Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Future Generation Computer Systems 19 (2003) 1197–1205

FGCS
FUTURE GENERATION COMPUTER SYSTEMS

www.elsevier.com/locate/future

On enumeration problems in Lie–Butcher theory

H. Munthe-Kaas*, S. Krogstad

Department of Computer Science, University of Bergen, N-5020 Bergen, Norway

Abstract

The algebraic structure underlying non-commutative Lie–Butcher series is the free Lie algebra over ordered trees. In this paper we present a characterization of this algebra in terms of balanced Lyndon words over a binary alphabet. This yields a systematic manner of enumerating terms in non-commutative Lie–Butcher series.

Keywords: Butcher theory; Trees; Enumeration; Lie group integrators; Runge–Kutta methods

1. Summary

Let $\mathfrak{g}$ be the free Lie algebra over ordered trees, with a grading induced by the number of nodes in the trees. In the first part of this paper we prove that $\mathfrak{g}_n$, the homogeneous component of degree n, has dimension:

$$\dim(\mathfrak{g}_n) = \frac{1}{2n}\sum_{d|n}\mu\left(\frac{n}{d}\right)\binom{2d}{d}.$$

For $n = 1, 2, 3, \ldots$ these numbers are 1, 1, 3, 8, 25, 75, 245, 800, 2700, 9225, 32 065, ... This sequence counts the number of order conditions in the Crouch–Grossman–Owren–Marthinsen methods for integrating differential equations on manifolds. It does also count the number of balanced Lyndon words over a binary alphabet.

In the last part of the paper we establish explicitly the 1–1 correspondence between balanced binary Lyndon words and Hall basis elements for the free Lie algebra of ordered trees.

2. Introduction to Lie–Butcher theory

Lie series, dating back to Sophus Lie (1842–1899), is a version of Taylor series adapted to general manifolds. Butcher series, invented by Butcher [4,5] and developed further in [8] (see also [9]) is an adaption of Taylor series to the study of order conditions of Runge–Kutta methods. In recent papers [6,10,12,14,15,17,19], efforts have been made to extend the theory of Runge–Kutta methods from $\mathbb{R}^n$ to Lie groups and homogeneous spaces. Some very interesting recent developments establish the connection between Butcher theory and renormalization theory in physics [3]. In the light of this development it seems important to further investigate the algebraic structure of non-commutative Lie–Butcher series on manifolds.

Without going into details we will in this section briefly review some basic results, yielding the connection between Lie–Butcher theory and the *free Lie algebra over ordered trees*.

Definition 1. The set of ordered trees (OT) is defined by the following rules:

(1) The one node tree is an ordered tree, $\bullet \in \mathrm{OT}$.

* Corresponding author.
E-mail addresses: hans@ii.uib.no (H. Munthe-Kaas), stein@ii.uib.no (S. Krogstad).

0167-739X/$ – see front matter © 2003 Published by Elsevier Science B.V.
doi:10.1016/S0167-739X(03)00045-1

(2) If $\tau_1, \tau_2, \dots, \tau_k \in \mathrm{OT}$, then the ordered k-tuple $\tau = (\tau_1, \tau_2, \dots, \tau_k) \in \mathrm{OT}$. We call τ_i the *branches* of τ.

We let $|\tau|$ denote the *degree* of the tree (number of nodes):

$$|\bullet| = 1, \qquad |(\tau_1, \dots, \tau_k)| = |\tau_1| + \cdots + |\tau_k| + 1. \tag{1}$$

The number of ordered trees of degree $n+1$ is given by the *Catalan numbers*:

$$C(n) = \frac{(2n)!}{n!(n+1)!}.$$

The number of ordered trees with 1, 2, 3, ... nodes is 1, 1, 2, 5, 14, 42, 132, ...

We will briefly introduce Lie–Butcher series, related to Runge–Kutta methods on manifolds. Let $\mathcal{M}$ be a homogeneous space acted upon transitively from left by a Lie group G, $p \mapsto g \cdot p$, where $g \in G$, $p \in \mathcal{M}$. Let $\mathfrak{g}$ be the Lie algebra of G and $\exp : \mathfrak{g} \to G$ the exponential map. For $V \in \mathfrak{g}$ and $p \in \mathcal{M}$, we let $V \cdot p \in T_p\mathcal{M}$ denote the tangent:

$$V \cdot p = \left.\frac{\partial}{\partial s}\right|_{s=0} \exp(sV) \cdot p.$$

A differential equation on $\mathcal{M}$ can always be written in the form:

$$y'(t) = f(y(t)) \cdot y(t) \tag{2}$$

for some function $f : \mathcal{M} \to \mathfrak{g}$.

