

# Universal Averages in Gauge Actions

Ruth Lawrence and Maor Siboni

Communicated by M. Moskowitz

**Abstract.** We give a construction of a universal average of Lie algebra elements whose exponentiation gives (when there is an associated Lie group) a totally symmetric geometric mean of Lie group elements (sufficiently close to the identity) with the property that in an action of the group on a space  $X$  for which  $n$  elements all take a particular point  $a \in X$  to a common point  $b \in X$ , also the mean will take  $a$  to  $b$ . The construction holds without the necessity for the existence of a Lie group and the universal average  $\mu_n(x_1, \dots, x_n)$  is a totally symmetric universal expression in the free Lie algebra generated by  $x_1, \dots, x_n$ . Its expansion up to three brackets is found explicitly and various properties of iterated averages are given. Although this is a purely algebraic result, it is expected to have applications in diverse fields. One known application is to the construction of explicit differential graded Lie algebra models of three dimensional cells and thereby to discretised differential geometry on cubulated manifolds. This work is based on the second author's minor thesis.

*Mathematics Subject Classification:* 17B01, 17B55, 55P62.

*Key Words:* DGLA, Maurer-Cartan, Baker-Campbell-Hausdorff formula, Karcher mean.

## 1. Introduction

Suppose that  $G$  is a Lie group with corresponding Lie algebra  $\mathfrak{g}$ . The exponential map provides a smooth diffeomorphism between a suitable convex centrally symmetric neighbourhood (say  $U$ ) of  $0 \in \mathfrak{g}$  and its image, a neighbourhood (which we denote  $V$ ) of  $1 \in G$ . The map  $x \mapsto -x$  on  $U$  induces the map  $g \mapsto g^{-1}$  on  $V$ . The map  $g \mapsto \frac{1}{2}g$  on  $U$  induces a map on  $V$  which we will call square-root and for which

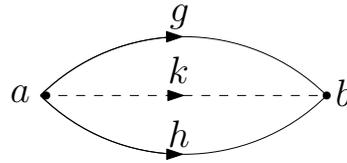
$$\sqrt{g^{-1}} = (\sqrt{g})^{-1}, \quad (\sqrt{g})^2 = g$$

Next observe that for  $g, h \in G$  sufficiently close to each other so that  $g^{-1}h \in V$  (and thus also  $h^{-1}g \in V$ ),

$$g \cdot \sqrt{g^{-1}h} = h \cdot \sqrt{h^{-1}g}$$

which we denote by  $k$ . Being symmetric in  $g, h$ , this expression can be used to define the geometric mean of  $g$  and  $h$ , which is neither  $\sqrt{gh}$  nor  $\sqrt{hg}$  (even when these are defined, being in general distinct from each other). Indeed there is a unique one-parameter subgroup ( $v \in V$ ) of form  $\gamma(t) = \exp(tv)$  for which  $g\gamma(1) = h$  so that  $g\gamma(t)$  traces a path from  $g$  to  $h$  whose midpoint is  $g\gamma(\frac{1}{2}) = k$ . Apart from its

symmetry, a characterising feature of this expression is that in any action of  $G$  on a space  $X$  for which  $g(a) = h(a) = b$ , some  $a, b \in X$ , also  $k(a) = b$ .



The aim of this paper is to extend this idea of the geometric mean to an arbitrary number of elements, and in particular, in terms of their logarithms, consider it as a universal average  $\mu_n$  at the Lie algebra level which we show exists as a totally symmetric element of the free Lie algebra on  $n$  generators, along with giving explicit formulae for the first few orders and general properties.

In Section 2, we translate the problem into a purely algebraic one at the Lie algebra level and, in terms of the Baker-Campbell-Hausdorff formula, give a closed formula for  $\mu_2(x, y)$  in the free Lie algebra on two generators. In Section 3, we give an algorithm for  $\mu_n$ , as the limit of an iterative procedure in terms of  $\mu_{n-1}$ , including a proof of its existence and an alternative algebraic characterisation. In Section 4, we use this characterisation to compute the expansion of  $\mu_n$  up to third order in brackets while in Section 5, a number of general properties of the  $\mu_n$  are discussed. In Section 6 we give an example on  $SL(2, \mathbb{R})$ . We conclude in Section 7 by indicating an application of universal averages to the explicit construction of symmetric DGLA models of simple cells.

## 2. Algebraic formulation of the problem and $\mu_2$

We make a non-standard definition of Lie algebra action. Denote by  $\text{BCH}(x, y)$  the *Baker-Campbell-Hausdorff formula* for the element of the free Lie algebra (over  $\mathbb{Q}$ ) in two variables  $x, y$  such that as formal series  $\exp(x) \cdot \exp(y) = \exp \text{BCH}(x, y)$  (see [4] for a short proof of existence and [3] for a computational formula). Let  $\mathfrak{g}$  be a Lie algebra over ground field  $\mathbb{Q}$ .

**Definition 2.1.** By a  $\mathfrak{g}$ -action we will mean a set  $M$  along with bijections  $u_x : M \rightarrow M$  for any  $x \in \mathfrak{g}$  such that

- (i)  $u_y(u_x(a)) = u_{\text{BCH}(x,y)}(a)$  for all  $x, y \in \mathfrak{g}$  and  $a \in M$ ;
- (ii)  $u_0 = \text{id}$ ;
- (iii) if  $u_x(a) = a$  for some  $x \in \mathfrak{g}$ ,  $a \in M$  then  $u_{\frac{1}{2}x}(a) = a$ .

Other consequences follow from the properties of BCH. For example,  $u_{-x} = (u_x)^{-1}$ . The main theorem of the paper is the following.

**Theorem 2.2.** *There exists a totally symmetric expression  $\mu_n(x_1, \dots, x_n)$ , in the free Lie algebra over  $\mathbb{Q}$  on the  $n$  generators,  $x_1, \dots, x_n$  with the following property. For any Lie algebra  $\mathfrak{g}$  and  $\mathfrak{g}$ -action on  $M$ , if  $x_1, \dots, x_n \in \mathfrak{g}$ ,  $a, b \in M$  satisfy  $u_{x_i}(a) = b$  for all  $i = 1, \dots, n$  then  $u_{\mu_n(x_1, \dots, x_n)}(a) = b$ .*

This will be proved constructively at the end of this section for  $n = 2$  and in the next section for general  $n$ .

**Example 2.3.** Suppose that  $G$  is a Lie group acting on a manifold  $M$  and  $\mathfrak{g}$  is its Lie algebra. Putting  $u_x(a) = (\exp(-x))(a)$  will define a  $\mathfrak{g}$ -action so long as condition (iii) is satisfied, for example if the action of  $G$  is free, although in this case the theorem would be vacuous.

**Example 2.4.** Suppose that  $E$  is a trivialised vector bundle over a base space  $B$  with fibre  $V$ . Let  $\Omega$  be the space of connections on  $E$ , that is  $\text{End } V$ -valued 1-forms on  $B$ ,

$$\omega = \sum_i \omega_i dx_i$$

in local coordinates  $(x_i)$  on  $B$ , with smooth functions  $\omega_i(x) \in \text{End } V$ . Then  $\omega \in \Omega$  defines a notion of flat sections of  $E$  with respect to  $\omega$ , by functions  $v : B \rightarrow V$  for which  $(d + \omega)v = 0$ , that is  $\frac{\partial v}{\partial x_i} + \omega_i v = 0$ .

