Canad. Math. Bull. Vol. 33 (3), 1990

ON A RELATIONSHIP BETWEEN ADJOINT ORBITS AND CONJUGACY CLASSES OF A LIE GROUP

BY
N. J. WILDBERGER

ABSTRACT. Let G be a Lie group with adjoint orbits θ_i and corresponding conjugacy classes $C_i = \exp\, \theta_i$, $i = 1, 2, 3$. We show that if G is nilpotent or compact, there is a neighbourhood U of 0 in $\mathfrak{g}$ such that if $\theta_i \in U$ then $\theta_3 \subset \theta_1 + \theta_2$ if and only if $C_3 \subset C_1C_2$.

0. **Introduction.** Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let U be a neighbourhood of 0 in $\mathfrak{g}$ on which the exponential map is a diffeomorphism. Let O_1, O_2, O_3 be adjoint orbits contained in U and let $C_i = \exp(O_i)$ be the corresponding conjugacy classes. We are interested in the relationship between the sets $O_1 + O_2$ and C_1C_2; specifically we conjecture that $O_3 \subseteq O_1 + O_2$ if and only if $C_3 \in C_1C_2$ if U is chosen appropriately.

We prove the conjecture in two cases. The first is when G is a nilpotent connected simply connected Lie group and $U = \mathfrak{g}$. The second is when G is compact and connected, where we prove the existence of some U with the conjectured property.

The techniques in these two cases are completely different. The first case is dealt with by establishing certain formal power series identities in a free Lie algebra on a set of two elements. The second case uses the Weyl functional equation for characters and the Kirillov character formula for compact Lie groups.

The result of §2 establishes the following conjecture of R. Thompson [3] in a special case.

Conjecture. If A and B are Hermitian matrices, then there will always exist unitary matrices U and V such that

$$e^{iA}e^{iB} = e^{i(UAU^{-1}+VBV^{-1})}.$$

Theorem 2.1, when applied to the case $G = U(n)$, yields the validity of the conjecture for all A, B in a certain neighbourhood of 0.

1. Let G be a nilpotent, connected, simply-connected Lie group with Lie algebra $\mathfrak{g}$. The group G acts on $\mathfrak{g}$ via the adjoint action Ad and the orbits of this action are called adjoint orbits. If O_1, $O_2 \subseteq \mathfrak{g}$ are two adjoint orbits, their sum

$$O_1 + O_2 = \{X_1 + X_2 | X_i \in O_i,\ i = 1,\ 2\}$$

Received by the editors December 8, 1988 and, in revised form, August 14, 1989.
AMS (1980) subject Classification: 22E15

is G-invariant and therefore a union of adjoint orbits. The group G also acts on G itself by conjugation and the orbits are conjugacy classes. If $C_1, C_2 \subseteq G$ are two conjugacy classes, their product

$$C_1C_2 = \{g_1g_2 | g_i \in C_i,\ i = 1,\ 2\}$$

is G-invariant and therefore a union of conjugacy classes. The exponential map $\exp : \mathfrak{g} \longrightarrow G$ is a diffeomorphism and commutes with the actions of G on $\mathfrak{g}$ and G.

THEOREM 1.1. *Let G be a nilpotent, connected, simply-connected Lie group. Let $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3 \subseteq \mathfrak{g}$ be adjoint orbits and let $C_i = \exp(\mathcal{O}_i)$, $i = 1, 2, 3$, be the corresponding conjugacy classes. Then $\mathcal{O}_3 \subseteq \mathcal{O}_1 + \mathcal{O}_2$ if and only if $C_3 \subseteq C_1C_2$.*

PROOF. For $X, A \in \mathfrak{g}$, we will denote the element

$$\mathrm{Ad}(\exp A)(X) = \sum_{n=0}^{\infty} \frac{\mathrm{ad}(A)^n(X)}{n!}$$

by X^A. Every element of the adjoint orbit through X is of this form for some $A \in \mathfrak{g}$.

For $X, Y \in \mathfrak{g}$, we will write $X \cdot Y$ for $[X, Y]$. The element

$$\log(\exp X \ \exp Y) = X + Y + \tfrac{1}{2}X \cdot Y + \dots$$

will be denoted by $X * Y$. This infinite series is given by the Baker-Campbell-Hausdorff formula (see Jacobson [1]).

Theorem 1.1 is equivalent to the following two statements.