Let $\mathcal{F}_\mathcal{V}(\mathcal{M}) = (\mathcal{M} \to \mathcal{V})$ denote all smooth functions from $\mathcal{M}$ to some vector space[1] $\mathcal{V}$. Any element $V \in \mathfrak{g}$ induces a (right invariant) derivation $V[\cdot] : \mathcal{F}_\mathcal{V}(\mathcal{M}) \to \mathcal{F}_\mathcal{V}(\mathcal{M})$ as

$$V[\phi](p) = \left.\frac{\partial}{\partial s}\right|_{s=0} \phi(\exp(sV) \cdot p) \quad \text{for any } \phi \in \mathcal{F}_\mathcal{V}(\mathcal{M}).$$

Similarly, higher order derivations are defined by iterating this definition:

$$(V \cdot W)[\phi] = V[W[\phi]].$$

Linear combinations of derivations, and the identity (0th order derivation) are defined in the obvious way.

The resulting linear space of all 0th and higher order right invariant derivations is denoted $\mathfrak{G}$, *the universal enveloping algebra* of $\mathfrak{g}$. We let $\exp : \mathfrak{g} \to \mathfrak{G}$ denote the formal exponential series:

$$\exp(V) = \sum_{j=0}^{\infty} \frac{V^j}{j!}.$$

Given a point $p \in \mathcal{M}$, the *Lie series* development of a function $\phi \in \mathcal{F}_\mathcal{V}(\mathcal{M})$ is given as

$$\phi(\exp(tV) \cdot p) = \exp(tV)[\phi](p).$$

In the *commutative case* where $G = \mathcal{M} = \mathbb{R}^n$, acting upon itself by translations, this is just Taylor series.

In the study of numerical methods for integrating (2), it is very useful to characterize a curve $z(t) \in \mathcal{M}$, $z(0) = p$, in terms of *elementary differentials* of f.

Definition 2. Given a point $p \in \mathcal{M}$ and a smooth $f : \mathcal{M} \to \mathfrak{g}$. To every $\tau \in \mathrm{OT}$ there corresponds an elementary differential $F_p(\tau) \in \mathfrak{g}$, defined as follows:

$$\begin{aligned} &F_p(\bullet) = f(p), \\ &F_p((\tau_1, \tau_2, \dots, \tau_k)) \\ &\quad = (F_p(\tau_1) \cdot F_p(\tau_2) \cdots F_p(\tau_k))[f](p). \end{aligned}$$

There are various ways of characterizing a curve $z(t) \in \mathcal{M}$, $z(0) = p$.

- The approach in [16] is based on a time dependent infinitesimal generator. Let a function $a : \mathrm{OT} \to \mathbb{R}$ define a curve $\nu(t, a, p) \in \mathfrak{g}$ as

  $$\nu(t, a, p) = \sum_{\tau \in \mathrm{OT}} \frac{t^{|\tau|-1}\alpha(\tau)}{(|\tau|-1)!} a(\tau) F_p(\tau).$$

  We characterize $z(t)$ as the solution of a differential equation:

  $$z'(t) = \nu(t, a, p) \cdot z(t), \qquad z(0) = p. \tag{3}$$

  We define $\alpha : \mathrm{OT} \to \mathbb{R}$ to be the function satisfying the analytical solution to (2) for $a(\tau) = 1$ for all τ. In [19] it is shown that $\alpha(\bullet) = 1$, and for $\tau = (\tau_1, \tau_2, \dots, \tau_k)$

  $$\alpha(\tau) = \prod_{i=1}^{k} \binom{\sum_{j=1}^{i} |\tau_j| - 1}{|\tau_i| - 1} \alpha(\tau_i) :$$

[1] In this paper we consider only $\mathcal{V} = \mathbb{R}$ or $\mathcal{V} = \mathfrak{g}$.

- A *pullback series* is the basis for the analysis of Crouch–Grossman [6] and Owren–Marthinsen [19]. Let a function $b : \mathrm{OT} \to \mathbb{R}$ define a higher order differential operator $\mu(t, b, p) \in \mathfrak{G}$ as

$$\mu(t,b,p) = \sum_{\tau=(\tau_1,\tau_2,\dots,\tau_k)\in\mathrm{OT}} \frac{t^{|\tau|-1}\alpha(\tau)}{(|\tau|-1)!} b(\tau) \times (F_p(\tau_1)\cdot F_p(\tau_2)\cdots F_p(\tau_k)).$$

