M.M. Schipper

The Campbell-Hausdorff theorem

Bachelor's thesis, June 26, 2014

Supervisor: Dr. L.D.J. Taelman

Mathematisch Instituut, Universiteit Leiden

Contents

1 Introduction

The map $\exp\colon \mathbf{C} \to \mathbf{C}^*$ is a group homomorphism. That is, for all $x, y \in \mathbf{C}$ we have

$$\exp(x+y) = \exp(x)\exp(y).$$

Additionally, one can define $\exp(A)$ for a complex $n \times n$ matrix A by the usual power series expansion

$$\exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!}.$$

Let A and B be complex $n \times n$ matrices. If we impose certain criteria on A and B, there exists a matrix C such that

$$\exp(C) = \exp(A)\exp(B).$$

In general, C is not equal to $A+B$, but we can take $C = \log(\exp(A)\exp(B))$ where log is defined by the usual power series expansion. We compute

$$\exp(A)\exp(B) = 1 + A + B + AB + \frac{A^2}{2!} + \frac{B^2}{2!} + \text{ higher-order terms.}$$

We take the logarithm to find the lowest-order terms of C. We have

$$\begin{aligned} C &= A + B + \frac{1}{2}(AB - BA) \\ &\quad + \frac{1}{12}(A^2B + AB^2 + BA^2 + B^2A - 2ABA - 2BAB) + \text{ higher-order terms} \\ &= A + B + \frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] - [B,[A,B]]) + \text{ higher-order terms,} \end{aligned}$$

where $[A,B]$ denotes the commutator $AB - BA$. The Campbell-Hausdorff theorem tells us that C is a series of which *all* terms are linear combinations of iterated commutators in A and B. The series itself does not depend on n, A or B. More naturally: C is an element of the free Lie algebra on $\{A, B\}$. The formula for C is even unique. See theorem 3.1 for the exact statement of the Campbell-Hausdorff theorem.

Campbell, Baker and Hausdorff created the first proof in 1906. In 1968 Eichler produced a totally different, purely algebraic proof. He shows by induction on n that all terms of order n are iterated commutators. Eichler proved the theorem in the context of matrices. In this thesis we will present Eichlers proof in the language of free Lie algebras.

If A and B commute all terms of C, except for $A+B$, equal zero. In this case we have $\exp(A+B) = \exp(A)\exp(B)$. Less trivially, there exist non-commuting matrices A, B where

$$\frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] - [B,[A,B]]) + \text{ higher-order terms} = 0$$

and thus $\exp(A+B) = \exp(A)\exp(B)$.

2 Associative and Lie algebras

2.1 Definitions

Let k be a field.

Definition 2.1. An *associative algebra* over k is a k-vector space A with a k-bilinear map $\cdot\colon A \times A \to A$ and an element $1_A \in A$ such that for all $a, b, c \in A$ we have

(A1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;

(A2) $1_A \cdot a = a \cdot 1_A = a$.

We will use A instead of $(A, \cdot, 1_A)$ to denote an associative algebra. For $x, y \in A$ we will use xy to indicate $x \cdot y$. Let A and A' be associative k-algebras. A map $f\colon A \to A'$ is called a homomorphism of associative algebras over k if it is k-linear, respects $\cdot$ and $f(1_A) = 1_{A'}$.

Example 2.2. The vector space $\mathrm{Mat}_n(k)$ of $n \times n$ matrices over k with matrix multiplication forms an associative k-algebra with unit $\mathrm{id}_{\mathrm{Mat}_n(k)}$.

Definition 2.3. A *Lie algebra* over k is a k-vector space L with a k-bilinear map $[-,-]\colon L \times L \to L$ such that for all $x, y, z \in L$ we have

(L1) $[x, x] = 0$;

(L2) $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$.

We will often denote a Lie algebra $(L, [-,-])$ simply by L. Let L and L' be Lie algebras over k. We will call a map $f\colon L \to L'$ a homomorphism of Lie algebras over k if f is k-linear and it respects $[-,-]$.

Lemma 2.4. *Let L be a Lie algebra over k with $x, y \in L$. We have $[x, y] = -[y, x]$.*

Proof. Since $[-,-]$ is k-bilinear, we have

$$[x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y].$$

By (L1) we have $[x, x] = [y, y] = [x + y, x + y] = 0$. Hence $[x, y] = -[y, x]$. □

Example 2.5. Let A be an associative k-algebra. Then A together with the operation

$$[-,-]\colon A \times A \to A\colon (a, b) \mapsto ab - ba$$

is a Lie algebra over k. The only non-trivial thing to check is (L2), which follows from the associativity of the product in A.

2.2 Free associative and free Lie algebras

Let T be a set.

Definition 2.6. A *free associative algebra on T over k* is a pair (A, f) with A an associative algebra over k and $f\colon T \to A$ a map such that for all associative k-algebras B and for all maps $g\colon T \to B$ there is a unique k-algebra homomorphism $h\colon A \to B$ such that the diagram

$$\begin{array}{ccc} T & \xrightarrow{f} & A \\ & {\scriptstyle g}\searrow & \downarrow{\scriptstyle h} \\ & & B \end{array}$$

commutes.

Proposition 2.7. *Let (A, f) and (A', f') be free associative algebras on T over k. There is a unique isomorphism of associative k-algebras $h\colon A \to A'$ with $h \circ f = f'$.*

Proof. By the universal property of A there is a unique k-algebra homomorphism $h\colon A \to A'$ such that $h \circ f = f'$. We will show that h is an isomorphism.

By the universal property of A' there is a homomorphism of associative k-algebras $h'\colon A' \to A$ such that $h' \circ f' = f$. We will show that h and h' are mutually inverse to each other. The map $h' \circ h\colon A \to A$ is a k-algebra homomorphism with $h' \circ h \circ f = h' \circ f' = f$. The homomorphism id_A has this property as well. According to the universal property of A we have $h' \circ h = \mathrm{id}_A$. Similarly one obtains $h \circ h' = \mathrm{id}_{A'}$. Thus h is the unique isomorphism as described. □

Since a free associative algebra on T over k is uniquely unique, it is allowable to speak of *the* free associative algebra on T over k.

Definition 2.8. Let $i, j \in \mathbf{N}$. Regard T^i as the set of words of length i. Let Ass_T^i be the vector space over k with basis T^i. The dimension of Ass_T^i is $|T|^i$.

Concatenation defines an associative map $\cdot\colon T^i \times T^j \to T^{i+j}$. It extends uniquely to a k-bilinear map $\cdot\, \mathrm{Ass}_T^i \times \mathrm{Ass}_T^j \to \mathrm{Ass}_T^{i+j}$. Let

$$\mathrm{Ass}_T = \bigoplus_{i \geq 0} \mathrm{Ass}_T^i .$$

Varying i and j, the system of maps $\cdot\colon\ \mathrm{Ass}_T^i \times \mathrm{Ass}_T^j \to \mathrm{Ass}_T^{i+j}$ gives a map $\cdot\colon\ \mathrm{Ass}_T \times \mathrm{Ass}_T \to \mathrm{Ass}_T$. This map is k-bilinear, associative and has a unit element, the empty word. Hence Ass_T is an associative algebra over k. The set T can be regarded as a subset of Ass_T, by identifying T with the set of words of length 1.

Proposition 2.9. *Let $\chi\colon T \hookrightarrow \mathrm{Ass}_T$ be the inclusion. Then (Ass_T, χ) is the free associative algebra on T over k.*

Proof. Let A be an associative k-algebra and let $g\colon T \to A$ be a map. Let $h\colon \mathrm{Ass}_T \to A$ be the k-linear map which sends a word $w = x_1x_2\cdots x_n$ with $n \in \mathbf{N}$ and $x_i \in T$ to $g(x_1)g(x_2)\cdots g(x_n)$. The map h is a k-algebra homomorphism with $h \circ \chi = g$.

Let $h'\colon \mathrm{Ass}_T \to A$ be a homomorphism of associative k-algebras with $h' \circ \chi = g$. For all $x \in T$ we have $h'(x) = h'(\chi(x)) = g(x)$. Hence such a map is fixed on T. Since Ass_T is generated by T as an associative algebra, h' is uniquely determined by its images on T. In conclusion $h\colon \mathrm{Ass}_T \to A$ is the unique k-algebra homomorphism with $h \circ \chi = g$. □

The associative algebra Ass_T can be regarded as the set of polynomials over k in non-commuting variables in the set T.

