broad applications. The weighted notion was introduced in [25].

Similarly, a differential operator (of weight \(\lambda\)) on a Lie algebra [28, 31, 35] is a linear map \(d : g \rightarrow g\) such that

\[
d[u, v]_g = [d(u), v]_g + [u, d(v)]_g + \lambda [d(u), d(v)]_g \quad \text{for all } u, v \in g.
\]

Let \(G\) be a group. For \(g \in G\), define the **adjoint action**

\[
Ad_g : G \rightarrow G, \quad Ad_g h := ghg^{-1} \quad \text{for all } h \in G.
\]

Then \(Ad_g\) is in \(\text{Aut}(G)\) and the resulting map

\[
Ad : G \rightarrow \text{Aut}(G), \quad Ad(g) := Ad_g \quad \text{for all } g \in G,
\]

is a group homomorphism.

Definition 2.2. A map \(D\) on a group \(G\) is called a **differential operator (of weight 1)** (resp. \(-1\)) if

\[
D(gh) = D(g)Ad_g D(h) \quad \text{(resp. } D(gh) = (Ad_g D(h))D(g)) \quad \text{for all } g, h \in G.
\]

Then the pair \((G, D)\) is called a **differential group of weight 1** (resp. \(-1\)).

It remains an open problem to find a suitable notion of differential operator of weight zero on a group. For the rest of the paper, we only work with differential group of weight 1, which we simply call a differential group.

As shown in [26], the derivation of a differential operator on a Lie group is a differential operator (of weight 1) on the corresponding Lie algebra. This notion is the special case of the classical notion of a crossed homomorphism or a 1-cocycle with non-abelian coefficients [50].

Definition 2.3. Let \(G\) be a group and let \(\Gamma\) be a group with a group action by \(G\), defined by \(\alpha \mapsto \alpha^x, \alpha \in \Gamma, x \in G\). A map \(f : G \rightarrow \Gamma\) is called a **crossed homomorphism** or a 1-cocycle if

\[
f(xy) = f(x)f(y)^x \quad \text{for all } x, y \in G.
\]

Crossed homomorphism is the key notion for group cohomology and Galois cohomology, especially for noncommutative groups. Taking \(\Gamma\) to be the group \(G\) itself equipped with the adjoint action of \(G\),
we obtain the notion of the differential operator on groups. In general, a crossed homomorphism can be interpreted as a relative differential operator on a group associated with a group representation [3].

The notion of a Rota-Baxter group was obtained in [26] in an effort to find the integration of a Rota-Baxter operator on a Lie algebra in order to study the factorization on a Lie group without having to integrate from a factorization on the Lie algebra [44, 45, 49].

**Definition 2.4.** A **Rota-Baxter group** (of weight 1) is a group $G$ with a map $B : G \rightarrow G$ satisfying the following equation
\[
B(g_1)B(g_2) = B(g_1 \text{Ad}_{B(g_1)} g_2) \quad \text{for all } g_1, g_2 \in G,
\]
called the **Rota-Baxter relation** for groups.

When $G$ is an abelian group, Eq. (2) means that $B$ is a group homomorphism.

A Rota-Baxter operator of weight 1 on a group is also the right inverse of a differential operator of weight 1, as in the case of operators on associative algebras and Lie algebras.

Let $(G, B)$ and $(G', B')$ be Rota-Baxter groups. A map $\Phi : G \rightarrow G'$ is a **Rota-Baxter group homomorphism** if $\Phi$ is a group homomorphism such that $\Phi B = B' \Phi$.

**Definition 2.5.** We define a **Rota-Baxter group of weight** $-1$ to be a group $G$ with a map $C : G \rightarrow G$ such that
\[
C(g_1)C(g_2) = C((\text{Ad}_{C(g_1)} g_2)g_1) \quad \text{for all } g_1, g_2 \in G.
\]

It still unknown how to give a suitable notion of Rota-Baxter operator of weight zero on a group such that the derivation is the Rota-Baxter operator of weight zero on a Lie algebra.

Let $B$ be a Rota-Baxter operator (of weight 1) on a group $G$. Define $C : G \rightarrow G$ by
\[
C(g) := B(g^{-1}).
\]

In Eq. (2), replacing $g_1$ and $g_2$ by $g_1^{-1}$ and $g_2^{-1}$, we obtain
\[
B(g_1^{-1})B(g_2^{-1}) = B(g_1^{-1} \text{Ad}_{B(g_1^{-1})} g_2^{-1}) = B\left(\left(\text{Ad}_{B(g_1^{-1})} g_1\right)^{-1}\right).
\]

This gives
\[
C(g_1)C(g_2) = C(\text{Ad}_{C(g_1)} g_2)g_1).
\]

Therefore, $C$ is a Rota-Baxter operator of weight $-1$.

Let $G_1$ and $G_2$ be subgroups of a group $G$ such that $G = G_1 G_2$ and $G_1 \cap G_2 = \{1\}$. Note that it is not necessary that $G$ is the direct product of $G_1$ and $G_2$. Let $C$ be the **projection to the first direct factor**, i.e.
\[
C(g_1g_2) := g_+.
\]

Then it is straightforward to deduce that $C$ is a Rota-Baxter operator of weight $-1$ [26].

*For the rest of the paper, by a Rota-Baxter group we mean a Rota-Baxter group of weight 1.*

### 2.2. Free operated groups

The notion of free operated groups can be defined in the language of universal algebra by a universal property. Here we give an explicit construction of a free operated group on a set by bracketed words.

**Definition 2.6.** Let $X$ be a set. The **free operated group** generated by $X$ is an operated group $(F(X), P_X)$ together with a map $j_X : X \rightarrow F(X)$ with the property that, for each operated group $(G, P)$ and map $f : X \rightarrow G$, there is a unique homomorphism $\tilde{f} : (F(X), P_X) \rightarrow (G, P)$ of operated groups such that $\tilde{f}j_X = f$. 
We first give some general notations. For any set $Y$, let $F(Y)$ denote the free group generated by $Y$, via the usual construction of reduced words in $Y \cup Y^{-1}$. So $F(Y)$ consists of the identity 1 and reduced words in the letter set $Y \cup Y^{-1}$, where a reduced word is an element of the form

$$w = w_1 \cdots w_k,$$

with $w_i \in Y \cup Y^{-1}$ so that no adjacent letters $w_i$ and $w_{i+1}$ are of the form $zz^{-1}$ for a $z \in Y \cup Y^{-1}$. Also let $\{Y\}$ denote a set in bijection with $Y$ but disjoint from $Y$. Elements in $\{Y\}$ are denoted by $[y], y \in Y$.

Now let $X$ be a given set. We define a directed system of free groups $G_n := G_n(X), n \geq 0$, by a recursion. First define the free group

$$G_0 := F(X).$$

Then define the free group

$$G_1 := F(X \cup \{F(X)\})$$

generated by $X \cup \{F(X)\}$ and let

$$i_{0,1} : G_0 \to G_1$$
denote the group homomorphism of free groups induced by the natural inclusion $X \to X \cup \{F(X)\}$.

Recursively, for any given integer $n \geq 1$, suppose that the groups $G_i, i \leq n$, have been constructed and the injective group homomorphisms

$$i_{i-1,i} : G_{i-1} \to G_i, i \leq n,$$

have been defined. Then define the free group

$$G_{n+1} := F(X \cup \{G_n\})$$

generated by the set $X \cup \{G_n\}$.

