Revue d'histoire des mathématiques,
5 (1999), p. 249–284.

POINCARÉ'S PROOF OF THE SO-CALLED BIRKHOFF-WITT THEOREM

Tuong Ton-That, Thai-Duong Tran (*)

In honor of the 100th birthday of the article, "Sur les groupes continus"

Abstract. — A methodical analysis of the research related to the article, "Sur les groupes continus", of Henri Poincaré reveals many historical misconceptions and inaccuracies regarding his contribution to Lie theory. A thorough reading of this article confirms the priority of his discovery of many important concepts, especially that of the *universal enveloping algebra* of a Lie algebra over the real or complex field, and the *canonical map* (*symmetrization*) of the symmetric algebra onto the universal enveloping algebra. The essential part of this article consists of a detailed discussion of his rigorous, complete, and enlightening proof of the so-called Birkhoff-Witt theorem.

Résumé. — La démonstration de Poincaré du théorème dit de Birkhoff-Witt. — Une analyse méthodique des travaux faits en connexion avec l'article, "Sur les groupes continus", de Henri Poincaré révèle des erreurs historiques et des jugements injustes en ce qui concerne sa contribution à la théorie de Lie. Une étude approfondie de cet article confirme l'antériorité de sa découverte de plusieurs concepts importants; notamment de *l'algèbre enveloppante universelle* d'une algèbre de Lie sur le corps réel ou le corps complexe, et de *l'application canonique* (*la symétrisation*) de l'algèbre symétrique sur l'algèbre enveloppante universelle. L'essentiel de cet article consiste en un examen approfondi de sa démonstration rigoureuse et complète du théorème de Birkhoff-Witt.

1. INTRODUCTION

In our research on the universal enveloping algebras of certain infinite-dimensional Lie algebras we were led to study in detail the original proofs

(*) Texte reçu le 18 février 1999, révisé le 18 mai 2000.
Tuong Ton-That and Thai-Duong Tran, Department of Mathematics, University of Iowa, Iowa City, IA 52242 (USA).
E-mail: tonthat@math.uiowa.edu and ttran@math.uiowa.edu.

of the so-called Birkhoff-Witt theorem (more recently, Poincaré-Birkhoff-Witt theorem). This, in turn, led us to the investigation of Poincaré's contribution to Lie theory (*i.e.*, the theory of Lie groups, Lie algebras, and their representations). To our great surprise we discovered many historical misconceptions and inaccuracies, even in some of the classics written by the leading authorities on the subject. This discovery has puzzled us for some time, and we have sought the opinions of several experts in the field. Their answers together with our thorough reading of several original articles on the subject shed some light on this mystery. We were astounded to find out that Poincaré was given credit neither for his fundamental discovery of the *universal enveloping algebra* of a Lie algebra over a field of characteristic zero, nor for his introduction of the *symmetrization map*, and only a cursory and belated acknowledgment of his contribution to the so-called *Birkhoff-Witt theorem*, of which he gave a rigorous, complete, beautiful, and enlightening proof. Indeed, in two of the most exhaustive treatises on universal enveloping algebra [Cohn 1981] and [Dixmier 1974], Poincaré [1900] was not mentioned. In many authoritative textbooks treating Lie theory such as [Chevalley 1955], [Cartan & Eilenberg 1956], [Kuros 1963], [Jacobson 1962], [Varadarajan 1984 (1974)], [Humphreys 1972], [Knapp 1986], . . ., Poincaré's discovery of the universal enveloping algebra and the symmetrization map was ignored. In some books his name was left off the Birkhoff-Witt theorem, and his fundamental article [Poincaré 1900] was not even quoted. In the *Encyclopaedia of Mathematics* [*Encyclopaedia* 1988–1994] under the rubric "Birkhoff-Witt theorem" it was written "*. . . The first variant of this theorem was obtained by H. Poincaré; the theorem was subsequently completely demonstrated by E. Witt* [1937] *and G.D. Birkhoff* [1][1937]. . . ". Clearly the author, T.S. Fofanova, did not read carefully [Poincaré 1900]; otherwise she would have realized that Poincaré had discovered and completely demonstrated this theorem at least thirty-seven years before Witt and Birkhoff. Why such slights can happen to one of the greatest mathematicians of all times, who published

[1] Actually Garrett Birkhoff (1911–1996), not G.D. (Birkhoff) which are the initials of George David Birkhoff (1884–1944), the father of Garrett. This inaccuracy only occurs in the translation, not in the original (Russian) version of the *Encyclopaedia*. We are grateful to Professor Sergei Silvestrov for elucidating this fact to us.

these results [Poincaré 1900] in one of the most prestigious scientific journals, *Transactions of the Cambridge Philosphical Society*, on the occasion of the jubilee of another great mathematician, Sir George Gabriel Stokes, is a most interesting mystery that we shall attempt to elucidate in this article. But before beginning our investigation we want to make it clear that our intention is to study thoroughly one of the most fundamental discoveries by one of the greatest minds in order to understand how important ideas are created, and not to rectify such injustices, for such a task is doomed to fail as the force of habit always prevails; a fact very clearly expressed in the following excerpt from [Gittleman 1975, p. 186], "*... l'Hospital's rule, Maclaurin's series, Cramer's rule, Rolle's theorem, and Taylor's series are familiar terms to calculus students. Actually, only one of these five mathematicians was the original discoverer of the result attributed to him, and that man was Rolle. The person who popularizes a result generally has his name attached to it, although later it may be learned that someone else had originally discovered the same result. For practical purposes names are not changed, but even so, the mistakes seem to compensate for one another. Although Maclaurin was credited with a series he did not discover, a rule which he did originate is now known as Cramer's rule...*". Besides, Poincaré is a member of the elite group of mathematicians to whom many important mathematical discoveries are attributed; indeed, in the *Encyclopaedia of Mathematics* [*Encyclopaedia* 1988–1994] 18 rubrics are listed under his name. Curiously, under the heading "*Poincaré last theorem*" the editorial comments state that "[this theorem] *is also known as the Poincaré-Birkhoff fixed-point theorem,*" and the author, M.I. Voĭtsekhovskiĭ, wrote "*...it was proved by him in a series of particular cases but he did not, however, obtain a general proof of this theorem*".[2] Misnaming mistakes seem to compensate one another after all.

[2] Voĭtsekhovskiĭ continues, "*The paper was sent by Poincaré to an Italian journal two weeks before his death, and the author expressed his conviction, in an accompanying letter to the editor, of the validity of the theorem in the general case.*" Indeed, on December 9, 1911, having some presentiments that he might not live long, Poincaré wrote a moving letter to Guccia, director and founder of the journal *Rendiconti del Circolo Matematico di Palermo* (*cf.* [Poincaré *Œuvres*, II, p. LXVII]), asking his opinion regarding what has become known as "Poincaré's Last Geometric Theorem" (see [Barrow-Green 1997, §7.4.2, pp. 169–174], for an English translation of the letter and an excellent discussion of the theorem). Mr. Guccia readily accepted the memoir for publication and it appeared on March 10, 1912, just a few months before Poincaré's

In his book [Bell 1937], E.T. Bell, who called Poincaré "*the Last Universalist*", considered the Last Geometric Theorem as Poincaré's "*unfinished symphony*" and wrote "... *And it may be noted that Poincaré turned his universality to magnificent use in disclosing hitherto unsuspected connections between distant branches of mathematics, for example, between continuous groups and linear algebra*". This is exactly the impression we had when reading his article, "Sur les groupes continus".

2. POINCARÉ'S WORK ON LIE GROUPS

To assess Poincaré's contribution to Lie theory in general we use two main sources [Poincaré *Œuvres*] and [Poincaré *in memoriam* 1921] and investigate in depth the references cited therein. We start with the article, "Analyse des travaux scientifiques de Henri Poincaré, faite par lui-même" [3] which was written by Poincaré himself in 1901 [Poincaré *in memoriam* 1921, pp. 3–135] at the request of G. Mittag-Leffler (*cf.* "Au lecteur" [Poincaré *in memoriam* 1921, pp. 1–2]). It is part of vol. 38 of the journal, *Acta Math.*, published in 1921 in memory of Henri Poincaré. (Actually, most of vol. 39 published in 1923 is also devoted to Poincaré's work). In the third part of the above-mentioned article, Section XII (Algèbre) and Section XIII (Groupes Continus) are devoted to his contribution to Lie theory. Actually, we think that because of Poincaré's impetus finite-dimensional continuous groups were eventually called Lie groups. Indeed, Poincaré expressed repeatedly his great admiration for Lie's work in [Poincaré 1899] and [Poincaré 1900] and wrote in *Rapport sur les*

death on July 17, 1912 (Sur un théorème de géométrie, *Rendiconti del Circolo Matematico di Palermo*, 33, pp. 375–407 = [Poincaré *Œuvres*, VII, pp. 499–538]). Ultimately it was G.D. Birkhoff who gave a complete proof of this theorem (Proof of Poincaré's geometric theorem, *Trans. Amer. Math. Soc.*, 14 (1913), pp. 14–22 = *Collected Mathematical Papers I*, pp. 673–681) and of its generalization to n dimensions (Une généralisation à n dimensions du dernier théorème de géométrie de Poincaré, *C. R. Acad. Sci. Paris*, 192 (1931), pp. 196–198 = *Collected Mathematical Papers II*, pp. 395–397).

[3] In [Poincaré *Œuvres*] this article is listed as published by *Acta Math.*, 30 (1913), pp. 90–92. In fact, it never existed as such; the editors of Poincaré's collected works probably found the manuscript of the article among his papers with his annotations regarding the journal and the date of publication but due to World War I it appeared eventually in [Poincaré *in memoriam* 1921]. This remark extends to all discrepancies between the intended and actual dates of publication of many of Poincaré's works in [Poincaré *in memoriam* 1921], for example, *Rapport sur les travaux de M. Cartan.*

travaux de M. Cartan [Poincaré *in memoriam* 1921, pp. 137–145]: "*Je commencerai par* les groupes continus et finis*, qui ont été introduits par Lie dans la science; le savant norvégien a fait connaître les principes fondamentaux de la théorie, et il a montré en particulier que la structure de ces groupes dépend d'un certain nombre de constantes qu'il désigne par la lettre c affectée d'un triple indice et entre lesquelles il doit y avoir certaines relations... une des plus importantes applications* des groupes de Lie...". So far as we know this is the first time that the name *Lie groups* was explicitly mentioned.

Poincaré's first encounter with Lie theory probably dated back to his article [Poincaré 1881] and its generalization [Poincaré 1883]. The problem he considered there can be phrased in modern language as follows:

For $X = (x_1, \ldots, x_n) \in \mathbb{C}^n$ let $\mathrm{GL}_n(\mathbb{C})$, the general linear group of all $n \times n$ invertible complex matrices, act on $\mathbb{C}^n$ via $(X, g) \mapsto Xg$, $g \in \mathrm{GL}_n(\mathbb{C})$. Let $F(X)$ denote a homogeneous form of degree m (*i.e.*, a homogeneous polynomial of degree m in n variables $(x_1, \ldots, x_n)$), find the subgroup G of $\mathrm{GL}_n(\mathbb{C})$ which preserves the form F; *i.e.*, $F(Xg) = F(X)$, for all $g \in G$. Conversely, given a subgroup G of $\mathrm{GL}_n(\mathbb{C})$ find all homogeneous forms that are left invariant by G. This is precisely the problem of polynomial invariants (*cf.* [Weyl 1946]).