1) For every $X, Y \in \mathfrak{g}$, $\exists A, B \in \mathfrak{g}$ such that

$$X * Y = X^A + Y^B.$$

2) For every $X, Y \in \mathfrak{g}$, $\exists C, D \in \mathfrak{g}$ such that

$$X + Y = X^C * Y^D.$$

These statements turn out to hold formally in any Lie algebra. To make this precise we will prove 1) and 2) in the free Lie algebra on a set $\{X, Y\}$ of two elements. We recall some notation and basic facts about such a free Lie algebra, following Jacobson [1].

Let F be a field of characteristic zero and set $M = FX \oplus FY$. Let

$$\mathcal{F} = F \cdot 1 \oplus M \oplus (M \otimes M) \oplus \dots$$

be the free algebra on the set $\{X, Y\}$, graded by setting elements in $M \otimes \cdots \otimes M$ (k times) to have degree k. Let $\mathcal{F}_L$ denote the associated Lie algebra and let $\mathcal{FL}$ denote the Lie algebra of $\mathcal{F}_L$ generated by X and Y. The elements $a \in \mathcal{F}\mathcal{L}$ are called Lie elements of $\mathcal{F}$, and $\mathcal{F}\mathcal{L}$ is the free Lie algebra on the set $\{X, Y\}$.

Let $\mathcal{F}\mathcal{L}_k$ denote the subspace of $\mathcal{F}\mathcal{L}$ of elements homogeneous of degree k. Then

$$\mathcal{F}\mathcal{L} = \sum_{k=1}^{\infty} \mathcal{F}\mathcal{L}_k.$$

If $A_1, \ldots, A_k$ are all either X or Y, the element

$$(\ldots(A_1 \cdot A_2) \cdot \ldots) \cdot A_k$$

is homogeneous of degree k and will be called a chain of length k. An induction using the Jacobi identity shows that if $a, b \in \mathcal{F}\mathcal{L}$ are chains of length k and l respectively, then $a \cdot b$ is a linear combination of chains of length $k+l$. It follows that $\mathcal{F}\mathcal{L}_k$ is the span of all chains of length k, and that $\mathcal{F}\mathcal{L}$ is the span of all chains. Also we have

$$\mathcal{F}\mathcal{L}_k \cdot \mathcal{F}\mathcal{L}_l \subseteq \mathcal{F}\mathcal{L}_{k+l}.$$

Now let $\overline{\mathcal{F}\mathcal{L}}$ denote the Lie algebra of formal power series. Thus a typical element of $\overline{\mathcal{F}\mathcal{L}}$ is

$$a = \sum_{k=1}^{\infty} a_k$$

$a_k \in \mathcal{F}\mathcal{L}_k$. $\overline{\mathcal{F}\mathcal{L}}$ is a Lie algebra in the obvious way. Let $\overline{\mathcal{F}\mathcal{L}}^{(i)}$ be the subspace of $\overline{\mathcal{F}\mathcal{L}}$ consisting of elements with no $\mathcal{F}\mathcal{L}_k$ component for $k < i$. For $a \in \overline{\mathcal{F}\mathcal{L}}^{(i)}$ and $a \notin \overline{\mathcal{F}\mathcal{L}}^{(i+1)}$, define

$$|a| = 2^{-i}.$$

This makes $\overline{\mathcal{F}\mathcal{L}}$ into a topological algebra. Then a series

$$\sum_{i=1}^{\infty} x_i$$

of elements $x_i \in \overline{\mathcal{F}\mathcal{L}}$ converges if and only if $\lim_{i\to\infty} |x_i| = 0$, which of course is exactly the condition one requires to be able to sum the series componentwise.

Now note that if $A \in \overline{\mathcal{F}\mathcal{L}}$, the series

$$X^A = X + A \cdot X + \frac{A \cdot (A \cdot X)}{2!} + \ldots$$

converges. Also note that

$$X * Y = X + Y + \tfrac{1}{2} X \cdot Y + \ldots$$

is an element of $\overline{\mathcal{F}\mathcal{L}}$; in fact the Baker-Campbell- Hausdorff formula expresses $X * Y$ as a series of chains (see-Jacobson [1]).

https://doi.org/10.4153/CMB-1990-048-4 Published online by Cambridge University Press

PROPOSITION 1.2. *There exist* A, $B \in \overline{\mathcal{FL}}$ *such that*

$$X * Y = X^A + Y^B. \tag{1.1}$$

PROOF. The argument is familiar and does not depend on the explicit form of $X * Y$ past the first few terms. Let $A = \sum_{k=1}^{\infty} a_k$, $B = \sum_{k=1}^{\infty} b_k$ with a_k, $b_k \in \mathcal{FL}_k$ to be determined. We will show that $(*)$ allows us to determine the a_k's and b_k's inductively.