We will now discuss the free Lie algebra.

Definition 2.10. A *free Lie algebra on T over k* is a pair (L, f) with L a Lie algebra over k and $f\colon T \to L$ a map such that for all maps $g\colon T \to M$ with M a Lie algebra over k, there is a unique map $h\colon L \to M$ a homomorphism of Lie algebras over k such that the diagram

$$\begin{array}{ccc} T & \xrightarrow{f} & L \\ & \searrow^{g} & \downarrow{h} \\ & & M \end{array}$$

commutes.

Theorem 2.11. 1. *Let (L, f) and (L', f') be free Lie algebras on T over k. There is a unique isomorphism of Lie algebras $h\colon L \to L'$ with $h \circ f = f'$.*

2. *A free Lie algebra on T over k exists.*

3. *Let (L, f) be a free Lie algebra. Then the map f is injective.*

4. *Let (L, f) be a free Lie algebra. We have $L = \bigoplus_{i\geq 1} L^i$ such that L^1 is the k-vector space with basis T and for all for all $i \geq 2$ we have*

$$L^i = \sum_{j=1}^{i-1} [L^j, L^{i-j}],$$

where $[L^i, L^{i-j}]$ denotes te k-vector space generated by $\{[a, b]\colon a \in L^j, b \in L^{i-j}\}$.

Proof. Since a free Lie algebra is defined by a universal property, it can be proven that it is uniquely unique analogously to the free associative algebra.

In [2, II.2.2, def.1] a Lie algebra $L(T)$ is constructed. In [2, II.2.2, prop.1] a map $\phi\colon T \to L(T)$ is given and it is proven that the pair $(L(T), \phi)$ is a free Lie algebra.

The map ϕ is injective, see [2, II.2.2, after cor.2]. By the uniqueness of the free Lie algebra we have $f = h \circ \phi$ for a certain isomorphism h. Hence f is injective.

The free Lie algebra $(L(T), \phi)$ has a vector space grading as described, see [2, II.2.6, eq.12 and above eq.12]. The vector space L has such a grading as well, since it is isomorphic to $L(T)$ as a Lie algebra over k. □

Since a free Lie algebra on T over k is uniquely unique, we will speak of *the* free Lie algebra on T over k. We will denote its Lie algebra by L_T and we will identify T with its image in L_T.

Notice that the vector spaces L^i in thm. 2.11.4 are uniquely determined. It is possible to give an explicit basis H_i of each L_T^i, but we will not do this. See [2, II.2.10, def.2 and prop.11] for a recursive construction of the sets H_i. In [2, II.2.11, thm.1] it is proven that $\bigcup_{i\geq 1} H_i$ forms a basis of L_T.

Example 2.12. Let $T = \{x, y\}$ with $x \neq y$. The construction of the first four sets H_i gives:

$$\begin{aligned}
H_1 &= \{x, y\} \\
H_2 &= \{[x, y]\} \\
H_3 &= \{[x, [x, y]], [y, [x, y]]\} \\
H_4 &= \{[x, [x, [x, y]]], [y, [x, [x, y]]], [y, [y, [x, y]]]\}.
\end{aligned}$$

For instance the element $[x, [y, [x, y]]] \in \mathrm{L}_T^4$ can be expressed in the basis H_4 as follows:

$$\begin{aligned}
[x, [y, [x, y]]] &= -[x, [[x, y], y]] \text{ by lemma 2.4} \\
&= [y, [x, [x, y]]] + [[x, y], [y, x]] \text{ by (L2)} \\
&= [y, [x, [x, y]]] \text{ by (L1) and lemma 2.4.}
\end{aligned}$$

Remark 2.13. Ass_T is a Lie algebra over k with Lie bracket $[a, b] = ab - ba$ for $a, b \in \mathrm{Ass}_T$. (See example 2.5.) By the universal property of L_T there exists a unique Lie homomorphism $h\colon \mathrm{L}_T \to \mathrm{Ass}_T$ with $h(x) = x$ for all $x \in \mathrm{L}_T$. The map h expands the nested Lie brackets into iterated commutators. For example for $x, y \in \mathrm{L}_T$ we have

$$h([x[x, y]]) = x(xy - yx) - (xy - yx)x.$$

Note that $h(\mathrm{L}_T^i) \subseteq \mathrm{Ass}_T^i$ for all i. At the end of the next section we will see that h is injective.

2.3 Enveloping algebra

Let L be a Lie algebra over k.

Definition 2.14. An *enveloping algebra of* L is a pair (U, f) with U an associative k-algebra and $f\colon L \to U$ a homomorphism of Lie algebras over k such that for all associative k-algebras A and for all Lie homomorphisms $g\colon L \to A$ there is a unique homomorphism of associative k-algebras $h\colon U \to A$ such that the diagram

$$\begin{array}{ccc}
L & \xrightarrow{f} & U \\
& {\scriptstyle g}\searrow & \downarrow{\scriptstyle h} \\
& & A
\end{array}$$

commutes.

Theorem 2.15. *1. Let (U, f) and (U', f') be enveloping algebras of L. There is a unique isomorphism of associative k-algebras $h\colon U \to U'$ with $h \circ f = f'$.*

2. An enveloping algebra of L exists.

3. Let (U, f) be an enveloping algebra of L. Then f is injective.

Proof. The proof of the uniqueness is analogous to the proof of the uniqueness of the free associative algebra.

In [1, III.1, after def.1.1] a pair (U_L, ϵ) is constructed. It is proven that this pair is an enveloping algebra of L, see [1, III.1, thm.1.2].

One should read between the lines to find the injectivity of the map ϵ. We will now clarify this. The objects T_L^1, T_L and I are defined in the text mentioned. The map ϵ is a composition of three maps, of which the first two are clearly injective. For all elements z in the ideal I of T_L, we have $z \notin T_L^1$. Hence for all $x, y \in L = T_L^1 \subseteq T_L$ with $x \neq y$ the classes of x and y in the quotient T_L/I are not the same. Hence ϵ is injective. By the uniqueness of an enveloping algebra, it follows that f is injective. □

Since an enveloping algebra of L is uniquely unique, we will speak of the enveloping algebra of L. We will denote it by (U_L, f).

Example 2.16. Let L be a Lie algebra over k with $[-,-]$ the zero map. Such a Lie algebra is called an abelian Lie algebra. Let $(e_i)_{i\in I}$ be a basis of L as k-vector space.

We will show that U_L is the polynomial ring on variables X_i for $i \in I$. The Lie algebra L is identified with the space of monomials via

$$f\colon L \to k[X_i\colon i \in I], e_i \mapsto X_i.$$

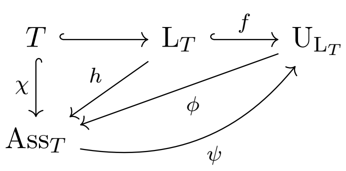
The map f is a Lie algebra homomorphism since $k[X_i\colon i \in I]$ is a commutative ring.

Let A be an associative algebra and let $g\colon L \to A$ be a Lie algebra homomorphism. Since the $g(e_i)$ commute, there is a unique algebra homomorphism $h\colon k[X_i\colon i \in I] \to A$ such that $X_i \mapsto g(e_i)$.

Let $h\colon \mathrm{L}_T \to \mathrm{Ass}_T$ be as in remark 2.13.

Theorem 2.17. *The pair $(\mathrm{U}_{\mathrm{L}_T}, f|_T)$ is the free associative algebra on T over k; there is a unique algebra isomorphism $\psi\colon \mathrm{Ass}_T \to \mathrm{U}_{\mathrm{L}_T}$ such that $\psi \circ \chi = f|_T$.*

Proof. Since $h\colon \mathrm{L}_T \to \mathrm{Ass}_T$ is a Lie homomorphism, there is a unique k-algebra homomorphism $\phi\colon \mathrm{U}_{\mathrm{L}_T} \to \mathrm{Ass}_T$ such that $\phi \circ f = h$, by the universal property of U_{L_T}.