The curve $z(t)$ is characterized by the requirement that

$$\phi(z(t)) = \mu(t,b,p)[\phi](p) \quad \text{for any } \phi \in \mathcal{F}_{\mathbb{R}}(\mathcal{M}).$$

In a formal sense, the pullback series is the exponential of the infinitesimal generator. It can be shown that for a given $z(t)$ we have

$$\mu(1,b,p) = \exp\left(\nu(s,a,p) + \frac{\partial}{\partial s}\right)\bigg|_{s=0},$$

where $\nu(s, a, p)$ is defined in (3). This yields

$$b((\tau_1,\tau_2,\dots,\tau_k)) = a(\tau_1)a(\tau_2)\cdots a(\tau_k).$$

From this we see that the analytical solution $y(t)$ is again characterized by $b(\tau) = 1$.

- The *development* of $z(t)$ is given as a curve $\sigma(t) \in \mathfrak{g}$ such that $z(t) = \exp(\sigma(t)) \cdot p$. The computation of $\sigma(t)$ from the infinitesimal generator $\nu(t, a, p)$ is accomplished by *Magnus series* [10,13,21]. In this computation, also commutators of elementary differentials appears. If a curve $z(t)$ can be represented as (3), then there exists a function $c : \mathrm{fla}(\mathrm{OT}) \to \mathbb{R}$ such that

$$\sigma(t) = \sum_{h\in H(\mathrm{fla}(\mathrm{OT}))} \frac{t^{|h|}\alpha(h)}{(|h|)!} c(h) F_p(h),$$

where $H(\mathrm{fla}(\mathrm{OT}))$ denotes a Hall basis for the free Lie algebra of ordered trees, to be defined below. (We do not discuss the extension of α from OT to $H(\mathrm{fla}(\mathrm{OT}))$ here.) Each $h \in H(\mathrm{fla}(\mathrm{OT}))$ is a (formal) commutator of trees, $F_p(h)$ denotes the corresponding commutator of elementary differentials and $|h|$ denotes the sum of the degree of all trees in the commutator. It should be noted that in the commutative theory all non-trivial commutators vanish

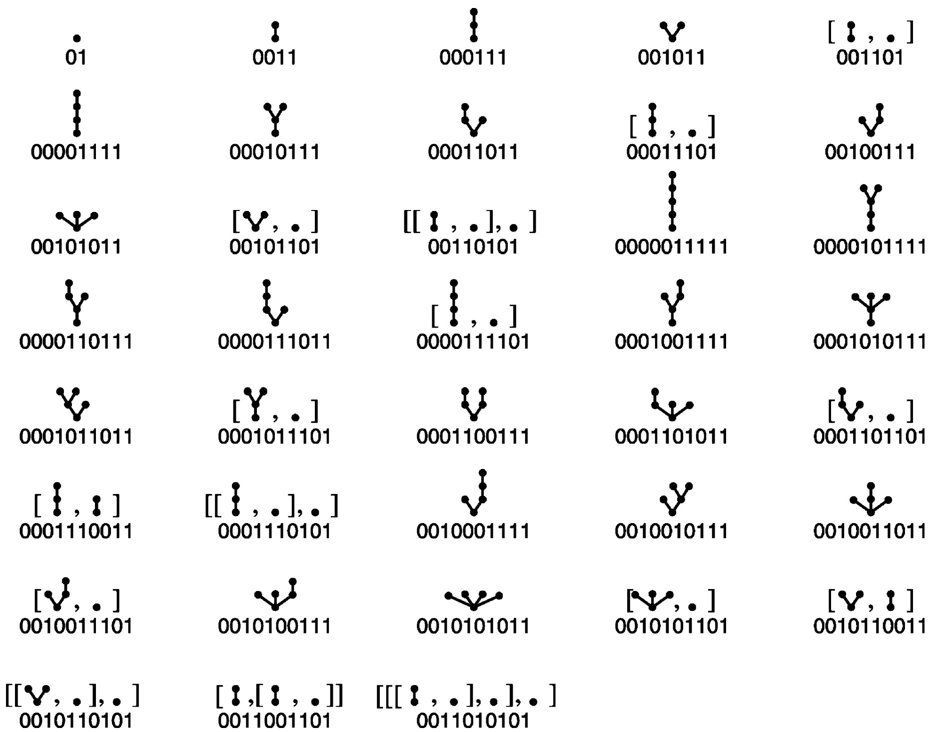

Fig. 1. Basis elements for the free Lie algebra of ordered trees up to order 5, and their corresponding binary balanced Lyndon words.

and $\sigma(t)$ reduces to the classical B-series as defined in [9].