In [Poincaré 1881] and [Poincaré 1883] he found all cubic ternary (of three variables) and quaternary (of four variables) forms that are preserved by certain Abelian groups (which he called "*faisceau de substitutions*"), and he also extended this result to the non-Abelian case. Conversely, he exhibited explicit groups that preserve quadratic and cubic ternary and quaternary forms. For example, in [Poincaré 1881, pp. 239–241] he found the subgroup of the unipotent group

$$\left\{ \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & \gamma \\ 0 & 0 & 1 \end{pmatrix} \; ; \; \alpha, \beta, \gamma \in \mathbb{C} \right\}$$

which preserves the quadratic form

$$\begin{aligned} &[\, x_1 \quad x_2 \quad x_3 \,] \begin{bmatrix} A_1 & B_3 & B_2 \\ B_3 & A_2 & B_1 \\ B_2 & B_1 & A_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &\qquad = A_1 x_1^2 + A_2 x_2^2 + A_3 x_3^2 + 2B_1 x_2 x_3 + 2B_2 x_1 x_3 + 2B_3 x_1 x_2, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad A_i, B_i \in \mathbb{C}\, ; \; 1 \leq i \leq 3. \end{aligned}$$

Conversely, he showed that all quadratic forms which are left invariant by the full unipotent group defined above must satisfy a certain partial differential equation. Lie theory also plays an important role in Poincaré's work on the conformal representation of functions of two variables which in turn leads to the theory of relativity. In [Poincaré 1906] and [Poincaré 1912] he studied the group of linear transformations which leave the Minkowski's metric $x^2 + y^2 + z^2 - t^2$ invariant, which he called the homogeneous Lorentz group, or as H.A. Lorentz wrote in [Lorentz 1921] "*groupe de relativité*", and discovered the Poincaré group which is the semidirect product of the four-dimensional translation group with the homogeneous Lorentz group.

In his analysis of his scientific accomplishments Poincaré classified his work in seven topics which range from "*Differential Equations*" to "*Philosophy of Science*". His accomplishments in any single one of these areas would already make him famous. Indeed Sir George H. Darwin (1845–1912), a physicist and son of the famous Charles Darwin (1809–1882), wrote in 1909 that Poincaré's celestial mechanics would be a vast mine for researchers for half a century [Boyer 1968, pp. 652]. Under rubric number three "*Questions diverses de Mathématiques pures*", Algebra, Arithmetic, Group theory, and Analysis Situs (combinatorial topology) are listed together, with Lie groups as a subsection of Group theory. This gives a false impression that he had only a slight interest in the subject. In fact with the exception of his first article on continuous groups [Poincaré 1899] the other three articles are quite long: [Poincaré 1900] (35 pages), [Poincaré 1901] (47 pages) and [Poincaré 1908] (60 pages). In all these articles he not only conveyed to the reader his keen interest in the subject but also some of the difficulties that preoccupied him over a ten-year period.

It was Lie's third theorem that motivated Poincaré to write [Poincaré 1899]. This theorem can be stated in *Poincaré's notations*[4] as follows:

If $\{X_1, \ldots, X_r\}$ is a system of infinitesimal transformations (*i.e.*, vector

[4] Poincaré used parentheses instead of brackets for the commutator products. To avoid confusion we replace the parentheses with the more conventional brackets. He also used the notation c_{iks} instead of the more convenient notation c^s_{ik} for the structure constants. We do not however replace this notation, which does not cause any confusion, to preserve as much as possible Poincaré's style and terminology.

fields) which satisfy the equation

$$[X_i, X_k] = \sum_{s=1,\ldots,r} c_{iks} X_s \tag{2.1}$$

then the structure constants c_{iks} must satisfy the relations

$$c_{kis} = -c_{iks}, \tag{2.2}$$

$$\sum_{k=1}^{r} (c_{iks}c_{jlk} + c_{\ell ks}c_{ijk} + c_{jks}c_{\ell ik}) = 0 \qquad (1 \le i, j, \ell, s \le r), \tag{2.3}$$

which follow immediately from the fact that the bracket $[\cdot\,,\,\cdot]$ is skew symmetric, and the Jacobi identity

$$\big[[X_a, X_b], X_c\big] + \big[[X_b, X_c], X_a\big] + \big[[X_c, X_a], X_b\big] = 0.$$

Conversely, if the coefficients c_{iks} satisfy equations (2.2) and (2.3) then there exists a system of infinitesimal transformations verifying equation (2.1), and hence a group of transformations with r parameters.

In [Poincaré 1900] he gave a different proof of this theorem, especially for the case of Lie algebras with non-trivial centers. His approach consists of reformulating Lie's construction of the *adjoint group*, deriving the differential equations associated with the *parametric group*, and then showing that these equations can be integrated. More specifically, if $X_1, \ldots, X_r$ form a basis of a Lie algebra of infinitesimal transformations, and

$$V = \sum v_i X_i, \quad T = \sum t_i X_i,$$

then one has the adjoint representation

$$T \longmapsto T' = \mathrm{e}^{-V} T \,\mathrm{e}^{V},$$

where $T' = \sum t'_i X_i$ and $\mathrm{e}^{tX} = \sum_{n=0}^{\infty} t^n/n! X^n$. The image of the adjoint representation is then called the adjoint group. By setting

$$\mathrm{e}^{V}\,\mathrm{e}^{T} = \mathrm{e}^{W}, \quad \text{where} \quad W = \sum w_i X_i,$$

it follows that the w are functions of the v and t, or, in other words the transformation e^{T} transforms e^{V} into e^{W}, *i.e.*, the v into the w; and the group thus defined in r variables is called the parametric group associated with the system $\{X_1, \ldots, X_r\}$. In [Poincaré 1901] and [Poincaré 1908] he studied in great detail these differential equations and the isomorphism between the adjoint and parametric groups. His research into expressions of W as a function of U and V in the formula $\mathrm{e}^U\mathrm{e}^V = \mathrm{e}^W$ resulted in a precursory form of the Baker-Campbell-Hausdorff formula (see [Schmid 1982] for a discussion regarding Poincaré's contribution to this theorem and, *e.g.*, [Varadarajan 1984 (1974), p. 114] or [Bourbaki 1972, Chap. II], for a more modern proof of this theorem).

3. POINCARÉ'S DISCOVERY OF THE UNIVERSAL ENVELOPING ALGEBRA AND THE SO-CALLED BIRKHOFF-WITT THEOREM

In this section we shall expound the main theme of this article, namely, Poincaré's priority in the discovery of the universal enveloping algebra and the so-called Birkhoff-Witt theorem. For this purpose we shall examine in detail his article [Poincaré 1900]. As a general rule we try to adhere faithfully to his exposition, notation, and style as much as possible. But in order to make our point we shall insert some comments, prove some claims which Poincaré considered self-evident but did not seem to be so obvious to us, and integrate his work into the more modern framework of Lie theory. At first reading [Poincaré 1900] seems to be hastily written, repetitive, and sometimes cryptic, and this might explain why not many people have read it; especially for the readers for whom French is not their first language. But by a careful analysis of [Poincaré 1900] one must conclude without a shade of doubt that Poincaré had discovered the concept of the universal algebra of a Lie algebra and gave a complete and rigorous proof of the so-called Birkhoff-Witt theorem. As we shall see, his entire proof of this theorem, with the exception of the claim that we will state as Theorem 3.3, is quite rigorous and modern in language. For these reasons we will translate the parts in [Poincaré 1900] that are relevant to our discussion for the benefit of the readers who are not familiar with French. But before going into the details we shall elaborate on why we consider his proof very enlightening. For example, his introduction of *the symmetrization map* is made quite natural by the observation that the most elementary "*regular*" (or "*symmetric*") polynomials are the linear polynomials and their powers. And, as it turns out, all symmetrized polynomials are linear combinations of those. From the symmetrization the notion of *equipollence* comes out naturally. This, in turn, leads to the notion of *equivalence*, and ultimately to the definition of the *universal enveloping algebra of a Lie algebra.* In the proof of the fact stated as Theorem 3.8 below, he introduced the notion of "*chains*" and cleverly showed that he could add *more* chains to paradoxically *reduce* the number of basic chains. This enabled him to proceed by induction (on the degree of the regular polynomials). This ingenious idea foreshadows some techniques used in the modern theory of *word problem.*

For the remainder of this section, in order to capture Poincaré's vivid flow of ideas we shall use the present tense to present his exposition. For convenience, we shall discuss the universal enveloping algebra first. Let $X_1, \dots, X_n$ be n elementary operators (Poincaré thinks of these operators as vector fields but never really uses this fact here). Let $\mathcal{L}$ be the Lie algebra over a field of characteristic zero $\mathbb{K}$ (Poincaré thinks of $\mathbb{K}$ as $\mathbb{R}$ or $\mathbb{C}$ but all concepts and proofs remain identical) generated by these n elementary operators which constitute a basis for $\mathcal{L}$. Let $\mathcal{A}$ denote the *non-commutative* algebra of polynomials in n variables $X_1, \dots, X_n$ with coefficients in $\mathbb{K}$. Consider the set of all elements of $\mathcal{A}$ of the form

$$P\big(XY - YX - [X, Y]\big)Q, \tag{3.1}$$

where P and Q are arbitrary polynomials in $\mathcal{A}$, and where $[X, Y]$ denotes the bracket product of $\mathcal{L}$. Define an *equivalence relation* $\sim$ in $\mathcal{A}$ by declaring that an element $A \in \mathcal{A}$ is equivalent to 0 if A is a linear combination of elements of the form (3.1) for some P and Q in $\mathcal{A}$, and $A \sim A'$, $A' \in \mathcal{A}$, if $A - A' \sim 0$. Then the quotient algebra (or residue ring) thus defined is now called the *universal enveloping algebra* of $\mathcal{L}$. In fact, this can be rephrased in modern language as follows:

Let $\mathcal{T}$ denote the tensor algebra over the underlying vector space of $\mathcal{L}$, then $\mathcal{T}$ is isomorphic to $\mathcal{A}$ (see, for example, [Lang 1965, Prop. 10, p. 423]). Let J denote the two-sided ideal of $\mathcal{T}$ generated by the tensors $X \otimes Y - Y \otimes X - [X, Y]$ where $X, Y \in \mathcal{L}$. Then the associative algebra $\mathcal{U} = \mathcal{T}/J$ is called the universal enveloping algebra of $\mathcal{L}$ (*cf.*, *e.g.*, [Bourbaki 1975, p. 12] or [Bourbaki 1960, p. 22]). Under the isomorphism between $\mathcal{T}$ and $\mathcal{A}$ the ideal J corresponds to the two-sided ideal of $\mathcal{A}$ spanned by all elements of the form (3.1). Actually this is exactly Harish-Chandra's approach to the universal enveloping algebra in [Harish-Chandra 1949].

Unaware that Poincaré had defined this notion in [Poincaré 1900], Harish-Chandra [1949, p. 900] wrote the following footnote: "*This algebra is the same as the one considered by Birkhoff* [1937] *and Witt* [1937], *though their method of construction is different".* He also wrote: "*In view of this* 1-1 *correspondence between representations of* $\mathcal{L}$ *and* $\mathcal{U}$ *it is appropriate to call* $\mathcal{U}$ *the general enveloping algebra of* $\mathcal{L}$". In his fundamental paper [Harish-Chandra 1951] on the role of the universal enveloping algebra of a semisimple Lie algebra in Lie theory, published two years later, he replaced the word "*general*" with "*universal*", probably under the influence of Birkhoff's work on universal algebras. Thus we can conclude that it was Harish-Chandra who named this algebra discovered by Poincaré "*universal enveloping algebra*".

In his very influential book [Chevalley 1955], C. Chevalley, one of the world's leading experts in Lie theory and a founding member of Bourbaki, attributed to Harish-Chandra the following theorem: *"There exists a one-to-one correspondence* (*but not multiplicative*!) *between elements of* $\mathcal{U}$ *and those of the symmetric algebra of* $\mathcal{L}$*; also if* $\mathcal{L}$ *is the Lie algebra of a Lie group* G*, then* $\mathcal{U}$ *is isomorphic to the algebra of right* (*or left*) *invariant differential operators over the algebra of analytic functions on* G" [Chevalley 1955, vol. III, chap. 5, §6]. Note that Chevalley and Harish-Chandra were colleagues at Columbia University during this period. It is also interesting to note that Chevalley also gives a proof of the Birkhoff-Witt theorem [Chevalley 1955, vol. III, Prop. 1, p. 163] without mentioning the work of Birkhoff and Witt. Anyhow, many authors seem not to acknowledge Poincaré's discovery of the fundamental notion of universal enveloping algebra; for example, in the encyclopaedic work [Dixmier 1974] Poincaré's work is not even referred to with regard to this algebra.