According to the universal property of Ass_T there is a unique homomorphism of associative algebras $\psi\colon\ \mathrm{Ass}_T \to \mathrm{U}_{\mathrm{L}_T}$ with $\psi \circ \chi = f|_T$. We obtain

$$\phi \circ \psi \circ \chi = \phi \circ f|_T = h|_T = \chi.$$

By the universal property of Ass_T we have $\phi \circ \psi = \mathrm{id}_{\mathrm{Ass}_T}$. Furthermore we find

$$\psi \circ \phi \circ f|_T = \psi \circ h|_T = \psi \circ \chi = f|_T.$$

Since $\psi \circ \phi$ can be interpreted as a Lie homomorphism, the universal property of L_T gives $\psi \circ \phi \circ f = f$. Then we have $\psi \circ \phi = \mathrm{id}_{\mathrm{U}_{\mathrm{L}_T}}$ by the universal property of $\mathrm{U}_{\mathrm{L}_T}$. Hence ψ and ϕ are mutually inverse bijections. So there is a unique isomorphism of associative algebras $\psi\colon\ \mathrm{Ass}_T \to \mathrm{U}_{\mathrm{L}_T}$ such that the diagram

$$\begin{array}{ccc} T & \xhookrightarrow{\;f|_T\;} & \mathrm{U}_{\mathrm{L}_T} \\ {\scriptstyle \chi}\big\downarrow & \overset{\sim}{\nearrow}{\scriptstyle \psi} & \\ \mathrm{Ass}_T & & \end{array}$$

is commutative. □

Since f is injective, $h = \phi \circ f$ is. We will use the injectivity of this map in the proof of the Campbell-Hausdorff theorem.

2.4 Completions and the exponential and logarithmic map

We aim to define an exponential and a logarithmic map by the usual power series expansions on the free associative algebra on T over k. These series need not be elements of Ass_T. In this section T is supposed to be a *finite* set.

Definition 2.18. We define the *completions* of the free Lie and free associative algebra to be

$$\hat{\mathrm{L}}_T := \prod_{i=1}^{\infty} \mathrm{L}_T^i,$$

and

$$\hat{\mathrm{Ass}}_T := \prod_{i=0}^{\infty} \mathrm{Ass}_T^i$$

respectively. The k-bilinear maps $[-,-]$ and $\cdot$ extend to k-bilinear maps on $\hat{\mathrm{L}}_T \times \hat{\mathrm{L}}_T$ and $\hat{\mathrm{Ass}}_T \times \hat{\mathrm{Ass}}_T$. This turns $\hat{\mathrm{L}}_T$ and $\hat{\mathrm{Ass}}_T$ into a Lie algebra and an associative algebra respectively. Furthermore define

$$\hat{\mathrm{m}}_T := \prod_{i=1}^{\infty} \mathrm{Ass}_T^i\,.$$

This is the ideal of $\hat{\text{Ass}}_T$ generated by T.

We will denote an element $(f_i)_i \in \hat{\text{Ass}}_T$ by $\sum_i f_i$. In this way we regard $\hat{\text{Ass}}_T$ as the set of power series over k in non-commuting variables in T. For $a \in \hat{\text{Ass}}_T$, we will use a_i to indicate the homogeneous term of degree i of a. If a is not the zero element, define $\text{ord}(a) := \min\{i \in \mathbf{N} \colon (a)_i \neq 0\}$. Otherwise, let $\text{ord}(a) = \infty$.

Definition 2.19. The injective Lie homomorphism $h \colon \text{L}_T \to \text{Ass}_T$ extends to a homomorphism of Lie algebras

$$h \colon \hat{\text{L}}_T \to \hat{\mathfrak{m}}_T, \sum_{i=1}^{\infty} f_i \mapsto \sum_{i=1}^{\infty} h(f_i).$$

By remark 2.13 we have $h(f_i) \in \text{Ass}_T^i$. Hence $\sum_{i=1}^{\infty} h(f_i) \in \prod_{i=1}^{\infty} \text{Ass}_T^i = \hat{\mathfrak{m}}_T$, so h is well defined. The map $h \colon \hat{\text{L}}_T \to \hat{\mathfrak{m}}_T$ is an injective Lie homomorphism since $h \colon \text{L}_T \to \text{Ass}_T$ is.

We will now define an exponential and a logarithmic function by the usual formulas on subspaces of $\hat{\text{Ass}}_T$. To avoid division by zero, we assume $\text{char}(k) = 0$.

Definition 2.20. Define the maps $\exp \colon \hat{\mathfrak{m}}_T \to 1 + \hat{\mathfrak{m}}_T$ and $\log \colon 1 + \hat{\mathfrak{m}}_T \to \hat{\mathfrak{m}}_T$ by the following formulas:

$$\exp(a) = \sum_{i=0}^{\infty} \frac{a^i}{i!}$$

and

$$\log(1 + a) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1} a^i}{i}$$

where $a \in \hat{\mathfrak{m}}_T$.

We will verify that exp is well defined. Let $a \in \hat{\mathfrak{m}}_T$ with $\alpha = \text{ord}(a)$. We find that $\text{ord}(a^i) = \alpha i \geq i$. Therefore we have

$$(\exp(a))_n = \left(\sum_{i=0}^{n} \frac{a^i}{i!} \right)_n \in \text{Ass}_T^n$$

for all $n \in \mathbf{N}$. Hence $\exp(a) \in \hat{\text{Ass}}_T$. Since $(\exp(a))_0 = 1$, exp is well defined. In the same way we can verify that log is well defined.

Lemma 2.21. exp *and* log *are mutually inverse bijections.*

Proof. Let $a \in \hat{\mathfrak{m}}_T$. It is known that log and exp are each others inverses in the power series ring $\mathbf{Q}[[X]]$. Since k is an extension field of $\mathbf{Q}$, there is a map of $\mathbf{Q}$-algebras $\psi \colon \mathbf{Q}[[X]] \to \hat{\text{Ass}}_T$ which transforms X into a. The power series $\sum \lambda_i X^i$ is mapped to $\sum \lambda_i a^i$, which lies in $\hat{\text{Ass}}_T$ since $\text{ord}(a^i) \geq i$.

If we apply ψ to the equality $\exp(\log(1 + X)) = 1 + X$, we get $\exp(\log(1 + a)) = 1 + a$. In the same way we find that exp is a right inverse of log. □

Remark 2.22. Let $x, y \in T$. Since $1 + \hat{\mathfrak{m}}_T$ is closed under multiplication $\exp(x)\exp(y) \in 1 + \hat{\mathfrak{m}}_T$. By lemma 2.21 there is a unique $z \in \hat{\mathfrak{m}}_T$ such that $\exp(z) = \exp(x)\exp(y)$.

3 The Campbell-Hausdorff theorem

In the previous chapter we have introduced all objects and maps we need to formulate the Campbell-Hausdorff theorem. In chapter 3, let T be the set $\{x, y\}$ with $x \neq y$. The ground field of the free associative and free Lie algebra is supposed to be of characteristic zero. Let $h\colon \hat{\mathrm{L}}_T \to \hat{\mathrm{m}}_T$ be as in definition 2.19. We will now formulate the main theorem of this thesis:

Theorem 3.1. *Let $z = \log(\exp(x)\exp(y)) \in \hat{\mathrm{m}}_T$. Then we have $z \in h(\hat{\mathrm{L}}_T)$.*

We will use the whole of chapter 3 to present Eichlers proof of the Campbell-Hausdorff theorem. (See [3] and [4, 7.7].)

The map $h\colon \hat{\mathrm{L}}_T \to \hat{\mathrm{m}}_T$ is an injective Lie homomorphism. So we can identify $h(\hat{\mathrm{L}_T})$ with $\hat{\mathrm{L}_T}$. We will call an element in $\hat{\mathrm{m}}_T$ Lie if it lies in $\hat{\mathrm{L}_T}$. We can regard the quotient space $\hat{\mathrm{m}}_T/\hat{\mathrm{L}}_T$. For $a, b \in \hat{\mathrm{m}}_T$ we will use $a \equiv b$ to indicate that the classes of a and b are the same in $\hat{\mathrm{m}}_T/\hat{\mathrm{L}}_T$. So we aim to prove that $z \equiv 0$. Since $z \in \hat{\mathrm{m}}_T$, it can be written uniquely as

$$z = \sum_{n=1}^{\infty} F_n \tag{3.1}$$

with $F_n \in \mathrm{Ass}_T^n$. We will prove by induction that $F_n \equiv 0$ for all n. Since $F_n \in \mathrm{Ass}_T^n$ and since $\mathrm{L}_T^n \subseteq \mathrm{Ass}_T^n$, we then find that $F_n \in \mathrm{L}_T^n$ and it follows that

$$z = \sum_{n=1}^{\infty} F_n \in \prod_{n\geq 1} \mathrm{L}_T^n = \hat{\mathrm{L}}_T.$$

3.1 Specialisations

To prove the theorem, we need to substitute other elements for x, y in the polynomials F_i. We will now make this rigorous.