As we see from this review, the free Lie algebra of ordered trees plays a central role in the Lie–Butcher theory. Our goal in this paper is to present a basis for this algebra (see Fig. 1 for a graphical presentation of the lowest order elements).

3. The free Lie algebra of ordered trees

Definition 3. A Lie algebra $\mathfrak{g}$ is *free* over the set I if:

(i) For every $i \in I$ there corresponds an element $X_i \in \mathfrak{g}$.
(ii) For any Lie algebra $\mathfrak{h}$ and any function $i \mapsto Y_i \in \mathfrak{h}$, there exists a unique Lie algebra homomorphism $\pi : \mathfrak{g} \to \mathfrak{h}$ satisfying $\pi(X_i) = Y_i$ for all $i \in I$.

This is written as $\mathfrak{g} = \mathrm{fla}(I)$.

A concrete representation of $\mathfrak{g}$ as a vector space over (formal) Lie brackets of elements in I, is discussed in Section 5.

The degree function (1) on OT is naturally extended to $\mathfrak{g}$, for any bracket $[\tau_1, \tau_2]$ we define

$$|[\tau_1, \tau_2]| = |\tau_1| + |\tau_2|.$$

Counting theorems for such *graded free Lie algebras* are developed in [18]. Let $\mathfrak{g}_j$ denote the homogeneous component of degree j, defined as the subspace of $\mathfrak{g}$ spanned by commutators of degree j. The following theorem is shown in [18]:

Theorem 4. *Let $\mathfrak{g}$ be the graded free Lie algebra generated by elements $i \in I$, each element with degree $|i| \in \mathbb{Z}^+$. Then*

$$\dim(\mathfrak{g}_n) = \nu_n = \frac{1}{n}\sum_{d|n}\left(\sum_j \lambda_j^{-d}\right)\mu\left(\frac{n}{d}\right), \tag{4}$$

where μ is the Möbius function and λ_j the roots of the polynomial:

$$p(x) = 1 - \sum_{i\in I} x^{|i|}. \tag{5}$$

In the case of $\mathrm{fla}(\mathrm{OT})$, there are $C(i-1)$ generators of grade i, thus

$$p(x) = 1 - \sum_{j=0}^{\infty} C(j)x^{j+1}.$$

The sum $\sum_j \lambda_j^{-d}$ should be interpreted as the trace of the nth power of the inverse of the companion matrix of $p(x)$, in our case given as

$$M = \begin{pmatrix} C(0) & C(1) & C(2) & \cdots \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix}.$$

By recursion one finds that the non-zero diagonal elements of M^d are given as

$$M^d_{i,i} = \mathrm{NCT}(d-1, d-1-i) \quad \text{for} \quad i = 0, \ldots, d-1,$$

where

$$\mathrm{NCT}(n, m) = \sum_{k=0}^{m} C(n-k)C(k)$$

is the new Catalan triangle (sequence A028364 in [11]). Thus

$$\sum_j \lambda_j^{-d} = \mathrm{trace}(M^d) = \sum_{i=0}^{d-1}\sum_{k=0}^{i} C(d-1-k)C(k) = \frac{1}{2}\binom{2d}{d}.$$

We arrive at the following lemma.

Lemma 5. *Let $\mathfrak{g}$ be the free Lie algebra over ordered trees, and let $\mathfrak{g}_n$ be its homogeneous component of grade n. Then*

$$\dim(\mathfrak{g}_n) = \frac{1}{2n}\sum_{d|n}\mu\left(\frac{n}{d}\right)\binom{2d}{d}.$$

For $n = 1, 2, 3, \ldots$ these numbers are 1, 1, 3, 8, 25, 75, 245, 800, 2700, 9225, 32 065, ... This integer sequence occurs in a number of contexts, it is the number of order conditions for Crouch–Grossman–Owren–Marthinsen methods, it counts the number of balanced binary Lyndon words of length $2n$, and there are even interesting connections to knot theory [2], and renormalization theory (see also sequence A022553 in [11]).

In the final part of this paper we will explicitly establish the connection between fla(OT) and balanced binary Lyndon words.

4. Lyndon words

Let $A = \{0, 1\}$ denote the binary alphabet. Let A^+ be the collection of all non-empty words over A. For a word $w \in A^+$ we define the *partial degree* $|w|_\ell$ with respect to a letter $\ell \in A$ the number of occurrences of ℓ in w, while the *balance* is given as

$$\mathrm{bal}(w) = |w|_1 - |w|_0.$$

A binary word is balanced if $\mathrm{bal}(w) = 0$.