Furthermore, from the injective group homomorphism $i_{n-1,n} : G_{n-1} \to G_n$, we obtain the injections

$$[G_{n-1}] \to [G_n], \quad [w] \mapsto [i_{n-1,n}(w)] \quad \text{for all } w \in G_{n-1}$$

and

$$X \cup [G_{n-1}] \to X \cup [G_n].$$

The latter then induces the injective homomorphism of free groups

$$i_{n,n+1} : G_n = F(X \cup [G_{n-1}]) \to G_{n+1} = F(X \cup [G_n]).$$

This completes the recursive construction of the directed system $\{G_n\}_{n \geq 0}$ of free groups. Finally, we define the direct limit of the directed system

$$G := \hat{G}(X) := \lim_{\to} G_n = \bigcup_{n \geq 0} G_n.$$

It is still a group which naturally contains $G_n, n \geq 0$, as subgroups via the structural injective group homomorphisms

$$i_n : G_n \to G.$$  \hspace{1cm} (4)

Denote

$$X^{\pm 1} := X \cup X^{-1}, \quad [G]^{-1} := \{w^{-1} \mid w \in [G]\} \text{ and } [G]^{\pm 1} := [G] \cup [G]^{-1}.$$  

For a $w \in G$, we may uniquely write

$$w = w_1 \cdots w_k \text{ with } w_i \in X^{\pm 1} \cup [G]^{\pm 1},$$  \hspace{1cm} (5)

with no adjacent factor being the inverse of each other. The factorization is called the **standard factorization** or **standard form** of $w$. Define the **breadth** of $w$ to be $\breath(w) := k$ and **depth** of $w$ to be

$$\depth(w) := \max\{\depth(w_i) \mid 1 \leq i \leq k\}, \text{ where } \depth(w_i) := \min\{n \mid w_i \in G_n \setminus G_{n-1}\}.$$
We next define an operator $P_X$ on $G$. For $u \in G$, we have $u \in G_n$ for some $n \geq 0$. We then define $P_X(u) = [u]$ which is in $G_{n+1}$ and hence in $G$. It is direct to check that $P_X(u)$ is independent of the choice of $n$ such that $u$ is in $G_n$. Thus, the pair $(G(X), P_X)$ is an operated group. Finally let

$$j_X : X \to G$$

be the natural inclusion.

**Theorem 2.7.** The operated group $(G(X), P_X)$ together with the map $j_X : X \to G(X)$ is a free operated group on $X$.

**Proof.** We just need to verify that the operated group $(G(X), P_X)$ together with $j_X$ satisfies the desired universal property of a free operated group on $X$. For this purpose, let $(G, P)$ be a given operated group and $f : X \to G$ be a map. We will apply a recursion to define a sequence of group homomorphisms

$$f_n : G_n \to G, n \geq 0,$$

that yield a group homomorphism $\hat{f} : G \to G$ by taking the direct limit.

First define $f_0 : G_0 = F(X) \to G$ by the universal property of the free group $F(X)$ on $X$ such that $f_0 j_X = f$. We then define a map

$$\hat{f}_0 : [G_0] \to G, \quad \hat{f}_0([w]) := P(f_0(w)) \quad \text{for all } w \in G_0.$$

We then define a group homomorphism $f_1 : G_1 := F(X \sqcup [G_0]) \to G$ by the universal property of the free group $G_1$ induced by the set map

$$X \sqcup [G_0] \to G, \quad x \mapsto f(x), [w] \mapsto \hat{f}_0([w]) \quad \text{for all } x \in X, [w] \in [G_0].$$

Then we have $f_1 i_{0,1} = f_0$.

Recursively, for any given integer $n \geq 1$, suppose that, for $i \leq n$, the group homomorphisms $f_i : G_i \to G$ have been defined such that $f_i i_{i-1,i} = f_{i-1}$. We then define

$$\hat{f}_n : [G_n] \to G, \quad [w] \mapsto P(f_n(w)) \quad \text{for all } w \in G_n,$$

and then extend it to a map

$$X \sqcup [G_n] \to G$$

by sending $x \in X$ to $f(x).$ We then apply the universal property of the free group $G_{n+1} := F(X \sqcup [G_n])$ and obtain the group homomorphism

$$f_{n+1} : G_{n+1} \to G.$$

Furthermore, to show that $f_{n+1} i_{n,n+1} = f_n$, we just need to check

$$f_{n+1} i_{n,n+1}(w) = f_n(w) \quad \text{for all } w \in X \sqcup [G_{n-1}].$$

(6)

This is clear for $w \in X$ since both sides are equal to $f(x)$. For $w \in [G_{n-1}]$, we write $w = [\overline{w}]$ for $\overline{w} \in G_{n-1}$. Then by Eq. (3) we obtain

$$i_{n,n+1}(w) = i_{n,n+1}([\overline{w}]) = [i_{n-1,n}(\overline{w})]$$

and hence

$$f_{n+1} i_{n,n+1}(w) = f_{n+1}([i_{n-1,n}(\overline{w})]) = \hat{f}_n([i_{n-1,n}(\overline{w})])$$

$$= P(f_i i_{i-1,n}(\overline{w})) = P(f_{i-1}(\overline{w})) = f_n([\overline{w}]) = f_n(w).$$

This proves Eq. (6), as needed.

Thus we can take the direct limit of $f_n, n \geq 0$, and obtain a group homomorphism

$$\hat{f} = \lim f_n : G \to G.$$
By construction, we have $f_{jX} = f$. The relation
\[ f_{n+1}(\lfloor w \rfloor) = \lfloor f_n(w) \rfloor \quad \text{for all } w \in \mathcal{G}_n, \]
gives $\tilde{f}P_X = P\tilde{f}$. Therefore $\tilde{f} : (\mathcal{G}(X), P_X) \to (G, P)$ is a homomorphism of operated groups.

Finally, by the construction of $\tilde{f}$, it is the unique homomorphism of operated group homomorphism such that $\tilde{f}j_X = f$. More precisely, let $\tilde{f}' : (\mathcal{G}, P_X) \to (G, P)$ be another homomorphism of operated algebra homomorphisms such that $\tilde{f}'j_X = f$. Then we recursively verify that $\tilde{f}_n = \tilde{f}'_n$, $n \geq 0$, for the group homomorphisms $i_n : \mathcal{G} \to \mathcal{G}$ in Eq. (4). This gives $\tilde{f}' = \tilde{f}$. Therefore the desired universal property is verified.

Free operated semigroups have been constructed in [22] in terms of bracketed words from $X$. One can think of elements of $\mathcal{G}(X)$ as reduced bracketed words from $X \cup X^{-1}$. Here a bracketed word is reduced if it does not have any pair of adjacent subwords that are the inverses of each other. It would be interesting to obtain combinatorial construction of free operated groups in terms of paths or trees as in the case of operated semigroups [22], or in terms of Cayley graphs for free groups, which are fundamental in geometric group theory [8, 9].

3. Free differential groups

With the notion of a differential group given in Definition 2.2, we now provide some properties of differential groups and then construct free differential groups.

**Lemma 3.1.** Let $(G, d)$ be a differential group.

(i) The derivation $d$ is compatible with the associativity of $G$. More precisely, for $g, h, k \in G$, the results of $d((gh)k)$ and $d(g(hk))$ agree. In fact, we have
\[ d(ghk) = d(g)gd(h)hd(k)h^{-1}g^{-1} \quad \text{for all } g, h, k \in G. \]  \hfill (7)

(ii) More generally, for $g_1, \ldots, g_n \in G$, $n \geq 2$, we have
\[ d(g_1 \cdots g_n) = \left( \prod_{i=1}^{n} d(g_i)g_i \right)(g_1 \cdots g_n)^{-1} \quad \text{for all } g_1, \ldots, g_n \in G. \]

**Proof.** (i) We directly check that
\[ d((gh)k) = d(gh)ghd(k)(gh)^{-1} = d(g)gd(h)g^{-1}ghd(k)(gh)^{-1} = d(g)gd(h)hd(k)h^{-1}k^{-1}. \]
Also
\[ d(g(hk)) = d(g)gd(hk)g^{-1} = d(g)gd(h)hd(k)h^{-1}g^{-1}. \]
Hence, Eq. (7) holds.