Definition 3.2. Let A be an associative algebra with $a, b \in A$. Let ψ be the unique homomorphism of associative algebras $\psi\colon \mathrm{Ass}_T \to A$ which maps x, y to a, b. (We use the universal property of Ass_T.) Define for each $F \in \mathrm{Ass}_T$ the element $F(a, b)$ by $F(a, b) := \psi(F) \in A$.

Let S be a set and let $a, b \in \hat{\mathrm{Ass}_S}$. Since $\hat{\mathrm{Ass}_S}$ is an associative algebra, we can substitute its elements.

Lemma 3.3. *Let $F \in \mathrm{Ass}_T$. Let $\alpha = \mathrm{ord}(a)$, $\beta = \mathrm{ord}(b)$ and $\zeta = \mathrm{ord}(F)$. Then $\mathrm{ord}(F(a, b)) \geq \zeta \min\{\alpha, \beta\}$.*

Proof. If $F \neq 0$, the statement is clearly true. Otherwise, the shortest word w of F has length ζ. If $w = u^\zeta$ where $u = x$ if $\alpha = \min\{\alpha, \beta\}$ and $u = y$ if $\beta = \min\{\alpha, \beta\}$, we find $\operatorname{ord}(w(a,b)) = \zeta \min\{\alpha, \beta\}$. This is the shortest word possible in $F(a,b)$. Hence $\operatorname{ord}(F(a,b)) \geq \zeta \min\{\alpha, \beta\}$. □

We have seen how to do substitutions in non-commutative polynomials. We can also do substitutions in power series with non-commuting variables in T. That is, we will define how to do substitutions in elements of $\hat{\mathrm{Ass}}_T$. Remember a_i denotes the homogeneous component of a, where $a \in \hat{\mathrm{Ass}}_T$.

Definition 3.4. Let ψ be as in definition 3.2 with $\hat{\mathrm{Ass}}_S$ substituted for A and assume $a, b \in \hat{\mathfrak{m}}_T$. Let

$$\hat{\psi}\colon \hat{\mathrm{Ass}}_T \to \hat{\mathrm{Ass}}_S, \sum_{i=0}^{\infty} f_i \mapsto \sum_{i=0}^{\infty} \psi(f_i).$$

By lemma 3.3 the map $\hat{\psi}$ is well defined. Define $f(a,b)$ by $f(a,b) := \hat{\psi}(f)$ for $f \in \hat{\mathrm{Ass}}_T$.

Lemma 3.5. *We have*

$$\log(\exp(a)\exp(b)) = \sum_{i=1}^{\infty} F_i(a,b).$$

Proof. By definition of $F_i(a,b)$, $\hat{\psi}$ and z we obtain

$$\sum_{i=1}^{\infty} F_i(a,b) = \sum_{i=1}^{\infty} \psi(F_i) = \hat{\psi}\left(\sum_{i=1}^{\infty} F_i\right) = \hat{\psi}(\log(\exp(x)\exp(y))).$$

Since ψ is an algebra homomorphism we have

$$\begin{aligned}\hat{\psi}(\log(\exp(x)\exp(y))) &= \sum_{j=1}^{\infty} \psi(\log(\exp(x)\exp(y))_j) \\ &= \sum_{j=1}^{\infty} \log(\exp(\psi(x))\exp(\psi(y)))_j \\ &= \log(\exp(a)\exp(b)).\end{aligned}$$

□

If we substitute Lie power series in a Lie polynomial, we expect to find a Lie power series. This turns out to be the case.

Lemma 3.6. *For $a, b \in \hat{\mathrm{L}}_T \subset \hat{\mathfrak{m}}_T$ and $F \in \mathrm{L}_T \subset \mathrm{Ass}_T$ we have $F(a,b) \in \hat{\mathrm{L}}_T$.*

Proof. Let ψ be as in definition 3.2 with $\hat{\mathrm{Ass}}_T$ substituted for A. Let $i\colon \hat{\mathrm{L}}_T \hookrightarrow \hat{\mathrm{Ass}}_T$ be the inclusion. Let ψ' be the unique homomorphism of Lie algebras $\psi'\colon \mathrm{L}_T \to \hat{\mathrm{L}}_T$ which sends x, y to a, b. (We use the universal property of L_T.)

$$\begin{array}{ccccc} T & \hookrightarrow & \mathrm{L}_T & \hookrightarrow & \mathrm{Ass}_T \\ & & \downarrow \psi' & & \downarrow \psi \\ & & \hat{\mathrm{L}}_T & \overset{i}{\hookrightarrow} & \hat{\mathrm{Ass}}_T \end{array}$$

Notice that ψ is a Lie homomorphism if we regard Ass_T and $\hat{\mathrm{Ass}}_T$ as Lie algebras. The Lie algebra homomorphism $\psi|_{\mathrm{L}_T}$ and $i \circ \psi'$ from L_T to $\hat{\mathrm{Ass}}_T$ are equal if we restrict them to T. So by the universal property of L_T the square commutes. Hence

$$F(a,b) = \psi(F) = (i \circ \psi')(F) = \psi'(F) \in \hat{\mathrm{L}}_T.$$

□

3.2 Proof of the Campbell-Hausdorff theorem

We will prove by induction on n that $F_n \equiv 0$ for all $n \in \mathbf{N}_{\geq 1}$.

By writing down the terms of z containing words of length 1 and 2 one finds $F_1 = x+y$ and $F_2 = \frac{1}{2}(xy - yx)$. Since we identify $h(\mathrm{L}_T^i)$ and L_T^i, the terms F_1 and F_2 lie in L_T^1 and L_T^2 respectively. In conclusion we have $F_1, F_2 \equiv 0$.

Let $N \in \mathbf{N}_{>2}$ and assume all polynomials F_n with $1 \leq n < N$ are Lie. We will use the rest of this chapter to prove that F_N is a Lie polynomial. The first step is to derive the following equation:

Lemma 3.7. *Let $a, b, c \in \mathrm{Ass}_T^1$. We have $F_N(a,b) + F_N(a+b,c) \equiv F_N(a,b+c) + F_N(b,c)$.*

Proof. The associativity of the product in $\hat{\mathrm{Ass}}_T$ gives

$$\Big(\exp(a)\exp(b)\Big)\exp(c) = \exp(a)\Big(\exp(b)\exp(c)\Big).$$

Simple as it is, this step is crucial in Eichlers proof. Since exp and log are mutually inverse and since $a, b, c \in \hat{\mathfrak{m}}_T$ we can apply lemma 3.5 to obtain

$$\exp\left(\sum_{j=1}^{\infty} F_j(a,b)\right)\exp(c) = \exp(a)\exp\left(\sum_{j=1}^{\infty} F_j(b,c)\right).$$

Note that $\sum_{j=1}^{\infty} F_j(a,b)$ and $\sum_{j=1}^{\infty} F_j(b,c)$ lie in $\hat{\mathfrak{m}}_T$. Taking the logarithm and applying lemma 3.5 a second time gives

$$\sum_{i=1}^{\infty} F_i\left(\sum_{j=1}^{\infty} F_j(a,b), c\right) = \sum_{i=1}^{\infty} F_i\left(a, \sum_{j=1}^{\infty} F_j(b,c)\right) \in \hat{\mathfrak{m}}_T. \qquad (3.2)$$

Let G_N be the homogeneous term of degree N of the left side of equation 3.2. By lemma 3.3 one finds that the terms with $i > N$ do not contribute to G_N. Hence we have

$$G_N = \left(\sum_{i=1}^{N} F_i\left(\sum_{j=1}^{\infty} F_j(a,b), c\right)\right)_N. \qquad (3.3)$$

We will now discuss which j, depending on i, contribute to G_N. We need the following two lemmas.