Finally we are coming to the main part of this article, namely, Poincaré's proof of the Birkhoff-Witt theorem. But before going into detail about the proof, we shall make some historical remarks. As discussed earlier, none of the leading experts in Lie theory seemed to be aware of the existence of Poincaré's work on the universal enveloping algebra and his proof of the so-called Birkhoff-Witt theorem prior to about 1956. Garrett Birkhoff and Ernst Witt certainly didn't mention Poincaré's work in [Birkhoff 1937] and [Witt 1937], respectively. M. Lazard [1952, 1954] generalized this theorem, which he called the Witt theorem, but did mention [Birkhoff 1937] and the work of Kourotchkine. So far as we know the authors who first noticed that the proof of the Birkhoff-Witt theorem already appeared in [Poincaré 1900] were H. Cartan and S. Eilenberg [1956]. Curiously, they called the theorem the Poincaré-Witt theorem and did not refer to Birkhoff's proof; moreover, they attributed the complete proof of the theorem to Witt. It appears that Bourbaki was the first to call this theorem the Poincaré-Birkhoff-Witt theorem in [Bourbaki 1960], a recognition acknowledged in arguably the most influential book on Lie algebras in the English language [Jacobson 1962]. From then on this is the prevalent name used for this theorem; however, many authors of serious books on Lie theory such as [Kuros 1963], [Cohn 1981] and more recently [Knapp 1986], etc., continue to call it the Birkhoff-Witt theorem.

Now let us carefully examine [Poincaré 1900], especially the portion relevant to our investigation, pp. 224–232.

The section heading is "Calcul des polynômes symboliques". Let $X, Y, Z, T, U, \dots,$ be n *elementary operators* (i.e., a basis for a Lie algebra over a commutative field $\mathbb{K}$ of characteristic zero). Consider the algebra of *symbolic* (*or formal*) *non-commutative polynomials* in these operators with coefficients in $\mathbb{K}$. Then as previously mentioned, we may identify this algebra with the tensor algebra $\mathcal{T}$.

Definition 3.1. — Two monomials are said to be *equipollent* if they differ only by the order of their factors. This definition extends obviously to two polynomials that are sums of pairwise equipollent monomials. Ex. XY^2, YXY, and Y^2X are equipollent monomials, and $3XY^2 + 3YZ^2 + 3ZX^2$ and $XY^2+YXY+Y^2X+YZ^2+ZYZ+Z^2Y+ZX^2+XZX+X^2Z$ are equipollent polynomials.

Definition 3.2. — A polynomial is said to be *regular* (or *normal*) if it can be expressed as a linear combination of powers of the form

$$(\alpha X + \beta Y + \gamma Z + \cdots)^p, \qquad p \in \mathbb{N},\ \alpha, \beta, \gamma \in \mathbb{K}. \tag{3.2}$$

Poincaré then makes several statements without bothering to prove them. (They must have seemed obvious to him; note that the same statements are made in [Poincaré 1899], which is an abridged version of [Poincaré 1900], where regular polynomials are called normal.) However, because of the importance of their implications we shall formulate these statements as a theorem and provide the reader with a proof which seems to be quite long, but we do not see how to shorten it. (We suspect that because of these unproved statements some authors did not consider Poincaré's proof rigorous.)

Theorem 3.3. — (i) *A necessary and sufficient condition for a polynomial to be regular is that if it contains among its terms a certain monomial then it must contain all monomials equipollent to that monomial and with the same coefficient.*

(ii) *Among all polynomials equipollent to a given polynomial there exists one and only one regular polynomial.*

Some preparatory work is needed for the proof of this theorem.

Let $\mathcal{P} \equiv \mathcal{P}(x_1, \ldots, x_n)$ denote the *commutative* algebra of polynomials in n indeterminates $x_1, \ldots, x_n$, with coefficients in the field $\mathbb{K}$. For an integer $m \geq 0$ let $\mathcal{P}^m \equiv \mathcal{P}^m(x_1, \ldots, x_n)$ denote the subspace of all homogeneous polynomials of degree m. If $(\alpha) = (\alpha_1, \ldots, \alpha_n)$ is a multi-index of non-negative integers set $x^{(\alpha)} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|(\alpha)| = \sum_{i=1}^n \alpha_i$. Then it is clear that the set $\{x^{(\alpha)}\}$ where (α) ranges over all multi-indices such that $|(\alpha)| = m$ forms a basis for $\mathcal{P}^m$. Let $\mathcal{A} \equiv \mathcal{A}(X_1, \ldots, X_n)$ denote the *algebra of non-commutative polynomials* in $X_1, \ldots, X_n$, then $\mathcal{A}$ is obviously graded. Let $\mathcal{A}^m \equiv \mathcal{A}^m(X_1, \ldots, X_n)$ denote the subspace of all homogeneous elements of $\mathcal{A}$ of degree m. Define the *symmetrization map* (Poincaré does not formulate this map explicitly, but it is obvious from the context that he must have it in mind) $\Phi_m : \mathcal{P}^m \to \mathcal{A}^m$ as follows:

For $1 \leq j \leq \alpha_1$ let $X'_j = X_1$, for $\alpha_1 + 1 \leq j \leq \alpha_1 + \alpha_2$ set $X'_j = X_2$, for $\alpha_1 + \alpha_2 + 1 \leq j \leq \alpha_1 + \alpha_2 + \alpha_3$, set $X'_j = X_3, \ldots,$ for $\alpha_1 + \cdots + \alpha_{n-1} + 1 \leq j \leq m$ set $X'_j = X_n$, with the convention that whenever $\alpha_i = 0$ then the term X_i does not appear. Let

$$\Phi_m(x^{(\alpha)}) = \mathrm{Sym}(X^{(\alpha)}) = \frac{1}{m!} \sum_{\sigma \in \Sigma_m} X'_{\sigma(1)} \cdots X'_{\sigma(m)} \tag{3.3}$$

and extend this definition by linearity to all elements of $\mathcal{P}^m$ (note that $X^{(\alpha)} = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$). For example, with $(\alpha) = (1, 2, 0, \ldots, 0)$ and $m = 3$, then $x^{(\alpha)} = x_1 x_2^2$, $X^{(\alpha)} = X_1 X_2^2$, $X'_1 = X_1$, $X'_2 = X_2$, $X'_3 = X_2$, and

$$\begin{aligned} \mathrm{Sym}(X^{(\alpha)}) = \frac{1}{6}(&X'_1X'_2X'_3 + X'_1X'_3X'_2 + X'_3X'_2X'_1 \\ &+ X'_2X'_1X'_3 + X'_2X'_3X'_1 + X'_3X'_1X'_2) \end{aligned} \tag{3.4}$$

$$
\begin{aligned}
&= \frac{1}{6}(X_1X_2^2 + X_1X_2^2 + X_2^2X_1 \\
&\qquad\qquad + X_2X_1X_2 + X_2^2X_1 + X_2X_1X_2) \\
&= \frac{1}{3}(X_1X_2^2 + X_2X_1X_2 + X_2^2X_1).
\end{aligned}
$$

Now the dimension of $\mathcal{P}^m$ is precisely the number of ways a n-tuple of integers $(\alpha_1, \dots, \alpha_n)$ can be chosen so that $|(\alpha)| = m$. If d_m denotes this number then a combinatorial formula gives $d_m = \binom{n+m-1}{m}$. Order this set Λ_m of multi-indices following the reverse lexicographic ordering as follows:

$$
(3.5) \quad \begin{cases} (\alpha) \prec (\beta) \text{ if for some } k,\ 1 \le k \le n,\ \alpha_k < \beta_k \\ \qquad\qquad\qquad\qquad \text{and } \alpha_{k+1} = \beta_{k+1}, \dots, \alpha_n = \beta_n; \\ (\alpha) \preceq (\beta) \text{ if either } (\alpha) \prec (\beta) \text{ or } (\alpha) = (\beta). \end{cases}
$$

Then $\preceq$ is obviously a total ordering; for example, $(m, 0, \dots, 0) \prec (m-1, 1, 0, \dots, 0) \prec (m-3, 3, 0, \dots, 0)$, and $(m, 0, \dots, 0)$ is the first (the least) element and $(0, 0, \dots, 0, m)$ is the last (the largest) element under this ordering. If p is an element of $\mathcal{P}^m$ of the form

$$
p = (c_1x_1 + \cdots + c_nx_n)^m, \quad \text{where } c_i \in \mathbb{K},\ 1 \le i \le n,
$$

then clearly

$$
p = \sum_{(\alpha)\in\Lambda_m} \binom{m}{\alpha} c^{(\alpha)} x^{(\alpha)},
$$

where the multinomial coefficient $\binom{m}{\alpha}$ is equal to $m!/\alpha_1! \dots \alpha_n!$, and $c^{(\alpha)} = c_1^{\alpha_1} \cdots c_n^{\alpha_n}$. If we denote by $\widetilde{X}^{(\alpha)}$ the image of $x^{(\alpha)}$ by Φ_m (*i.e.*, $\widetilde{X}^{(\alpha)} = \operatorname{Sym}(X^\alpha)$), then by linearity

$$
\Phi_m(p) = \sum_{(\alpha)\in\Lambda_m} \binom{m}{\alpha} c^{(\alpha)} \widetilde{X}^{(\alpha)}.
$$

On the other hand an easy computation shows that

$$
(3.6) \quad \binom{m}{\alpha} \widetilde{X}^{(\alpha)} = \text{sum of all distinct elements} \\ \text{of } \mathcal{A}^m \text{ equipollent to } X^{(\alpha)}.
$$

By expanding $(c_1X_1 + \cdots + c_nX_n)^m$ and taking into account the non-commutativity of the products of the X_i we see that

$$\begin{aligned}(3.6)\quad &(c_1X_1+\cdots+c_nX_n)^m \\ &= \sum_{(\alpha)\in\Lambda_m} c^{(\alpha)}(\text{sum of all distinct elements of } \mathcal{A}^m \text{ equipollent to } X^{(\alpha)}) \\ &= \sum_{(\alpha)\in\Lambda_m} \binom{m}{\alpha} c^{(\alpha)} \widetilde{X}^{(\alpha)}.\end{aligned}$$

Thus

$$\Phi_m\big((c_1x_1+\cdots+c_nx_n)^m\big) = (c_1X_1+\cdots+c_nX_n)^m.$$

It follows by linearity that the image of a regular polynomial in $\mathcal{P}^m$ is the regular polynomial in $\mathcal{A}^m$, obtained by substituting the variable x_i by the variable X_i; moreover, *regular elements of $\mathcal{A}^m$ are already symmetrized.* Let $\Phi\colon\mathcal{P}\to\mathcal{A}$ denote the linear map obtained by setting

$$\Phi(p) = \sum_{m\geq 0}\Phi_m(p_m), \tag{3.8}$$

where p is decomposed into homogeneous elements as $p=\sum_{m\geq 0}p_m$.