(ii) The general equation follows from a simple induction on $n$. \hfill \Box

**Lemma 3.2.** Let $(G, d)$ be a differential group. Then we have
\[ d(1) = 1, \quad d(g^{-n}) = (g^{-1}d(g)^{-1})^ng^n \quad \text{for all } g \in G, n \geq 1. \]

**Proof.** Denote $d(1) = a$. Then from
\[ a = d(1) = d(1^2) = d(1)d(1) = a^2, \]
we obtain $a = 1$. To prove the second equality, we apply induction on $n \geq 1$. Let $n = 1$. Then for $g \in G$, we have
\[ 1 = d(1) = d(gg^{-1}) = d(g)gd(g^{-1})g^{-1}, \]
which yields
\[ d(g^{-1}) = g^{-1}d(g)^{-1}g. \]
For \( k \geq 1 \), assume
\[ d(g^{-k}) = (g^{-1}d(g)^{-1})^k g^k. \]
Then we obtain
\[
\begin{align*}
    d(g^{-(k+1)}) &= d(g^{-1}g^{-k}) \\
    &= d(g^{-1})g^{-1}d(g^{-k})g \\
    &= g^{-1}d(g)^{-1}g^{-1}(g^{-1}d(g)^{-1})^k g^{k+1} \\
    &= (g^{-1}d(g)^{-1})^{k+1} g^{k+1}.
\end{align*}
\]
This completes the induction. \( \square \)

We now construct free differential groups by adding a differential structure to the classical construction of free groups.

**Definition 3.3.** The **free differential group** on a set \( X \) is a differential group \((F(X), D)\) together with a set map \( i = i_X : X \rightarrow F(X) \) with the property that, for any differential group \((G, d)\) and set map \( f : X \rightarrow G \), there is a unique homomorphism of differential groups \( \bar{f} : (F(X), D_X) \rightarrow (G, d) \) such that \( \bar{f}i = f \).

As a side remark, the notion of a free endo group can be given in the same way and can be constructed in the same way as below.

Let \( X \) be a set. Denote
\[
\Delta X := X \times \mathbb{N} = \{x^{(n)} := (x, n) \mid x \in X, n \in \mathbb{N}\}.
\]

Let
\[
F(X) := F(\Delta X)
\]
denote the free group on the set \( \Delta X \), as usual in terms of reduced words in the alphabet \( \Delta X \) and their formal inverses. We define an operator \( D := D_X \) on \( F(X) \) in two steps. First define
\[
D : \Delta X \rightarrow \Delta X, \quad x^{(n)} \mapsto x^{(n+1)}, \quad x^{(n)} \in \Delta X = X \times \mathbb{N}.
\]
Then extend \( D \) to \( F(X) = F(\Delta X) \) by applying Eq. (1). More precisely, an element \( x \) of \( F(\Delta X) \) can be uniquely expressed as a reduced word
\[
x := x_1 \cdots x_k, \quad x_i \in \Delta X \cup (\Delta X)^{-1},
\]
with the convention that the length of the empty word \( x = 1 \) is 0. When \( k = 0 \), we define \( D(1) = 1 \). When \( k = 1 \), we have \( x \in \Delta X \cup (\Delta X)^{-1} \). We accordingly define
\[
D(x) = D(x^{(n)}) := x^{(n+1)} \text{ or } D(x) = D(x^{(n)})^{-1} := (x^{(n)})^{-1} (x^{(n+1)})^{-1} x^{(n)} , \quad n \geq 0.
\]
For a given integer \( m \geq 1 \), assume that \( D(x) \) has been defined for all reduced words \( x \) of length \( k = m \). Then for a reduced word \( x = x_1 \cdots x_{m+1} \) of length \( m + 1 \), we recursively define
\[
D(x) = D(x_1)D(x_2 \cdots x_{m+1})^{-1}.
\]

Let
\[
i_X : X \rightarrow \Delta X \rightarrow F(X), \quad x \mapsto x^{(0)}, x \in X,
\]
be the canonical inclusion.
**Theorem 3.4.** Let $X$ be a set.

(i) The pair $(F[X], D_X)$ is a differential group.

(ii) The triple $(F[X], D_X, i_X)$ is a free differential group on $X$.

**Proof.** (i) We just need to verify that $D$ is a differential operator on the group $F(X)$, that is, Eq. (1) holds for all $g, h \in F(X)$. This can be done by an induction on the sum $m + n$ of the lengths $m$ and $n$ of $g$ and $h$ as follows.

First consider $m + n = 0$. Then $g = h = 1$ and so

$$D(gh) = D(1) = 1 = D(g)D(h)g^{-1}.$$ 

Thus, Eq. (1) holds. Let $k \geq 0$ be given. Assume that Eq. (1) holds for all $g, h \in F(X)$ with $m + n = k$ and consider $g, h \in F(X)$ with $m + n = k + 1$. So either $m > 0$ or $n > 0$. First suppose $m > 0$. Then $g = \tau_1g_1$ for $\tau_1 \in \Delta X, g_1 \in F(X)$. Then

$$D(gh) = D(\tau_1g_1h) = D(\tau_1)D(g_1)\tau_1^{-1}.$$ 

Also,

$$D(g)D(h)g^{-1} = D(\tau_1)D(g_1)\tau_1^{-1}D(h)g^{-1} = D(\tau_1)D(g_1)D(h)g_1^{-1}\tau_1^{-1} = D(\tau_1)D(g_1)\tau_1^{-1},$$

where the last step utilizes the induction hypothesis. The case of $n > 0$ follows from the same argument. This completes the induction.

(ii) We just need to verify that the differential group $(F[X], D)$ together with $i_X : X \to F(X)$ satisfies the expected universal property stated in the theorem. Now let a differential group $(G, d)$ and a map $f : X \to G$ be given. Define a differential group homomorphism $\tilde{f} : F[X] \to G$ by first taking

$$\tilde{f} : \Delta X \to G, \quad x^{(n)} \mapsto d^n(f(x)) \quad \text{for all } x^{(n)} \in \Delta X,$$

and then taking

$$\tilde{f} : F[X] = F(\Delta X) \to G$$

to be the unique group homomorphism from $\tilde{f}$ by the universal property of the free group $F(\Delta X)$ on $\Delta X$.

We are left to verify that $\tilde{f}$ is a homomorphism of differential groups satisfying $\tilde{f}i_X = f$ and is the unique one with this property. For the first property, we just need to check that $\tilde{f}$ is compatible with the differential operators: $\tilde{f}D = df$ which is proved by an induction on the length by applying Eq. (8). Indeed, it is direct to check that $\tilde{f}D(g) = df(g)$ if $g = 1$ or if $g$ is in $\Delta X$. Further if $g = \tau_1g_1$ with $\tau_1 \in \Delta X$, then applying the induction hypothesis to $g_1$, we obtain

$$\tilde{f}D(\tau_1g_1) = \tilde{f}(D(\tau_1)\tau_1D(g_1)\tau_1^{-1}) = \tilde{f}D(\tau_1)\tilde{f}(\tau_1)\tilde{f}D(g_1)\tilde{f}(\tau_1)^{-1} = D(\tilde{f}(\tau_1))D(\tilde{f}(g_1))\tilde{f}(\tau_1)^{-1} = D(\tilde{f}(\tau_1))D(\tilde{f}(g_1)) = D\tilde{f}(g).$$

This completes the induction.

The construction of $\tilde{f}$ shows that it is the only way to obtain a differential group homomorphism, showing the uniqueness of $\tilde{f}$. □

4. **Free Rota-Baxter groups**

We now construct the free Rota-Baxter group generated by a set. To be specific, we first give the definition.

**Definition 4.1.** Let $X$ be a set. The free **Rota-Baxter group** generated by $X$ is a Rota-Baxter group $(\mathfrak{R}(X), B_X)$ together with a map $j_X : X \to \mathfrak{R}(X)$ such that, for any Rota-Baxter group $(G, B)$ and map $f : X \to G$, there is a unique homomorphism $f : (\mathfrak{R}(X), B_X) \to (G, B)$ of Rota-Baxter groups such that $fj_X = f$. 

From the construction of free operated groups obtained in Section 2.2, the free Rota-Baxter group can be obtained by taking a quotient as follows.

Let a set $X$ be given and let $(G(X), P_X)$ be the free operated group generated by $X$ constructed in Theorem 2.7. Let $N := N_X$ be the normal operated subgroup of $G(X)$ generated by elements of the form

$$P_X(u)P_X(v)(P_X(uAd_{P_X(v)}v))^{-1} \text{ for all } u, v \in G(X).$$

Then the quotient $G(X)/N$, with its operator induced from $P_X$ modulo $N$, is automatically a free Rota-Baxter group generated by $X$. We next give a direct and explicit construction of the free Rota-Baxter group on $X$, by displaying a canonical subset of $G(X)$ and equipping it with the desired Rota-Baxter group structure. The construction is inspired by the construction of free Rota-Baxter (associative) algebras [23].

