Math. Struct. in Comp. Science (2008), *vol.* 18, *pp.* 189–204.
doi:10.1017/S0960129508006622 Printed in the United Kingdom

Model theory of exponentials on Lie algebras

ANGUS MACINTYRE

Department of Mathematical Sciences, Queen Mary University of London, Mile End, London E1 4NS, United Kingdom
Email: angus@dcs.qmul.ac.uk

Received 30 July 2007

In Memory of Sauro Tulipani

This paper presents an analysis of definitions and decidability for exponential functions on various matrix algebras. The main idea is to show that, generically, the entries of the exponential (or logarithm) of a matrix are Pfaffian functions of the entries of the matrix.

1. Introduction

In the last fifteen years model theorists have reached a very satisfactory understanding of the exponential function both on $\mathbb{R}$ and $\mathbb{C}$. This paper contributes to a more general study of exponentials on finite-dimensional Lie algebras over $\mathbb{R}$ and $\mathbb{C}$, and shows that most of the earlier results can be widely generalised.

For the classical exponential on $\mathbb{R}$, the model theory is that of the structure

$$< \mathbb{R}, +, \cdot, -, <, 0, 1, \exp > .$$

The main qualitative result is due to Wilkie (Wilkie 1996), who showed that the structure is model-complete and o-minimal.The analysis was later constructivised, and axioms found, leading to the proof of decidability, relative to Schanuel's Conjecture, in Macintyre and Wilkie (1996).

For the complex exponential, the model theory is that of

$$< \mathbb{C}, +, \cdot, -, <, 0, 1, \exp > .$$

Tarski knew in the 1930's that this is undecidable, using the interpretability of $2\pi i\mathbb{Z}$. I observed about fifteen years ago that the theory is not model-complete. A more recent, deeper understanding contrasts these 'negative' results with the fascinating insights of Zilber (2004), which suggest that one may be able to reach a satisfying analysis of the definable sets, modulo the basic Godelian clutter coming from $2\pi i\mathbb{Z}$.

It is particularly striking that Schanuel's Conjecture is crucial in both the real and complex cases, though in different ways.

In 2003 I noticed that the Weierstrass $\wp$-functions on their fundamental domains have a model-complete theory very similar to that of the real exponential (Macintyre 2005a).

I developed the basic ideas of Bianconi (Bianconi 1990) to obtain constructive model-completeness results and (relative) decidability results, assuming André's Conjecture on transcendence in the theory of 1-motives (Bertolin 2002).

These recent results are really about the exponential map on the tangent space of a 1-dimensional complex torus, and suggest that one should now consider the general setting of model theory for exponentials of finite-dimensional Lie algebras over $\mathbb{R}$ and $\mathbb{C}$.

2. Formalism

K will be one of $\mathbb{R}$ or $\mathbb{C}$ throughout the rest of the paper, and I will specify in context any restrictions that have to be made.

Serre (1965) will be our basic reference for Lie theory over K.

Our formalism will have to account for at least some of the following:

(1) a Lie group G (a K-analytic group);
(2) a Lie algebra $\mathcal{G}$ over K, typically the tangent space $T_e(G)$ with the usual Lie algebra structure;
(3) the exponential map from $T_e(G)$ to G, or perhaps its restriction to some natural subset of $T_e(G)$;
(4) a logarithm map from a natural subset of G to $T_e(G)$.

From the standpoint of the classical 'affine' Tarskian formalism, the Lie algebra $\mathcal{G}$ is a more natural structure than G. If $\mathcal{G}$ has dimension n over K, we can simply regard it as K^n, qua vector space over K $(=< K, +, ., - >$, with an extra map $[\,]$ from K^n to itself, satisfying the usual Lie axioms.

G, on the other hand, is simply an n-dimensional analytic group, which is not in general affine, and so may have to be interpreted via a (possibly infinite) covering by affine pieces. This does not fit the Tarskian foundations so neatly, especially when G is not compact.

One natural way to obtain an 'affine' piece of G is through the following sequence of ideas (Serre 1965):

(1) From an n-dimensional $\mathcal{G}$ over K, construct $CH(\mathcal{G})$, that is, its Campbell–Hausdorff formal group law (in $K[[x_1, \ldots, x_n]]$).
(2) Show that $CH(\mathcal{G})$ gives K^n the structure of an analytic group chunk, also denoted $CH(\mathcal{G})$, whose tangent space at the identity is naturally isomorphic to $\mathcal{G}$.
(3) If $\mathcal{G} = T_e(G)$, where G is an analytic group chunk, show that there is a unique local isomorphism exp from $CH(\mathcal{G})$ to G, which induces (via the functor T_e) the identity on $\mathcal{G}$ (under the identification of $T_e(CH(\mathcal{G})$ with $\mathcal{G}$).

The main point here is that exp is defined on all of $CH(\mathcal{G})$, and is a local isomorphism. As Serre shows in Serre (1965, Section $LG5.35$), exp extends to $\mathcal{G}$ through

$$\exp(x) = \exp\left(\frac{x}{m}\right)^m$$

for sufficiently large positive integers m.