If $\sum_{(\alpha)\in\Lambda_m}\lambda_{(\alpha)}x^{(\alpha)}$ is an arbitrary element of $\mathcal{P}^m(\lambda_{(\alpha)}\in\mathbb{K}$, for all $(\alpha)\in\Lambda_m)$, then from equations (3.3) and (3.6) it follows that

$$\begin{aligned}\Phi_m\Big(\sum_{(\alpha)\in\Lambda_m}\lambda_{(\alpha)}x^{(\alpha)}\Big) &= \sum_{(\alpha)\in\Lambda_m}\lambda_{(\alpha)}\widetilde{X}^{(\alpha)} \\ &= \sum_{(\alpha)\in\Lambda_m}\lambda_{(\alpha)}\frac{1}{\binom{m}{\alpha}}\,(\text{sum of all distinct elements of } \mathcal{A}^m \text{ equipollent to } X^{(\alpha)}).\end{aligned}$$

Since the non-ordered monomials of degree m (*i.e.*, the set of all distinct elements of $\mathcal{A}^m$ equipollent to $X^{(\alpha)}$ for all $(\alpha)\in\Lambda_m$) form a basis for $\mathcal{A}^m$, it follows that if

$$\Phi_m\Big(\sum_{(\alpha)\in\Lambda_m}\lambda_{(\alpha)}x^{(\alpha)}\Big) = 0,$$

then $\lambda_{(\alpha)}=0$ for all $(\alpha)\in\Lambda_m$. Thus Φ_m is one-to-one, and it follows that Φ is a *monomorphism* of vector spaces (but *not* an algebra homomorphism). Let $\mathcal{R}\equiv\mathcal{R}(X_1,\ldots,X_n)$ denote the subspace of all regular polynomials of $\mathcal{A}$; then among the consequences of Theorem 3.3 one can infer that Φ *is a vector space isomorphism of the vector space* $\mathcal{P}$ *onto the vector space* $\mathcal{R}$, *and moreover, regular elements in* $\mathcal{A}$ *are already symmetrized.*

Lemma 3.4. — *For any positive integer* r *there exists an* n*-tuple* $(c) = (c_1, \dots, c_n)$ *of positive integers such that*

$$(c)^{(\beta)} \geq r\, c^{(\alpha)} \quad \text{whenever } (\alpha) \prec (\beta),\ \forall\, (\alpha), (\beta) \in \Lambda_m.$$

Proof. — Given $r \in \mathbb{N}^*$ choose $c_1 = 1$ and define c_k inductively by $c_k = r(c_{k-1})^m$ for $2 \leq k \leq n$. Then clearly $c_1 \leq c_2 \leq \cdots \leq c_n$ and $c_k = r(c_{k-1})^m \geq r c_1^{\alpha_1} c_2^{\alpha_2} \cdots c_{k-1}^{\alpha_{k-1}}$ since $\alpha_1 + \cdots + \alpha_n = m$. Hence if $(\alpha) \prec (\beta)$, *i.e.*, $\alpha_k < \beta_k$ and $\alpha_{k+1} = \beta_{k+1}, \dots, \alpha_n = \beta_n$ then

$$\begin{aligned} c^{(\beta)} \geq c_k^{\beta_k} c_{k+1}^{\beta_{k+1}} \cdots c_n^{\beta_n} &\geq c_k c_k^{\alpha_k} c_{k+1}^{\alpha_{k+1}} \cdots c_n^{\alpha_n} \\ &\geq (r c_1^{\alpha_1} c_2^{\alpha_2} \cdots c_{k-1}^{\alpha_{k-1}})(c_k^{\alpha_k} c_{k+1}^{\alpha_{k+1}} \cdots c_n^{\alpha_n}) = r c^{(\alpha)}. \ \square \end{aligned}$$

Lemma 3.5. — *The polynomials* $\widetilde{X}^{(\alpha)} \equiv \Phi_m(x^{(\alpha)}) \equiv \mathrm{Sym}(X^\alpha)$ *are regular for all* $(\alpha) \in \Lambda_m$.

Proof. — We prove by induction that for each $(\beta) \in \Lambda_m$ the following statement holds:

(3.9) *For every* $(\alpha) \preceq (\beta)$, $\widetilde{X}^{(\alpha)}$ *can be expressed as*

$$\widetilde{X}^{(\alpha)} = f_{(\alpha)}^{(\beta)} + \sum_{(\gamma) \succ (\beta)} \lambda_{(\alpha)(\gamma)}^{(\beta)} \widetilde{X}^{(\gamma)},$$

where $f_{(\alpha)}^{(\beta)}$ *is a regular element of* $\mathcal{A}^m$ *and the constants* $\lambda_{(\alpha)(\gamma)}^{(\beta)}$ *are rational numbers.*

First observe that if

$$(\beta) = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{th slot}}}{m}, 0, \dots, 0),$$

then $x^{(\beta)} = x_i^m$ and $\Phi_m(x^{(\beta)}) = \widetilde{X}^{(\beta)} = X_i^m$, which is by definition regular. Thus the first element in this reverse lexicographical ordering is $(\beta) = (m, 0, \dots, 0)$ and the statement (3.9) holds trivially with $\widetilde{X}^{(m,0,\dots,0)} = X_1^m$, where

$$f_{(m,0,\dots,0)}^{(m,0,\dots,0)} = X_1^m \quad \text{and} \quad \lambda_{(m,0,\dots,0),(\gamma)}^{(m,0,\dots,0)} = 0$$

for all $(\gamma) \succ (m, 0, \ldots, 0)$.

Now assume the statement holds for (β). Let (β') denote the immediate successor to (β), and consider the element $g_{(\beta')} \in \mathcal{A}^m$ of the form

$$g_{(\beta')} = (c_1 X_1 + \cdots + c_n X_n)^m = \sum_{(\alpha)\in\Lambda_m} \binom{m}{(\alpha)} c^{(\alpha)} \widetilde{X}^{(\alpha)},$$

where the n-tuple $(c) = (c_1, \ldots, c_n)$ is yet to be determined. This implies that

$$\binom{m}{(\beta')} c^{(\beta')} \widetilde{X}^{(\beta')} = g_{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \widetilde{X}^{(\alpha)} - \sum_{(\gamma)\succ(\beta')} \binom{m}{(\gamma)} c^{(\gamma)} \widetilde{X}^{(\gamma)}.$$

From (3.9) it follows that

$$\begin{aligned}
\binom{m}{(\beta')} c^{(\beta')} \widetilde{X}^{(\beta')} &= g_{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \Big(f^{(\beta)}_{(\alpha)} + \sum_{(\gamma)\succ(\beta)} \lambda^{(\beta)}_{(\alpha)(\gamma)} \widetilde{X}^{(\gamma)} \Big) \\
&\qquad - \sum_{(\gamma)\succ(\beta')} \binom{m}{(\gamma)} c^{(\gamma)} \widetilde{X}^{(\gamma)} \\
&= \Big(g_{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} f^{(\beta)}_{(\alpha)} \Big) \\
&\qquad - \sum_{(\alpha)\preceq(\beta)} \sum_{(\gamma)\succ(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\gamma)} \widetilde{X}^{(\gamma)} \\
&\qquad - \sum_{(\gamma)\succ(\beta')} \binom{m}{(\gamma)} c^{(\gamma)} \widetilde{X}^{(\gamma)}.
\end{aligned}$$

Since (β') is right after (β), the multi-indices $(\gamma) \succ (\beta)$ consist of $(\gamma) = (\beta')$ and $(\gamma) \succ (\beta')$. Hence equation (3.10) can be written as

$$\begin{aligned}
\binom{m}{(\beta')} c^{(\beta')} \widetilde{X}^{(\beta')} &= \Big(g_{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} f^{(\beta)}_{(\alpha)} \Big) \\
&\qquad - \Big(\sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\beta')} \Big) \widetilde{X}^{(\beta')} \\
&\qquad - \sum_{(\gamma)\succ(\beta')} \Big\{ \Big(\sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\gamma)} \Big) + \binom{m}{(\gamma)} c^{(\gamma)} \Big\} \widetilde{X}^{(\gamma)}.
\end{aligned}$$

This implies that

$$\Big[\binom{m}{(\beta')} c^{(\beta')} + \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\beta')} \Big] \widetilde{X}^{(\beta')} \tag{3.11}$$

$$= \Big(g_{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} f^{(\beta)}_{(\alpha)} \Big) \\ - \sum_{(\gamma)\succ(\beta')} \Big\{ \Big(\sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\gamma)} \Big) + \binom{m}{(\gamma)} c^{(\gamma)} \Big\} \widetilde{X}^{(\gamma)}.$$

In equation (3.11) we can solve for $\widetilde{X}^{(\beta')}$ provided that

$$\binom{m}{(\beta')} c^{(\beta')} + \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\beta')}$$

is not zero. To insure this we now determine $(c) = (c_1, \dots, c_n)$ as in Lemma 3.4 by choosing the integer r such that

$$r > d_m \max_{(\alpha)\preceq(\beta)} \Big\{ \binom{m}{(\alpha)} \big| \lambda^{(\beta)}_{(\alpha)(\beta')} \big| \Big\}, \tag{3.12}$$

where $d_m = \binom{m+n-1}{m}$ is the cardinality of Λ_m. Then

$$\binom{m}{(\beta')} c^{(\beta')} \geq c^{(\beta')} > \sum_{(\alpha)\preceq(\beta)} \frac{c^{(\beta')}}{d_m} \\ \geq \sum_{(\alpha)\preceq(\beta)} \frac{r c^{(\alpha)}}{d_m} > \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} \big| \lambda^{(\beta)}_{(\alpha)(\beta')} \big| c^{(\alpha)}.$$

It follows that

$$\binom{m}{(\beta')} c^{(\beta')} + \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \lambda^{(\beta)}_{(\alpha)(\beta')} \\ \geq \binom{m}{(\beta')} c^{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} \big| \lambda^{(\beta)}_{(\alpha)(\beta')} \big| > 0.$$

Thus we have shown that the n-tuple $(c) = (c_1, \dots, c_n)$ of positive integers can be chosen so that the coefficient of $\widetilde{X}^{(\beta')}$ in equation (3.11) is a positive rational number, and the coefficients of $\widetilde{X}^{(\gamma)}$ in the sum $\sum_{(\gamma)\succ(\beta')}$ are rational numbers. Obviously,

$$g_{(\beta')} - \sum_{(\alpha)\preceq(\beta)} \binom{m}{(\alpha)} c^{(\alpha)} f^{(\beta)}_{(\alpha)}$$

is a regular polynomial, so by dividing both sides of equation (3.11) by the coefficient of $\widetilde{X}^{(\beta')}$ we can write

$$\widetilde{X}^{(\beta')} = f^{(\beta')}_{(\beta')} + \sum_{(\gamma)\succ(\beta')} \lambda^{(\beta')}_{(\beta')(\gamma)} \widetilde{X}^{(\gamma)}, \tag{3.13}$$

where $f^{(\beta')}_{(\beta')}$ is regular and the constants $\lambda^{(\beta')}_{(\beta')(\gamma)}$ are rational. For $(\alpha) \prec (\beta')$, *i.e.*, $\alpha \preceq (\beta)$ equation (3.9) can be written as

$$\begin{aligned}
\widetilde{X}^{(\alpha)} &= f^{(\beta)}_{(\alpha)} + \lambda^{(\beta)}_{(\alpha)(\beta')}\widetilde{X}^{(\beta')} + \sum_{(\gamma)\succ(\beta')} \lambda^{(\beta)}_{(\alpha)(\gamma)}\widetilde{X}^{(\gamma)} \\
&= \big(f^{(\beta)}_{(\alpha)} + \lambda^{(\beta)}_{(\alpha)(\beta')} f^{(\beta')}_{(\beta')}\big) \\
&\qquad + \sum_{(\gamma)\succ(\beta')} \big(\lambda^{(\beta)}_{(\alpha)(\beta')}\lambda^{(\beta')}_{(\beta')(\gamma)} + \lambda^{(\beta)}_{(\alpha)(\gamma)}\big)\widetilde{X}^{(\gamma)}.
\end{aligned}\tag{3.14}$$

Set $f^{(\beta')}_{(\alpha)} = f^{(\beta)}_{(\alpha)} + \lambda^{(\beta)}_{(\alpha)(\beta')} f^{(\beta')}_{(\beta')}$ and $\lambda^{(\beta')}_{(\alpha)(\gamma)} = \lambda^{(\beta')}_{(\alpha)(\beta')}\lambda^{(\beta')}_{(\beta')(\gamma)} + \lambda^{(\beta)}_{(\alpha)(\gamma)}$, then it follows from equations (3.13) and (3.14) that for all $(\alpha) \preceq (\beta')$,

$$\widetilde{X}^{(\alpha)} = f^{(\beta')}_{(\alpha)} + \sum_{(\gamma)\succ(\beta')} \lambda^{(\beta')}_{(\alpha)(\gamma)}\widetilde{X}^{(\gamma)},$$

where $f^{(\beta')}_{(\alpha)}$ is obviously regular and the coefficients $\lambda^{(\beta')}_{(\alpha)(\gamma)}$ are obviously rational. Hence we have completed the induction. Now for the proof of the lemma in the statement (3.9), choose $(\beta) = (\beta)_{\max} = (0, \dots, 0, m)$ to be the last element of Λ_m. Then (3.9) reads:

"For every $(\alpha) \preceq (\beta)_{\max}$, $\widetilde{X}^{(\alpha)} = f^{(\beta)_{\max}}_{(\alpha)}$, where $f^{(\beta)_{\max}}_{(\alpha)}$ is a regular element of $\mathcal{A}^m$". This is exactly what the lemma affirms. ∎

From the fact that $\Phi_m : \mathcal{P}^m \to \mathcal{A}^m$ is a monomorphism it follows that the system $\{\widetilde{X}^{(\alpha)}, (\alpha) \in \Lambda_m\}$ is linearly independent. Therefore if $\mathcal{R}^m$ denotes the subspace of $\mathcal{R}$ of all homogeneous non-commutative regular polynomials of degree m in the indeterminates $X_1, \dots, X_n$, then Lemma 3.5 and equation (3.7) imply that *the system* $\{\widetilde{X}^{\alpha}, (\alpha) \in \Lambda_m\}$ *forms a basis for* $\mathcal{R}^m$. It follows immediately from the discussion preceding Lemma 3.4 that $\Phi_m : \mathcal{P}^m \to \mathcal{R}^m$ is an isomorphism and hence, Φ *is an isomorphism of* $\mathcal{P}$ *onto* $\mathcal{R}$ (clearly from equation (3.7) $\mathcal{R}^m \subset \Phi_m(\mathcal{P}^m)$, Lemma 3.5 shows that $\mathcal{R}^m = \Phi_m(\mathcal{P}^m)$). Now each $\widetilde{X}^{(\alpha)}$, being a regular homogeneous polynomial of degree m, is therefore a linear combination of polynomials of the form $(c_1X_1 + \cdots + c_nX_n)^m$. Let S be the set of such polynomials, then S is a finite set of vectors spanning the vector space $\mathcal{R}^m$. From a general fact in linear algebra (see, for example, [Hoffman & Kunze 1971, Corollary 2, p. 44]) we can deduce the following.

Corollary 3.6. — (i) *The vector space* $\mathcal{R}^m$ *admits a basis consisting of vectors of the form*

$$f_i = (c^i_1X_1 + c^i_2X_2 + \cdots + c^i_nX_n)^m, \quad 1 \le i \le d_m,\ c^i_j \in \mathbb{K}. \tag{3.15}$$

(ii) *The same conclusion holds with* $\mathcal{P}^m$ *replacing* $\mathcal{R}^m$ *and* x_j, $1 \leq j \leq n$, *replacing* X_j.

Proof of Theorem 3.3. — Clearly since $\mathcal{A}$ is a graded algebra, it suffices to prove the theorem for $\mathcal{A}^m, m \geq 0$.

(i) If f is a regular element of $\mathcal{A}^m$ then since $\{\widetilde{X}^{(\alpha)},\ (\alpha) \in \Lambda_m\}$ forms a basis for $\mathcal{R}^m$,

$$f = \sum_{(\alpha)\in\Lambda_m} \lambda_{(\alpha)} \widetilde{X}^{(\alpha)}, \quad \lambda_{(\alpha)} \in \mathbb{K}.$$

Thus if $\lambda_{(\alpha)} \neq 0$ then since $\widetilde{X}^{(\alpha)}$ contains all distinct monomials equipollent to $X^{(\alpha)}$, therefore f contains all monomials equipollent to $X^{(\alpha)}$ with the same coefficient

$$\frac{1}{\binom{m}{(\alpha)}} \lambda_{(\alpha)}.$$

Conversely, if f is a polynomial in $\mathcal{A}^m$ which contains all monomials equipollent to a fixed monomial, which we may assume without loss of generality to be $X^{(\alpha)}$, with the same coefficient, then f must contain $\mu_{(\alpha)}\widetilde{X}^{(\alpha)}$ with $\mu_{(\alpha)}$ a non-zero constant. Hence f is of the form

$$\sum_{(\alpha)\in\Lambda_m} \lambda_{(\alpha)} \widetilde{X}^{(\alpha)}, \quad \lambda_{(\alpha)} \in \mathbb{K},$$

and therefore is a regular polynomial.

(ii) First observe that if $X_{i_1} \cdots X_{i_m}$, $1 \leq i_j \leq n$, $1 \leq j \leq m$, is a monomial in $\mathcal{A}^m$, then it is equipollent to a unique monomial $X^{(\alpha)} = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ for some $(\alpha) \in \Lambda_m$. Then from the definition of $\widetilde{X}^{(\alpha)} \equiv \mathrm{Sym}(X^\alpha)$ and part (i) of this theorem, $\widetilde{X}^{(\alpha)}$ is the unique regular polynomial of $\mathcal{A}^m$ that is equipollent to $X^{(\alpha)}$, and hence to $X_{i_1\ldots i_m}$ (for example, $X_2X_1X_2$ is equipollent to $X^{(1,2,0,\ldots,0)} = X_1X_2^2$, which from equation (3.4) is equipollent to the regular polynomial $\widetilde{X}^{(1,2,0,\ldots,0)} = \frac{1}{3}(X_1X_2^2 + X_2X_1X_2 + X_2^2X_1) = \frac{1}{6}[(X_1+X_2)^3 + (X_1-X_2)^3 - 2X_1^3])$.

Let

$$p = \sum_{i_1,\ldots,i_m} \lambda_{i_1\cdots i_m} X_{i_1} \cdots X_{i_m},$$

where the sum is over all distinct non-commutative homogeneous monomials of degree m and the coefficients $\lambda_{i_1\ldots i_m}$ are uniquely determined. Then since each $X_{i_1}\cdots X_{i_m}$ is equipollent to a unique regular polynomial $\widetilde{X}^{(\alpha)}$ for some $(\alpha)\in\Lambda_m$, p is equipollent to the unique regular polynomial

$$\sum_{(\alpha)\in\Lambda_m}\mu_{(\alpha)}\widetilde{X}^{(\alpha)},$$

where $\mu_{(\alpha)}$ is the sum of all $\lambda_{i_1\ldots i_m}$ for which $X_{i_1\ldots i_m}$ is equipollent to $\widetilde{X}^{(\alpha)}$. ∎

Remark 3.7. — It follows from Corollary 3.6 (i) and equation (3.7) that a polynomial p of $\mathcal{A}$ is regular if and only if $\mathrm{Sym}(p)=p$. Now define (as in [Godement 1982, 5.6.1]) a polynomial

$$p=\sum_{i_1\ldots i_m}\lambda_{i_1\ldots i_m}X_{i_1}\cdots X_{i_m},\quad 1\le i_j\le n,\ 1\le j\le m,$$

to be *symmetric* if all its coefficients $\lambda_{i_1\ldots i_m}$ are symmetric, *i.e.*, for all $\sigma\in\Sigma_m$, $\lambda_{i_{\sigma(1)}\ldots i_{\sigma(m)}}=\lambda_{i_1\ldots i_m}$. Then Theorem 3.3 (i) implies that a polynomial in $\mathcal{A}$ is regular if and only it is symmetric, since a monomial is equipollent to $X_{i_1}\cdots X_{i_m}$, if and only if it is of the form $X_{i_{\sigma(1)}}\cdots X_{i_{\sigma(m)}}$ for some $\sigma\in\Sigma_m$. Thus in this context regular is synonymous with symmetric, and this is probably what [Bourbaki 1969] must have had in mind when he affirmed that Poincaré gave a proof of algebraic nature that the associative algebra generated by the X_i, $1\le i\le n$, has as basis certain symmetric functions in X_i. In fact, in [Godement 1982, 5.6.1], for example, this fact is used to define the vector space isomorphism $\beta\colon S(\mathfrak{g})\to\mathcal{U}(\mathfrak{g})$ of the symmetric algebra of polynomial functions on the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$ onto the universal enveloping algebra $\mathfrak{g}$. Obviously, $\mathcal{P}$ is isomorphic to $S(\mathfrak{g})$ and as we shall see $\mathcal{R}$ is isomorphic to $\mathcal{U}(\mathfrak{g})$; thus the map β is basically Φ.

Now let us return to [Poincaré 1900]. Let $X_1,\ldots,X_r$ be elementary operators (*i.e.*, infinitesimal transformations) which form a basis for a Lie algebra $\mathcal{L}$. Define the Lie bracket as

$$[X,Y]=XY-YX;\quad X,Y\in\mathcal{L}. \tag{3.16}$$

Two polynomials in $\mathcal{A}$ are said to be *equivalent* if one can be reduced to the other in taking into account relation (3.16).

For example, the product $P(XY - YX - [X, Y])Q$ as defined in equation (3.1) (where the first and the last factors P and Q are two arbitrary monomials in $\mathcal{A}$) is equivalent to zero, and obviously so are linear combinations of products of that form (*i.e.*, P and Q may be taken to be polynomials). Products of the form (3.1) are called *trinomial products.*

The difference of two monomials which differ only by the order of two consecutive factors is equivalent to a polynomial of lesser degree. Indeed, let X and Y be those two consecutive factors. Then our monomials are written as

$$PXYQ \quad \text{and} \quad PYXQ,$$

P and Q being two arbitrary monomials, and their difference

$$P(XY - YX)Q$$

is equivalent to $P[X, Y]Q$, which has degree one less, since $[X, Y]$ is of first degree, while XY and YX are of second degree.

Now let M and M' be two arbitrary equipollent monomials; that is, they only differ by the order of their factors. One can find a sequence of monomials

$$M, M_1, M_2, \dots, M_p, M',$$

in which the first and the last terms are the given monomials and any term in the sequence differs only from the preceding by the order of two consecutive factors. The difference $M - M'$, which is the sum of the differences $M - M_1, M_1 - M_2, \dots, M_p - M'$, is therefore again equivalent to a polynomial of lesser degree.

More generally, the difference of two equipollent polynomials is equivalent to a polynomial of lesser degree. We now claim the following.

Theorem 3.8. — *In the algebra $\mathcal{A}$ any arbitrary polynomial is equivalent to a unique regular polynomial.*

Proof. — First let us show that this equivalence relation[5] is *additive*, *i.e.*, if $p \sim p'$ and $q \sim q'$ then $p + q \sim p' + q'$. This is obvious since this is equivalent to $p - p' \sim 0$ and $q - q' \sim 0$, and hence $(p + q) - (p' + q') = (p - p') + (q - q') \sim 0 + 0 = 0$.

[5] Poincaré used the symbol $=$ to denote this equivalence relation. To avoid confusion we adopt the more conventional symbol $\sim$.

Now let P_n be an arbitrary polynomial of degree n; then by Theorem 3.3 (ii) P_n is equipollent to a unique regular polynomial P'_n of the same degree n, and by the remark preceding this theorem, $P_n - P'_n$ is equivalent to a polynomial P_{n-1} of lesser degree (which we may assume, without loss of generality, of degree $n-1$). Therefore, $P_n \sim P'_n + P_{n-1}$, and P_{n-1} is in turn equipollent to a regular polynomial P'_{n-1}, and hence

$$\begin{aligned} P_n \sim P'_n + P_{n-1} &= P'_n + P'_{n-1} + (P_{n-1} - P'_{n-1}) \\ &\sim P'_n + P'_{n-1} + (P_{n-2}), \ldots, \end{aligned}$$

and so on; one finally arrives to a polynomial of degree zero which is obviously regular. Thus one can conclude that

$$P_n \sim P'_n + P'_{n-1} + P'_{n-2} + \cdots,$$

where the second member is a regular polynomial. We therefore have a means to reduce any polynomial to a regular polynomial by making use of the relations (3.16). It remains to find out if this reduction can be done uniquely.