Let  $G$  be the gauge group, consisting of smooth maps  $g : B \rightarrow GL(V)$ . It acts on sections by taking  $v(x)$  to  $g(x) \cdot v(x)$ . Correspondingly,  $G$  acts on  $\Omega$  by

$$g \cdot \omega = -dg \cdot g^{-1} + g\omega g^{-1}$$

or in coordinates, by  $(g \cdot \omega)_i = -\frac{\partial g}{\partial x_i} g^{-1} + g\omega_i g^{-1}$ . Define  $M \subset \Omega$  to be the space of flat connections, solutions of the Maurer-Cartan equation

$$M = \left\{ \omega \in \Omega \mid d\omega + \frac{1}{2}[\omega, \omega] = 0 \right\}$$

that is,  $\omega_i(x)$  such that  $\frac{\partial \omega_j}{\partial x_i} - \frac{\partial \omega_i}{\partial x_j} + [\omega_i, \omega_j] = 0$ . The action of  $G$  preserves  $M$ . The infinitesimal gauge group action is that of  $\mathfrak{g} = \text{Maps}(B, \text{End } V)$  while  $f \in \mathfrak{g}$  defines a vector field on  $\Omega$  (and on  $M$ ) by

$$f \cdot \omega = -df + [f, \omega]$$

Define  $u_f : M \rightarrow M$  to be the flow of the infinitesimal action of  $-f$  in unit time, that is, it is the map  $\omega(0) \mapsto \omega(1)$  for solutions of the equation  $\dot{\omega} = df - [f, \omega]$ . This can be written

$$a \mapsto e^{-\text{ad}_f} a + \frac{1 - e^{-\text{ad}_f}}{\text{ad}_f} df = (\exp(-f))a(\exp f) - d(\exp(-f)) \cdot \exp f$$

and is the same as the action of  $\exp(-f)$ . Note that the term  $\frac{1 - e^{-\text{ad}_f}}{\text{ad}_f}$  is to be interpreted as a power series in  $\text{ad}_f$  and is thus invertible with inverse  $\sum_{n=0}^{\infty} \frac{B_n}{n!} (-\text{ad}_f)^n$  where  $B_n$  are Bernoulli numbers, for sufficiently small  $f$  (the series has radius of convergence  $2\pi$ ). Modulo convergence issues (which arise both from the previous sentence and the BCH series), the collection  $\{u_f\}$  defines a  $\mathfrak{g}$ -action on  $M$ , because  $u_f(a) = a$  precisely when  $df = \text{ad}_f a$ , which is a linear condition on  $f$ . Despite this limitation, this is not a problem for the constructions of this paper, because starting in a sufficiently small neighbourhood in  $\mathfrak{g}$ , the iterative constructions will remain there.

**Remark 2.5.** The previous example can be generalised to be constructed from any pair  $(E, \nabla)$  of a smooth vector bundle  $E$  with flat connection  $\nabla$ . The space of *all* flat connections is now identified (relative to  $\nabla$ ) as solutions of a Maurer-Cartan equation on which there is an infinitesimal gauge group action.

**Example 2.6.** For a regular cell complex  $X$ , it is possible to associate a DGLA model  $A = A(X)$  over  $\mathbb{Q}$  satisfying the following conditions

- (i) as a Lie algebra,  $A(X)$  is freely generated by a set of generators, one for each cell in  $X$  and whose grading is one less than the geometric degree of the cell;
- (ii) vertices (that is 0-cells) in  $X$  give rise to generators  $a$  which satisfy the Maurer-Cartan equation  $\partial a + \frac{1}{2}[a, a] = 0$  (a flatness condition);
- (iii) for a cell  $x$  in  $X$ , the part of  $\partial x$  without Lie brackets is the geometric boundary  $\partial_0 x$  (where an orientation must be fixed on each cell);
- (iv) (locality) for a cell  $x$  in  $X$ ,  $\partial x$  lies in the Lie algebra generated by the generators of  $A(X)$  associated with cells of the closure  $\bar{x}$ .

The existence and general construction of such a model was demonstrated by Sullivan in the appendix to [15]. By [2], there exist consistent (even symmetric) towers of models of simplices, and such towers are unique up to (exact) DGLA isomorphism. The model of an interval is unique [10]. In [5], an explicit symmetric model of the bi-gon (exhibiting the dihedral symmetry of the bi-gon) was given, the main intermediate step being the construction of a ‘symmetric point’ in the model of the boundary of the bi-gon, invariant under the full symmetries of the bi-gon. Similarly in [6], an explicit construction of a model of a single triangle which is invariant under the action of the symmetry group  $S_3$  of the triangle is given, and again the main intermediate step is the construction of a totally symmetric central ‘point’ (solution of Maurer-Cartan).

The geometry in this example is now seen on the (infinite-dimensional) space of solutions  $M \subset A_{-1}$  of the Maurer-Cartan equation  $dx + \frac{1}{2}[x, x] = 0$  on which  $f \in A_0$  acts infinitesimally by  $\dot{x} = df - [f, x]$ . Thus in particular the geometry of the original cell complex  $X$  is reproduced inside  $M$ , with vertices in  $X$  defining elements of  $M$  and edges in  $X$  defining elements of  $A_0$  whose unit time infinitesimal flow takes the starting endpoint to the concluding one, while intermediate (rational) times of the flow also describe points (elements) of  $M$ .

The fact that  $A(X)$  is a free Lie algebra enables a grading by the number of Lie brackets and thus avoids all issues of convergence and injectivity radius which dogged previous examples given above.

While the inspiration for the construction of such models came from rational homotopy theory ([12], [14]), their application may be to diverse fields where such infinity structures enter, from deformation theory to discretisation of differential equations, to be discussed in future work [9].

The motivation for the constructions of this paper is precisely this example, in the case of the bi-gon for  $\mu_2$  and an appropriate three-dimensional cell for  $\mu_n$ ,  $n > 2$ . For a banana-shaped cell, with two points  $a$  and  $b$  between which there are  $n$  edges and a single three-dimensional cell, the central point of the cell is found at the midpoint of the ‘diagonal’ from  $a$  to  $b$  given by the mean  $\mu_n(x_1, \dots, x_n)$ , see [7].

**Remark 2.7.** (Properties of BCH) We list here a few properties of BCH which we use in the arguments of the rest of this section:

(a) The first few terms of  $\text{BCH}(x, y)$  are

$$\begin{aligned} \text{BCH}(x, y) &= x + y + \frac{1}{2}[x, y] + \frac{1}{12}(X^2y + Y^2x) - \frac{1}{24}XYXy \\ &\quad - \frac{1}{720}(X^4y + Y^4x) + \frac{1}{120}(X^2Y^2x + Y^2X^2y) + \frac{1}{360}(XY^3x + YX^3y) + \cdots \end{aligned}$$

where  $X, Y$  denote  $\text{ad}_x, \text{ad}_y$ , respectively.