An affine piece of G can be obtained by considering a restriction of exp giving an isomorphism of open neighbourhoods U and V of 0 and e in $CH(\mathcal{G})$ and G, respectively.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

In turn, this leads us to consider logarithmic maps from open subsets of G to $\mathscr{G}$. A major nuisance for systematic first-order work on the Lie group G is the non-explicit nature of the 'logarithmic regions', that is, natural domains for the logarithm.

One particular way of looking at the model theory of G in light of the preceding discussion is to work with sorts for:

(1) a chunk of G in K^n;
(2) the Lie algebra $L(G) = \mathscr{G}$ on K^n;
(3) exp from $CH(\mathscr{G})$ into the chunk of G.

The sorting is not crucial since everything is happening in K^n.

Note that if we restrict to a chunk with compact closure (living on a reasonable set), the above formulation effectively puts us in an o-minimal situation. For $K = \mathbb{R}$ this is clear by interpretation in $\mathbb{R}_{an}$. For $K = \mathbb{C}$ one passes via the interpretation of $\mathbb{C}$ as $\mathbb{R}^2$, using the real and imaginary parts of the analytic functions involved.

However, it is much less obvious what one can prove about model-completeness and decidability. This issue, in full generality, will be addressed in a later paper. In the present paper, I restrict consideration to matrix groups G, which are explicitely given affinely. Thus, if G is a closed subgroup of $GL_m(K)$, we construe G as a subset of K^{m^2}. Then $T_e(G)$ is naturally construed as a subalgebra of $M_m(K)$, and it is well known that near (the matrix) 0 exp is given by the standard infinite series, and similarly for $\log(\mathbb{I} + x)$.

Thus, I now consider G a closed subgroup of $GL_m(K)$ and $\mathscr{G}$ the corresponding Lie algebra of matrices, and will subsequently consider formulations involving exp or log on suitable subsets.

3. Elementary results

I fix n and work with G a subgroup of the units of the ring $M_n(K)$, and the corresponding $\mathscr{G}$ a Lie subalgebra of the usual Lie algebra substructure on the matrix ring. The exponential function exp is defined by the usual convergent series.

I work in the usual field structure (or ordered field structure in the case of the reals) on K, and have a primitive for exp.

So for $n = 1$ and $K = \mathbb{R}$ we have the results of Wilkie (1996) and Macintyre and Wilkie (1996), and for $n = 1$ and $K = \mathbb{C}$ we have the results and perspective of Zilber (2004).

For $n = 2$ and $K = \mathbb{R}$ we already move out of o-minimality. We identify $\mathbb{R}$ inside the matrix ring as the set of all $r.\mathbb{I}$ for $r \in \mathbb{R}$, and thus as the centre. Now consider any matrix A with $A^2 = -\mathbb{I}$. Then $\mathbb{R} + \mathbb{R}A$ is a field, isomorphic to $\mathbb{C}$ and is the centraliser of A. Moreover, it is closed under exp, and thus we have (in terms of A) an interpretation of the complex exponential, and so of $\mathbb{Z}$. Note, too, that $< \mathbb{C}, exp >$ is easily interpreted in $< M_2(\mathbb{C}), exp >$ using the centre, so again there can be no o-minimality (using the interpretation of $\mathbb{C}$ as $\mathbb{R}^2$).

The issue of interpretability the other way is more delicate.

Theorem 3.1.

(a) $< M_2(\mathbb{C}), \exp >$ is interpretable in $< \mathbb{C}, \exp >$.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

(b) $< M_2(\mathbb{R}), \exp >$ is not interpretable in $< \mathbb{R}, \exp >$.
(c) $< M_2(\mathbb{R}), \exp >$ is not interpretable in $< \mathbb{C}, \exp >$ if Zilber's Conjecture (Zilber 2004) is true.
(d) $< M_2(\mathbb{R}), \exp >$ is interpretable in $< \mathbb{C}, \mathbb{R}, \exp >$.

Proof.

(a) Let $A \in M_2(\mathbb{C})$. Then A can be written uniquely (Jordan form) in the form $a = B + C$, with B diagonalisable and C nilpotent (in fact $C^2 = 0$), and $BC = CB$.
So

$$\exp(A) = \exp(B)\exp(C) = \exp(B)(\mathbb{I} + C).$$

Now, since B is diagonalisable, we have

$$D^{-1}BD = diag(\lambda_1, \lambda_2)$$

for some D and λ_1, λ_2. Then

$$\exp(B) = D\ diag(\exp(\lambda_1), \exp(\lambda_2))D^{-1}.$$

(b) This is immediate from an earlier remark about o-minimality.
(c) It suffices, by the preceding discussion, to show that $< \mathbb{R}, \exp >$ is not interpretable. But this follows from the infinitary ω-stability of $< \mathbb{C}, \exp >$ conjectured by Zilber.
(d) This is essentially the same argument as in (a). We use the predicate for $\mathbb{R}$ to interpret $M_2(\mathbb{R})$ inside $M_2(\mathbb{C})$, and the rest of the proof in (a) applies. □

4. A remark on the complex logarithm

In the real case, the exponential and the logarithm are bi-interpretable. However, the situation is quite different for the complex case. Here one may naturally consider the complex field carrying the extra structure of the branch of the logarithm one gets by putting $log(1) = 0$ and having as domain of definition the plane with the non-positive real axis deleted. Let us write this structure as

$$< \mathbb{C}, \log > .$$

Theorem 4.1. $< \mathbb{C}, \log >$ is o-minimal and decidable if Schanuel's Conjecture is true.