Suppose that a_k, b_k have been determined for $k = 1, \ldots, m-1$. Rewrite (1.1) as

$$\tfrac{1}{2}X \cdot Y + \cdots = A \cdot X + \frac{A \cdot (A \cdot X)}{2!} + \cdots + B \cdot Y + \frac{B \cdot (B \cdot Y)}{2!} + \ldots \tag{1.2}$$

and consider the component in $\mathcal{FL}_{m+1}$. The only a_k's and b_k's which occur in this component are ones with $k \leqq m$, and a_m, b_m only occur in one term each, namely $A \cdot X$ and $B \cdot Y$. We thus get

$$a_m \cdot X + b_m \cdot Y = \text{some element in } \mathcal{FL}_{m+1}.$$

Now any element of $\mathcal{FL}_{m+1}$ can be written as a linear combination of chains of length $m+1$ and so in the form

$$a' \cdot X + b' \cdot Y$$

where a', b' are in $\mathcal{FL}_m$. Thus set $a_m = a'$, $b_m = b'$.

To begin the induction, note that after cancellation there are no terms in (1.1) of degree 1 so that

$$a_1 \cdot X + b_1 \cdot Y = \tfrac{1}{2}X \cdot Y$$

and thus $a_1 = 0$, $b_1 = \frac{1}{2}X$. □

Before proving formula 2), we must observe that the Baker-Campbell-Hausdorff formula for $X * Y$ allows its extension to $\overline{\mathcal{FL}}$. That is,

$$A * B = A + B + \tfrac{1}{2}A \cdot B + \ldots$$

is a convergent series for any $A, B \in \overline{\mathcal{FL}}$.

PROPOSITION 1.3 *There exist* C, $D \in \overline{\mathcal{FL}}$ *such that*

$$X + Y = X^C * Y^D. \tag{1.3}$$

PROOF. The equation (1.3) can be expanded as

$$\begin{aligned} X + Y &= \left(X + C \cdot X + \frac{C \cdot (C \cdot X)}{2!} + \ldots\right) * \left(Y + D \cdot Y + \frac{D \cdot (d \cdot Y)}{2!} + \ldots\right) \\ &= X + Y + \tfrac{1}{2}X \cdot Y + C \cdot X + D \cdot Y + \text{higher degree terms.} \end{aligned}$$

As in the proof of the preceding proposition, the linear terms cancel and then the homogeneous components of C and D may be computed inductively. □

We remark that the elements $A, B, C, D \in \overline{\mathcal{FL}}$ obtained in the previous two propositions are not unique. One way to see this is that the chains of length k are not independent.

We can now complete the proof of the theorem by noting that for $\mathfrak{g}$ nilpotent (and $F = R$), all elements in $\overline{\mathcal{FL}}$ are finite sums so the elements A, B, C, D obtained in Propositions 1.2 and 1.3 can be regarded as elements of $\mathfrak{g}$ to prove statements 1) and 2). □

We remark that the difficulty in extending this line of reasoning to more general Lie algebras is that we have no information about the convergence of the formal series for A, B, C, D obtained above.

2. Propositions 1.2 and 1.3 suggest that an analogue of Theorem 1.1 should hold for other Lie groups. In this section we consider the case that G is a compact connected Lie group. Even for $SU(2)$ it is clear that the analogue of Theorem 1.1 can not hold globally (that is, for arbitrary adjoint orbits). Nevertheless, the following result shows that it holds in a neighbourhood of 0.

THEOREM 2.1. *Let G be a compact connected Lie group. Then there exists a neighbourhood U of 0 in $\mathfrak{g}$ such that if $\mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3 \subseteq U$ are adjoint orbits and $C_i = \exp(\mathcal{O}_i)$, $i = 1, 2, 3$, the corresponding conjugacy classes, then $\mathcal{O}_3 \subseteq \mathcal{O}_1 + \mathcal{O}_2$ if and only if $C_3 \subseteq C_1C_2$.*

PROOF. We will use both a formula of Weyl on the functional equation satisfied by characters and the Kirillov character formula. Let us recall these results.

Normalize the Haar measure dg on G so that

$$\int_G dg = 1.$$

Let $C(G)$ denote the space of continuous functions on G.