Definition 6. Let $<$ denote the following ordering of A^+:

- If $\mathrm{bal}(v) < \mathrm{bal}(w)$ then $v < w$.
- If $\mathrm{bal}(v) = \mathrm{bal}(w)$ then $v < w$ if and only if v comes before w in lexicographical order (the order in which the words appear in a dictionary).

Definition 7. A *Lyndon word* $w \in A^+$ is a word

$$w = \ell_1\ell_2\cdots\ell_r, \quad \ell_i \in A,$$

which is smaller than all its non-trivial right factors, i.e.:

$$w < \ell_j\ell_{j+1}\cdots\ell_r \quad \text{for all } 1 < j \leq r.$$

The set of all Lyndon words is denoted $L(A)$.

A balanced word $w \in A^+$ is called *non-negative* if all the non-trivial right factors have non-negative balance, and it is called *positive* if they all have a positive balance. Note that due to our ordering of A^+, any Lyndon word is non-negative, and that every positive balanced word is a Lyndon word.

Lemma 8. *Every non-negative balanced word w has a unique factorization in positive balanced words. We call this the positive factorization*:

$$\mathrm{pfac}(w) = w_1w_2\cdots w_k,$$

where w_i are positive.

Proof. Let $w = \ell_1\ell_2\cdots\ell_r$ where $\ell_i \in A$. We obtain the positive factorization by splitting w at all the points j where $\mathrm{bal}(\ell_j\ell_{j+1}\cdots\ell_r) = 0$. □

Lemma 9. *There is a* 1–1 *correspondence between positive balanced binary words and trees $\tau \in$* OT.

Proof. First we define the map $\mathrm{bin} : \mathrm{OT} \to A^+$ recursively as

$$\mathrm{bin}(\bullet) = 01,$$
$$\mathrm{bin}((\tau_1, \ldots, \tau_k)) = 0\,\mathrm{bin}(\tau_1)\cdots\mathrm{bin}(\tau_k)1. \tag{6}$$

By recursion we see that $\mathrm{bal}(\mathrm{bin}(\tau)) = 0$ and that all non-trivial right factors of $\mathrm{bin}(\tau)$ are non-balanced, hence $w = \mathrm{bin}(\tau)$ is always a positive balanced binary Lyndon word.

Now let w be a positive balanced word. We must have $w = 0\tilde{w}1$ where $\tilde{w}$ is non-negative. Let $w_1w_2\cdots w_k = \mathrm{pfac}(\tilde{w})$. We find that every positive w can be written $w = 0w_1w_2\cdots w_k1$ where $w_1, \ldots, w_k$ are positive and $k \geq 0$. Define the map tree( ) taking a positive w to a tree as follows:

$$\mathrm{tree}(w) = \begin{cases} \bullet & \text{if } w = 01, \\ (\mathrm{tree}(w_1), \ldots, \mathrm{tree}(w_k)) & \text{if } w = 0w_1\cdots w_k1, \text{ for positive } w_i. \end{cases} \tag{7}$$

It follows from recursion that $\mathrm{tree}(\mathrm{bin}(\tau)) = \tau$ for all $\tau \in \mathrm{OT}$. □

Consider OT as an alphabet. From the ordering of binary words defined above, we induce an ordering of OT. We extend this to the lexicographical ordering of all words over OT.

Lemma 10. *Let*

$$\tau = \tau_1\tau_2\cdots\tau_k, \quad \tau_i \in \mathrm{OT}$$

be a word over OT. *Then τ is a Lyndon word with respect to lexicographical ordering of* OT *if and only if the word*

$$w = \mathrm{bin}(\tau_1)\mathrm{bin}(\tau_2)\cdots\mathrm{bin}(\tau_k)$$

is a balanced binary Lyndon word with respect to the ordering of Definition 6.

Proof. Consider the word $w = \mathrm{bin}(\tau_1)\mathrm{bin}(\tau_2)\cdots\mathrm{bin}(\tau_k)$ where $\tau_i \in \mathrm{OT}$. Since every $\mathrm{bin}(\tau_i)$ is a positive balanced word, w is a non-negative balanced word, and by Lemma 8 it is represented in its unique positive factorization. Thus w is a Lyndon word with respect to the ordering of Definition 6 if and only if it is lexicographically less than all its non-trivial balanced positive right factors, or equivalently $\tau = \tau_1\tau_2\cdots\tau_k$ is a Lyndon word with respect to the lexicographical ordering of OT. □

We have established a natural 1–1 correspondence between balanced binary Lyndon words and Lyndon words over the (infinite) alphabet OT. The construction of a basis for the free Lie algebra fla(OT) now follows from the standard theory of Hall bases, to be reviewed in the next section.