Proof. It is clear that this structure is interpretable in the real exponential field enriched by primitives for sine and cosine on $[0, 2\pi]$. This gives o-minimality via the usual $\mathbb{R}_{an}$ interpretation. The decidability modulo Schanuel can be proved along the lines of Macintyre and Wilkie (1996), using the fact that sine and cosine are Pfaffian away from $\pi\mathbb{Z}$. Finally, one uses the complex Schanuel Conjecture, rather than the real one as in Macintyre and Wilkie (1996). (I will recall the basics of Pfaffian functions later in this paper).

It is worth noting that Bianconi showed that the restricted trigonometric functions are not interpretable in $< \mathbb{R}, \exp >$ (Bianconi 1997). □

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

5. Some readily accessible positive results beyond $< \mathbb{R}, \exp >$

5.1. $so(2)$

$so(2)$ is the real Lie algebra of skew-symmetric elements of $M_2(\mathbb{R})$. The elements are of the form

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

$$\exp \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \quad \in SO(2),$$

where $SO(2)$ is the special orthogonal group.

So $\exp : so(2) \longrightarrow SO(2)$ is surjective, by inspection.

In this particular case, the Lie algebra is abelian, and

$$\exp(A + B) = \exp(A) \cdot \exp(B).$$

Moreover, the kernel of exp is readily identifiable as the set of all matrices of the form

$$\begin{pmatrix} 0 & 2n\pi \\ -2n\pi & 0 \end{pmatrix}, \quad n \in \mathbb{Z},$$

so the Lie algebra formulation is obviously undecidable. If we pass to a logarithmic formulation, the situation changes (*cf.* Section 4).

Note that $so(2)$ and $\mathbb{R}$ are isomorphic real Lie algebras, but the following Lie groups, all with the preceding as Lie algebra, are not isomorphic:

(1) $\mathbb{R}^*$;
(2) $\mathbb{R}^+$;
(3) $SO(2)$.

The first is not connected; the second is connected and simply connected, but not compact; and the third is connected, but not simply connected, and is compact. In each case, the group chunk $CH(\mathcal{G})$ is $\mathbb{R}$ with a fragment of $+$, and that is all one can say. exp is onto $SO(2)$, but not in the other cases. Corresponding to the logarithm for $\mathbb{R}_+$, defined on the non-negative elements of the group, we have a logarithm for $SO(2)$ defined as follows:

$$\log \begin{pmatrix} u & v \\ -v & u \end{pmatrix} = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}$$

where

$$x = \arg(u + iv) \quad u + iv \neq -1$$

where arg is given by the imaginary part of the branch of the complex logarithm mentioned in Section 4.

Now, just as the real logarithm is Pfaffian on its domain of definition, the above logarithm is Pfaffian (as a function of two real variables u, v with $u^2 + v^2 = 1$ and $(u, v) \neq (1, 0)$) except at the four points

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

From this observation, and the compactness of $SO(2)$, one can use the method of Macintyre and Wilkie (1996) or Gabrielov and Vorobjov (2004) to prove, just as for the complex logarithm in Section 4, that the structure of $SO(2)$ with the above logarithm (interpreted as a function of two variables) is constructively model-complete, and decidable relative to Schanuel's Conjecture. One now asks, naturally, if this is true for the higher orthogonal groups. This is in fact so, but harder to prove. For $n = 3$, however, we can begin an analysis along the lines of what we have just done.

5.2. $SO(3)$

For $G = SO(3)$, we have $T_e(G)$ is $so(3)$, the Lie algebra of skew-symmetric elements in $M_2(\mathbb{R})$. G is compact and connected, but not simply connected.

A typical A in $so(3)$ is of the form

$$\begin{pmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{pmatrix}.$$

$\exp : so(3) \longrightarrow SO(3)$ is given by Rodrigues' formula (Barut *et al.* 1994):

$$\exp(A) = \mathbb{I} + \frac{\sin(\theta)}{\theta}A + \frac{1-\cos(\theta)}{\theta^2}A^2 \tag{1}$$

where $\theta = \sqrt{x^2 + y^2 + z^2}$ (and, for $\theta = 0$, $\frac{\sin(\theta)}{\theta} = 1$ and $\frac{1-\cos(\theta)}{\theta^2} = -1$).

Since G is compact and connected, exp is surjective. exp is not injective, as one can easily see by considering exp on the subalgebra of all A as above with $x = y = 0$. This observation gives undecidability for the Lie algebra formulation.

Note that exp is clearly Pfaffian, in the sense that its entries are, except on $\pi\mathbb{Z}$.