Since both the equivalence relation $\sim$ and the notion of regular polynomials are additive, this problem is equivalent to the following:

Can a non-identically zero regular polynomial be equivalent to zero? Or equivalently, *can we find a sum of trinomial products of the form* (3.1) *which is a non-identically zero regular polynomial?* All sums of such products are indeed equivalent to zero and vice-versa. If we define a *regular sum* to be a sum of trinomial products of the form (3.1) which is also a regular polynomial then the answer (*negative*) to this question (and hence to the question above regarding uniqueness) can be stated as follows:

Lemma 3.9. — *Every regular sum is identically zero.*

Proof of the lemma. — The degree of a trinomial product (3.1) is clearly $d^0(P) + d^0(Q) + 2$. Thus we call the degree of a sum S of trinomial products the highest of all the degrees of the products in S *even though as we shall see when S is a regular sum the terms of highest degree in these different products mutually cancel each other.*

The trinomial product (3.1) can be considered as the sum of two products, the *binomial product*

$$P(XY - YX)Q, \tag{3.17}$$

where we call $PXYQ$ the *positive monomial* and $-PYXQ$ the *negative monomial*; and the product

$$-P[X,Y]Q, \tag{3.18}$$

which we call the *complementary product.*

Thus if S is an arbitrary sum of trinomial products of degree p and of degree $< p$ then we can write

$$S = S_p - T_p + S_{p-1} - T_{p-1} + \cdots + S_k - T_k + \cdots + S_2 - T_2, \tag{3.19}$$

where S_k, $2 \leq k \leq p$, is a sum of homogeneous binomial products of degree k, whereas $-T_k$ is the sum of the corresponding complementary products. First observe that if S is a regular sum then every homogeneous component of S is also regular since regularity is graded; in particular *S_p is regular.* Since S_p is a sum of binomial products of degree p of the form $PXYQ - PYXQ$ and since equipollence is an additive equivalence relation it follows immediately that S_p is equipollent to zero. But zero is a regular polynomial and Theorem 3.3 (ii) implies that two regular polynomials cannot be equipollent without being identical, and therefore S_p must be identically zero.

Remark 3.10. — From the discussion above it follows that *the degree of a regular sum as we defined it is actually at least one more than the classical degree of a polynomial in $\mathcal{A}$.*

Thus in particular when S is a regular sum of degree 3 (actual degree 2) then

$$S = S_3 - T_3 + S_2 - T_2. \tag{3.20}$$

Since S_3 is homogeneous of degree 3, a typical binomial product of S_3 must be of the form

$$(XY - YX)Z \quad \text{or} \quad Z(XY - YX).$$

Since S_3 is regular, hence symmetric, Theorem 3.3 (i) implies that if the binomial product $(XY - YX)Z = XYZ - YXZ$ occurs in S_3, all six monomials equipollent to XYZ (resp. $-YXZ$) must occur in S_3 with the same coefficient.

Thus S_3 must be a sum of terms of the form

$$\sum(XY - YX)Z - \sum Z(XY - YX), \tag{3.21}$$

where the sign $\sum$ means that one must sum over the term which is explicitly expressed under the sign and the other two terms obtained by cyclically permuting the three letters X, Y, Z. Note that one can verify directly from equation (3.21) that S_3 is identically zero. It follows from equation (3.21) that the sum of the complementary products $-T_3$ contains terms of the form

$$-\left(\sum[X,Y]Z - \sum Z[X,Y]\right) \tag{3.22}$$

Since $S_2 - T_3$ is homogeneous of degree two and S is regular, it follows that $S_2 - T_3$ is also regular, and hence symmetric. Since it contains $-[X,Y]Z + Z[X,Y]$, Theorem 3.3 (i) again implies that it must contain permutations of these terms with the same coefficients, *i.e.*,

$$\begin{aligned}&-[X,Y]Z - [Y,X]Z - [Z,Y]X - [X,Z]Y - [Z,X]Y - [Y,Z]X\\ &\qquad + Z[X,Y] + X[Z,Y] + Y[Z,X] + Z[Y,X] + Y[X,Z] + X[Y,Z],\end{aligned}$$

which can be regrouped in the following form:

$$-\left(\sum[X,Y]Z - \sum Z[X,Y]\right) + \left(\sum[X,Y]Z - \sum Z[X,Y]\right), \tag{3.23}$$

using the fact that the bracket product $[\ ,\ ]$ is anti-symmetric. From equations (3.22) and (3.23) it follows that

$$S_2 = \sum[X,Y]Z - \sum Z[X,Y] \tag{3.24}$$

and $S_2 - T_3 = 0$. Since S_2 is a sum of terms of the form $WZ - ZW$ with $W = [X,Y]$ it follows that the complementary polynomial is of degree one and is a sum of terms of the form $[W,Z]$. Thus we have

$$T_2 = \sum\big[[X,Y],Z\big],$$

where $\sum$ has the same meaning as above. Thus T_2 is a sum of terms of the form

$$\big[[X,Y],Z\big] + \big[[Y,Z],X\big] + \big[[Z,X],Y\big]. \tag{3.25}$$

It follows from equation (2.1) that T_2 is a polynomial of first degree which is obviously symmetric, and hence regular. Therefore, if T_2 is not identically zero, the sum S would be a regular polynomial which is not identically zero.

Therefore, in order that a polynomial can be reduced in a unique fashion to a regular polynomial, it is necessary that the expression (3.25) is identically zero. But one recognizes there the Jacobi identities which play such an important role in Lie theory. It remains to show that this condition is sufficient.

At this juncture, it is important to make the following remark:

It follows from Remark 3.10 *that we have actually proved that every polynomial of degree* 0, 1, *or* 2 *is equivalent to a unique regular polynomial of degree* 0, 1, *or* 2, *respectively.*

Now by induction suppose that the lemma has been proven for regular sums of degree $1, 2, \ldots, p-1$ and propose to extend it to regular sums of degree p.

Thus, let $S = S_p - T_p + S_{p-1} - T_{p-1} + \cdots$ be a sum of trinomial products. Let us call $S_p - T_p$ the *head* (or leading terms) of the sum S. We say that a sum of trinomial products form a *chain* if the negative monomial of each product is equal to, and of opposite sign of, the positive monomial of the product that follows. The positive monomial of the first product and the negative monomial of the last one are called extreme monomials of the chain. Examples of chains:

$$\begin{aligned} C_1\colon\ & XZ(XY)W - XZ(YX)W - XZ[X,Y]W + X(ZY)XW \\ & \qquad - X(YZ)XW - X[Z,Y]XW + XY(ZX)W \\ & \qquad\quad - XY(XZ)W - XY[Z,X]W, \end{aligned}$$

$$\begin{aligned} C_2\colon\ & XZ(XY)W - XZ(YX)W - XZ[X,Y]W + XZY(XW) \\ & \qquad - XZY(WX) - XZY[X,W] + X(ZY)WX - X(YZ)WX \\ & \qquad\quad - X[Z,Y]WX + XYZ(WX) - XYZ(XW) - XYZ[W,X] \\ & \qquad\qquad + XY(ZX)W - XY(XZ)W - XY[Z,X]W. \end{aligned}$$

Remark 3.11. — *It results from the definition that all positive monomials* (*and hence, all negative monomials*) *of the same chain can only differ by the order of their factors.*

A chain is said to be *closed* if its extreme monomials are equal and of opposite sign. If $S_p - T_p$ is a closed chain of trinomial products it is clear that S_p is identically zero since the positive and negative monomials cancel each other two by two.

We have seen that if S is a regular sum, S_p is identically zero. It follows therefore that *the head of a regular sum must always consist of one or more closed chains.*

If two chains have the same extreme monomials, then their difference is a closed chain. For example,

$$\begin{aligned} C_1 - C_2\colon\ & XYZ(XW) - XYZ(WX) - XYZ[X,W] \\ & + X(YZ)WX - X(ZY)WX - X[Y,Z]WX \\ & + XZY(WX) - XZY(XW) - XZY[W,X] \\ & + X(ZY)XW - X(YZ)XW - X[Z,Y]XW. \end{aligned}$$

We shall use this remark to show that a closed chain can always be decomposed in many ways into two or more closed chains. An arbitrary closed chain can be in many ways regarded as the difference of two chains C and C' having the same extreme monomials. Let C'' be a third chain having the same extreme monomials, then the chain $C - C'$ is then decomposed into two other closed chains $C - C''$ and $C'' - C'$.

Now remark first that if a regular sum of degree p is identically zero, it must be the same for all regular sums of degree p *which have the same head.* The difference of these two sums will be indeed a regular sum of degree $p-1$ which will be identically zero according to our inductive hypothesis. *Therefore it suffices for us to form all closed chains of degree* p *and prove that each one of them can be considered as the head of an identically zero regular sum.* Indeed, each regular sum S of order p has as head one or more of those closed chains. Let S' be one of those closed chains, then if we show that there exists an identically zero regular sum having S' as head, it follows immediately from the remark above that S must be identically zero. Thus by induction we suppose this statement holds for all closed chains of degree $\leq p-1$ and we will show that it is true for all closed chains of degree p.

To establish this assertion, we are going to decompose the closed chain in question into several closed chains. It is clear that it suffices to prove the proposition for each component.

A chain is called *simple of the first kind* if the first factor of all of its monomials either positive or negative is everywhere the same. A chain is called *simple of the second kind* if the last factor of all of its monomials either positive or negative is everywhere the same. Moreover, a simple chain can be either closed or *open* (not closed).

Since p is larger than three, it is clear that every closed chain can be regarded as the sum of a certain number of simple chains, alternatively of the first and second kinds or vice-versa.

Thus let S be a closed chain, $C_1, C_2, \dots, C_n$ be simple chains of the first kind, $C'_1, C'_2, \dots, C'_n$ be simple chains of the second kind, such that

$$S = C_1 + C'_1 + C_2 + C'_2 + \cdots + C_n + C'_n,$$

the extreme negative monomial of each chain being, of course, equal and of opposite sign to the extreme positive monomial of the next chain, and the extreme negative monomial of C'_n being equal and of opposite sign to the extreme positive monomial of C_1. Note that, *a priori*, C_1 or C'_n, can be the zero chain, for example, if S starts with $XYQ - YXQ - \cdots$, where $d^0(Q) > 1$; then $C_1 = 0$, but then we can consider C_1 as the zero simple closed chain of the form $XYQ - \cdots - XYQ$, and similarly for C'_n.

Let X be the first factor of all the monomials of C_1, Z the last factor of all the monomials of C'_1, Y the first factor of all the monomials of C_2, and T the last factor of all the monomials of C'_2 (we do not exclude the case where two of the operators X, Y, Z, T are identical).

Let C'' be a simple chain of the second kind having its extreme positive monomial equal and of opposite sign to the extreme negative monomial of C'_2, and in which all monomials have the last factor equal to T, and moreover, the extreme negative monomial has X as its first factor.

Let C''' be a simple chain of the first kind such that all monomials in it have X as the first factor, and moreover, the extreme monomials are respectively equal and of opposite signs to the extreme negative monomial of C'' and to the extreme positive monomial of C_1.

Schematically we have the following diagram:

$$\begin{aligned}(3.26)\qquad &\underbrace{X\square\cdots - X\square Z}_{C_1} + \underbrace{X\square Z\cdots - Y\square Z}_{C'_1} + \underbrace{Y\square Z\cdots - Y\square T}_{C_2}\\ &\quad + \underbrace{Y\square T\cdots - \square T}_{C'_2} + \underbrace{\square T\cdots - X\square T}_{C''} + \underbrace{X\square T\cdots - X\square}_{C'''},\end{aligned}$$

where each box $\square$ represents certain unspecified monomial which does not have any effect on our discussion.