(b) Uniqueness implies  $\text{BCH}(\text{BCH}(x, y), z) = \text{BCH}(x, \text{BCH}(y, z))$  for any symbols  $x, y, z$ , that is, associativity of BCH. Denote the combined BCH of  $n$  symbols  $x_1, \dots, x_n \in A$  by  $\text{BCH}(x_1, \dots, x_n)$  which will be a formal sum of terms, the zeroth order being  $x_1 + \cdots + x_n$  and higher orders being rational linear combinations of (repeated) Lie brackets of the  $x_i$ 's.

(c)  $\text{BCH}(x, -x) = 0$  while  $\text{BCH}(-x_1, \dots, -x_n) = -\text{BCH}(x_n, \dots, x_1)$ .

(d)  $\text{BCH}(x, y, -x) = (\exp \text{ad}_x)y$ .

(e)  $\text{BCH}(\exp(\text{ad}_e)x, \exp(\text{ad}_e)y) = \exp(\text{ad}_e)\text{BCH}(x, y)$ .

**Proof of Theorem 2.2 for  $n = 2$ .** Set  $\mu_2(x, y) = \text{BCH}(x, \frac{1}{2}\text{BCH}(-x, y))$ . There are two requirements to check in order to verify that this is a solution.

(a) If  $u_x(a) = u_y(a) = b$  then  $u_{\text{BCH}(-x, y)}(b) = u_y(u_{-x}(b)) = u_y(a) = b$ . Thus  $u_{\frac{1}{2}\text{BCH}(-x, y)}(b) = b$  and hence combining with  $u_x(a) = b$  we get  $u_{\mu_2(x, y)}(a) = b$ .

(b) Interchanging  $x$  and  $y$  in the formula for  $\mu_2(x, y)$ , we get

$$\text{BCH}(y, \frac{1}{2}\text{BCH}(-y, x)) = \text{BCH}(y, \text{BCH}(-y, x), -\frac{1}{2}\text{BCH}(-y, x))$$

Since  $-\text{BCH}(-y, x) = \text{BCH}(-x, y)$ , this simplifies, as required, to the expression  $\text{BCH}(x, \frac{1}{2}\text{BCH}(x, -y)) = \mu_2(x, y)$ . ■

Using the first few terms in the expansion of BCH, we get

$$\mu_2(x, y) = \frac{1}{2}(x + y) - \frac{1}{48}[x, [x, y]] - \frac{1}{48}[y, [y, x]] + \cdots$$

up to the second order in Lie brackets.

**Remark 2.8.** The formula for  $\mu_2(x, y)$  first appeared in [5]. It is unique satisfying the conditions of the theorem, as follows from [5], since in the example of the Lie algebra and action coming from the bi-gon, there is only a one-parameter family of flows from  $a$  to  $b$ , namely  $\text{BCH}(x, t\text{BCH}(-x, y))$  and only the value  $t = \frac{1}{2}$  (which gives  $\mu_2$ ) is symmetric in  $x$  and  $y$ .

**Lemma 2.9.** (i)  $\mu_2(-x, -y) = -\mu_2(x, y)$

(ii)  $\mu_2(\text{BCH}(z, x), \text{BCH}(z, y)) = \text{BCH}(z, \mu_2(x, y))$

(iii)  $\mu_2(\text{BCH}(x, z), \text{BCH}(y, z)) = \text{BCH}(\mu_2(x, y), z)$

**Proof.** (i) By definition and Remark 2.7(c) above,

$$\mu_2(-x, -y) = \text{BCH}(-x, \frac{1}{2}\text{BCH}(x, -y)) = -\text{BCH}(\frac{1}{2}\text{BCH}(y, -x), x)$$

But by Remark 2.7(b),(c),(d),  $\text{BCH}(y, -x) = \text{BCH}(x, -x, y, -x) = e^X \text{BCH}(-x, y)$  while  $e^X x = x$  where  $X = \text{ad}_x$ . Thus by (e)

$$\mu_2(-x, -y) = -\text{BCH}(e^X \frac{1}{2} \text{BCH}(-x, y), e^X x) = -e^X \text{BCH}(\frac{1}{2} \text{BCH}(-x, y), x)$$

which by Remark 2.7(d) simplifies to  $-\text{BCH}(x, \frac{1}{2} \text{BCH}(-x, y)) = -\mu_2(x, y)$ .

(ii) This follows immediately from the definition, since

$$\text{BCH}(-\text{BCH}(z, x), \text{BCH}(z, y)) = \text{BCH}(-x, y)$$

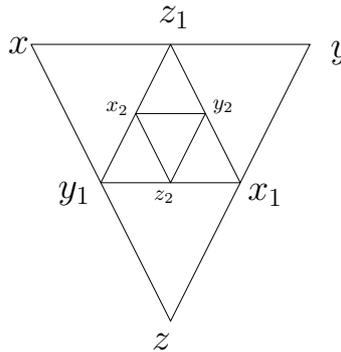
(iii) Follows by combining (i), (ii). ■

### 3. An algorithm for $\mu_n$

Knowing how to average (symmetrically) pairs of objects using  $\mu_2$  from Section 2, one can average triples of objects by iteratively averaging pairs. Thus, starting with  $x, y, z$ , define three sequences  $(x_k), (y_k), (z_k)$  in the free Lie algebra  $L[x, y, z]$  on three generators  $x, y, z$  by

$$x_{k+1} = \mu_2(y_k, z_k), \quad y_{k+1} = \mu_2(x_k, z_k), \quad z_{k+1} = \mu_2(x_k, y_k)$$

with initial conditions  $x_0 = x, y_0 = y, z_0 = z$ . Pictorially, representing  $x_k, y_k, z_k$  by points (though they would be edges in the representation of Example 2.5) and  $\mu_2(x, y)$  by the midpoint of a line drawn between  $x$  and  $y$ , we get



Below we will prove that the three sequences do indeed all converge to a common element of  $L[x, y, z]$ , which we denote  $\mu_3(x, y, z)$ . Convergence here means the convergence of the truncated expressions with  $\leq K$  Lie brackets, for all  $K \in \mathbb{N}$ . The case of  $K = 0$  is precisely the geometric picture above with convergence to the centroid.

In the same manner, one can inductively construct each universal average from the previous one, defining  $\mu_n(x_1, \dots, x_n) \in L[x_1, \dots, x_n]$ , in terms of  $\mu_{n-1}$ .

**Lemma 3.1.** *There is a unique sequence of totally symmetric elements  $\mu_n, n > 2$  in the free Lie algebra  $L[x_1, \dots, x_n]$  over  $\mathbb{R}$  such that for each  $n > 2$ , the  $n$  sequences  $(x_1^k), \dots, (x_n^k)$  defined iteratively by*

$$x_i^{k+1} = \mu_{n-1}(x_1^k, \dots, \widehat{x_i^k}, \dots, x_n^k) \tag{1}$$

with initial conditions  $x_i^0 = x_i, 1 \leq i \leq n$ , will all converge to  $\mu_n$  as  $k \rightarrow \infty$ , in the sense that their truncations with  $\leq K$  Lie brackets will converge, for all  $K \in \mathbb{N}$ . The part of  $\mu_n$  without Lie brackets is  $\frac{1}{n}(x_1 + \dots + x_n)$ .