What about a logarithm? It is not at all obvious what its domain should be, and whether it is Pfaffian. The main result of this paper (Theorem 6.2) will find a natural domain, and show that the logarithm is Pfaffian. This will lead to a series of positive results along the lines of those in the preceding subsection.

6. Model theory of eigenvalues

6.1. *Going beyond the Rodrigues formula*

While it is well known that the computation of exp, even for real matrices, is naturally approached using a computation of exp of their complex eigenvalues (an idea that has already been used in this paper in the proof of Theorem 1 (a)), it requires considerable effort to get from this to formulas (such as that of Rodrigues) for the entries of exp of an explicitly given matrix. It seems that the needs of physics have been a major incentive to finding such formulas explicitly. Google led me to the paper Barut *et al.* (1994), which contains ideas and calculations very convenient for the modeltheoretic problems I address in this paper. The calculations in Barut *et al.* (1994) are done only for $n \leqslant 6$, and I have not seen explicit formulas for higher n, but the basic ideas (as hinted in Barut *et al.* (1994)) apply for all n, and the lack of explicitness is not a problem for my purposes.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

6.2. *The basics of Pfaffian functions*

The basic results are due to Khovanski, but for my purposes, the recent papers by Gabrielov and Vorobjov are very convenient (Gabrielov and Vorobjov 2004), and I take the definitions below from them.

A Pfaffian chain of order $r \geqslant 0$ and degree $\alpha \geqslant 1$ on an open set $U \subset \mathbb{R}^n$ (respectively, $U \subset \mathbb{C}^n$) is a sequence of real (respectively, analytic) functions $f_1, \ldots, f_r$ on U satisfying the Pfaffian equations

$$df_j = df_J(x) = g_{1j}dx_1 + .. + g_{ij}dx_i + \cdots + g_{nj}dx_n,$$

where the $g_{ij} = g_{ij}(x, f_1(x), \ldots, f_j(x))$ are polynomials in $x = (x_1, \ldots, x_n)$ of degree not exceeding α. A function $f(x) = P(x, f_1(x), \ldots, f_r(x))$ where $P(x, y_1, \ldots, y_r)$ is a polynomial of degree not exceeding β is a Pfaffian function of order r and degree (α, β).

The basic example of a Pfaffian function (on any U) is exp. sin and cos are not Pfaffian on the whole space, but are on $\mathbb{C}$ with the set $\pi\mathbb{Z}$ removed.

The essential difference between the real and complex cases is that in the real case, but not in the complex case, one has major finiteness theorems for the number of connected components of zero sets and positivity sets, with bounds uniform in families and depending only on order and degree. These results of Khovanski have been of immense significance for model theory, first in the work of Wilkie and then as part of the general theory of 0-minimality.

An important cautionary note is that, in general, the real and imaginary parts of a complex Pfaffian function need not be Pfaffian in the real sense. An important, though very restricted, case where one does get real Pfaffian functions in this way is to be found in Macintyre (2005b).

Algebraic functions (and their real and imaginary parts) are Pfaffian on the appropriate domains. We will see examples below.

Sums and products of Pfaffian functions are Pfaffian, and the class of Pfaffian functions is closed under division by non-vanishing Pfaffian functions. These properties will be used systematically below.

6.3. *Around Cayley–Hamilton*

I work in $M_n(\mathbb{C})$, though $M_n(\mathbb{R})$ is my primary interest. The topology is that of $\mathbb{C}^{n^2}$ or, equivalently $\mathbb{R}^{2n^2}$. Indeed, the formulation in real terms will be crucial as I aim to prove that certain functions are Pfaffian. I think of complex numbers as pairs of reals, and complex functions as pairs of real functions, and so on.

Consider the Zariski open set W consisting of the A with distinct eigenvalues. W is Zariski open since W can be defined by the non-vanishing of the discriminant of the characteristic polynomial

$$det(A - z\mathbb{I})$$

of A.

Note that W is a dense open set.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

The basic problem is to define on W the functions

$$\lambda_1, \ldots, \lambda_n$$

giving the distinct eigenvalues. Bearing in mind that each complex number is given by a pair of real numbers, namely its real and imaginary parts, the obvious definition (in the semi-algebraic category) is to give the induced lexicographic order to the eigenvalues, and to take the

$$\lambda_1, \ldots, \lambda_n$$

as giving the eigenvalues in this lexicographic order.

These functions are pairs of real-valued functions of $2n^2$ real variables, but for convenience I write them as functions of the variable A, that is, as

$$\lambda_1(A), \ldots, \lambda_n(A).$$

The first objective is to find open sets on which these functions are Pfaffian. First observe that by applying the Implicit Function Theorem (in the analytic category) to

$$det(A - z\mathbb{I})$$

at a point $(A, \lambda_i(A))$, one gets that the λ_i are analytic on W.

We now write $det(A - z\mathbb{I})$ as

$$det(A - z\mathbb{I}) = (-1)^n \left(z^n + c_1(A)z^{n-1} + \cdots + c_n(A)\right),$$

where the c_j are polynomials over $\mathbb{Z}$ in the entries of A.