Thus the closed chain S is decomposed into a sum of two closed chains as $S' + S''$, where

$$S' = (C''' + C_1) + C_1' + C_2 + (C_2' + C''),$$
$$S'' = -C'' + C_3 + \cdots + C_n - C'''.$$

The closed chain S' contains only four simple chains, since $(C''' + C_1)$ and $(C_2' + C^{''})$ are simple chains; S'' contains two simple chains less than S. Continuing this scheme we end up decomposing S into closed components which consist of only four simple chains. *Thus it suffices to consider the case of closed chains S formed by four simple chains as, for example, the form S'.*

Therefore, it follows from (3.26) that the extreme positive monomials of the four chains that form S' have respectively for first and last factors:

for $C''' + C_1$	X and T,
for C_1'	X and Z,
for C_2	Y and Z,
for $C_2' + C''$	Y and T.

Let M_1, M_1', M_2, M_2' denote these four monomials.

From Remark 3.11 it follows that all these monomials are equipollent to each other and are equipollent to a certain monomial which we will call $XYPZT$. Set

$$Q_1 = XYPZT, \qquad Q_1' = XYPTZ,$$
$$Q_2 = YXPTZ, \qquad Q_2' = YXPZT.$$

We are going to construct a series of simple chains which will constitute a decomposition of S' as follows:

Name of the chain	Extreme positive monomial	Extreme negative monomial
$C''' + C_1$	$M_1 = X\Box T$	$-M_1' = -X\Box Z$
C_1'	M_1'	$-M_2 = -Y\Box Z$
C_2	M_2	$-M_2' = -Y\Box T$
$C_2' + C''$	M_2'	$-M_1$
D_1	M_1	$-Q_1$
D_1'	M_1'	$-Q_1'$

D_2	M_2	$-Q_2$
D_2'	M_2'	$-Q_2'$
E_1	Q_1	$-Q_1'$
E_1'	Q_1'	$-Q_2$
E_2	Q_2	$-Q_2'$
E_2'	Q_2'	$-Q_1$

We can suppose that every monomial of the chain D_1 has as first factor X and last factor T; thus D_1 is both a simple chain of the first and second kind, and similarly for other D and D' chains. Furthermore, we can suppose that the E and E' chains are reduced to a single trinomial product, for example,

$$E_1 = XYP\big(ZT - TZ - [Z,T]\big).$$

The closed chain $S' = (C''' + C_1) + C_1' + C_2 + (C_2' + C'')$ can be decomposed into five closed chains as follows:

$$U_1 = \underbrace{(M_1 \cdots - M_1')}_{C'''+C_1} + \underbrace{M_1' \cdots - Q_1'}_{D_1'} + \underbrace{XYP(TZ - ZT - [T,Z])}_{-E_1} + \underbrace{Q_1 \cdots - M_1}_{-D_1},$$

$$U_1' = \underbrace{M_1 \cdots - M_2}_{C_1'} + \underbrace{M_2 \cdots - Q_2}_{D_2} + \underbrace{(YX - XY - [Y,X])PTZ}_{-E_1'} + \underbrace{Q_1' \cdots - M_1'}_{-D_1'},$$

$$U_2 = \underbrace{M_2 \cdots - M_2'}_{C_2} + \underbrace{M_2' \cdots - Q_2'}_{D_2'} + \underbrace{YXP(ZT - TZ - [Z,T])}_{-E_2} + \underbrace{Q_2 \cdots - M_2}_{-D_2},$$

$$U_2' = \underbrace{M_2' \cdots - M_1}_{C_2+C''} + \underbrace{M_1 \cdots - Q_1}_{D_1} + \underbrace{(XY - YX - [X,Y])PZT}_{-E_2'} + \underbrace{Q_2' \cdots - M_2'}_{-D_2'},$$

$$V = \underbrace{XYP(ZT - TZ - [Z,T])}_{E_1} + \underbrace{(XY - YX - [X,Y])PTZ}_{E_1'} + \underbrace{(YXP(TZ - ZT - [T,Z])}_{E_2} + \underbrace{(YX - XY - [Y,X])PZT}_{E_2'}.$$

Clearly, $U_1 + U_1' + U_2 + U_2' + V = (C''' + C_1) + C_1' + C_2 + (C_2' + C'') = S$. We must show that each one of the five closed chains above is the head of an identically zero regular sum. The first four chains are of the form

$$U_1 = XH_1, \quad U_1' = H_1'Z, \quad U_2 = YH_2, \quad U_2' = H_2'T,$$

where each chain H_1, H_1', H_2, H_2' is a closed chain of degree $p-1$; therefore by induction, each is the head of an identically zero regular sum. It follows that U_1, U_1', U_2, and U_2' are identically zero, and therefore each of them can be considered as the head of an identically zero regular sum of degree p.

Finally for V, it is the head of the sums

$$\begin{aligned}&XYP(ZT-TZ-[Z,T])+(XY-YX-[X,Y])PTZ\\ &\quad -YXP(-TZ+ZT-[Z,T])-(-YX+XY-[X,Y])PZT\\ &\qquad -[X,Y]P(ZT-TZ-[Z,T])-(XY-YX-[X,Y])P[T,Z],\end{aligned}$$

which can be expanded and rearranged as

$$\begin{aligned}&XYPZT-XYPTZ+XYPTZ-YXPTZ+YXPTZ\\ &\quad -YXPZT+YXPZT-XYPZT-XYP[Z,T]-[X,Y]PTZ\\ &\qquad +YXP[Z,T]+[X,Y]PZT-[X,Y]PZT+[X,Y]PTZ\\ &\qquad\ -XYP[T,Z]+YXP[T,Z]+[X,Y]P[Z,T]+[X,Y]P[T,Z],\end{aligned}$$

which is identically zero. Since 0 is a regular sum, it follows that V is the head of an identically zero regular sum of degree p.

Note that our analysis remains unchanged when two or more of the operators X, Y, Z, T are identical. For example, when $X = Y$, then $E_1' = E_2' = 0$, and we set $Q_1 = Q_2' = X(XP)ZT = XP'ZT$, $Q_2 = Q_1' = X(XP)TZ = XP'TZ$. The definition of the various chains remains the same, and we can immediately verify that V is identically zero. Finally, in order that this proof is valid, p must be greater than three since the chain V must have at least four factors. But this was the assumption in our inductive hypothesis. Thus the proof of Lemma 3.9, and hence of Theorem 3.8, is achieved. ∎

Corollary 3.12 (The so-called Birkhoff-Witt Theorem). — *Let $\mathcal{U}(\mathcal{L})$ denote the universal enveloping algebra of a Lie algebra over a (commutative) field of characteristic zero. If $\{X_1, \ldots, X_n\}$ is a basis of $\mathcal{L}$ and if $(\alpha) = (\alpha_1, \ldots, \alpha_n)$ denotes an n-tuple of integers ≥ 0, set*

$$X^{(\alpha)} = X_1^{\alpha_1} \cdots X_n^{\alpha_n}, \quad \widetilde{X}^{(\alpha)} = \mathrm{Sym}(X^{\alpha}), \quad \big|(\alpha)\big| = \alpha_1 + \cdots + \alpha_n.$$

Then the set $\{\widetilde{X}^{(\alpha)}\}_{(\alpha)}$, for all (α) such that $|(\alpha)| \geq 0$, forms a vector space basis for $\mathcal{U}(\mathcal{L})$. Moreover, any set of elements of $\mathcal{U}(\mathcal{L})$ of the form

$\{X_{i_1} \cdots X_{i_m},\ 1 \leq i_j \leq n,\ 1 \leq j \leq m,\ m \geq 0\}$, *where each* $X_{i_1} \cdots X_{i_m}$ *is a representative of an equipollence class* $X^{(\alpha)}$ *for all distinct* (α) *such that* $|(\alpha)| \geq 0$*, is a basis of* $\mathcal{U}(\mathcal{L})$*; in particular, the set of ordered monomials* $\{X^{(\alpha)},\ |(\alpha)| \geq 0\}$ *forms a basis for* $\mathcal{U}(\mathcal{L})$.

Proof. — From our discussion pertaining to Poincaré's discovery of the universal enveloping algebra of a Lie algebra, it follows that the quotient algebra of the polynomial algebra $\mathcal{A}$ modulo the equivalence relation $\sim$ can be regarded as the universal enveloping algebra of the Lie algebra $\mathcal{L}$ generated by $X_1, \ldots, X_n$.

Define a map from $\mathcal{A}$ to $\mathcal{R}$, the vector space of all regular polynomials in $\mathcal{A}$, by assigning to each polynomial A in $\mathcal{A}$ the unique regular polynomial $\bar{A}$ equivalent to A as defined by Theorem 3.8. From the proof of Theorem 3.8 it follows that this map is linear, and that it is surjective since the unique regular polynomial equivalent to a given regular polynomial is itself. The kernel of this homomorphism is, by the definition of the equivalence relation $\sim$, the vector space spanned by all trinomials of the form $P(XY - YX - [X,Y])Q$ for arbitrary P and Q in $\mathcal{A}$. Let $\mathcal{I}$ denote this kernel, then obviously $\mathcal{I}$ is a two-sided ideal of $\mathcal{A}$. It follows from the first isomorphism theorem that $\mathcal{A}/\mathcal{I}$ is isomorphic to $\mathcal{R}$ as vector spaces.

From the remark following Lemma 3.5, it follows that the set $\{\widetilde{X}^{(\alpha)}\}_{(\alpha)}$, $|(\alpha)| \geq 0$, forms a basis for $\mathcal{R}$, and hence a basis for $\mathcal{U}(\mathcal{L}) \cong \mathcal{A}/\mathcal{I}$ via the isomorphism above. Note that we have shown following Lemma 3.5 that $\mathcal{P}$ is isomorphic to $\mathcal{R}$ via the isomorphism Φ, therefore $\mathcal{P}$ is isomorphic to $\mathcal{U}(\mathcal{L})$. For the second part of the theorem, we remark that it follows from Theorem 3.3 that each $X_{i_1} \cdots X_{i_m}$ is equipollent to a unique regular polynomial $\widetilde{X}^{(\alpha)}$ for some $(\alpha) \in \Lambda_m$. Thus it suffices to consider the set $\{X^{(\alpha)}, |(\alpha)| \geq 0\}$. We also remark that it suffices to show that the set $\{X^{(\alpha)}, (\alpha) \in \Lambda_m\}$ is linearly independent in $\mathcal{U}(\mathcal{L})$ for all $m \geq 0$, since $\mathcal{U}(\mathcal{L})$ is a filtered algebra. From the proof of Theorem 3.8 it follows that each $X^{(\alpha)}, (\alpha) \in \Lambda_m$, is equivalent to a unique regular polynomial of the form $\widetilde{X}^{(\alpha)} + P_{(\alpha)}$, where $P_{(\alpha)}$ is a regular polynomial of degree $< m$. Thus if for some scalars $\lambda_{(\alpha)} \in \mathbb{K}$ such that $\sum_{(\alpha)\in\Lambda_m} \lambda_{(\alpha)} X^{(\alpha)}$ is zero in $\mathcal{U}(\mathcal{L})$ (*i.e.*, equivalent to 0), then since the equivalence relation $\sim$ is linear it follows that the regular polynomial $\sum_{(\alpha)\in\Lambda_m} \lambda_\alpha(\widetilde{X}^{(\alpha)} + P_\alpha)$ is equivalent to 0. It follows from Theorem 3.8 (or more precisely Lemma 3.9) that

$$\sum_{(\alpha)\in\Lambda_m} \lambda_{(\alpha)} \widetilde{X}^{(\alpha)} + \sum_{(\alpha)\in\Lambda_m} \lambda_{(\alpha)} P_{(\alpha)}$$

must be identically zero. Since $d^0(\widetilde{X}^{(\alpha)}) = m$ and $d^0(P_{(\alpha)}) < m$ for all $(\alpha) \in \Lambda_m$, it follows that

$$\sum_{(\alpha)\in\Lambda_m} \lambda_{(\alpha)} \widetilde{X}^{(\alpha)} = 0;$$

and hence by the first part of the proof of the theorem, it follows that $\lambda_{(\alpha)} = 0$ for all $(\alpha) \in \Lambda_m$. This completes the proof of the theorem. ∎