**Definition 4.2.** An element $w \in G(X)$ is called a **Rota-Baxter group word** on $X$ if $w$ contains no subword $[u][v]$ with $u, v \in G(X)$. Denote by $\mathcal{R}(X)$ the set of all Rota-Baxter group word on $X$.

In terms of the standard factorization in Eq. (5), each Rota-Baxter group word may be uniquely written as

$$w = w_1 \cdots w_k$$

(9)

where $w_i \in X^{\pm 1} \cup [G(X)]^{\pm 1}$ and no adjacent $w_i$ and $w_{i+1}$ are both in $[G(X)]$ or both in $[G(X)]^{-1}$. Note that it is okay if one of $w_i$ or $w_{i+1}$ is in $[G(X)]$ and the other one in $[G(X)]^{-1}$. We call Eq. (9) the **standard factorization** of $w$.

We are going to define a multiplication $\diamond$ on $\mathcal{R}(X)$. First we define 1 to be the identity: $1 \diamond u = u \diamond 1 = u$ for all $u \in \mathcal{R}(X)$. For $u, v \in \mathcal{R}(X) \setminus \{1\}$, we apply induction on $(\text{dep}(u), \text{dep}(v)) \geq (0,0)$ lexicographically.

When $(\text{dep}(u), \text{dep}(v)) = (0,0)$, we have $u, v \in G_0$ and define $u \diamond v := uv$. Let $m, n \geq 0$ be given such that $(m, n) > (0,0)$ lexicographically. Suppose that the product $u \diamond v$ have been defined for $u, v \in \mathcal{R}(X)$ with $(\text{dep}(u), \text{dep}(v)) < (m, n)$ lexicographically and consider $u, v \in \mathcal{R}(X)$ with $(\text{dep}(u), \text{dep}(v)) = (m, n)$. We have the following four cases to check, with the first three cases dealing with $\text{bre}(u) = \text{bre}(v) = 1$.

**Case 1.** Let $u = [\overline{u}]$ and $v = [\overline{v}]$ for some $\overline{u}$ and $\overline{v} \in \mathcal{R}(X)$. Write $\overline{v} = \overline{v}_1 \cdots \overline{v}_k$ in the standard factorization. Then the expression

$$\overline{w} \text{Ad}_u\overline{v} := \left\{ \begin{array}{ll}
(u \diamond \overline{v}_1)\overline{v}_2 \cdots \overline{v}_{k-1}(u \diamond \overline{v}_k)^{-1}, & \text{if } \overline{v}_k = (\overline{v}_k')^{-1} \in [\mathcal{R}(X)]^{-1}, \\
(u \diamond \overline{v})u^{-1}, & \text{otherwise},
\end{array} \right.$$  

(10)

is well defined by the induction hypothesis on $(\text{dep}(u), \text{dep}(v))$. In particular if $k = 1$, then

$$\overline{w} \text{Ad}_u\overline{v} := \left\{ \begin{array}{ll}
(u \diamond \overline{v})^{-1}, & \text{if } \overline{v} = (\overline{v})^{-1} \in [\mathcal{R}(X)]^{-1}, \\
(u \diamond \overline{v})u^{-1}, & \text{otherwise},
\end{array} \right.$$

Then we define

$$u \diamond v := [\overline{u} \diamond \overline{w} \text{Ad}_u\overline{v}].$$

(11)

Here since $\text{dep}(u) > \text{dep}(\overline{u})$, $\overline{u} \diamond \overline{w} \text{Ad}_u\overline{v}$ is defined by induction hypothesis.

**Case 2.** Let $u = [\overline{u}]^{-1}$ and $v = [\overline{v}]^{-1}$ for some $\overline{u}, \overline{v} \in \mathcal{R}(X)$. Then define

$$u \diamond v := ([\overline{v}] \diamond [\overline{u}])^{-1},$$

(12)

where $[\overline{v}] \diamond [\overline{u}]$ is reduced to Case 1.

**Case 3.** Let $u$ and $v$ have breadth 1, but be not in the above two cases. Then the concatenation $uv$ is still in $\mathcal{R}(X)$. Then we can define

$$u \diamond v := uv.$$  

(13)
Case 4. In general, write \( u = u_1 \cdots u_k \) and \( v = v_1 \cdots v_\ell \) with \( k, \ell \geq 1 \) in the standard factorizations. Define
\[
   u \odot v := u_1 \cdots u_{k-1} (u_k \odot v_1) v_2 \cdots v_\ell,
\]
(14)
where \( u_k \odot v_1 \) is defined by Cases 1, 2, 3. Notice that \( 1 \odot u = u = u \odot 1 \) for \( u \in \mathcal{G} \).

Note that by the construction of \( \mathcal{R}(X) \), the set \( \mathcal{R}(X) \) is closed under taking the bracket: for \( w \in \mathcal{R}(X) \), we have \([w] \in \mathcal{R}(X)\).

**Theorem 4.3.** With the above definition of \( \odot \), the triple \((\mathcal{R}(X), \odot, [\cdot])\) is a \( \text{Rota-Baxter} \) group of weight 1.

**Proof.** By the construction of \( \mathcal{R}(X) \), we have \( 1 \in \mathcal{R}(X) \) and \( w^{-1} \in \mathcal{R}(X) \) provided \( w \in \mathcal{R}(X) \). So we are left with checking the associativity of \( \odot \) and the Rota-Baxter relation in Eq. (2). first verify the associativity of \( \odot \):
\[
   (u \odot v) \odot w = u \odot (v \odot w) \quad \text{with } u, v, w \in \mathcal{R}(X). \tag{15}
\]
which will be accomplished by an induction on \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) \geq 3 \).

**1. The initial step on \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) \).** For the initial step of \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) = 3 \), we have \( \text{bre}(u) = \text{bre}(v) = \text{bre}(w) = 1 \). In this case, we apply the induction on \( \text{dep}(u), \text{dep}(v), \text{dep}(w) \) \( \geq (0, 0, 0) \) lexicographically. If \( \text{dep}(u), \text{dep}(v), \text{dep}(w) = (0, 0, 0) \), then \( u, v, w \) are in the free group \( \mathcal{G}_0 \) and Eq. (15) holds by the associativity of the free group \( \mathcal{G}_0 \).

For the inductive step, let \( p, q, r \geq 0 \) be given such that \( (p, q, r) > (0, 0, 0) \) lexicographically. Assume that Eq. (15) has been proved for \( u, v, w \in \mathcal{R}(X) \) with \( \text{bre}(u) = \text{bre}(v) = \text{bre}(w) = 1 \) and \( \text{dep}(u), \text{dep}(v), \text{dep}(w) < (p, q, r) \) lexicographically, and consider \( u, v, w \in \mathcal{R}(X) \) with \( \text{bre}(u) = \text{bre}(v) = \text{bre}(w) = 1 \) and \( \text{dep}(u), \text{dep}(v), \text{dep}(w) = (p, q, r) \). Then according to whether all of \( u, v, w \) are in \( [\mathcal{R}(X)] \) or not, we consider the following four cases.

**Case 1.** Suppose that all \( u, v, w \) are in \( [\mathcal{R}(X)] \). Then we take \( u = [\overline{u}], v = [\overline{v}], w = [\overline{w}] \) for some \( \overline{u}, \overline{v}, \overline{w} \in \mathcal{R}(X) \). Write \( \overline{v} = \overline{v}_1 \cdots \overline{v}_m \) and \( \overline{w} = \overline{w}_1 \cdots \overline{w}_n \) in the standard factorizations. We divide this case to four subcases according to whether \( \overline{v}_m \) and \( \overline{w}_n \) are in \( [\mathcal{R}(X)]^{-1} \) or not.