Now let w be a variable for one of the entries of A. We have

$$\lambda_i(A)^n + c_1(A) \cdot \lambda_i(A)^{n-1} + \cdots + c_n(A) = 0,$$

so

$$n\lambda_i(A)^{n-1}\frac{\partial \lambda_i}{\partial w} + (n-1)\lambda_i(A)^{n-2}c_i(A)\frac{\partial \lambda_i}{\partial w} + \cdots + c_{n-1}(A)\frac{\partial \lambda_i}{\partial w} + \frac{\partial c_1}{\partial w}\lambda_i(A)^{n-1} + \frac{\partial c_2}{\partial w}\lambda_i(A)^{n-2} + \cdots = 0$$

giving

$$\frac{\partial \lambda_i}{\partial w} = -\frac{\left(\sum_j \frac{\partial c_j}{\partial w}\lambda_i(A)^{n-j}\right)}{\frac{\partial}{\partial z}det(A - z\mathbb{I})(\lambda_i)},$$

provided the denominator is not 0. But the denominator is not zero on W, precisely because the discriminant is not zero.

The idea is to write the denominator as a polynomial in w with coefficients rational in A. This is 19th century algebra, but can also be done by elementary model theory.

Lemma 6.1. For each n there are polynomials

$$F_1, \ldots, F_k$$

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

in $\mathbb{Q}[w_1, \ldots, w_n]$, and polynomials

$$\begin{matrix} G_{0,0}, & . & . & . & G_{0,n-1} \\ . & . & . & . & . \\ . & . & . & . & . \\ G_{k,0}, & . & . & . & G_{k,n-1} \end{matrix}$$

in $\mathbb{Q}[w_1, \ldots, w_n]$, such that if

$$H(w_1, \ldots, w_n, z) = z^n + w_1 z^{n-1} + \cdots + w_n$$

and $H'(w_1, \ldots, w_n, z)$ is $\frac{\partial}{\partial z} H(w_1, \ldots, w_n, z)$, then, for any algebraically closed field K of characteristic 0 and $\beta_1, \ldots, \beta_n \in K$ such that the discriminant of $H(\overline{\beta}, z)$ is not 0, then some $F_j(\overline{\beta}) \neq 0$, and in that case

$$\left(\sum_l G_{j,l}(\bar{\beta}) z^l \right) H'(\bar{\beta}, z) \equiv F_j(\bar{\beta}) \text{mod } \ H(\bar{\beta}, z)$$

in K[z].

Proof. We use a straightforward Herbrand argument for witnessing the fact that $H(\overline{\beta}, z)$ and $H'(\overline{\beta}, z)$ are coprime over any field of characteristic 0 if the discriminant is non-zero. □

Now we return to the equation for $\partial\lambda/\partial w$. Using the preceding result, we see that if $F_j(c_1(A), \ldots, c_n(A)) \neq 0$, then

$$\frac{\partial \lambda}{\partial w} = -\left(\sum \frac{\partial c_j}{\partial w} \lambda(A)^{n-j} \right) \cdot \left(\sum G_{j,l}\left(c_1(A), \ldots, c_n(A)\right) w^l \right),$$

showing (since the c_j are polynomials) that each λ is Pfaffian on

$$W \cap \{A : F_j(c_1(A), \ldots, c_n(A)) \neq 0\}.$$

Note: It can certainly happen, for example, for $so(n)$, that all A in the Lie algebra under consideration satisfy

$$F_j(c_1(A), \ldots, c_n(A)) = 0$$

for a fixed j. This could happen if we had chosen $F_j(w_1, \ldots, w_n)$ as w_1, which is an entirely natural choice if one follows the standard procedure for finding the greatest common divisor of two polynomials.

Nothing difficult has been proved in the preceding. The λ are algebraic on W, and we have shown that there is a finite cover of W by (explicitly computable) open sets such that λ is complex Pfaffian on each. Something similar is done in Macintyre (2005b). In fact, a close look at the proof of Theorem 4.1 of that paper yields the following theorem.

Theorem 6.1. On the open sets of the cover given above, the real and imaginary parts of the λ are real Pfaffian.

We are still some distance from our goal of finding open sets U where the entries of $exp(A)$ for $A \in U$ are Pfaffian. The preceding Theorem is our starting point.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

6.4. *exp is Pfaffian*

I will now use the main idea from Barut *et al.* (1994).

By Cayley–Hamilton, we have that for each n there are polynomials $s_{j,n}$, for $1 \leqslant j \leqslant n$, in $\mathbb{Z}[x_1, \ldots, x_n]$, that are symmetric and homogeneous of degree j and such that for any element A of $M_n(K)$, with K any field,

$$A^n = s_1(\lambda_1, \ldots, \lambda_n) \cdot A^{n-1} + \cdots + s_n(\lambda_1, \ldots, \lambda_n),$$

where $\lambda_1, \ldots, \lambda_n$ is a list (which may, of course, include repetitions) of the eigenvalues of A. (Until further notice, this listing has nothing to do with that in the previous subsection). It is convenient to spell out that $s_{j,n}$ is $(-1)^j$ times the usual jth symmetric function of n variables. Note that, by Cayley–Hamilton, the $s_{j,n}(\lambda_1, \ldots, \lambda_n)$ are uniformly polynomial functions of the entries of A. It is convenient for the formulation of Lemma 6.2 and Theorem 6.2 below to set $s_{j,n} = 1$ if $j = 0$.