5. Hall bases for free Lie algebras

Let $A = \mathrm{OT}$ be the alphabet of ordered trees and let A^+ be all non-empty words with lexicographical ordering. It is well known that the set of all Lyndon words in A^+ induces a basis for the free Lie algebra generated by A. In order to explain this result we need some definitions. For a thorough exposition of the general theory (see [20]).

Recall that the *free magma*, $M(A)$, consists of A and all binary bracketed expressions in A. Formally we say that $M(A)$ consists of all well-formed expressions, where

(1) Any letter $\tau \in A$ is a well-formed expression.
(2) If τ_1 and τ_2 are well-formed expressions, then $[\tau_1, \tau_2]$ is well formed.

The *foilage* is the mapping $f : M(A) \to A^+$ obtained by removing all the brackets:

$$f(a) = a \quad \text{for all } a \in A,$$
$$f([\tau_1, \tau_2]) = f(\tau_1)f(\tau_2) \quad \text{for all } \tau_1, \tau_2 \in M(A).$$

The lexicographical ordering of A^+ induces an ordering of $M(A)$. For $m_1, m_2 \in M(A)$ we say that $m_1 < m_2$ if $f(m_1) < f(m_2)$.

The free Lie algebra over the alphabet A and the field $\mathbb{R}$ is a vector space $\mathfrak{g}$ obtained by taking all (finite) $\mathbb{R}$-linear combinations of elements in $M(A)$ and extending the bracket $[\cdot,\cdot]$ to $\mathfrak{g}$ by the following rules:

(1) Skew symmetry: $[V, W] = -[W, V]$ for all $V, W \in \mathfrak{g}$.
(2) Bilinearity: $[V, W + Z] = [V, W] + [V, Z]$ and $[V, rW] = r[V, W]$ for all $V, W, Z \in \mathfrak{g}$, $r \in \mathbb{R}$.
(3) Jacobi identity: $[V, [W, Z]] + [W, [Z, V]] + [Z, [V, W]] = 0$ for all $V, W, Z \in \mathfrak{g}$.

Obviously $M(A)$ is spanning $\mathfrak{g}$. To construct a basis we must systematically remove terms which are linearly dependent. This is achieved by a *Hall basis*, defined as a subset $H \subset M(A)$ according to

(1) $A \subset H$.
(2) For any $m = [m_1, m_2] \in H\backslash A$ one has $m_2 \in H$ and

$$m < m_2.$$

(3) For any $m = [m_1, m_2] \in M(A)\backslash A$, one has $m \in H$ if and only if

$$m_1, m_2 \in H \quad \text{and} \quad m_1 < m_2$$

and

$$\text{either } m_1 \in A \quad \text{or} \quad m_1 = [m_3, m_4] \quad \text{and} \quad m_4 \geq m_2.$$

Elements of H are called *Hall elements*. In this paper we have based the construction on the lexicographical ordering of $M(A)$. Different orderings give rise to different Hall bases. Thus, e.g. the Hall basis defined in [18] is obtained from this definition, using a different ordering of $M(A)$. To obtain a basis of $\mathfrak{g}$ from the Lyndon words in A^+, we must turn Lyndon words into Hall elements, by inserting brackets in the right places. We seek a bracketing map $b : L(A) \to M(A)$ such that $f(b(\ell)) = \ell$ and $b(L(A)) = H \subset M(A)$ is a Hall basis.

A sequence of Hall elements $\{h_1, h_2, \ldots, h_n\} \subset H \subset M(A)$ is called *a standard Hall sequence* if for any i, either h_i is a letter, or $h_i = [h'_i, h''_i]$ where $h''_i \geq h_{i_1}, \ldots, h_n$. The sequence $\{h_1, h_2, \ldots, h_n\}$ is *decreasing* if $h_i \geq h_{i+1}$ for $i = 1, \ldots, n-1$, otherwise let h_j be the rightmost element such that

$h_j < h_{j+1}$. We define a mapping between standard sequences:

$$\lambda(\{h_1, h_2, \ldots, h_n\}) = \begin{cases} \{h_1, h_2, \ldots, h_n\} & \text{if the sequence is decreasing,} \\ \{h_1, h_2, \ldots, [h_j, h_{j+1}], \ldots, h_n\} & \text{otherwise.} \end{cases} \tag{8}$$

By repeatedly applying the map λ, the result becomes a decreasing standard sequence of Hall words. Let λ^* denote iterated application of λ until a decreasing sequence is obtained. Note that a sequence of letters is a standard Hall sequence. We define the *standard bracketing* of a word $w = \ell_1\ell_2\cdots\ell_n$, where $\ell_i \in A$ are letters, as the decreasing standard sequence:

$$b(w) = \lambda^*(\{\ell_1, \ldots, \ell_n\}). \tag{9}$$

The sequence $b(w) = \{h_1, \ldots, h_r\}$ can be uniquely characterized as the decreasing sequence of Hall elements $\{h_1, \ldots, h_r\}$ such that

$$w = f(h_1)f(h_2)\cdots f(h_r).$$

It can be shown that if $w \in L(A)$ is a Lyndon word, then $b(w)$ is containing a single Hall element. Furthermore, the image under b of $L(A)$ is a Hall basis for the free Lie algebra over A (see [20]). Thus we arrive at the main theorem of this paper.