**Case 1.1.** Suppose \( \overline{v}_m = (\overline{v}_m)^{-1}, \overline{w}_n = (\overline{w}_n)^{-1} \in [\mathcal{R}(X)]^{-1} \) for some \( \overline{v}_m, \overline{w}_n \in [\mathcal{R}(X)] \). Then on the one hand,
\[
   (u \odot v) \odot w = ([\overline{u}] \odot [\overline{v}]) \odot [\overline{w}] \tag{Eqs. (11) and (10)}
\]
\[
   = \left[ \overline{u} \odot (([\overline{u}] \odot \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} (\overline{u} \odot \overline{v}_m)^{-1}) \right] \odot [\overline{w}] \tag{Eqs. (11) and (10)}
\]
\[
   = \left[ \overline{u} \odot (([\overline{u}] \odot \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} (\overline{u} \odot \overline{v}_m)^{-1}) \odot (\overline{u} \odot ([\overline{u}] \odot \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} (\overline{u} \odot \overline{v}_m)^{-1} \odot \overline{w}_1) \right.
\]
\[
   \overline{w}_2 \cdots \overline{w}_{n-1} \left( \overline{u} \odot ([\overline{u}] \odot \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} (\overline{u} \odot \overline{v}_m)^{-1} \odot \overline{w}_n^{-1} \right) \right] \tag{Eqs. (13) and (14)).
\]
On the other hand,
\[
   u \odot (v \odot w) = ([\overline{u}] \odot ([\overline{v}] \odot [\overline{w}]) \tag{Eqs. (11) and (10)}
\]
\[
   = ([\overline{u}] \odot (\overline{v} \odot (([\overline{u}] \odot \overline{w}_1) \overline{w}_2 \cdots \overline{w}_{n-1} (\overline{v} \odot \overline{w}_n)^{-1})) \tag{Eqs. (11) and (10)}
\]
Case 1.2. Suppose $\bar{v}_m = (x^{m-1}, \ldots, x^1, x^0) \in [R(x)]^{-1}$ for some $\bar{v}_m \in [R(x)]$ and $\bar{w}_m \notin [R(x)]^{-1}$. Then

So Eq. (15) holds.

(2) Suppose $\bar{w}_m = (y^{m-1}, \ldots, y^1, y^0) \in [R(x)]^{-1}$ for some $\bar{w}_m \in [R(x)]$ and $\bar{v}_m \notin [R(x)]^{-1}$. Then

By induction on depth and Eqs. (11), (10), and (12), and (14).
\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) (\overline{v}_2 \cdots \overline{v}_m \circ ([\overline{u}]^{-1} [\overline{u}]) \circ ([\overline{v}] \circ [\overline{w}]) (\overline{u}) \circ [\overline{v}]\right)^{-1} \right) \quad (1 = [\overline{u}]^{-1} [\overline{u}])
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) \left( \overline{v}_2 \cdots \overline{v}_m \circ [\overline{u}]^{-1} \right) \left( \overline{u} \circ ([\overline{v}] \circ [\overline{w}]) \right) ([\overline{u}] \circ [\overline{v}]^{-1}) \right) \quad \text{(Eq. (14))}
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) \overline{v}_2 \cdots \overline{v}_{m-1} ([\overline{u}] \circ \overline{v}_m)^{-1} \right) \left( \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} ([\overline{u}] \circ \overline{v}_m)^{-1} \right) \right) \circ \overline{w}
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} ([\overline{u}] \circ \overline{v}_m)^{-1} \right) \left( \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{m-1} ([\overline{u}] \circ \overline{v}_m)^{-1} \right) \right)^{-1} \quad (\overline{v}_m = (\overline{v}_m')^{-1} \text{ and Eqs. (11), (10), (14)),}
\end{align*}
\]

showing that Eq. (15) is valid in this case.

**Case 1.3.** Suppose \( \overline{v}_m \not\in [\mathcal{R}(X)]^{-1} \) and \( \overline{w}_n = (\overline{w}_n')^{-1} \in [\mathcal{R}(X)]^{-1} \) for some \( \overline{w}_n \in [\mathcal{R}(X)] \). Then we check

\[
\begin{align*}
(u \circ v) \circ w
\end{align*}
\]

\[
\begin{align*}
= ([u] \circ [v]) \circ [w]
\end{align*}
\]

\[
\begin{align*}
= \left[ \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1}) \right) \right] \circ \overline{w} & \quad \text{(Eqs. (11) and (10))}
\end{align*}
\]

\[
\begin{align*}
\overline{w}_2 \cdots \overline{w}_{m-1} \left( \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1}) \right) \right) \circ \overline{w}_{n}'^{-1} & \quad \text{(Eqs. (11) and (10))}
\end{align*}
\]

\[
\begin{align*}
\overline{w}_2 \cdots \overline{w}_{m-1} \left( \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1}) \right) \right) \circ \overline{w}_{n}'^{-1} & \quad \text{(Eqs. (13) and (14)),}
\end{align*}
\]

and

\[
\begin{align*}
u \circ (v \circ w)
\end{align*}
\]

\[
\begin{align*}
= [u] \circ ([v] \circ [w])
\end{align*}
\]

\[
\begin{align*}
= \left[ \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) \overline{w}_2 \cdots \overline{w}_{n-1} ([\overline{v}] \circ \overline{w}_n')^{-1} \right) \right] & \quad \text{(Eqs. (11) and (10))}
\end{align*}
\]

\[
\begin{align*}
= \left[ \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) \overline{w}_2 \cdots \overline{w}_{n-1} ([\overline{v}] \circ \overline{w}_n')^{-1} \right) \right] & \quad \text{(Eqs. (11), (10) and (14))}
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1} [\overline{u}]) \circ ([\overline{v}] \circ \overline{w}_1) \overline{w}_2 \cdots \overline{w}_{n-1} \right) & \quad \text{(inserting 1 = [\overline{u}]^{-1} [\overline{u}])}
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1} [\overline{u}]) \circ ([\overline{v}] \circ \overline{w}_1) \overline{w}_2 \cdots \overline{w}_{n-1} \right) & \quad \text{(Eq. (14))}
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1} [\overline{u}]) \circ ([\overline{v}] \circ \overline{w}_1) \overline{w}_2 \cdots \overline{w}_{n-1} \right) & \quad \text{(induction on depth)}
\end{align*}
\]

\[
\begin{align*}
\overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1} [\overline{u}]) \circ \overline{w}_1 \right)
\end{align*}
\]

\[
\begin{align*}
\overline{w}_2 \cdots \overline{w}_{m-1} \left( \overline{u} \circ \left( ([\overline{u}] \circ \overline{v}) ([\overline{u}]^{-1}) \right) \circ \overline{w}_{n}'^{-1} \right) & \quad \text{(Eqs. (11), (10), and (14)).}
\end{align*}
\]

Hence Eq. (15) holds.
Case 1.4. Suppose \( \overline{v}_m, \overline{w}_n \not\in [\mathcal{R}(X)]^{-1} \). Then we have
\[
(u \odot v) \circ w
= (\lfloor u \rfloor \circ \lfloor v \rfloor) \circ \lfloor w \rfloor
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \lfloor w \rfloor
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \left[ \left( \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \overline{w} \right]
\]
(Eqs. (11) and (10))
\[
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(Eqs. (11) and (10))
\[
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(inserting 1 = \( [\overline{u}]^{-1} \overline{u} \))
\[
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(Eq. (14))
\[
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(induction on depth)
\[
= \left[ \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right] \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(Eqs. (11) and (10))
\[
= \left( \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right) \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(Eq. (14))
\[
= \left( \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right) \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(Eq. (14))
\[
= \left( \overline{u} \circ \left( \overline{[u]} \circ \overline{v}[\overline{u}]^{-1} \right) \right) \circ \left( [u] \circ \overline{v}[u]^{-1} \right] \circ \overline{w} \right]
\]
(Eqs. (11), (10), and (14)).

Therefore, Eq. (15) is valid.

Case 2. Suppose \( u \not\in [\mathcal{R}(X)] \). Then \( u \in \mathcal{X} \) or \( u \in [\mathcal{R}(X)]^{-1} \). If \( u \in \mathcal{X} \), then
\[
(u \odot v) \circ w = (uv) \circ w = u(\circ w) = u \odot (v \circ w).
\]
Suppose \( u \not\in \mathcal{X} \), i.e., \( u \in [\mathcal{R}(X)]^{-1} \). If \( v \not\in [\mathcal{R}(X)]^{-1} \), then Eq.(16) holds again. Assume \( v \in [\mathcal{R}(X)]^{-1} \).