Now let Δ_n be the discriminant

$$\prod_{i<j} (\lambda_i - \lambda_j),$$

which is a square root of a polynomial in the $s_{j,n}$, for $1 \leqslant j \leqslant n$.

Note that Δ_n is homogeneous in $\lambda_1, \ldots, \lambda_n$ of degree $\frac{n(n-1)}{2}$.

Consider $\Delta_n s_{r,n}$, which is homogeneous of degree $\frac{n(n-1)}{2} + r$. A particular formula for this polynomial will turn out to be crucial for the proofs that exp and log are Pfaffian.

Construe the $\lambda_1, \ldots, \lambda_n$ as independent variables. Note the natural action of S_n on the ring $\mathbb{Z}[\lambda_1, \ldots, \lambda]$ in fixing symmetric polynomials. The alternating group A_n fixes Δ_n, but transpositions send Δ_n to $-\Delta_n$.

The highest power to which any λ_j occurs in (a monomial of) Δ_n is $n-1$, and in $\Delta_n s_{r,n}$ is n. Now, in the expansion of $\Delta_n s_{r,n}$, consider some

$$\overline{\lambda}^{\bar{\gamma}}$$

(with the obvious multi-index notation).

If at least two γ_i in $\bar{\gamma}$ are zero, then any σ switching exactly two corresponding λ fixes the above monomial but sends $\Delta_n s_{r,n}$ to $-\Delta_n s_{r,n}$. It follows that the above monomial does not occur in the expansion of $\Delta_n s_{r,n}$.

A similar argument works for a monomial in which two of the exponents are the same.

This leaves us to consider only monomials of one of two forms: those where there are n non-zero exponents, and those where there are exactly $n-1$.

In the first case, the sum of the exponents is $\frac{n(n+1)}{2}$, so if the coefficient is non-zero, we must have

$$\frac{n(n+1)}{2} = \frac{n(n-1)}{2} + r,$$

so $n = r$.

In the second case, just one m with $1 \leqslant m \leqslant n$ does not occur as a γ_i, and the sum of the exponents is $\frac{n(n+1)}{2} - m$, so if the coefficient is non-zero, we get

$$\frac{n(n+1)}{2} - m = \frac{n(n-1)}{2} + r,$$

so $r + m = n$.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

So $r = n$ gives us the first case, and here some λ occurs to power n (the maximum possible) since otherwise the sum of the exponents is too small. For $1 \leqslant m \leqslant n$, $n-m < n$, so in the second case $r < n$ and here again some λ occurs to power n, since otherwise the sum of the exponents is too small.

Moreover, again by counting sums of exponents, in all cases in a given monomial with non-vanishing coefficient, just one λ occurs to power n. Now just look for such terms and one gets a proof of the following lemma.

Lemma 6.2. For$[1 \leqslant r \leqslant n]$,

$$\begin{aligned}\Delta_n s_{r,n} \;=\; & \theta(r)(\lambda_1^n \cdot \Delta_{n-1}(\lambda_2,\ldots,\lambda_n)s_{r-1,n-1}(\lambda_2,\ldots,\lambda_n)\\ & -\lambda_2^n \cdot \Delta_{n-1}(\lambda_1,\lambda_3,\ldots,\lambda_n)s_{r-1,n-1}(\lambda_1,\lambda_3,\ldots,\lambda_n)\\ & +\lambda_3^n \cdot \Delta_{n-1}(\lambda_1,\lambda_2,\lambda_4,\ldots,\lambda_n)s_{r-1,n-1}(\lambda_1,\lambda_2,\lambda_4,\ldots,\lambda_n) - \ldots.\end{aligned}$$

where $\theta(r) = 1$ if $r = 1$, and $\theta(r) = -1$ if $r \neq 1$.

Note that $\Delta_{n-1}s_{r-1,n-1}$ has degree $\frac{n(n-1)}{2} + n - r$.

Now we extend the above in the style of Barut *et al.* (1994). We had

$$A^n = s_{1,n}A^{n-1} + \cdots + s_{n,n}\mathbb{I}$$

(here we suppress the λ, as we will do also in the following when convenient).

By recursion, we define $s_{j,k,n}$ for $k \geqslant 0$ using

$$A^{n+k} = s_{1,k,n}A^{n-1} + \cdots + s_{n,k,n}\mathbb{I},$$

where $s_{j,0,n} = s_{j,n}$, $1 \leqslant j \leqslant n$, and

$$\begin{aligned} s_{1,k+1,n} &= s_{1,0,n} \cdot s_{1,k,n} + s_{2,k,n}\\ s_{2,k+1,n} &= s_{2,0,n} \cdot s_{1,k,n} + s_{3,k,n}\\ &\cdots\\ s_{n,k+1,n} &= s_{n,0,n} \cdot s_{1,k,n}.\end{aligned}$$

Note that the $s_{j,k,n}$ are homogeneous of degrees $j + k$.