Theorem 11. *There is a natural* 1–1 *correspondence between the set of balanced binary Lyndon words and a Hall basis H for the free Lie algebra* fla(OT). *The correspondence is given by the following algorithm, taking a balanced binary Lyndon word $w \in \{0,1\}^+$ to a basis element $h \in H$*:

(1) *Compute the positive factorization of* w:

$$w_1w_2\cdots w_k = \mathrm{pfac}(w).$$

(2) Compute the standard bracketing of $\{w_1, \ldots, w_k\}$:

$$h = b(w_1, w_2, \ldots, w_k).$$

This theorem simplifies the construction of software for Lie–Butcher series, since the set of all balanced binary Lyndon words is very easy to generate and represent. In Fig. 1 we see the first terms in fla(OT), sorted according to the grade.

6. Generating balanced Lyndon words

There are numerous approaches for generating Lyndon words, say of length $\leq n$, over an alphabet A. One option is to use the algorithm of Duval [7] (also described in [20]). Given any Lyndon word of length $\leq n$, this algorithm always finds the next word w.r.t. lexicographical ordering in the desired subset of Lyndon words. This algorithm can be modified to the set of balanced Lyndon words, but it seems more efficient to make use of the following elimination property. Let z denote the last letter in the alphabet A. Then

$$L(A) = L(A') \cup \{z\}, \tag{10}$$

where $A' = (A\backslash z)z^*$, and $z^* = z^+ \cup \{\emptyset\}$. This property is used in [1] where an algorithm for computing all Lyndon words of a given multidegree is presented.

We follow the lines of Section 4. First we generate the positive balanced words of degree $\leq n$ (where the degree of a word w is $|w|_1$), and then generate all Lyndon words of degree $\leq n$ over the alphabet consisting of positive balanced words of degree $\leq n$. The positive balanced words are easily obtained by the following recursion step that starts with the word $w = 1$:

(1) if w is balanced, then stop and add w to the list of balanced words.
(2) if w has degree n, then repeat the step for $0w$.
(3) otherwise, repeat the step for $0w$ and $1w$.

Let $L_n(A)$ denote the Lyndon words of degree $\leq n$ over the alphabet A. The property (10) then applies:

$$L_n(A) = L_n(A') \cup \{z\}, \tag{11}$$

where in A', we only need to consider the words of degree $\leq n$. Using this property repeatedly, we can find all Lyndon words, but we can only eliminate one generator per step. To reduce the number of steps, we make the following simple observation. Let $w \in A$.

If the sum of the degree of w and the smallest degree over all words in A is greater than n, then

$$L_n(A) = L_n(A \setminus w) \cup \{w\}. \tag{12}$$

Example 12. Let a and b denote the positive balanced words 0011 and 01 respectively, and suppose we wish to find all Lyndon words of degree ≤ 5 over $A = \{a, b\}$. Using the above two properties, we get

$$\begin{aligned} L_5(A) &= L_5(\{a, ab, ab^2, ab^3\}) \cup \{b\} \\ &= L_5(\{a, ab\}) \cup \{ab^2, ab^3, b\} \\ &= L_5(\{a, a^2b\}) \cup \{ab, ab^2, ab^3, b\} \\ &= L_5(\{a\}) \cup \{a^2b, ab, ab^2, ab^3, b\} \\ &= \{a, a^2b, ab, ab^2, ab^3, b\}. \end{aligned}$$

The approach in this example can easily be applied to construct fast algorithms for generating all binary balanced Lyndon words of a given maximal degree.

We note also that this procedure implicitly gives the standard bracketing of the words involved. For an alphabet A with largest letter z, we have (analogous to (10)):

$$\begin{aligned} H(\mathrm{fla}(A)) = H(&\mathrm{fla}(\{(-\mathrm{ad}_z)^k(a) | k \geq 0, \quad a \in A \setminus z\})) \\ &\cup \{z\}. \end{aligned}$$

Thus in Example 12, the bracketings of the words in $L_5(A) \setminus b$ are generated by the elements a, $[a, b]$, $[[a, b], b]$ and $[[[a, b], b], b]$, etc. We have used here the equivalence of Hall sets and Lazard sets, and refer to [20] for further reading.