If \( w \not\in [\mathcal{R}(X)]^{-1} \), then
\[
u \circ (v \circ w) = u \circ (vw) = (u \circ v)w = (u \circ v) \circ w.
\]

If \( w \in [\mathcal{R}(X)]^{-1} \), let \( u = (u')^{-1} \), \( v = (v')^{-1} \) and \( w = (w')^{-1} \) for some \( u', v', w' \in [\mathcal{R}(X)] \). On the one hand,
\[
(u \odot v) = \left( (u')^{-1} \circ (v')^{-1} \right) \circ (w')^{-1} = (v' \circ u')^{-1} \circ (w')^{-1} \in \left( w' \circ (v' \circ u') \right)^{-1}.
\]
On the other hand,
\[ u \circ (v \circ w) = (u')^{-1} \circ ((v')^{-1} \circ (w')^{-1}) = (u')^{-1} \circ (w' \circ v')^{-1} = ((w' \circ v') \circ u')^{-1}. \]

Then the proof can be reduced to Case 1.

**Case 3.** Suppose \( w \notin [\mathcal{R}(X)] \). The proof is similar to Case 2.

**Case 4.** Suppose \( v \notin [\mathcal{R}(X)] \). Then \( v \in X \) or \( v \in [\mathcal{R}(X)]^{-1} \). This case is similar to Case 2 by symmetry. This completes the initial step of the induction on \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) \).

**II. The inductive step on \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) \).** For a given \( k \geq 3 \), suppose that Eq. (15) has been verified for all \( u, v, w \in \mathcal{R}(X) \) with \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) \leq k \) and consider the case when \( \text{bre}(u) + \text{bre}(v) + \text{bre}(w) = k + 1 \). Then \( k + 1 \geq 4 \) and so at least one of \( \text{bre}(u), \text{bre}(v) \) or \( \text{bre}(w) \) is greater than one. We accordingly have the following three cases to verify.

**Case 1.** Suppose \( \text{bre}(u) \geq 1 \). Then we write \( u = u_1 \cdots u_m \) in the standard factorization with \( m \geq 2 \). Then by Eq (14) and the induction hypothesis we have
\[
(u \circ v) \circ w = ((u_1 \cdots u_m) \circ v) \circ w
= ((u_1 \cdots u_{m-1})(u_m \circ v)) \circ w
= (u_1 \cdots u_{m-1})((u_m \circ v) \circ w)
= (u_1 \cdots u_{m-1})(u_m \circ (v \circ w))
= (u_1 \cdots u_{m-1}u_m) \circ (v \circ w)
= u \circ (v \circ w).
\]

**Case 2.** Suppose \( \text{bre}(w) \geq 2 \). Then the verification is similar to the previous case.

**Case 3.** Suppose \( \text{bre}(v) \geq 2 \). Write \( v = v_1 \cdots v_n \) in the standard factorization with \( n \geq 2 \). Then repeatedly applying Eq. (14), we obtain
\[
(u \circ v) \circ w = (u \circ (v_1 \cdots v_n)) \circ w
= ((u \circ v_1)v_2 \cdots v_n) \circ w
= (u \circ v_1)(v_2 \cdots v_{n-1})(v_n \circ w)
= u \circ (v_1 \cdots v_{n-1}v_n \circ w)
= u \circ (v_1 \cdots v_n) \circ w
= u \circ (v \circ w).
\]

Now we have completed the inductive proof of the associativity in Eq. (15).

Finally, we verify Eq. (2). Let \( u = [\overline{u}] \) and \( v = [\overline{v}] \) be in \( [\mathcal{R}(X)] \). Write \( \overline{v} = \overline{v}_1 \cdots \overline{v}_k \) in the standard factorization. If \( \overline{v}_k = (\overline{v}_k')^{-1} \in [\mathcal{R}(X)]^{-1} \), we have
\[
\overline{\text{Ad}}_u \overline{v} = (u \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{k-1} (u \circ \overline{v}_k)^{-1}
= (u \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{k-1} (\overline{v}_k \circ u^{-1})
= ((u \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_{k-1} \overline{v}_k) \circ u^{-1}
= u \circ v \circ u^{-1}
\quad \text{(Eq. (14))}
\]

If \( \overline{v}_k \notin [\mathcal{R}(X)]^{-1} \), we obtain
\[
\overline{\text{Ad}}_u \overline{v} = (u \circ \overline{v})u^{-1}
= ((u \circ \overline{v}_1) \overline{v}_2 \cdots \overline{v}_k)u^{-1}
\quad \text{(Eq. (14))}
\]
\begin{align}
&= (u \circ v_1) v_2 \cdots v_{k-1} (v_k \circ u^{-1}) \\
&= (u \circ v_1) v_2 \cdots v_k \circ u^{-1} \\
&= u \circ v \circ u^{-1}
\end{align}
(Eq. (13))
(Eq. (14))
(Eq. (14)).

Hence
\[ u \circ v = [u \circ \text{Ad}_v] = [u \circ v \circ u^{-1}], \]
and so Eq. (2) holds. This completes the proof. \hfill \Box

**Theorem 4.4.** Let \( X \) be a set. The triple \((\mathcal{R}(X), [\cdot, \cdot], \circ)\), together with the inclusion \( i_X : X \to \mathcal{R}(X) \), is a free Rota-Baxter group of weight 1 on \( X \).

**Proof.** Let \((G, P)\) be a Rota-Baxter group and \( f : X \to G \) a set map. It suffices to show that there is a unique Rota-Baxter group homomorphism
\[ \tilde{f} : \mathcal{R}(X) \to G, w \mapsto \tilde{f}(w) \]
such that \( f = \tilde{f} i_X \).

**(Existence)** We use induction on \( \text{dep}(w) \geq 0 \) to define \( \tilde{f}(w) \).

For the initial step of \( \text{dep}(w) = 0 \), if \( \text{bre}(w) \leq 1 \), then \( w = 1 \) or \( w \in X \cup X^{-1} \). Then define
\begin{equation}
\tilde{f}(w) := \begin{cases}
1_G, & \text{if } w = 1, \\
f(w), & \text{if } w \in X,
\end{cases}
\end{equation}
(17)
If \( \text{bre}(w) \geq 2 \), write \( w = w_1 \cdots w_n \) in the standard factorization with \( n \geq 2 \) and define
\begin{equation}
\tilde{f}(w) := \tilde{f}(w_1 \cdots w_n) := \tilde{f}(w_1) \cdots \tilde{f}(w_n),
\end{equation}
(18)
where each \( \tilde{f}(w_i) \) is defined in Eq. (17).

For a given \( k \geq 0 \), assume that \( \tilde{f}(w) \) have been defined for \( w \in \mathcal{R}(X) \) with \( \text{dep}(w) \leq k \) and consider \( w \in \mathcal{R}(X) \) with \( \text{dep}(w) = k + 1 \).

First consider \( w \) with \( \text{bre}(w) = 1 \). From \( \text{dep}(w) = k + 1 \geq 1 \), we have \( w \in [\mathcal{R}(X)] \) or \( w \in [\mathcal{R}(X)]^{-1} \).

In the first case, write \( w = [\overline{w}] \) for some \( \overline{w} \in \mathcal{R}(X) \) with \( \overline{w} = k \). So the induction hypothesis allows us to define
\begin{equation}
\tilde{f}(w) := \tilde{f}([\overline{w}]) := P(\tilde{f}(\overline{w})).
\end{equation}
(19)
In the second case, write \( w = [\overline{w}]^{-1} \) for some \( \overline{w} \in \mathcal{R}(X) \) with \( \overline{w} = k \). Then the inductive hypothesis allows us to define
\begin{equation}
\tilde{f}(w) := \tilde{f}([\overline{w}]^{-1}) := P(\tilde{f}(\overline{w}))^{-1}.
\end{equation}
(20)

Next consider \( w \) with \( \text{bre}(w) \geq 2 \). Write \( w = w_1 \cdots w_n \) in the standard factorization with \( n \geq 2 \) and define
\begin{equation}
\tilde{f}(w) := \tilde{f}(w_1 \cdots w_n) := \tilde{f}(w_1) \cdots \tilde{f}(w_n),
\end{equation}
(21)
where each \( \tilde{f}(w_i) \) is defined in Eqs. (17), (19) or (20).