I will now prove for the $s_{j,k,n}$ a natural extension of what was just proved for the $s_{j,n}$.

It will be convenient to use the notation

$$(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n)$$

in the calculations below.

We begin with the $\Delta_n s_{1,k,n}$. For $k = 0$ the preceding lemma gives

$$\begin{aligned}\Delta_n s_{1,0,n} \;=\; & \lambda_1^n\Delta_{n-1}(\lambda_2,\ldots,\lambda_n)\\ & -\lambda_2^n \cdot \Delta_{n-1}(\lambda_1,\lambda_3,\ldots,\lambda_n)\\ & +\lambda_3^n \cdot \Delta_{n-1}(\lambda_1,\lambda_2,\lambda_4,\ldots,\lambda_n) - \ldots.\end{aligned}$$

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

Now

$$\begin{aligned}
\Delta_n s_{1,1,n} &= \Delta_n(s_{1,0,n}s_{1,0,n} + s_{2,0,n}) = \Delta_n s_{1,0,n}s_{1,0,n} + \Delta_n s_{2,0,n} \\
&= \sum_j (-1)^{j+1}\lambda_j^n \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n)(\lambda_1 + \cdots + \lambda_n) \\
&\quad + (-1)\sum_j (-1)^{j+1}\lambda_j^n \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n)s_{1,n}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n) \\
&= \sum_j (-1)^{j+1}\lambda_j^{n+1} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n) \\
&\quad + \sum_j (-1)^{j+1}\lambda_j^{n+1} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n)(\lambda_2 + \cdots + \lambda_n) \\
&\quad - \sum_j (-1)^{j+1}\lambda_j^{n} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n)(\lambda_2 + \cdots + \lambda_n) \\
&= \sum_j (-1)^{j+1}\lambda_j^{n+1} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots \lambda_n).
\end{aligned}$$

I have given this example purely to illustrate the inductive technique used to prove the following important theorem.

Theorem 6.2.

$$\Delta_n s_{l,k,n} = \sum_j (-1)^{j+1}\lambda_j^{n+k} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots) s_{l-1,k,n}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots).$$

Proof. We have it for $k = 0$, and l, n. So, if $l < n$

$$\Delta_n s_{l,k+1,n} = \Delta_n(s_{l,0,n} \cdot s_{1,k,n} + s_{l+1,k,n})$$

while

$$\Delta_n s_{n,k+1,n} = \Delta_n s_{n,0,n} \cdot s_{1,k,n}.$$

In the first case, for $l = 1$

$$\begin{aligned}
\Delta_n s_{1,k+1,n} &= \Delta_n \left(s_{1,0,n} \cdot s_{1,k,n} + s_{2,k,n}\right) \\
&= s_{1,0,n} \left(\sum_j (-1)^{j+1}\lambda_j^{n+k} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots, \lambda_n) \right) \\
&\quad - \sum_j (-1)^{j+1}\lambda_j^{n+k} (\Delta_{n-1} \cdot s_1)(\lambda_1, \ldots, \hat{\lambda}_j, \ldots, \lambda_n) \\
&= \sum_j (-1)^{j+1}\lambda_j^{n+k+1} \Delta_{n-1}(\lambda_1, \ldots, \hat{\lambda}_j, \ldots, \lambda_n).
\end{aligned}$$

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

If $1 < l < n-1$,

$$\begin{aligned}\Delta_n s_{l,k+1,n} &= \Delta_n(s_{l,0,n}s_{1,k,n} + s_{l+1,k,n}) \\ &= s_{l,0,n}\left(\sum_j (-1)^{j+1}\lambda_j^{n+k}\Delta_{n-1}(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n)\right) \\ &\quad -\sum_j (-1)^{j+1}\lambda_j^{n+k}(\Delta_{n-1}\cdot s_l)(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n) \\ &= -\sum_j (-1)^{j+1}\lambda_j^{n+k+1}(\Delta_{n-1}\cdot s_l)(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n)\end{aligned}$$

by the usual cancellation.