Lemma 3.8. *Let* $A, B \in \hat{\mathrm{m}}_T$*, let* $i \in \mathbf{N}_{\geq 2}$*, let* $n \in \mathbf{N}_{\geq 1}$*. Let* $\epsilon \in \bigoplus_{i \geq n} \mathrm{Ass}_T^i$*. For all* $F \in \mathrm{Ass}_T^i$ *we have*

$$(F(A+\epsilon, B))_n = (F(A, B))_n.$$

Proof. We will give a proof by induction on i. Assume $\epsilon \neq 0$.
Step 1. Let $F \in \mathrm{Ass}_T^2$. Then F is a linear combination of the words x^2, xy, yx and y^2. We have $\mathrm{ord}(\epsilon A) = \mathrm{ord}(A\epsilon) \geq n+1$ and $\mathrm{ord}(\epsilon^2) \geq 2n$. So we get

$$((A+\epsilon)^2)_n = (A^2 + A\epsilon + \epsilon A + (\epsilon)^2)_n = (A^2)_n.$$

In the same way we obtain $((A+\epsilon)(B))_n = (AB)_n$. The statement follows for $i = 2$.
Step 2. Let $I \in \mathbf{N}_{>2}$ and assume the statement is true for all $2 \leq i < I$. It is sufficient to prove the statement for monomials in Ass_T^I. Let $F \in \mathrm{Ass}_T^I$ be a monomial. Then $F = Hx$ or $F = Hy$ for a certain $H \in \mathrm{Ass}_T^{I-1}$. We will prove the statement for the first case. Using lemma 3.3 we find $\mathrm{ord}(H(A+\epsilon, B)\epsilon) \geq n+1$. Since $(A)_0 = 0$ and by the induction hypothesis we obtain

$$\begin{aligned}(H(A+\epsilon, B)(A+\epsilon))_n &= (H(A+\epsilon, B)A)_n \\ &= \sum_{j=0}^{n-1} (H(A, B))_j (A)_{n-j} \\ &= (H(A, B)A)_n.\end{aligned}$$

Hence $(F(A+\epsilon, B))_n = (F(A, B))_n$. □

Lemma 3.9. *Let* $A, B \in \hat{\mathrm{m}}_T$ *and let* $\epsilon \in \bigoplus_{i \geq 2} \mathrm{Ass}_T^i$*. Let* $n \in \mathbf{N}$*. For all* $F \in \mathrm{Ass}_T^n$ *we have*

$$(F(A+\epsilon, B))_n = (F(A, B))_n.$$

Proof. We will prove this lemma by induction on n. Assume $\epsilon \neq 0$.
Step 1. For $n = 0$ and for all $F \in \mathrm{Ass}_T^0$ we have $F(A+\epsilon, B) = F = F(A, B)$.
Step 2. Let $N \in \mathbf{N}$ and assume the statement holds for all $n \in \mathbf{N}$ with $n < N$. It is sufficient to prove the statement for monomials in Ass_T^N. Let $F \in \mathrm{Ass}_T^N$ be a monomial. Then $F = Hx$ or $F = Hy$ for a certain $H \in \mathrm{Ass}_T^{N-1}$. We will prove the statement for the first case. Using lemma 3.3, one finds $\mathrm{ord}(H(A+\epsilon, B)\epsilon) \geq N+1$. Hence

$$(H(A+\epsilon, B)(A+\epsilon))_N = (H(A+\epsilon, B)A)_N.$$

Since $(A)_0 = 0$ and since $\mathrm{ord}(H(A+\epsilon, B)) \geq N-1$ we have

$$(H(A+\epsilon, B)A)_N = (H(A+\epsilon, B))_{N-1}(A)_1.$$

By the induction assumption the right side equals

$$(H(A, B))_{N-1}(A)_1 = (H(A, B)A)_N.$$

It follows that $(F(A+\epsilon, B))_N = (F(A, B))_N$. □

We will now go back to equation 3.3. For $i = 1$ we have

$$\left(F_1\left(\sum_{j=1}^{\infty} F_j(a,b), c\right)\right)_N = \left(\sum_{j=1}^{\infty} F_j(a,b) + c\right)_N = F_N(a,b)$$

since $F_j(a,b)$ lies in Ass_T^j and since c is homogeneous of degree $1 \neq N$. For $2 \leq i \leq N-1$ we use lemma 3.8 to obtain

$$\left(\sum_{i=2}^{N-1} F_i\left(\sum_{j=1}^{\infty} F_j(a,b), c\right)\right)_N = \left(\sum_{i=2}^{N-1} F_i\left(\sum_{j=1}^{N-1} F_j(a,b), c\right)\right)_N .$$

For $i = N$ we use lemma 3.9 to find

$$\left(F_N\left(\sum_{j=1}^{\infty} F_j(a,b), c\right)\right)_N = (F_N(F_1(a,b), c))_N = F_N(F_1(a,b), c).$$

Hence

$$G_N = F_N(a,b) + \left(\sum_{i=2}^{N-1} F_i\left(\sum_{j=1}^{N-1} F_j(a,b), c\right)\right)_N + F_N(a+b, c).$$

We will now determine the class of G_N in $\hat{\mathfrak{m}}_T/\hat{\mathrm{L}}_T$. This is where we use the induction assumption. For $2 \leq j < N$ we have $F_j \in \mathrm{L}_T$ by this assumption. Since $a, b \in \mathrm{Ass}_T^1 = \mathrm{L}_T^1$ we can apply lemma 3.6 to obtain $F_j(a,b) \in \mathrm{L}_T$. Applying the induction assumption and lemma 3.6 a second time, we find $\sum_{i=2}^{N-1} F_i(\sum_{j=1}^{N-1} F_j(a,b), c) \in \mathrm{L}_T$. In conclusion

$$G_N \equiv F_N(a,b) + F_N(a+b, c).$$

We can do the same computation for the right side of equation 3.2 to obtain the important result

$$F_N(a,b) + F_N(a+b, c) \equiv F_N(a, b+c) + F_N(b,c). \tag{3.4}$$

This finishes the proof of lemma 3.7. □

We will now substitute several variables for a, b and c in equation 3.4 to finally obtain $F_N \equiv 0$. Before we start substituting, we will derive some useful facts.

Lemma 3.10. *Let $A, B \in \hat{\mathfrak{m}}_T$ with $AB = BA$. We have* $\exp(A)\exp(B) = \exp(A+B)$.

Proof. Since A and B commute, we can apply Newton's binomial theorem to find

$$\exp(A+B) = \sum_{i=0}^{\infty} \frac{(A+B)^i}{i!} = \sum_{i=0}^{\infty}\sum_{j=0}^{i} \frac{A^{i-j}B^j}{j!(i-j)!}.$$

Furthermore we have

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

The coefficient of $A^{i-j}B^j$ with $0 \leq j \leq i$ is $\frac{1}{j!(i-j)!}$ for both $\exp(A+B)$ and $\exp(A)\exp(B)$. Since the terms of both series are equal, we get $\exp(A+B) = \exp(A)\exp(B) \in 1 + \hat{\mathfrak{m}}_T$. □

We will use the following three facts while performing substitutions.

Lemma 3.11. *Let $r, s \in k$. We have*

1. $F_N(ra, sa) = 0$.
2. $F_N(a, 0) = F_N(0, a) = 0$.
3. $F_N(ra, rb) = r^N F_N(a, b)$.

Proof. Since ra and sa commute, we have $\exp(ra)\exp(sa) = \exp(ra + sa)$ by lemma 3.10. We take the logarithm and use lemma 3.5 to get

$$\sum_{i=1}^{\infty} F_i(ra, sa) = F_1(ra, sa).$$

Because a lies in Ass_T^1, the polynomial $F_i(ra, sa)$ is homogeneous of degree i. By the uniqueness of the power series in $\hat{\mathrm{Ass}}_T$ we have $F_i(ra, sa) = 0$ for all $i > 1$. Hence $F_N(ra, sa) = 0$.