For functions $\varphi_1, \varphi_2 \in C(G)$ define $\varphi_1 * \varphi_2$ by

$$\varphi_1 * \varphi_2(g) = \int_G \varphi_1(g')\varphi_2(g'^{-1}g)dg'.$$

Let $\hat{G}$ denote the set of (equivalence classes of) irreducible unitary representations of G. For $\rho \in \hat{G}$, let χ_ρ denote its character. This is a smooth function on G given by

$$\chi_\rho(g) = tr\,\rho(g) \qquad \forall g \in G.$$

We may also consider χ_ρ as a distribution on G, that is, for $\varphi \in C(G)$,

$$\begin{aligned}\chi_\rho(\varphi) &= tr\,\rho(\varphi)\\ &= \int_G \chi_\rho(g)\varphi(g)dg.\end{aligned}$$

https://doi.org/10.4153/CMB-1990-048-4 Published online by Cambridge University Press

Recall that a function $\varphi \in C(G)$ is called central if

$$\varphi(x) = \varphi(yxy^{-1}) \qquad \forall x,\ y \in G.$$

The Weyl functional equation states that if $\varphi_1,\ \varphi_2 \in C(G)$ are central, $\rho \in \hat{G}$ and $d_\rho = \chi_\rho(e)$, then

$$\chi_\rho(\varphi_1 * \varphi_2) = \frac{1}{d_\rho}\, \chi_\rho(\varphi_1)\chi_\rho(\varphi_2). \tag{2.1}$$

Actually this formula is perhaps more often found in a different but equivalent form, namely, as

$$\int_G \chi_\rho(gxg^{-1}y)dg = \frac{1}{d_\rho}\, \chi_\rho(x)\chi_\rho(y) \qquad \forall x,\ y \in G$$

(see Weyl [5]).

Equation (2.1) may also be derived from the well-known equations (see Weil [4])

$$\chi_\rho * \chi_{\rho'} = \begin{cases} \dfrac{1}{d_\rho}\, \chi_\rho & \text{if } \rho = \rho' \\ 0 & \text{if } \rho \neq \rho' \end{cases}$$

together with the fact that the characters χ_ρ form an orthonormal basis of the central functions in $L^2(G,\ dg)$. To see this, write $\varphi_1 = \sum a_\rho \bar{\chi}_\rho$ and $\varphi_2 = \sum b_\rho \bar{\chi}_\rho$. Then

$$\varphi_1 * \varphi_2 = \sum \frac{a_\rho b_\rho}{d_\rho}\, \bar{\chi}_\rho,$$

so that

$$\chi_\rho(\varphi_1 * \varphi_2) = \frac{a_\rho b_\rho}{d_\rho} = \frac{1}{d_\rho}\, \chi_\rho(\varphi_1)\chi_\rho(\varphi_2).$$

The Kirillov character formula gives an expression for χ_ρ as the Fourier transform of the invariant measure on a certain co-adjoint orbit. Let $\mathfrak{g}^*$ be the linear dual of $\mathfrak{g}$. Then G acts on $\mathfrak{g}^*$ by the co-adjoint representation and the orbits of this action are called co-adjoint orbits. Choose a Lebesgue measure dX on $\mathfrak{g}$ and define, for $\psi \in C_c^\infty(\mathfrak{g})$, its Fourier transform $\psi^\wedge$ on $\mathfrak{g}^*$ by

$$\psi^\wedge(f) = \int_{\mathfrak{g}} \psi(X)e^{if(X)}dX \qquad \forall f \in \mathfrak{g}^*.$$

Every co-adjoint orbit Ω has a G-invariant measure $d\mu$, unique up to a constant. The Kirillov character formula states that there exists a neighbourhood U_0 of 0 in $\mathfrak{g}$ on which the exponential map is a diffeomorphism and a non-zero, real function p defined on U_0 satisfying $p(0) = 1$ with the following property. For a function φ with support in exp U_0, let $\tilde{\varphi}$ denote its lift to U_0; that is,

$$\tilde{\varphi}(X) = \varphi(\exp X) \qquad \forall X \in U_0.$$

Then there exists a co-adjoint orbit $\Omega_\rho \subseteq \mathfrak{g}^*$ such that for all $\varphi \in C_c^\infty(\exp U_0)$,

$$\chi_\rho(\varphi) = \int_{\Omega_\rho} (\tilde{\varphi} p^{-1})^\wedge(f)\, d\mu(f). \tag{2.2}$$

For a proof and additional discussion see Kirillov [2].