**Proof.** The proof is by induction on  $n$ , starting with  $n = 3$ . As it stands, the relations (1) are highly non-linear, since  $\mu_{n-1}$  is non-linear. It is a notoriously difficult problem to iterate non-linear operations. However, since  $(x_1^k, \dots, x_n^k)$  are obtained from  $(x_1^0, \dots, x_n^0)$  by iterating  $k$  times the fixed relations (1), it can also be seen that they will be obtained from  $(x_1^1, \dots, x_n^1)$  by  $(k-1)$  applications of the same iteration procedure. This means that if in the formulae for  $x_1^{k-1}, \dots, x_n^{k-1}$  we replace  $x_1, \dots, x_n$  by  $x_1^1, \dots, x_n^1$ , respectively, then we will obtain  $x_1^k, \dots, x_n^k$ . Let  $T_n$  denote the operation on the free Lie algebra  $L[x_1, \dots, x_n]$ , specified by substituting  $x_1^1, \dots, x_n^1$  for  $x_1, \dots, x_n$ , respectively, that is,

$$x_i \rightarrow \mu_{n-1}(x_1, \dots, \widehat{x}_i, \dots, x_n), \quad 1 \leq i \leq n$$

This is a linear map and by the above argument,  $x_i^k = T_n(x_i^{k-1})$  and thus  $x_i^k = (T_n)^k(x_i)$ .

For  $K \geq 0$ , let  $V^{[\leq K]}$  denote the vector subspace of  $V = L[x_1, \dots, x_n]$  spanned by Lie monomials with  $\leq K$  brackets. Since  $L[x_1, \dots, x_n]$  is free and the Jacobi relation preserves the number of brackets, there is a well-defined projection map  $V \rightarrow V^{[\leq K]}$  annihilating all Lie monomials with  $> K$  brackets. Let  $T_n^{[\leq K]}$  denote the truncation of  $T_n$  to  $V^{[\leq K]}$ , so that the truncation of  $x_i^k$  to  $V^{[\leq K]}$  is  $(T_n^{[\leq K]})^k(x_i)$ . To understand what happens in the limit  $k \rightarrow \infty$ , we need to investigate the eigenvalues of  $T_n^{[\leq K]}$ .

A basis for  $V^{[\leq K]}$  can be found which is a subset of the (finite) set of all Lie monomials in  $x_1, \dots, x_n$  with  $\leq K$  brackets. With respect to this basis, the matrix of  $T_n^{[\leq K]}$  will be block lower triangular, the blocks being determined by the numbers of brackets, since under substitution in a monomial with  $r$  brackets, all terms will have at least  $r$  brackets. The blocks on the diagonal in this matrix come from that part of the substitution which retains the same number of brackets, that is

$$x_i \rightarrow \frac{1}{n-1}(x_1 + \dots + \widehat{x}_i + \dots + x_n), \quad 1 \leq i \leq n \quad (2)$$

In particular, the  $(0,0)$  block will be the action on terms with no brackets, that is on the span of  $x_1, \dots, x_n$ , and will be the matrix induced by (2), namely an  $n \times n$  matrix all of whose entries are  $\frac{1}{n-1}$  except for zeroes on the diagonal. This matrix is diagonalizable and its eigenvalues are 1 (simple) and  $-\frac{1}{n-1}$  (with multiplicity  $n-1$ ). Choose a diagonalizing basis, say  $y_1 = x_1 + \dots + x_n$  and  $y_i = x_{i-1} - x_i$ ,  $i = 2, \dots, n$ .

The  $(r,r)$  block of the matrix for  $T_n^{[\leq K]}$  will be given by the action of (2) on linear combinations of monomials with exactly  $r$  brackets. Change basis to a new basis which are monomials in  $y_1, \dots, y_n$ . The  $(r,r)$  block of the matrix for  $T_n^{[\leq K]}$  in this new basis will be given by the action of the substitution,

$$y_1 \rightarrow y_1, \quad y_i \rightarrow -\frac{1}{n-1}y_i \quad (2 \leq i \leq n)$$

Under this substitution any Lie monomial in the  $y_i$  with  $r$  brackets, will scale by a factor  $(-\frac{1}{n-1})^{r+1-s}$  where  $s$  is the number of times that  $y_1$  appears in the monomial. Such a monomial with  $r$  brackets contains  $r+1$  (not necessarily distinct)  $y_i$ 's; so the exponent here is always non-negative, while it can only vanish if  $s = r+1$ , that is a monomial containing only  $y_1$ . Since  $[y_1, y_1] = 0$ , such a monomial can only be a basis element if  $r = 0$ . In conclusion, the  $(r,r)$  block of the matrix for  $T_n^{[\leq K]}$  in this new basis is diagonal with entries which are powers of  $(-\frac{1}{n-1})$ , with strictly

positive exponents for  $r > 0$ . But a lower triangular block matrix whose diagonal blocks are all diagonal matrices is just a lower triangular matrix. The entries on the diagonal are all non-negative integer powers of  $(-\frac{1}{n-1})$  and the entry 1 (exponent zero) occurs only from  $y_1$  in the  $(0,0)$  block. We conclude that  $T_n^{[\leq K]}$  has all its eigenvalues which are non-negative integer powers of  $(-\frac{1}{n-1})$ , while the eigenvalue 1 appears without multiplicity.

While the matrix for  $T_n^{[\leq K]}$  may not be diagonalizable, it can be expressed in Jordan block form and the diagonal entries of those blocks must be the eigenvalues, that is there is a single size one Jordan block with eigenvalue 1 and the remaining Jordan blocks all have eigenvalues which are positive powers of  $(-\frac{1}{n-1})$  with varying multiplicities and block sizes. The powers of any Jordan block with eigenvalue  $\lambda$ ,  $|\lambda| < 1$ , converge to zero, and so in a basis with respect to which  $T_n^{[\leq K]}$  is in Jordan normal form (say with the Jordan block of eigenvalue 1 first),  $(T_n^{[\leq K]})^k$  converges as  $k \rightarrow \infty$  to a matrix identically zero except for the  $(1,1)$  position which is 1. Denote the first basis element by  $v_1 \in V^{[\leq K]}$ ; it is an eigenvector of  $T_n^{[\leq K]}$  of eigenvalue 1. It depends on  $n$ , but we omit this dependence for ease of notation. The conclusion is that for any  $x \in V^{[\leq K]}$ ,

$$(T_n^{[\leq K]})^k x \longrightarrow av_1 \quad \text{as } k \rightarrow \infty$$

where  $a$  is the first component of  $x$  in the new basis (the coefficient of  $v_1$ ).