If we take $r = 1$ and $s = 0$ we find $F_N(a, 0) = 0$ and analogously $F_N(0, a) = 0$. Since F_N is homogeneous of degree N, we have $F_N(ra, rb) = r^N F_N(a, b)$. □

We will now start substituting in equating 3.4. Bear in mind that we aim to derive $F_N \equiv 0$. We will first derive a relation between $F_N(a, b)$ and $F_N(b, a)$.

Lemma 3.12. *We have*

$$F_N(a, b) \equiv (-1)^{N+1} F_N(b, a).$$

Proof. First of all, substitute $-b$ for c in equation 3.4. (Note that $-b \in \mathrm{Ass}_T^1$.) We find

$$\begin{aligned} F_N(a, b) + F_N(a + b, -b) &\equiv F_N(a, 0) + F_N(b, -b) \\ &\equiv 0, \end{aligned}$$

where the last equivalence holds by fact 2 and 1 of lemma 3.11. Therefore we have

$$F_N(a, b) \equiv -F_N(a + b, -b). \tag{3.5}$$

Next, substitute $-a$ for b in equation 3.4 and use fact 1 and 2 of lemma 3.11 to find

$$\begin{aligned} F_N(a, -a) + F_N(0, c) &\equiv F_N(a, -a + c) + F_N(-a, c) \\ &\equiv 0. \end{aligned}$$

Then, replacing a, c by $-a, b$ respectively gives

$$F_N(a, b) \equiv -F_N(-a, a + b). \tag{3.6}$$

With equation 3.5 and 3.6 we can find a relation between $F_N(a,b)$ and $F_N(b,a)$. We have

$$\begin{aligned} F_N(a,b) &\equiv -F_N(a+b,-b) \\ &\equiv F_N(-a-b,a) \\ &\equiv -F_N(-b,-a) \\ &\equiv -(-1)^N F_N(b,a), \end{aligned}$$

where the last equivalence follows from lemma 3.11.3. Hence the relation between $F_N(a,b)$ and $F_N(b,a)$ is

$$F_N(a,b) \equiv (-1)^{N+1} F_N(b,a). \tag{3.7}$$

□

Equation 3.4 has more to offer.

Lemma 3.13. *We have*

$$(1-2^{1-N})F_N(a,b) \equiv (1+(-1)^N)2^{-N}F_N(a,a+b).$$

Proof. Now we substitute $-b/2$ for c. (Note that $-b/2 \in \mathrm{Ass}^1_T$.) This gives

$$\begin{aligned} F_N(a,b) + F_N(a+b,-b/2) &\equiv F_N(a,b/2) + F_N(b,-b/2) \\ &\equiv F_N(a,b/2), \end{aligned}$$

where the last equivalence follows from lemma 3.11.1. So we have

$$F_N(a,b) \equiv F_N(a,b/2) - F_N(a+b,-b/2). \tag{3.8}$$

Next, substituting $-b/2$ for a in equation 3.4 and using lemma 3.11.1 gives

$$\begin{aligned} F_N(-b/2,b) + F_N(b/2,c) &\equiv F_N(-b/2,b+c) + F_N(b,c) \\ &\equiv F_N(b/2,c). \end{aligned}$$

We replace b, c by a, b respectively, to obtain

$$F_N(a,b) \equiv F_N(a/2,b) - F_N(-a/2,a+b). \tag{3.9}$$

With equation 3.8 and 3.9 we can pass from polynomials in A, B to polynomials in $A/2, B/2$. This enables us to find a relation between $F_N(a,b)$ and itself. We will use 3.9 to rewrite the two terms on the right side of 3.8.

$$\begin{aligned} F_N(a,b/2) &\equiv F_N(a/2,b/2) - F_N(-a/2,a+b/2) \\ &\equiv F_N(a/2,b/2) + F_N(a/2,a/2+b/2) \\ &\equiv 2^{-N}(F_N(a,b) + F_N(a,a+b)), \end{aligned}$$

where the last two equivalences follow from 3.6 and fact 3 of lemma 3.11.3 respectively. The second term of 3.8 is rewritten by 3.9 as follows:

$$\begin{aligned} F_N(a+b,-b/2) &\equiv F_N(a/2+b/2,-b/2) - F_N(-a/2-b/2,a+b/2) \\ &\equiv -F_N(a/2,b/2) + F_N(a/2+b/2,a/2) \\ &\equiv 2^{-N}(-F_N(a,b) + F_N(a+b,a)), \end{aligned}$$

using 3.5, 3.6 and lemma 3.11.3. So 3.8 becomes

$$F_N(a,b) \equiv 2^{1-N}F_N(a,b) + 2^{-N}F_N(a,a+b) - 2^{-N}F_N(a+b,a).$$

We use 3.7 to simplify this equation to

$$(1-2^{1-N})F_N(a,b) \equiv (1+(-1)^N)2^{-N}F_N(a,a+b). \tag{3.10}$$

□

If N is odd, division by $1-2^{1-N}$ gives the result $F_N(a,b) \equiv 0$. (We use the fact that $N > 1$.) If N is even, we need to do a little work to derive the same.

Lemma 3.14. *If N is even, we have $F_N(a,b) \equiv 0$.*

Proof. We replace b by $b-a$ in equation 3.10, obtaining

$$(1-2^{1-N})F_N(a,b-a) \equiv 2^{1-N}F_N(a,b). \tag{3.11}$$

Applying 3.6 on the left side of 3.11 gives

$$-F_N(-a,b) \equiv \frac{2^{1-N}}{1-2^{1-N}}F_N(a,b). \tag{3.12}$$

This equation allows us to find a relation between $F_N(a,b)$ and itself. We substitute $-a$ for a in equation 3.12 and use the equation itself to obtain

$$\begin{aligned} -F_N(a,b) &\equiv \frac{2^{1-N}}{1-2^{1-N}}F_N(-a,b) \\ &\equiv -\left(\frac{2^{1-N}}{1-2^{1-N}}\right)^2 F_N(a,b). \end{aligned} \tag{3.13}$$

Since $N > 2$, we have

$$\left(\frac{2^{1-N}}{1-2^{1-N}}\right)^2 \neq 1.$$

Thus we have $F_N(a,b) \equiv 0$. □

Finally, we substitute x, y for a, b respectively (which is possible since $x, y \in \mathrm{Ass}^1_T$), to find the desired result.

$$F_N = F_N(x,y) = F_N(a,b) \equiv 0.$$

Eichler has used slightly different substitutions in his proof. This 'freedom' suggests there is a shorter way to obtain $F_N \equiv 0$.

4 Campbell-Hausdorff for matrices

We will show that the Campbell-Hausdorff theorem in the context of matrices is implied by theorem 3.1. We need to think about convergence of series in $\mathrm{Mat}_n(\mathbf{C})$ before we start substituting matrices for the variables of the free associative algebra.

4.1 Convergence of matrix series

Definition 4.1. Let $|\cdot|$ be the map

$$|\cdot|\colon \mathrm{Mat}_n(\mathbf{C}) \to \mathbf{R}_{\geq 0}, (a_{ij})_{i,j} \mapsto \sqrt{\sum_{i,j} |a_{ij}|^2}$$

where $|a_{ij}|$ is the norm of a_{ij} on $\mathbf{C}$.

The map $|\cdot|$ is a norm on $\mathrm{Mat}_n(\mathbf{C})$ and $(\mathrm{Mat}_n(\mathbf{C}), |\cdot|)$ is isomorphic to $\mathbf{R}^{2n^2}$ with the euclidean norm. The norm $|\cdot|$ is submultiplicative:

Lemma 4.2. *We have $|AB| \leq |A||B|$ for all $A, B \in \mathrm{Mat}_n(\mathbf{C})$.*

Proof. If $A = (a_{ij})_{i,j}$ and $B = (b_{ij})_{i,j}$, then we have by the triangle inequality and the multiplicative property of the norm on $\mathbf{C}$

$$\begin{aligned} |(AB)_{ij}| &= |\sum_{l=1}^{n} a_{il} b_{lj}| \\ &\leq \sum_{l=1}^{n} |a_{il}||b_{lj}|. \end{aligned}$$

By the Cauchy-Schwarz inequality we get

$$\sum_{l=1}^{n} |a_{il}||b_{lj}| \leq \sqrt{\sum_{l=1}^{n} |a_{il}|^2} \sqrt{\sum_{l=1}^{n} |b_{lj}|^2}.$$

It follows that

$$\begin{aligned} |AB| &= \sqrt{\sum_{i,j} |(AB)_{ij}|^2} \\ &\leq \sqrt{\sum_{i} \sum_{l=1}^{n} |a_{il}|^2 \sum_{j} \sum_{l=1}^{n} |b_{lj}|^2} \\ &= |A||B|. \end{aligned}$$

□

Let $T = \{x, y\}$ and take the ground field k for Ass_T and L_T equal to $\mathbf{C}$. Let $F \in \mathrm{Ass}_T$ and let $A, B \in \mathrm{Mat}_n(\mathbf{C})$. Then $F(A, B) \in \mathrm{Mat}_n(\mathbf{C})$ makes sense. (See definition 3.2.) In this chapter we will denote an element $f \in \hat{\mathrm{Ass}}_T$ by $\sum_i f_i$ where each f_i is a monomial or by $\sum a_w w$ where we sum over all words w with $a_w \in \mathbf{C}$.