Acknowledgements

This work was sponsored by the Norwegian Research Council under contract no. 127582/432 and 142955/431, through the SYNODE project.

References

[1] P. Andary, Finely homogeneous computations in free Lie algebras, Discrete Math. Theor. Comput. Sci. 1 (1997) 101–114.
[2] D.J. Broadhurst, On the enumeration of irreducible k-fold Euler sums and their roles in knot theory and field theory, J. Math. Phys., in press.
[3] C. Brouder, Runge–Kutta methods and renormalization, Eur. Phys. J. C 12 (2000) 521–534.
[4] J. Butcher, Coefficients for the study of Runge–Kutta integration processes, J. Aust. Math. Soc. 3 (1963) 185–201.
[5] J. Butcher, The Numerical Analysis of Ordinary Differential Equations, Runge–Kutta and General Linear Methods, Wiley, Chichester, 1987.
[6] P.E. Crouch, R. Grossman, Numerical integration of ordinary differential equations on manifolds, J. Nonlin. Sci. 3 (1993) 1–33.
[7] J.-P. Duval, Génération d'une section des classes de conjugaison et arbre des mots de Lyndon de longueur bornée, Theor. Comput. Sci. 60 (1988) 255–283.
[8] E. Hairer, G. Wanner, On the Butcher group and general multi-value methods, Computing 13 (1974) 1–15.
[9] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations. I. Nonstiff Problems, 2nd ed., Springer Series in Computational Mathematics, Springer, Berlin, 1993, p. 8.
[10] A. Iserles, S.P. Nørsett, On the solution of linear differential equations on Lie groups, Phil. Trans. Roy. Soc. Lond. A 357 (1999) 983–1019.
[11] The On-line Encyclopedia of Integer Sequences. http://www.research.att.com/~njas/sequences/.
[12] D. Lewis, J.C. Simo, Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups, J. Nonlin. Sci. 4 (1994) 253–299.
[13] W. Magnus, On the exponential solution of the of differential equations for a linear operator, Comm. Pure Appl. Math. VII (1954) 649–673.
[14] H. Munthe-Kaas, Lie–Butcher theory for Runge–Kutta methods, BIT 35 (1995) 572–587.
[15] H. Munthe-Kaas, A. Zanna, Numerical integration of differential equations on homogeneous manifolds, in: F. Cucker, M. Shub (Eds.), Foundations of Computational Mathematics, Springer, Berlin, 1997, pp. 305–315.
[16] H. Munthe-Kaas, Runge–Kutta methods on Lie groups, BIT 38 (1) (1998) 92–111.
[17] H. Munthe-Kaas, High order Runge–Kutta methods on manifolds, J. Appl. Numer. Math. 29 (1999) 115–127.
[18] H. Munthe-Kaas, B. Owren, Computations in a free Lie algebra, Phil. Trans. Roy. Soc. Lond. A 357 (1999) 957–981.
[19] B. Owren, A. Marthinsen, Runge–Kutta methods adapted to manifolds and based on rigid frames, BIT 39 (1) (1999) 116–142.
[20] C. Reutenauer, Free Lie Algebras, London Mathematical Society Monographs, New Series, Academic Press, New York, 1993, p. 7.
[21] A. Zanna, Collocation and relaxed collocation for the Fer and the Magnus expansions, SIAM J. Numer. Anal. 36 (1999) 1145–1182.

H. Munthe-Kaas received his PhD in Numerical Analysis from the Norwegian University of Science and Technology (NTNU) in 1989. In 1990 he worked as a Post. Doc. Research Scientist at CERFACS, Toulouse. Since 1991 he has been affiliated with the Department of Computer Science, University of Bergen (as Professor since 1997). He has also held an Adjoint Professorship in Mathematics, NTNU (1996–2000). In 1995 he was awarded the Carl-Erik Fröberg prize in numerical analysis for his work on Runge–Kutta methods on manifolds. In the academic year 2002–2003 Munthe-Kaas is leading a research program on geometric integration at the Centre for Advanced Study in Oslo together with Professor Brynjulf Owren, NTNU.

S. Krogstad received his Masters in Algebra from the Norwegian University of Science and Technology (NTNU) in 1997. He is currently a PhD student in Numerical Analysis at the Department of Computer Science, University of Bergen. His research interests include numerical methods for ordinary differential equations and geometric integration.