Applying this for  $x = x_i$ , we obtain that the truncation of  $x_i^k$  to  $V^{[\leq K]}$ , has a limit as  $k \rightarrow \infty$ , and this limit is a multiple of the same eigenvector  $v_1$ , say  $a_i v_1$ , where  $a_i$  is the first component of  $x_i$  with respect to the new basis. In order to find  $a_i$ , it suffices to truncate to the first block  $V^{[0]}$  on which we know that  $T_n^{[0]}$  is diagonal with respect to the basis  $\{y_i\}$ , and the eigenvector with eigenvalue 1 is  $y_1$ . However,

$$x_1 = \frac{1}{n}y_1 + \frac{1}{n} \sum_{j=2}^n (n+1-j)y_j, \quad x_i = x_1 - \sum_{j=2}^i y_j$$

and so the coefficient of  $y_1$  in any  $x_i$  is  $\frac{1}{n}$ . Thus  $a_i = \frac{1}{n}$  for all  $i$  and so the truncations to  $\leq K$  Lie brackets of the sequences  $x_i^k$  have a common limit as  $k \rightarrow \infty$ , whose zero bracket part is  $\frac{1}{n}y_1$ , for all  $k \in \mathbb{N}$ . Hence also the sequences  $x_i^k$  themselves have a common limit, which we denote by  $\mu_n(x_1, \dots, x_n)$  and whose zero bracket part is  $\frac{1}{n}(x_1 + \dots + x_n)$ .

Finally, inductively we see that  $\mu_n$  is totally symmetric under interchange of the  $x_i$ 's. This holds for  $n = 2$  by Section 2. Assuming  $\mu_{n-1}$  is symmetric (for some  $n \geq 3$ ), the symmetry of the construction of the sequences  $x_i^k$  means that their limit must also be symmetric. To be more precise, any permutation of  $x_1, \dots, x_n$  will induce an identical permutation of  $x_1^k, \dots, x_n^k$ , for each  $k$ , and so knowing that they share a common limit, it must be invariant under the symmetric group. ■

In the course of the proof of Lemma 3.1, we showed that at every truncation  $K$ , the operator  $T_n^{[\leq K]}$  has exactly one eigenvalue 1 with all the rest of modulus strictly less than 1, while the iteration converges to a multiple of this eigenvector, that is to a fixed point of  $T_n^{[\leq K]}$ . It follows that  $\mu_n$  itself is a fixed point of  $T_n$ , and that this condition determines  $\mu_n$  up to scaling.

**Lemma 3.2.**  $\mu_n(x_1, \dots, x_n)$  as constructed in Lemma 3.1 is defined uniquely up to scaling, by the property

$$\begin{aligned} \mu_n \left( \mu_{n-1}(\widehat{x}_1, \dots, x_n), \dots, \mu_{n-1}(x_1, \dots, \widehat{x}_i, \dots, x_n), \dots, \mu_{n-1}(x_1, \dots, \widehat{x}_n) \right) \\ = \mu_n(x_1, \dots, x_n) \end{aligned}$$

It is this fact which we use to compute the first few coefficients in an explicit expansion of  $\mu_n$  in the next section.

**Proof of Theorem 2.2.** Two facts remain to be verified, namely that  $\mu_n$  as constructed in Lemma 3.1 has rational coefficients, and that they satisfy the central diagonal property.

*Rationality:* Inductively we show that  $\mu_n(x_1, \dots, x_n)$  lies in the free Lie algebra over  $\mathbb{Q}$  generated by  $x_1, \dots, x_n$ . For  $n = 2$  this follows from the explicit formula for  $\mu_2$  in Section 2 and rationality of BCH. Assuming rationality for  $\mu_{n-1}$ , the matrix elements of  $T_n^{[\leq K]}$  in the proof of Lemma 3.1, are rational and so the unique eigenvector (up to scaling) of eigenvalue 1 also has rational components. The proof of Lemma 3.1 identifies this eigenvector with  $\mu_n^{[\leq K]}$ , up to scaling. Since the zeroth block in  $\mu_n$  (that is the part without Lie brackets) is  $\mu_n(x_1, \dots, x_n)^{[0]} = \frac{1}{n}(x_1 + \dots + x_n)$ , which has rational coefficients, thus all the coefficients in  $\mu_n^{[\leq K]}$  are rational, completing the inductive step.

*Central diagonal property:* Using the construction of  $\mu_n$  in Lemma 3.1, we prove inductively that

$$(u_{x_i}(a) = b \text{ for } 1 \leq i \leq n) \implies u_{\mu_n(x_1, \dots, x_n)}(a) = b$$

For  $n = 2$  it is known by Section 3. Assume it is true for  $n - 1$ , some  $n \geq 3$ . Suppose that  $x_1, \dots, x_n \in \mathfrak{g}$ ,  $a, b \in X$  are such that  $u_{x_i}(a) = b$  for all  $i = 1, \dots, n$ . In the notation of Lemma 3.1, we see by induction on  $k$  (starting from  $k = 0$  which is the initial assumption) that  $u_{x_i^k}(a) = b$  for  $i = 1, \dots, n$  and all  $k \in \mathbb{N}$ . Since  $u_x(a)$  depends continuously on  $x$  and  $x_i^k \rightarrow \mu_n(x_1, \dots, x_n)$  as  $k \rightarrow \infty$ , the inductive step follows. ■

**Lemma 3.3.**  $\mu_n(-x_1, \dots, -x_n) = -\mu_n(x_1, \dots, x_n)$

**Proof.** This follows by induction from Lemma 2.9(i) and Lemma 3.1. ■

As a corollary, the expansion of  $\mu_n(x_1, \dots, x_n)$  will involve only odd numbers of symbols and therefore even numbers of brackets.

#### 4. Expansion of $\mu_n$ in Lie brackets

In this section we obtain formula for the expansion of  $\mu_n(x_1, \dots, x_n)$  in Lie brackets, up to the third order. From Lemma 3.3, only even orders are present in  $\mu_n$ . The zeroth order is the average  $\mu_n^{[0]} = \frac{1}{n}(x_1 + \dots + x_n)$  by Lemma 3.1.

Let  $V^{[r]}$  denote the piece of the free Lie algebra  $L[x_1, \dots, x_n]$  spanned by Lie monomials with exactly  $r$  brackets. In order to obtain a spanning set, it is sufficient to enumerate Lie monomials of the form  $[x_{i_1}, [x_{i_2}, \dots, [x_{i_r}, x_{i_{r+1}}] \dots]]$ .

The symmetric group  $S_n$  permutes the  $x_i$ 's and thus acts on  $V^{[r]}$ ; denote the invariant subspace under this action by  $\bar{V}^{[r]}$ . It is spanned by the images of Lie monomials with exactly  $r$  brackets under the symmetrization map  $\bar{\cdot}$  defined by  $\bar{w} \equiv \frac{1}{n!} \sum_{\sigma \in S_n} \sigma(w)$ .