By Eq. (19), we have \( \overline{f} \circ [\cdot] = P \circ \overline{f} \). So we are left to prove that \( \tilde{f} \) is a group homomorphism:
\begin{equation}
\tilde{f}(u \circ v) = \tilde{f}(u) \tilde{f}(v) \quad \text{for all } u, v \in \mathcal{R}(X),
\end{equation}
(22)
which will be achieved by an induction on \( (\text{dep}(u), \text{dep}(v)) \geq (0, 0) \) lexicographically.

If \( (\text{dep}(u), \text{dep}(v)) = (0, 0) \), then \( u, v \in G_0 \) and Eq. (22) follows from Eq. (18). Next fix \( m, n \geq 0 \) with \( (m, n) > (0, 0) \) lexicographically. Assume that Eq. (22) is valid for \( u, v \in \mathcal{R}(X) \) with \( (\text{dep}(u), \text{dep}(v)) <
\[(m, n)\] lexicographically, and consider \(u, v \in \mathcal{R}(X)\) with \((\text{dep}(u), \text{dep}(v)) = (m, n)\). According to the four cases in the definition of \(\circ\), we have the following four cases to verify.

**Case 1.** Suppose that \(u = [\bar{u}]\) and \(v = [\bar{v}]\) for some \(\bar{u}, \bar{v} \in \mathcal{R}(X)\). Write \(\bar{v} = \bar{v}_1 \cdots \bar{v}_k\) in the standard factorization. According to \(\bar{v}_k\) in \([\mathcal{R}(X)]^{-1}\) or not, we consider two subcases.

**Subcase 1.1.** \(\bar{v}_k = (\bar{v}'_k)^{-1} \in [\mathcal{R}(X)]^{-1}\) with \(\bar{v}'_k \in [\mathcal{R}(X)]\). Then

\[
\begin{align*}
\bar{f}(u \circ v) &= \bar{f}([\bar{u}] \circ [\bar{v}]) \\
&= \bar{f}\left([\bar{u}] \circ \left((u \circ \bar{v}_1)\bar{v}_2 \cdots \bar{v}_{k-1}(u \circ \bar{v}'_k)^{-1}\right)\right) & \text{(Eqs. (11) and (10))} \\
&= P\left(\bar{f}\left([\bar{u}] \circ \left((u \circ \bar{v}_1)\bar{v}_2 \cdots \bar{v}_{k-1}(u \circ \bar{v}'_k)^{-1}\right)\right)\right) & \text{(Eq. (19))} \\
&= P\left(\bar{f}(\bar{u})\bar{f}\left((u \circ \bar{v}_1)\bar{v}_2 \cdots \bar{v}_{k-1}(u \circ \bar{v}'_k)^{-1}\right)\right) & \text{(induction on depth)} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(u \circ \bar{v}'_k)^{-1}\right) & \text{(Eq. (21))} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(\bar{v}'_k)^{-1}\right) & \text{(induction on depth and Eq. (20))} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(u \circ \bar{v}'_k)^{-1}\right) & \text{(induction on depth)} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(u \circ \bar{v}'_k)^{-1}\right) & \text{(Eq. (20))} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(\bar{v}'_k)^{-1}\right)(\bar{f}(\bar{u})\bar{f}(u \circ \bar{v}'_k)^{-1}) & \text{\((\bar{v}_k = (\bar{v}'_k)^{-1})\)} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(u \circ \bar{v}'_k)^{-1}\right) & \text{\((\bar{v}_k = (\bar{v}'_k)^{-1})\)} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v}_1)\bar{f}(\bar{v}_2) \cdots \bar{f}(\bar{v}_{k-1})\bar{f}(u \circ \bar{v}'_k)^{-1}\right) & \text{\((\bar{v} = \bar{v}_1 \cdots \bar{v}_k)\)} \\
&= P\left(\bar{f}(\bar{u})\bar{f}(u)\bar{f}(\bar{v})\right) & \text{\(((G, P)\ being\ a\ Rota-Baxter\ group))} \\
&= \bar{f}(\bar{u})\bar{f}(\bar{v}) & \text{\((\text{Eq. (19))}) \\
&= \bar{f}(\bar{u})\bar{f}(\bar{v}).
\end{align*}
\]

**Subcase 1.2.** \(\bar{v}_k \notin [\mathcal{R}(X)]^{-1}\). Then

\[
\begin{align*}
\bar{f}(u \circ v) &= \bar{f}([\bar{u}] \circ [\bar{v}]) \\
&= \bar{f}\left([\bar{u}] \circ ((u \circ \bar{v})u^{-1})\right) & \text{(Eqs. (11) and (10))} \\
&= P\left(\bar{f}\left([\bar{u}] \circ ((u \circ \bar{v})u^{-1})\right)\right) & \text{(Eq. (19))} \\
&= P\left(\bar{f}(\bar{u})\bar{f}((u \circ \bar{v})u^{-1})\right) & \text{(induction on depth)}
\end{align*}
\]
\[ P \left( f(u) f(v) \right) = P \left( f(u) \right) P \left( f(v) \right) \]

\textbf{Case 2.} Suppose \( u = \lfloor u \rfloor^{-1} \) and \( v = \lfloor v \rfloor^{-1} \) for some \( u, v \in \mathcal{R}(X) \). Then

\[ f(u \diamond v) = f(\lfloor u \rfloor \diamond \lfloor v \rfloor) = f(\lfloor u \rfloor) f(\lfloor v \rfloor). \]

\textbf{Case 3.} Suppose \( b r e(u) = b r e(v) = 1 \), but \( u, v \) are not in the previous two cases. Then by definition, \( u \diamond v = uv \) is the concatenation. It follows from Eq. (21) that

\[ f(u \diamond v) = f(u) f(v). \]

\textbf{Case 4.} Suppose in general \( b r e(u), b r e(v) \geq 1 \). Let \( u = u_1 \cdots u_m \) and \( v = v_1 \cdots v_n \) be in the standard factorizations. Then

\[ f(u \diamond v) = f(u_1 \cdots u_{m-1} (u_m \diamond v_1)v_2 \cdots v_n) = f(u_1) \cdots f(u_{m-1}) f(u_m \diamond v_1) f(v_2) \cdots f(v_n). \]

In summary, the map \( \bar{f} \) is a Rota-Baxter group homomorphism.

\textbf{(Uniqueness)} Notice that Eqs. (17)–(21) give the only possible way to define \( \bar{f}(w) \) in order for \( \bar{f} \) to be a Rota-Baxter group homomorphism that extends \( f \). This proves the uniqueness.

Now the proof of Theorem 4.4 is completed.

\textbf{Acknowledgments}

The authors thank the referees for their helpful reports.
Data availability statement
The manuscript has no associated data.

Disclosure statement
On behalf of all authors, the corresponding author states that there is no conflict of interest.

Funding
This work is supported by NSFC (12071191 and 12171211), the Natural Science Foundation of Jiangxi Province (20192ACB21008) and the Natural Science Foundation of Shandong Province (ZR2020MA002).

References

[1] Angiono, I., Galindo, C., Vendramin, L. (2017). Hopf braces and Yang-Baxter operators. *Proc. Am. Math. Soc.* 145:1981–1995.

[2] Bai, C., Guo, L., Ni, X. (2010). Nonabelian generalized Lax pairs, the classical Yang-Baxter equation and PostLie algebras. *Commun. Math. Phys.* 297:553–596.

[3] Bai, C., Guo, L., Sheng, Y., Tang, R. (2023). Post-groups, (Lie)-Butcher groups and the Yang-Baxter equation. *Math. Ann.* DOI: 10.1007/s00208-023-02592-z.

[4] Bardakov, V. G., Gubarev, V. (2022). Rota-Baxter groups, skew left braces, and the Yang-Baxter equation. *J. Algebra* 596:328–351.

[5] Bardakov, V. G., Gubarev, V. (2023). Rota-Baxter operators on groups. *Proc. Math. Sci.* 133:4. DOI: 10.1007/s12044-023-00723-9.

[6] Bardakov, V. G., Neshchadim, M. V., Yadav, M. K. (2022). Symmetric skew braces and brace systems. arXiv:2204.12247.