Remark 3.13. — We note that throughout this section the basis for the Lie algebra $\mathcal{L}$ can be the *infinite* set $\{X_1, \ldots, X_n, \ldots\}$. The tensor algebra $\mathcal{T}$ remains isomorphic to the non-commutative algebra $\mathcal{A}$ of polynomials in infinitely many variables $X_1, \ldots, X_n, \ldots$ (see, *e.g.*, [Schwartz 1998 (1975), Prop. (2.4), p. 40], and every argument remains the same. As a special case, let $\mathcal{L}_n$ (resp. $\mathcal{A}_n$) denote the Lie algebra (resp. the non-commutative polynomial algebra) generated by $X_1, \ldots, X_n$. Let $\mathcal{L}$ (resp. $\mathcal{A}$) denote the *inductive limit* of $\mathcal{L}_n$ (resp. $\mathcal{A}_n$); then all theorems in this section can be easily generalized. For example, let

$$X_{ij} = x_i \frac{\partial}{\partial x_j}, \quad 1 \le i, j \le n;$$

then

$$[X_{ij}, X_{k\ell}] = \delta_{jk} X_{i\ell} - \delta_{\ell i} X_{kj},$$

and $\{X_{ij}, 1 \le i, j \le n\}$ generates the Lie algebra $\mathfrak{gl}_n$ and the associative polynomial algebra $\mathcal{A}_n$, respectively. By letting $n \to \infty$ we get the Lie algebra $\mathfrak{gl}_\infty$ and $\mathcal{A}_\infty$, respectively. Another example is the Heisenberg Lie algebra $\mathcal{H}_n$ spanned by the vector fields

$$P_j = \frac{i\sqrt{2}}{2}\Big(x_j + \frac{\partial}{\partial x_j}\Big), \quad Q_j = \frac{\sqrt{2}}{2}\Big(-x_j + \frac{\partial}{\partial x_j}\Big), \quad 1 \le j \le n,$$

$i = \sqrt{-1}$, and $R = iI$, where I is the identity operator. Then we have the commutation relations

$$[P_j, P_k] = [Q_j, Q_k] = 0, \quad [P_j, R] = [Q_j, R] = 0,$$
$$[P_j, Q_k] = -\delta_{jk} R, \qquad 1 \le j, k \le n.$$

Let $\mathcal{A}_n$ denote the algebra of non-commutative polynomials in the vector fields $P_1, \ldots, P_n$, $Q_1, \ldots, Q_n$ and R. Then when $n \to \infty$ we obviously have the generalization of the theorems in this section to the Heisenberg Lie algebra $\mathcal{H}_\infty$, and hence to $\mathcal{A}_\infty$.

4. CONCLUSION

A. Einstein said "*A good idea is very rare*". We reckon that there are, at the very least, three "good" ideas in [Poincaré 1900], namely, the universal enveloping algebra of a Lie algebra, the symmetrization map, and the proof of the so-called Birkhoff-Witt theorem. And in our opinion, none of these were properly appreciated and recognized. We have gone to great length and sometimes with repetitive arguments to try to convince the mathematics community of what a great feat Poincaré achieved in [Poincaré 1900]. But even if we failed, we would be much wiser by our reading a masterpiece by a great master.

We leave the reader with the following thought of Paul Painlevé, another great master, in the obituary written for the newspaper *Le Temps* (and reprinted in [Painlevé 1921]), on July 18, 1912, the day after Poincaré died:

"*Henri Poincaré n'a pas été seulement un grand créateur dans les sciences positives. Il a été un grand philosophe et un grand écrivain. Certains de ses aphorismes font songer à Pascal*: *'La pensée n'est qu'un éclair entre deux longues nuits, mais c'est cet éclair qui est tout'. Son style traduit la démarche même de sa pensée*: *des formules brèves et saisissantes, paradoxales parfois quand on les isole, réunies par des explications hâtives, qui rejettent des détails faciles pour ne dire que l'essentiel. C'est pourquoi des critiques superficiels lui ont reproché d'être 'décousu'*: *la vérité, c'est que, sans éducation scientifique préalable, une telle démarche logique est difficile à égaler*: *le lion ne fait pas des enjambées de souris.*"[6]

---

[6] This can be roughly translated as follows: "*Henri Poincaré was not only a great creator in the positive sciences. He was a great philosopher and a great writer. Some of his aphorisms make us think of Pascal*: *'Thought is just a flash of lightning in the middle of two long nights, but it is this lightning that is everything'. His style reflects the very development of his thought*: *brief and startling formulae, sometimes paradoxical when one isolates them, joined together by hasty explanations, which reject easy details in order just to express the essential. That is why superficial critiques reproach him as being 'incoherent'*: *the truth is that, without prerequisite education, such logical development is difficult to match*: *the lion does not take a mouse's paces.*"

Acknowledgments

The authors wish to thank the referees of the *Revue d'Histoire des Mathématiques* whose comments helped to improve the overall quality of this article. They also wish to thank the technical staff of the Department of Mathematics of the University of Iowa, Mr. Brian Treadway and Ms. Cymie Wehr, for their impeccable job and for their patience in preparing this manuscript.

BIBLIOGRAPHY

Encyclopaedia

[1988–1994] *Encyclopaedia of Mathematics*, ed. by I.M. Vinogradov, Dordrecht-Boston: Kluwer Academic Publishers, 1988–1994, translated from Математическая Энциклопедия, главный редактор И.М. Виноградов, 495, Москва: Совет·ская Энциклопедия, 1977 (the rubric "Birkhoff-Witt theorem" is in Tome 1, p. 495).

Barrow-Green (June)

[1997] *Poincaré and the Three-Body Problem*, Providence: American Mathematical Society, 1997.

Bell (Eric T.)

[1937] *Men of Mathematics*, New York: Simon & Schuster, 1937.

Birkhoff (Garrett)

[1937] Representability of Lie algebras and Lie groups by matrices, *Annals of Mathematics*, 38 (April 1937), pp. 526–532.

Bourbaki (Nicolas)

[1960] *Éléments de mathématiques, Fascicule XXVI. Groupes et algèbres de Lie*, Chap. 1, Paris: Hermann, 1960.

[1969] *Éléments d'histoire des mathématiques*, Paris: Hermann, 1969.

[1972] *Éléments de mathématiques, Fascicule XXXVII. Groupes et algèbres de Lie*, Paris: Hermann, 1972.

[1975] *Elements of Mathematics. Lie Groups and Lie Algebras,* Part I, Chapter I-3, Reading, MA: Addison-Wesley, 1975; Paris: Hermann, 1971–73.

Boyer (Carl)

[1968] *A History of Mathematics*, Princeton: Princeton University Press, 1968.

Cartan (Henri) & Eilenberg (Samuel)

[1956] *Homological Algebra*, Princeton: Princeton University Press, 1956.

Chevalley (Claude)

[1955] *Théorie des groupes de Lie*, vol. III, Paris: Hermann, 1955.

Cohn (Paul M.)

[1981] *Universal Algebra*, Dordrecht: Reidel, 1981.

Dixmier (Jacques)

[1974] *Algèbres enveloppantes*, Paris: Gauthier-Villars, 1974.

Gittleman (Arthur)

[1975] *History of Mathematics*, Columbus, OH: C.E. Merrill Publishing Company, 1975.

Godement (Roger)

[1982] *Introduction à la théorie des groupes de Lie*, 2 vols., Publications mathématiques de l'Université de Paris VII, Paris: Univ. Paris VII, 1982.

HARISH-CHANDRA

[1949] On Representations of Lie Algebras, *Annals of Mathematics*, 50 (October 1949), pp. 900–915.

[1951] On Some Applications of the Universal Enveloping Algebra of a Semisimple Lie Algebra, *Transactions of the American Mathematical Society*, 70–71 (1951), pp. 28–96.

HOFFMAN (Kenneth) & KUNZE (Ray)

[1971] *Linear Algebra*, second ed., Englewood Cliffs, NJ: Prentice-Hall, 1971.

HUMPHREYS (James E.)

[1972] *Introduction to Lie Algebras and Representation Theory*, New York: Springer-Verlag, 1972.

JACOBSON (Nathan)

[1962] *Lie Algebras*, New York: John Wiley & Sons, 1962.

KNAPP (Anthony)

[1986] *Representation Theory of Semisimple Groups*, Princeton: Princeton University Press, 1986.

KUROŚ (Aleksandrevich Gennadievich)

[1963] *Lectures on General Algebras*, New York: Chelsea, 1963.

LANG (Serge)

[1965] *Algebra*, Reading, MA: Addison-Wesley, 1965.

LAZARD (Michel)

[1952] Sur les algèbres enveloppantes universelles de certaines algèbres de Lie, *Comptes rendus hebdomadaires des séances de l'Académie des sciences de Paris Sér. I* (18 Feb. 1952), pp. 788–792.

[1954] Sur les algèbres enveloppantes de certaines algèbres de Lie, *Publications scientifiques de l'Université d'Alger, Sér. A*, 1 (1954), pp. 281–294.

LORENTZ (Hendrik Antoon)

[1921] Deux mémoires de Henri Poincaré sur la physique mathématique, *Acta mathematica*, 38 (1921), pp. 293–308.

PAINLEVÉ (Paul)

[1921] Henri Poincaré, *Acta Math.*, 38 (1921), pp. 309–402.

POINCARÉ (Henri)

[*Œuvres*] *Œuvres de Henri Poincaré*, 11 vols., Paris: Gauthier-Villars, 1916– .

[1881] Formes cubiques ternaires et quaternaires, *Journal de l'École polytechnique Paris*, XXXI (1881), pp. 199–253.

[1883] Sur la reproduction des formes, *C. R. Acad. sci. Paris* (29 oct. 1883), pp. 949–951.

[1899] Sur les groupes continus, *C. R. Acad. sci. Paris*, 128 (1899), pp. 1065–1069.

[1900] Sur les groupes continus, *Transactions of the Cambridge Philosophical Society*, 18 (1900), pp. 220–255 = *Œuvres de Henri Poincaré*, vol., III, Paris: Gauthier-Villars (1934), pp. 173–212.

[1901] Quelques remarques sur les groupes continus, *Rendiconti del Circolo matematico di Palermo*, 15 (1901), pp. 213–260.

[1906] Sur la dynamique de l'électron, *Rend. Circ. Mat. Palermo*, 21 (1906), pp. 129–176.

[1908] Nouvelles remarques sur les groupes continus, *Rend. Circ. Mat. Palermo*, 25 (1908), pp. 261–321.

[1912] Sur la théorie des quanta, *Journal de physique théorique et appliquée 5e sér.*, 2 (1912), pp. 5–34.

[*In memoriam* 1921] Henri Poincaré *in memoriam*, *Acta Math.*, 38 (1921).

SCHMID (Wilfried)
[1982] Poincaré and Lie groups, *Bulletin of the American Mathematical Society (N.S.)*, 6 (1982), pp. 175–186. This article is also reprinted in Felix E. Browder (ed.), *The Mathematical Heritage of Henri Poincaré*, Proceedings of Symposia in Pure Mathematics, Vol. 39, Providence, R.I.: American Mathematical Society, 1983, Part 1, pp. 157–168.

SCHWARTZ (Laurent)
[1998 (1975)] *Les tenseurs*, Paris: Hermann, 1998; 1e éd. 1975.

VARADARAJAN (V.S.)
[1984 (1974)] *Lie Groups, Lie Algebras and Their Representations*, New York: Springer-Verlag, 1984; Englewood Cliffs, NJ: Prentice-Hall, 1974.

WEYL (Hermann)
[1946] *The Classical Groups, Their Invariants and Representations*, second ed., Princeton: Princeton Univ. Press, 1946.

WITT (Ernst)
[1937] Treue Darstellung Liescher Ringe, *Journal für die reine und angewandte Mathematik*, 177 (1937), pp. 152–160.