**Second order:** The space  $\bar{V}^{[2]}$  is spanned by the symmetrizations of the different types of Lie monomials with two brackets  $[x_i, [x_j, x_k]]$ , where the type is determined by coincidences or otherwise in the list  $i, j, k$ . Thus there are *prima facie* two generators, coming from symmetrizations of

$$[x_i, [x_j, x_k]], [x_i, [x_i, x_j]]$$

where in this list the indices  $(i, j, k)$  are all considered distinct. However, on interchange of  $j$  and  $k$  the first expression changes sign and so its symmetrization vanishes. Thus  $\bar{V}^{[2]}$  is one-dimensional, generated by

$$v^{[2]} \equiv \sum_{i,j,i \neq j} [x_i, [x_i, x_j]] = \sum_{i,j,i < j} [x_i - x_j, [x_i, x_j]]$$

In particular,  $\mu_n^{[2]}$  is some multiple of this element, so that

$$\mu_n^{[\leq 2]} = \frac{1}{n} \sum_i x_i + c_n \sum_{i,j,i \neq j} [x_i, [x_i, x_j]]$$

for some scalar  $c_n$ . To find  $c_n$ , use the defining property of  $\mu_n$  from Lemma 3.2. The part with at most two Lie brackets in the left hand side of the equality in Lemma 3.2 is

$$\begin{aligned} & \frac{1}{n} \sum_i \left( \frac{1}{n-1} \sum_{j \neq i} x_j + c_{n-1} \sum_{j \neq i, k \neq i \text{ distinct}} [x_j, [x_j, x_k]] \right) \\ & + c_n \sum_{i,j, i \neq j} \left[ \frac{1}{n-1} \sum_{k \neq i} x_k, \left[ \frac{1}{n-1} \sum_{l \neq i} x_l, \frac{1}{n-1} \sum_{m \neq j} x_m \right] \right] \\ & = \frac{1}{n(n-1)} \sum_{i, j \neq i} x_j + \frac{c_{n-1}}{n} \sum_{i,j,k \text{ distinct}} [x_j, [x_j, x_k]] \\ & + \frac{c_n}{(n-1)^3} \sum_{i,j,k,l,m, i \neq j,k,l, j \neq m} [x_k, [x_l, x_m]] \\ & = \frac{1}{n(n-1)} \sum_j \left( \sum_{i \neq j} 1 \right) x_j + \frac{c_{n-1}}{n} \sum_{j,k \text{ distinct}} \left( \sum_{i \neq j,k} 1 \right) [x_j, [x_j, x_k]] \\ & + \frac{c_n}{(n-1)^3} \sum_{k,l,m} \left( \sum_{i,j \text{ distinct}, i \neq k,l, j \neq m} 1 \right) [x_k, [x_l, x_m]] \end{aligned}$$

However the set of  $i, j$  distinct with  $i \neq k, l$  and  $j \neq m$  has order

$$\begin{aligned} & (n-3)(n-2) + (n-1) \quad \text{for } k, l, m \text{ distinct} \\ & (n-2)^2 + (n-1) \quad \text{for } k = l \neq m \\ & (n-2)^2 \quad \text{for } k = m \neq l \\ & (n-2)^2 \quad \text{for } l = m \neq k \\ & (n-1)(n-2) \quad \text{for } k = l = m \end{aligned}$$

Moreover, the sum of  $[x_k, [x_l, x_m]]$  over  $k, l, m$  with the various conditions in the five cases above, vanishes in the first, fourth and fifth cases (because of antisymmetry of the expression and symmetry of the condition under interchanging  $l$  and  $m$ ). In the second and third cases, these sums are,

$$\sum_{k,m \text{ distinct}} [x_k, [x_k, x_m]], \quad \sum_{k,l \text{ distinct}} [x_k, [x_l, x_k]]$$

respectively, which are identical except for a sign. The conclusion is that

$$\sum_{k,l,m} \left( \sum_{i,j \text{ distinct}, i \neq k,l, j \neq m} 1 \right) [x_k, [x_l, x_m]] = (n-1) \sum_{k,m \text{ distinct}} [x_k, [x_k, x_m]]$$

and thus the  $\leq 2$  Lie bracket part of the LHS of Lemma 3.2 simplifies to

$$\frac{1}{n} \sum_j x_j + \frac{n-2}{n} c_{n-1} \sum_{j,k \text{ distinct}} [x_j, [x_j, x_k]] + \frac{c_n}{(n-1)^2} \sum_{k,m \text{ distinct}} [x_k, [x_k, x_m]]$$

Identifying this with  $\mu_n^{[\leq 2]}$  leaves

$$c_n = \frac{n-2}{n} c_{n-1} + \frac{c_n}{(n-1)^2}$$

which simplifies to  $n^2 c_n = (n-1)^2 c_{n-1}$  for  $n > 2$ . This implies that  $n^2 c_n$  is independent of  $n$  and thus shares the value at  $n = 2$  which is  $-\frac{1}{12}$ , since  $c_2 = -\frac{1}{48}$  by Section 2. Hence  $c_n = -\frac{1}{12n^2}$  and

$$\mu_n^{[\leq 2]} = \frac{1}{n} \sum_i x_i - \frac{1}{12n^2} \sum_{i,j,i \neq j} [x_i, [x_i, x_j]]$$

### 5. Properties of $\mu_n$

The core properties of  $\mu_n$  used to identify its first few coefficients in Section 5 are that (i) it is totally symmetric (Lemma 3.1), (ii) it involves only even numbers of Lie brackets (Lemma 3.3), (iii) its zeroth order part is the usual average (Lemma 3.1), and (iv) it is invariant under the substitution map  $T_n$  of Section 4 (Lemma 3.2). These uniquely determine  $\mu_n$ . In addition there are two properties which follow inductively from Lemma 2.9(ii),(iii) and the construction of  $\mu_n$  as the limit of a sequence.

**Lemma 5.1.**

- (i)  $\mu_n(\text{BCH}(z, x_1), \dots, \text{BCH}(z, x_n)) = \text{BCH}(z, \mu_n(x_1, \dots, x_n))$
- (ii)  $\mu_n(\text{BCH}(x_1, z), \dots, \text{BCH}(x_n, z)) = \text{BCH}(\mu_n(x_1, \dots, x_n), z)$

**Non-uniqueness of  $\mu_n$ :** The expressions  $\mu_n$  for  $n > 2$  are not uniquely determined by the universality condition in Theorem 2.2 (without the condition of Lemma 3.2). Indeed,

$$x = \text{BCH}(y, \mu_n(x_1, \dots, x_n))$$

will also satisfy the conditions, for any  $y$  totally symmetric in the  $x_i$ 's for which  $u_y(a) = a$  necessarily follows from  $u_{x_i}(a) = b$  for all  $i$ . If it is additionally known

that the set  $V_a$  of  $y$  satisfying  $u_y(a) = a$  is a vector space, then the fact that it is closed under BCH while  $\text{BCH}(x_i, -x_1) \in V_a$  ensures that  $V_a$  is closed under Lie bracket and the non-uniqueness follows from the following lemma.

**Lemma 5.2.** *The  $S_n$ -invariant part of the sub-Lie algebra of  $L[x_1, \dots, x_n]$  generated by  $\text{BCH}(x_i, -x_1)$ ,  $i = 2, \dots, n$ , is non-trivial for  $n > 2$ .*

A generalization of Lemma 3.2 is the following.