[7] Baxter, G. (1960). An analytic problem whose solution follows from a simple algebraic identity. *Pacific J. Math.* 10:731–742.

[8] Bestivina, M., Feighn, M., Handel, M. (2000). The Tits alternative for Out(Fn). I. Dynamics of exponentially-growing automorphisms. *Ann. Math.* 151:517–623.

[9] Bestivina, M., Feign, M., Handel, M. (2005). The Tits alternative for Out(Fn). II. A Kolchin type theorem. *Ann. Math.* 161:1–59.

[10] Bourbaki, N. (1974). *Elements of Mathematics: Algebra I*. Paris: Hermann, chapters 1–3.

[11] Carinena, J., Grabowski, J., Marmo, G. (2000). Quantum bi-Hamiltonian systems. *Int. J. Modern Phys. A.* 15:4797–4810.

[12] Catino, F., Mazzotta, M., Stefanelli, P. (2023). Rota-Baxter operators on Clifford semigroups and the Yang-Baxter equation. *J. Algebra* 622(15):587–613.

[13] Caranti, A., Stefanelli, L. (2023). Skew braces from Rota-Baxter operators: A cohomological characterisation and an example. *Annali di Matematica* 202:1–13.

[14] Cedó, F., Jespers, E., Okniński, J. (2014). Braces and the Yang-Baxter equation. *Commun. Math. Phys.* 327:101–116.

[15] Connes, A., Kreimer, D. (2000). Renormalization in quantum field theory and the Riemann-Hilbert problem. I. *Commun. Math. Phys.* 210:249–273.

[16] Das, A. (2020). Deformations of associative Rota-Baxter operators. *J. Algebra* 560:144–180.

[17] Etingof, P., Schedler, T., Soloviev, A. (1999). Set-theoretical solutions to the quantum Yang-Baxter equation. *Duke Math. J.* 100:169–209.

[18] Gao, X., Guo, L. (2017). Rota's Classification Problem, rewriting systems and Gröbner-Shirshov bases. *J. Algebra* 470:219–253.

[19] Gao, X., Guo, L., Zhang, H. (2022). Rota's program on algebraic operators, rewriting systems and Gröbner-Shirshov bases. *Adv. Math. (China)* 51:1–31.

[20] Goncharov, M. (2021). Rota-Baxter operators on cocommutative Hopf algebras. *J. Algebra* 582:39–56.

[21] Gubarev, V. (2018). Universal enveloping associative Rota–Baxter algebras of preassociative and postassociative algebra. *J. Algebra* 516:298–328.

[22] Guo, L. (2009). Operated semigroup, Motzkin paths and rooted tree. *J. Algebraic Combinatorics* 29:35–62.

[23] Guo, L. (2012). *An introduction to Rota-Baxter Algebras, Surveys of Modern Mathematics 4*. Somerville: International Press; Beijing: Higher Education Press.
[24] Guo, L., Gustavson, R., Li, Y. (2022). An algebraic study of Volterra integral equations and their operator linearity. J. Algebra 595:398–433.
[25] Guo, L., Keigher, W. (2008). On differential Rota-Baxter algebras. J. Pure Appl. Algebra 212:522–540.
[26] Guo, L., Lang, H., Sheng, Y. (2021). Integration and geometrization of Rota-Baxter operators. Adv. Math. 387:107834.
[27] Guo, L., Sit, W., Zhang, R. (2013). Differential type operators and Gröbner-Shirshov bases. J. Symbolic Comput. 52:97–123.
[28] Jiang, J., Sheng, Y. (2022). Representations and cohomologies of relative Rota-Baxter Lie algebras and applications. J. Algebra 602:637–670.
[29] Jiang, J., Sheng, Y., Zhu, C. (2021). Lie theory and cohomology of relative Rota-Baxter operators. arXiv:2108:02627.
[30] Kampé de Fériet, J. (1934). L’etat actuel du probleme de la turbulaence (I and II). La Sci. Aérienne 3:9–34; 4(1935):12–52.
[31] Karimjanov, I., Kaygorodov, I., Ladra, M. (2020). Rota-type operators on null-filiform associative algebras. Linear Multilinear Algebra 68:205–219.
[32] Kolchin, E. R. (1973). Differential Algebra and Algebraic Group. New York: Academic Press.
[33] Kupershmidt, B. A. (1999). What a classical r-matrix really is. J. Nonlinear Math. Phys. 6:448–488.
[34] Lang, H., Sheng, Y. (2023). Factorizable Lie bialgebras, quadratic Rota-Baxter Lie algebras and Rota-Baxter Lie bialgebras. Commun. Math. Phys. 397:763–791.
[35] Liu, X., Guo, L., Guo, X. (2019). λ-Differential operators and differential modules for the Virasoro algebra. Linear Multilinear Algebra 67:1308–1324.
[36] Lu, J., Yan, M., Zhu, Y. (2000). On the set-theoretical Yang-Baxter equation. Duke Math. J. 104:1–18.
[37] Mackenzie, K. (2005). General Theories of Lie Groupoids and Lie Algebroids. Cambridge: Cambridge University Press.
[38] MacLane, S. (1998). Categories for the Working Mathematician (2nd ed.), New York: Springer.
[39] Munthe-Kaas, H. Z., Lundervold, A. (2013). On post-Lie algebras, Lie-Butcher series and moving frames. Found. Comput. Math. 13:583–613.
[40] Munthe-Kaas, H. Z., Wright, W. M. (2008). On the Hopf algebraic structure of Lie group integrators. Found. Comput. Math. 8:227–257.
[41] Nijenhuis, A. (1951). X_n−1-forming sets of eigenvectors. Indag. Math. 13:200–212.
[42] Noether, E. (1927). Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern. Math. Ann. 96:26–61.
[43] van der Put, M., Singer, M. (1997). Galois Theory of Difference Equations. Berlin: Springer.
[44] Reyman, A. G., Semenov-Tian-Shansky, M. A. (1979). Reduction of Hamilton systems, affine Lie algebras and Lax equations. Invent. Math. 54:81–100.
[45] Reyman, A. G., Semenov-Tian-Shansky, M. A. (1981). Reduction of Hamilton systems, affine Lie algebras and Lax equations II. Invent. Math. 63:423–432.
[46] Ritt, J. F. (1934). Differential Equations from the Algebraic Standpoint. American Mathematical Society Colloquium Publications, Vol. 14, New York: American Mathematical Society.
[47] Rota, G.-C. (1969). Baxter algebras and combinatorial identities, I, II. Bull. Amer. Math. Soc. 75:325–329.
[48] Rota, G.-C. (1995). Baxter operators, an introduction. In: Kung, J. P. S., ed. Gian-Carlo Rota on Combinatorics, Introductory Papers and Commentaries. Boston: Birkhäuser.
[49] Semenov-Tian-Shansky, M. A. (1983). What is a classical r-matrix? Funct. Anal. Appl. 17(4):259–272.
[50] Serre, J.-P. (1997). Galois Cohomology. Berlin: Springer.
[51] Stefanello, L. (2022). Skew braces that do not come from Rota-Baxter operators, preprint.
[52] van der Waerden, B. L. (1949). Modern Algebra. Vol. 1. New York: Frederick Ungar.
[53] Zhang, T., Gao, X., Guo, L. (2021). Reynolds algebras and their free objects from bracketed words and rooted trees. J. Pure. Appl. Algebra 225:106766.
[54] Zhang, Y., Gao, X., Guo, L. (2020). Matching Rota-Baxter algebras, matching dendriform algebras and matching pre-Lie algebras. J. Algebra 552:134–170.
[55] Zhang, Y., Gao, X., Guo, L. (2022). Hopf algebra of multidecorated rooted forests, free matching Rota-Baxter algebras and Gröbner-Shirshov bases. Pac. J. Math. 317:441–475.
[56] Zheng, Y. Y., Gao, X., Manchon, D. (2020). Free (tri)dendriform family algebras. J. Algebra 547:456–493.
[57] Zheng, S., Gao, X., Guo, L., Sit, W. Rota-Baxter type operators, rewriting systems and Gröbner-Shirshov bases. J. Symb. Comput., accepted, arXiv:1412.8055v1.