Lemma 4.3. *Let $\sum_i f_i \in \hat{\mathrm{Ass}}_T$. Assume $\sum_i f_i(|A|, |B|)$ converges absolutely in $\mathbf{R}$, then $\sum_i f_i(A, B)$ converges in $\mathrm{Mat}_n(\mathbf{C})$.*

Proof. Since $\sum_i |f_i(|A|, |B|)|$ converges in $\mathbf{R}$, the sequence of partial sums

$$\Big(\sum_{i=0}^{n} \Big|f_i(|A|, |B|)\Big|\Big)_n$$

is Cauchy. Let $\epsilon \in \mathbf{R}_{>0}$. Let $N \in \mathbf{N}$ be such that for all $n, m \in \mathbf{N}_{>N}$ with $n \geq m$ we have

$$\sum_{i=m+1}^{n} \Big|f_i(|A|, |B|)\Big| = \Bigg|\sum_{i=m+1}^{n} \Big|f_i(|A|, |B|)\Big|\Bigg| < \epsilon.$$

Then we have by the triangle inequality and the submultiplicative property of $|\cdot|$ on $\mathrm{Mat}_n(\mathbf{C})$ the inequality

$$\Bigg|\sum_{i=m+1}^{n} f_i(A, B)\Bigg| \leq \sum_{i=m+1}^{n} \Big|f_i(|A|, |B|)\Big|.$$

We use the fact that each f_i is a monomial. For $n, m \in \mathbf{N}_{>N}$ with $n \geq m$ the last term is smaller than ϵ. Hence $(\sum_i f_i(A, B))_n$ is a Cauchy sequence in $\mathrm{Mat}_n(\mathbf{C})$. By the completeness, the sequence converges. □

4.2 Matrix exponential and logarithmic map

Remark 4.4. The series $\exp(\lambda) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}$ and $\log(1 + \lambda) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}\lambda^i}{i}$ are absolutely convergent for all $\lambda \in \mathbf{R}$ and for all $\lambda \in (-1, 1)$ respectively. By lemma 4.3 $\sum_{i=0}^{\infty} \frac{A^i}{i!}$ converges and $\sum_{i=1}^{\infty} \frac{(-1)^{i+1}A^i}{i}$ converges if $|A| < 1$ with $A \in \mathrm{Mat}_n(\mathbf{C})$.

Definition 4.5. Define the *matrix exponential* and *logarithmic map* to be

$$\exp\colon \mathrm{Mat}_n(\mathbf{C}) \to \mathrm{Mat}_n(\mathbf{C}), A \mapsto \sum_{i=0}^{\infty} \frac{A^i}{i!},$$

$$\log\colon D \to \mathrm{Mat}_n(\mathbf{C}), 1 + A \mapsto \sum_{i=1}^{\infty} \frac{(-1)^{i+1}A^i}{i}$$

with $D = \{1 + A \in \mathrm{Mat}_n(\mathbf{C})\colon \sum_{i=1}^{\infty} \frac{(-1)^{i+1}A^i}{i} \text{ converges}\}$.

Lemma 4.6. *Let* $A, B \in \mathrm{Mat}_n(\mathbf{C})$.

1. *If* A *and* B *commute, we have* $\exp(A+B) = \exp(A)\exp(B)$.
2. $\exp(A)$ *is invertible with inverse* $\exp(-A)$.

Proof. The first statement is proven as in lemma 3.10. Since A and $-A$ commute we have $\exp(A)\exp(-A) = \exp(0) = 1$. Similarly $\exp(-A)$ is a left inverse of $\exp(A)$. □

Remark 4.7. By lemma 4.6 the codomain of the matrix-exponential map can be taken equal to the unit group of the domain (analogous to the exponential map on $\mathbf{C}$). We shall denote the unit group of $\mathrm{Mat}_n(\mathbf{C})$ by $\mathrm{GL}_n(\mathbf{C})$. The map $\exp\colon \mathbf{C} \to \mathbf{C}^*$ is a group homomorphism, whereas $\exp\colon \mathrm{Mat}_n(\mathbf{C}) \to \mathrm{GL}_n(\mathbf{C})$ is not if $n > 1$.

Let $A, B \in \mathrm{Mat}_n(\mathbf{C})$.

Definition 4.8. Let $\psi\colon \mathrm{Ass}_T \to \mathrm{Mat}_n(\mathbf{C})$ be the map as in definition 3.2 which replaces x, y by A, B. Define

$$\hat{\psi}\colon \Big\{ \sum_i f_i \in \hat{\mathrm{Ass}}_T \text{ such that } \sum_i f_i(A,B) \text{ converges} \Big\} \to \mathrm{Mat}_n(\mathbf{C}), \sum_i f_i \mapsto \sum_i \psi(f_i).$$

Lemma 4.9. *Let* $\sum_i f_i = \log(\exp(x))$ *and* $\sum_i g_i = \exp(\log(1+x))$.

1. *If* $\sum_i f_i(|A|,|B|)$ *converges absolutely we have* $\log(\exp(A)) = A$.
2. *If* $\sum_i g_i(|A|,|B|)$ *converges absolutely we have* $\exp(\log(1+A)) = 1 + A$.
3. *If* $|\exp(|A|) - 1| < 1$ *the series* $\sum_i f_i(|A|,|B|)$ *converges absolutely.*
4. *If* $|A| < 1$ *the series* $\sum_i g_i(|A|,|B|)$ *converges absolutely.*

Proof. If $\sum_i f_i(|A|,|B|)$ converges absolutely, the terms of $\sum_i f_i(A,B)$ can be interchanged. Hence we have

$$\hat{\psi}\Big(\sum_i f_i\Big) = \sum_i f_i(A,B) = \log(\exp(A)).$$

We apply $\hat{\psi}$ to the equality $\log(\exp(x)) = x$ (see lemma 2.21) to find $\log(\exp(A)) = A$. The second statement is proven analogously.

Let $h = \sum h_w w \in \hat{\mathfrak{m}}_T$ with $h_w \geq 0$ for all words w. We can write

$$-\log(1-h) = \sum_{i=1}^{\infty} \frac{h^i}{i} = \sum a_w w,$$
$$\log(1+h) = \sum_{i=1}^{\infty} \frac{(-1)^{i+1} h^i}{i} = \sum b_w w.$$

By the triangle inequality and since $a_w \geq 0$ we have $|b_w| \leq a_w$ for all w. So if $\sum a_w w(|A|,|B|)$ converges, the series $\sum_w b_w w(|A|,|B|)$ converges absolutely.

We can take h equal to $\exp(x)$. If we assume $|\exp(|A|) - 1| < 1$, it follows that $\sum_w b_w w(|A|,|B|)$ converges absolutely. The fourth statement is derived similarly starting with $\exp(-\log(1-x)) = \sum a_w w$ and $\exp(\log(1+x)) = \sum b_w w$. □

Let $A, B \in \mathrm{Mat}_n(\mathbf{C})$ with $|A|, |B| < \log(\sqrt{2})$.