**Conjecture 5.3.** For  $1 \leq m \leq n$ ,

$$\mu_{\binom{n}{m}}(\mu_m(x_S) \mid S \subset \{1, \dots, n\}, |S| = m) = \mu_n(x_1, \dots, x_n) \tag{3}$$

This is trivial for  $m = 1$  (setting  $\mu_1$  to be the identity) and is Lemma 3.2 for  $m = n - 1$ .

**Comparison up to second order:** Using the results of Section 4, up to second order the left hand side of (3) becomes

$$\begin{aligned} & \frac{1}{\binom{n}{m}} \sum_{|S|=m} \left( \frac{1}{m} \sum_{j \in S} x_j - \frac{1}{12m^2} \sum_{i,j \in S, i \neq j} [x_i, [x_i, x_j]] \right) \\ & - \frac{1}{12 \binom{n}{m}^2} \sum_{|S|=|T|=m, S \neq T} \frac{1}{m^3} \left[ \sum_{i \in S} x_i, \left[ \sum_{j \in S} x_j, \sum_{k \in T} x_k \right] \right] \\ & = \frac{1}{m \binom{n}{m}} \sum_j \left( \sum_{S, |S|=m, S \ni j} 1 \right) x_j - \frac{1}{12m^2 \binom{n}{m}} \sum_{i,j, i \neq j} \left( \sum_{S, |S|=m, S \ni i,j} 1 \right) [x_i, [x_i, x_j]] \\ & - \frac{1}{12m^3 \binom{n}{m}^2} \sum_{i,j,k} \left( \sum_{S,T, |S|=|T|=m, S \neq T, S \ni i,j, T \ni k} 1 \right) [x_i, [x_j, x_k]] \end{aligned}$$

The last term breaks up into parts (as in the calculation of Section 5) according to the coincidences amongst the indices  $i, j, k$ . Since  $[x_i, [x_j, x_k]]$  is antisymmetric in  $j, k$ , the only non-trivial sums appearing will be when  $i = j \neq k$  or  $i = k \neq j$ , and here these sums will be identical except for sign. In these two cases, the possible distinct sets  $S, T$  of order  $m$  for which  $S \ni i, j$  and  $T \ni k$  are divided into cases

$$\begin{aligned} & (S \ni i, S \not\ni k, T \ni k) \cup (S \ni i, k, T \ni i, k, S \neq T) \cup (S \ni i, k, T \ni k, T \not\ni i) \\ & (S \ni i, j, T \ni i, j, S \neq T) \cup (S \ni i, j, T \ni i, T \not\ni j) \end{aligned}$$

for  $i = j \neq k$  and  $i = k \neq j$ , respectively. The difference in these enumerations is thus  $\binom{n-2}{m-1} \binom{n-1}{m-1}$  and so

$$\sum_{i,j,k} \sum_{\substack{S,T, |S|=|T|=m \\ S \neq T, S \ni i,j, T \ni k}} [x_i, [x_j, x_k]] = \binom{n-2}{m-1} \binom{n-1}{m-1} \sum_{i,j, i \neq j} [x_i, [x_i, x_j]]$$

so that the LHS of (3) up to second order simplifies to

$$\begin{aligned} & \frac{\binom{n-1}{m-1}}{m\binom{n}{m}} \sum_j x_j - \frac{\binom{n-2}{m-2}}{12m^2\binom{n}{m}} \sum_{i,j, i \neq j} [x_i, [x_i, x_j]] - \frac{\binom{n-2}{m-1}\binom{n-1}{m-1}}{12m^3\binom{n}{m}^2} \sum_{i,j, i \neq j} [x_i, [x_i, x_j]] \\ &= \frac{1}{n} \sum_j x_j - \frac{m-1}{12mn(n-1)} \sum_{i,j, i \neq j} [x_i, [x_i, x_j]] - \frac{n-m}{12mn^2(n-1)} \sum_{i,j, i \neq j} [x_i, [x_i, x_j]] \\ &= \frac{1}{n} \sum_j x_j - \frac{1}{12n^2} \sum_{i,j, i \neq j} [x_i, [x_i, x_j]] \end{aligned}$$

which coincides with  $\mu_n^{[\leq 2]}$ . Thus the conjecture holds up to second (or third) order in Lie brackets.

## 6. An example: $SL(2, \mathbb{R})$

Consider the action of  $G = SL(2, \mathbb{R})$  on  $X = \mathbb{H}$ , the upper-half plane, defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}$$

The corresponding infinitesimal action is of  $\mathfrak{g}$  by the vector field

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \cdot z \equiv \frac{d}{dt} \left( \begin{pmatrix} 1 + ta & tb \\ tc & 1 - ta \end{pmatrix} \cdot z \right) \Big|_{t=0} = -cz^2 + 2az + b,$$

the corresponding flow on  $\mathbb{H}$  being given by  $\dot{z} = -cz^2 + 2az + b$ . This has one fixed point in  $\mathbb{H}$  whenever  $a^2 + bc < 0$ . Conversely, for each point  $z \in \mathbb{H}$ , there is a unique element of  $\mathfrak{g}$  up to scaling which fixes it. This action is simple enough that the flow can be easily explicitly solved. From this, it can be determined that those elements of  $\mathfrak{g}$  for which the flow in unit time fixes a point  $z \in \mathbb{H}$  can be described as a union of two sets, a one-dimensional subspace along with elements for which  $a^2 + bc$  is a negative integer multiple of  $\pi^2$ , for the latter the action of flow by unit time is the identity. For example, those elements which flow  $i$  to itself in unit time (but for which the flow by unit time is not the identity) are precisely

$$\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}, \quad b \in \mathbb{R}.$$

Given two points  $z, w \in \mathbb{H}$ , there is a one-parameter family of elements of  $\mathfrak{g}$  which flows  $z$  to  $w$  in unit time. For example, those elements of  $G$  which take  $-1 + i$  to  $1 + i$  are

$$\begin{pmatrix} \cos t + \sin t & 2 \cos t \\ \sin t & \cos t + \sin t \end{pmatrix}, t \in \mathbb{R}$$

By directly solving the flow equation, one finds that the elements of  $\mathfrak{g}$  which flow  $-1 + i$  to  $1 + i$  in unit time are all of the form

$$\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix},$$

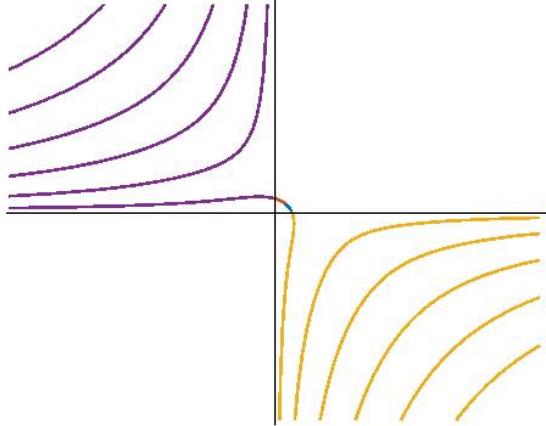
where  $(b, c) = (2, 0), (0, 1)$  or for  $bc > 0$  it is parametrised by  $(bc = t^2)$ ,

$$b = t(\coth t + \sqrt{\operatorname{csch}^2 t - 1}), \quad c = \frac{1}{2}t(\coth t - \sqrt{\operatorname{csch}^2 t - 1}) \quad (0 < \sinh |t| < 1)$$

while for  $bc < 0$  it is parametrised by ( $bc = -t^2$ ),

$$b = t(\cot t + \sqrt{(\csc t)^2 + 1}), \quad c = \frac{1}{2}t(\cot t - \sqrt{(\csc t)^2 + 1}), \quad t \notin \pi\mathbb{Z}$$

In either parametrisation, as  $t \rightarrow 0+$ ,  $(b, c) \rightarrow (2, 0)$ , while as  $t \rightarrow 0-$ ,  $(b, c) \rightarrow (0, 1)$ . Denote the subset of  $\mathfrak{g}$  so defined by  $W$ . A plot of the associated points  $(b, 2c)$  in the plane  $\mathbb{R}^2$  is shown below.