Introduce a G-invariant Euclidean structure on $\mathfrak{g}$ so that any adjoint orbit will be contained in a sphere centered at 0. Choose $r > 0$ so that the ball of radius r centered at 0 is contained in U_0. Let U be the interior of the ball of radius $r/2$.

Let ψ_1, ψ_2 be G-invariant, smooth functions on $\mathfrak{g}$ with compact support contained in U. Then

$$\psi_3 = \psi_1 * \psi_2,$$

the (abelian) convolution of ψ_1 and ψ_2, will also be G-invariant and will have compact support in U_0. Let $\rho \in \hat{G}$. Then since $\int_{\Omega_\rho} d\mu = d_\rho$,

$$\begin{aligned} \int_{\Omega_\rho} \psi_3^\wedge d\mu &= \int_{\Omega_\rho} \psi_1^\wedge \psi_2^\wedge d\mu \\ &= \frac{1}{d_\rho} \int_{\Omega_\rho} \psi_1^\wedge d\mu \int_{\Omega_\rho} \psi_2^\wedge d\mu. \end{aligned} \tag{2.3}$$

Now let $\varphi_i \in C_c^\infty(\exp U)$ be uniquely determined by the condition

$$\tilde{\varphi}_i p^{-1} = \psi_i, \qquad i = 1,\ 2,\ 3.$$

Then Kirillov's character formula (2.2) together with (2.3) shows that

$$\chi_\rho(\varphi_3) = \frac{1}{d_\rho} \chi_\rho(\varphi_1)\chi_\rho(\varphi_2).$$

On the other hand, consider the group convolution of the two central functions φ_1 and φ_2,

$$\varphi = \varphi_1 * \varphi_2.$$

Then (2.1) states that

$$\chi_\rho(\varphi) = \frac{1}{d_\rho} \chi_\rho(\varphi_1)\chi_\rho(\varphi_2).$$

Thus φ and φ_3 are smooth functions on G satisfying

$$\chi_\rho(\varphi) = \chi_\rho(\varphi_3)$$

for all irreducible unitary representations ρ. Since the characters form an orthogonal basis of the central functions on G, we conclude that

$$\varphi = \varphi_3.$$

https://doi.org/10.4153/CMB-1990-048-4 Published online by Cambridge University Press

Now since $\psi_1,\ \psi_2$ were arbitrary, we may conclude that for arbitrary smooth central functions $\varphi_1,\ \varphi_2$ of compact support in $\exp\ U$,

$$(\varphi_1\ *\ \varphi_2)^{\sim}p^{-1} = (\tilde{\varphi}_1 p^{-1})\ *\ (\tilde{\varphi}_2 p^{-1}),$$

where the convolutions of the left and right sides of the equations occur on the group G and the Lie algebra $\mathfrak{g}$ respectively.

Now suppose that $O_1,\ O_2$ are adjoint orbits in U and $C_i = \exp\,(O_i)$ the corresponding conjugacy classes. Find $\varphi_1,\ \varphi_2 \in C_c^\infty(\exp\ U)$ both central, real and non-negative with support close to $C_1,\ C_2$ respectively. Then the support of $\varphi_1\ *\ \varphi_2$ will be close to C_1C_2. On the other hand, $\psi_i = \tilde{\varphi}_i p^{-1}$ will have support close to O_i and $\psi_1\ *\ \psi_2$ will have support close to $O_1 + O_2$. It follows that

$$\exp(O_1 + O_2) = C_1C_2,$$

which is equivalent to the theorem. □

References

1. N. Jacobson, *Lie Algebras*, John Wiley and Sons (1962), New York.
2. A. A. Kirillov, *The characters of unitary representations of Lie groups*, Funct. Anal. and its Applications **2, No. 2** (1968), 133–146.
3. R. Thompson, *Author vs Referee: A Case History for Middle Level Mathematicians*, Amer. Math. Monthly **90, No. 10** (1983), 661–668.
4. A. Weil, *L'integration dans les groups topologiques et ses application.*, Actualités Sci. et Ind. **869, 1145** (1941), Hermann and Cie, Paris.
5. H. Weyl, *Gruppen theorie und Quantenmechanik 2. Aufl.*, S. Hirzel (1931), Leipzig.

University of New South Wales
Kensington, NSW
2033 Australia

https://doi.org/10.4153/CMB-1990-048-4 Published online by Cambridge University Press