Lemma 4.10. *For $\sum_i f_i = \log(\exp(x)\exp(y))$, the series $\sum_i f_i(|A|, |B|)$ converges absolutely.*

Proof. The proof is analogous to the proof of lemma 4.9.1 where we take h equal to $\exp(x)\exp(y)$ and use the fact that $|\exp(|A|)\exp(|B|) - 1| < \exp(\log(\sqrt{2}))^2 - 1 = 1$. $\square$

Remark 4.11. By lemma 4.10 the series $\sum_i f_i(A, B)$ converges and its terms can be interchanged. Therefore we can define the matrix $C := \log(\exp(A)\exp(B)) = \sum_i f_i(A, B)$. By lemma 4.9, C is the unique matrix with $|\exp(|C|) - 1| < 1$ and $\exp(C) = \exp(A)\exp(B)$.

Theorem 4.12. *The matrix C is a series of which all terms are linear combinations of iterated commutators in A and B.*

Proof. As in lemma 3.5 we can prove $C = \sum_{n=1}^{\infty} F_n(A, B)$ where F_n is as in equation 3.1. The last equality in this lemma holds since it is allowed to interchange the terms of $\log(\exp(A)\exp(B))$. Each $F_i(A, B)$ is a linear combination of iterated commutators in A and B by theorem 3.1. $\square$

5 Commuting matrices

Let $A, B \in \mathrm{Mat}_n(\mathbf{C})$. If A and B commute, that is $[A, B] = 0$, we have $\exp(A + B) = \exp(A)\exp(B)$. In general, the converse implication does not hold.

Lemma 5.1. *There exists an $n \in \mathbf{N}$ and matrices $A, B \in \mathrm{Mat}_n(\mathbf{C})$ such that $[A, B] \neq 0$ and $\exp(A + B) = \exp(A)\exp(B)$.*

Proof. Take $n = 2$,

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 2\pi i \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 \\ 0 & 2\pi i \end{pmatrix}.$$

Because A is diagonal, we compute

$$\exp(A) = \begin{pmatrix} \exp(0) & 0 \\ 0 & \exp(2\pi i) \end{pmatrix} = 1_{\mathrm{Mat}_n(\mathbf{C})}.$$

We can compute $\exp(B)$ using the Jordan Normal Form of B. Since 0 and $2\pi i$ are eigenvalues of B, this form is diagonal. We have

$$B = Q \cdot \begin{pmatrix} 0 & 0 \\ 0 & 2\pi i \end{pmatrix} \cdot Q^{-1} = Q \cdot A \cdot Q^{-1}$$

for a certain matrix $Q \in \mathrm{GL}_n(\mathbf{C})$. Hence $\exp(B) = Q \cdot \exp(A) \cdot Q^{-1} = 1_{\mathrm{Mat}_n(\mathbf{C})}$. In the same way one finds $\exp(A + B) = 1_{\mathrm{Mat}_n(\mathbf{C})}$. In conclusion we have $\exp(A + B) = \exp(A)\exp(B)$, but

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & -4\pi^2 \end{pmatrix} \neq \begin{pmatrix} 0 & 2\pi i \\ 0 & -4\pi^2 \end{pmatrix} = BA.$$

□

However, for a certain subspace of $\mathrm{Mat}_n(\mathbf{C})$ the converse implication is true.

Definition 5.2. Define for all $l \in \mathbf{N}$ with $1 \leq l \leq n$ the subspace

$$T_{n,l} := \{X = (x_{ij})_{i,j} \in \mathrm{Mat}_n(\mathbf{C})\colon x_{ij} = 0 \text{ if } j < i + l\}.$$

$T_{n,l}$ is the subspace of matrices where all non-zero entries lie above l-th diagonal. For example, for $A \in T_{4,2}$ the matrix has the following shape

$$A = \begin{pmatrix} 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We will often denote $T_{n,1}$ simply by T_n. This is the subspace of upper triangular matrices with zeros on the main diagonal.

Remark 5.3. We have the inclusions of subspaces $T_n = T_{n,1} \supset T_{n,2} \supset .... \supset T_{n,n} = \{0\}$. For all $l, m \in \mathbf{N}$ with $1 \leq l, m \leq n$ we have

$$T_{n,l}T_{n,m} = \{AB \colon A \in T_{n,l}, B \in T_{n,m}\} \subseteq T_{n,l+m}. \tag{5.1}$$

Proposition 5.4. *The maps* $\exp \colon T_n \to 1 + T_n$ *and* $\log \colon 1 + T_n \to T_n$ *are mutually inverse bijections.*

Proof. Let $A \in T_n$. By remark 5.3 we have $\exp(A) = 1 + \sum_{i=1}^{n-1} \frac{A^i}{i!} \in 1 + T_n$. Since A is nilpotent, $\log(A)$ is well defined and $\log(A) \in T_n$. For $\sum_i f_i = \exp(\log(x))$, the series $\sum_i f_i(|A|, |B|)$ converges absolutely since it concerns a finite sum. By lemma 4.9.1 we have $\exp(\log(A)) = A$. We will now prove that $\exp \colon T_n \to 1 + T_n$ is injective.

Let $A, B \in T_n$ with $\exp(A) = \exp(B)$. I will prove by induction that $A - B \in T_{n,i}$ for all $1 \leq i \leq n$. We clearly have $A - B \in T_{n,1}$. Let $I \in N$ and assume $A - B \in T_{n,i}$ for all $1 \leq i < I$. Since $\exp(A) = \exp(B)$ we have

$$A - B = \sum_{j=2}^{n-1} \frac{B^j - A^j}{j!}.$$

Since

$$B^j - A^j = B(B^{j-1} - A^{j-1}) - (A - B)A^{j-1},$$

we find that $B^j - A^j \in T_{n,I}$ if $B^{j-1} - A^{j-1} \in T_{n,I-1}$ using the induction assumption and remark 5.3. Since $B - A \in T_{n,I-1}$, it follows inductively that $B^j - A^j \in T_{n,I}$ for all $2 \leq j \leq n - 1$. Hence we obtain $A - B \in T_{n,n} = \{0\}$. Since $A = B$, the map exp is injective. It then follows that exp and log are mutually inverse to each other. □

Theorem 5.5. *Let* $A, B \in T_n$. *We have* $[A, B] = 0$ *if and only if* $\exp(A + B) = \exp(A)\exp(B)$.

Proof. We will prove the implication from right to left. Assume $\exp(A+B) = \exp(A)\exp(B)$. We claim that $[A, B] \in T_{n,i}$ for all $1 \leq i \leq n$. We will prove this by induction on i.
Step 1. For $i = 1$ we have $[A, B] \in T_{n,2} \subset T_{n,1}$ by remark 5.3.
Step 2. Let $I \in \mathbf{N}$, suppose the claim is true for all $1 \leq i < I$. Let F_j be as in equation 3.1. Applying equation 5.1 $j - 1$ times, we obtain $F_j(A, B) \in T_{n,j}$. Hence for all $j \geq n$ we have $F_j(A, B) = 0$. According to the proof of theorem 4.12 we have

$$\exp(A)\exp(B) = \exp(A + B + F_2(A, B) + F_3(A, B) + ... + F_{n-1}(A, B)).$$

Since $\exp(A + B) = \exp(A)\exp(B)$ and by proposition 5.4 one finds

$$0 = F_2(A, B) + F_3(A, B) + ... + F_{n-1}(A, B).$$

For $j > 2$, we have $F_j(A, B) \in T_{n,I}$ by equation 5.1 and the induction assumption. In conclusion, we have $F_3(A, B) + ... + F_{n-1}(A, B) \in T_{n,I}$. If $F_2(A, B) \notin T_{n,I}$ there is a j with $j < i + I$ and $(F_2(A, B))_{ij} \neq 0$. We then have

$$F_2(A, B) + F_3(A, B) + ... + F_n(A, B) \neq 0.$$

This is a contradiction. Hence $[A, B] = 2F_2(A, B) \in T_{n,I}$.

It follows that $[A, B] \in T_{n,n} = \{0\}$. □

Bibliography

[1] Jean-Pierre Serre. *Lie algebras and Lie groups.* W.A. Benjamin, Inc., New York-Amsterdam, 1965.

[2] Nicolas Bourbaki. *Lie groups and Lie algebras, Part I.* Addison-Wesley, Massachusetts, 1975.

[3] M. Eichler. A new proof of the Baker-Campbell-Hausdorff formula. *Journal of the Mathematical Society of Japan.* 20:23-25, 1968.

[4] John Stillwell. *Naive Lie Theory.* Springer, New York, 2008.