The connection between the infinitesimal action and the action of  $G$  is described by the formula

$$\exp \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} = \cos \sqrt{-bc} \cdot I + \frac{\sin \sqrt{-bc}}{\sqrt{-bc}} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}$$

The result of this paper defines maps  $\mu_n: W_i^n \rightarrow W_i$  for each component  $W_i$  of  $W$ .

Since  $w_0 \equiv \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \in W$  and  $\begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \in \mathfrak{g}$  fixes  $-1 + i$ , hence

$$\text{BCH} \left( \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, s \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \right) \in W$$

for  $s \in \mathbb{R}$  defines one component. In this parametrisation, according to Lemma 5.1(i),  $\mu_n$  will act as the arithmetic mean on the parameters  $s$ .

## 7. Conclusions and applications

After this work was completed, our attention was brought to the field of Karcher means and related constructions [11]. In one setting, these are constructions of weighted means of families of positive definite matrices or operators, while the original Karcher mean [8] was a centre of mass on a Riemannian manifold defined as a minimizer of a sum of least squares. One of the constructions for weighted means of positive matrices [1] uses an inductive symmetrization procedure like that used here

in Section 3. When using the trace metric on matrices,  $d(A, B) = \|\log A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\|$  these different definitions are known to agree. Our mean will also coincide with these when interpreted on the same space, e.g. in Section 6; however we are working in this paper mainly in the context of free Lie algebras.

In Section 4, we only managed to determine  $\mu_n$  up to third order in brackets (this is the same as up to the second order, the third order itself vanishing by Lemma 3.3). For specific values of  $n$ , it is possible to use Lemma 3.2, to computationally determine the sequences of coefficients in  $\mu_n$ , using a Hall basis for the free Lie algebra. This was carried out up to third order in [13], but in principle the same technique could be used to higher orders. Knowing the existence of recurrence relations with polynomial coefficients for these sequences, formulae can then be derived for them as functions of  $n$ . It is unclear which method is faster to determine higher order coefficients, a direct analytic technique or using computations for specific  $n$ . Note that the dimension of the totally symmetric part of the four bracket part of the free Lie algebra is greater than 1 (apparently 10), so that unlike in Section 4, the recurrence relations obtained even for the 4-bracket part will be *systems* of homogeneous linear recurrence relations with polynomial coefficients.

As noted in Section 2, the motivating example for this paper is that of DGLA models of cell complexes. For a cubical cell, pick two opposite vertices and consider the six minimal paths from one to the other. By the functorial nature of the dependence of  $A(X)$  on  $X$ , a model of the cube can be obtained from that of a six-faceted banana; see [7]. Starting from a symmetric model of the six-faceted banana as generated from  $\mu_6$ , the generated model of the cube will share those symmetries of the cube which preserve the chosen diagonal. This model of the cube is used in [9] to construct a discrete version of differential geometry on cubulated manifolds.

It is expected that there may be other applications of universal averages, particularly as they seem to satisfy various interesting relations as in Section 5.

**Acknowledgements:** This research was supported in part by grant 2016219 from the United States-Israel Binational Science Foundation (BSF). The authors have no competing interests to declare. The authors wish to thank D. Kazhdan for the suggestion to look at the example of  $SL_2(\mathbb{R})$  and a referee for the suggestion to emphasise the rationality of the coefficients. The authors wish to thank one of the referees for many useful remarks and for bringing to their attention work on the Karcher mean.

## References

- [1] T. Ando, C.-K. Li, R. Mathias: *Geometric means*, Lin. Alg. Appl. 385 (2004) 305–334.
- [2] U. Buijs, Y. Félix, A. Murillo, D. Tanré: *Maurer-Cartan elements in the Lie models of finite simplicial complexes*, Canad. Math. Bull. 60 (2017) 470–477.
- [3] E. Dynkin: *Calculation of the coefficients in the Campbell-Hausdorff formula (Russian)*, Dokl. Akad. Nauk USSR 57 (1947) 323–326.
- [4] M. Eichler: *A new proof of the Baker-Campbell-Hausdorff formula*, J. Math. Soc. Japan 20 (1968) 23–25
- [5] N. Gadish, I. Griniasty, R. Lawrence: *An explicit symmetric DGLA model of a bi-gon*, J. Knot Theory Ramif. 28/11 (2019), art. no. 1940008.

- [6] I. Griniasty, R. Lawrence: *An explicit symmetric DGLA model of a triangle*, Higher Structures 3/1 (2019) 1–16.
- [7] I. Griniasty, R. Lawrence: *Explicit symmetric DGLA models of 3-cells*, Canadian Math. Bull., doi:10.4153/S0008439520000508, to appear.
- [8] H. Karcher: *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math. 30/5 (1977) 509–541
- [9] R. Lawrence, N. Ranade, D. Sullivan: *Discrete analogues of differential geometry on cubulated manifolds*, in preparation.
- [10] R. Lawrence, D. Sullivan: *A formula for topology/deformations and its significance*, Fund. Math. 225 (2014) 229–242.
- [11] J. Lawson, Y. Lim: *Weighted means and Karcher equations of positive operators*, Proc. Natl. Acad. Sci. USA 110/39 (2013) 15626–15632.
- [12] D. Quillen: *Rational homotopy theory*, Ann. of Math. (2), 90 (1969) 205–295.
- [13] M. Siboni: *Universal Averages in Gauge Actions*, Minor Thesis, Hebrew University, Jerusalem (2019).
- [14] D. Sullivan: *Infinitesimal computations in topology*, Inst. Hautes Études Sci. Publ. Math. 47 (1977) 269–331.
- [15] T. Tradler, M. Zeinalian: *Infinity structure of Poincaré duality spaces*, Algebr. Geom. Topol. 7 (2007) 233–260.

Ruth Lawrence, Einstein Institute of Mathematics, Hebrew University of Jerusalem, Israel;  
ruthel.naimark@mail.huji.ac.il.

Maor Siboni, Racah Institute of Physics, Hebrew University of Jerusalem, Israel;  
maor.siboni@mail.huji.ac.il.

Received May 12, 2020  
and in final form August 25, 2020