If $l = n-1$,

$$\begin{aligned}\Delta_n s_{n-1,k+1,n} &= \Delta_n(s_{n-1,0,n}s_{1,k,n} + s_{n,k,n}) \\ &= s_{n-1,0,n}\left(\sum_j (-1)^{j+1}\lambda_j^{n+k}\Delta_{n-1}(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n)\right) + s_{n,0,n}\Delta_n s_{1,k,n} \\ &= -\sum_j (-1)^{j+1}\lambda_j^{n+k+1}(\Delta_{n-1}\cdot s_{n-2})(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n) \\ &\quad +\sum_j (-1)^{j+1}\lambda_j^{n+k}(\Delta_{n-1}\cdot s_{n-1})(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n) \\ &\quad + s_n\left(\sum_j (-1)^{j+1}\lambda_j^{n+k+1}\Delta_{n-1}\right).\end{aligned}$$

For the final case ($l = n$) we have

$$\Delta_n \cdot s_n = (-1)\sum_j (-1)^{j+1}\lambda_j^n s_{n-1}(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n),$$

and

$$\begin{aligned}\Delta_n \cdot s_{n,1,n} &= \Delta_n \cdot s_n \cdot s_{1,k,n} \\ &= s_n \sum_j (-1)^{j+1}\lambda_j^{n+k}\Delta_{n-1}(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n) \\ &= -\sum_j (-1)^{j+1}\lambda_j^{n+k+1}\Delta_{n-1}(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n).\end{aligned}$$

□

6.5. *The applications*

The preceding result is very useful in getting explicit representations of $\exp(A)$ and, more generally, $f(A)$ for certain analytic functions.

We will shortly have to revert to the special choice (lexicographic ordering) of the eigenvalues $(\lambda_1,\ldots,\hat{\lambda}_j,\ldots,\lambda_n)$, and restrict our matrices A to be in the set W_n where these eigenvalues are distinct. Note that this assumption was not made in the preceding

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

computations. Our perspective will be that the A are parametrised by their entries, and the eigenvalues are (Pfaffian) functions of these entries.

The role of the discriminant (as a function of the eigenvalues, or of the entries) is a bit mysterious. Only in its presence will we have intelligible formulas for the powers of A and thus for the entries of power series in A. We do this simply by adding up the products of the discriminant with the powers of A, and then, if possible, rearranging to obtain a Pfaffian representation.

In particular, for $A \in M_n(\mathbb{C})$

$$\Delta_n \exp(A) = \Delta_n \left(\mathbb{I} + A + \frac{A^2}{2!} + \cdots + \frac{A^{n-1}}{(n-1)!} \right) + \sum_{k \geqslant 0} \frac{1}{(n+k)!} \left(\sum_j (-1)^{j+1} \lambda_j^{n+k} (\Delta_{n-1} \cdot \hat{s}_{l-1})(\lambda_1, \ldots, \hat{\lambda}_j, \ldots, \lambda_n) \right),$$

giving

$$\Delta_n(A) \exp(A) = \Delta_n \left(\mathbb{I} + A + \frac{A^2}{2!} + \cdots + \frac{A^{n-1}}{(n-1)!} \right) + \sum_{k \geqslant 0} \sum_j \frac{(-1)^{j+1} \lambda_j^{n+k}}{(n+k)!} \sum_{l=1}^{n} \theta(l) (\Delta_{n-1} \cdot \hat{s}_{l-1})(\lambda_1, \ldots, \hat{\lambda}_j, \ldots, \lambda_n),$$

where $\theta(l) = 1$ for $l = 1$, and $\theta(l) = -1$ otherwise.

But this gives us $\Delta_n(A) \exp(A)$ as a polynomial in $A, \Delta_n(A)$, the exponentials of the eigenvalues, Δ_{n-1}, and the $\hat{s}_{l-1}$. We already know that most of these ingredients are Pfaffian. The exceptions are the exponentials of the eigenvalues. But we showed earlier that the eigenvalues are Pfaffian, and it is a trivial calculation to show that the exponential of a Pfaffian function is Pfaffian. So $\Delta_n(A) \exp(A)$ is Pfaffian, and it follows that $\exp(A)$ is also on the open set where the discriminant does not vanish.

This is true for A real or complex, and it provides a very general method for showing that the entries of $\exp(A)$, for A in certain Lie algebras $\mathcal{G}$ of matrices, are Pfaffian on certain natural open sets (for example, where $\Delta_n \neq 0$).

Earlier we showed that on the open where the discriminant does not vanish the eigenvalues are complex Pfaffian functions of the entries, and then, by Macintyre (2005b), their real and imaginary parts are Pfaffian. However, as A ranges in $\mathcal{G}$, the $Im(\lambda)$ may have non-compact range, so that neither the real nor imaginary parts of $\exp(\lambda)$ may be Pfaffian. If we restrict A so that the range is compact, there will only be finitely many values of $Im\lambda$ where $\exp(\lambda)$ fails to be Pfaffian, and we will get the entries of $\exp(A)$ to have real and imaginary parts Pfaffian. For this we need the trivial exercise that sin or cos of a real Pfaffian function f are Pfaffian at points where the value of f is not an integer multiple of π.

Note that the zero matrix is a point where the discriminant vanishes, which is a mild nuisance since we would like to work with *exp* on an open neighbourhood of 0.

What about a logarithm? Consider $\log(\mathbb{I} + B)$ for $||B|| < 1$. If $(\mathbb{I} + B) \in SO(n)$ and λ is an eigenvalue of B, then $1 + \lambda = \exp(i\theta)$ for some $\theta \in \mathbb{R}$. This time we apply Theorem 6.2

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

to the Taylor series for $\log(\mathbb{I}+B)$, for $||B|| < 1$, to get that $\log(\mathbb{I}+B)$ is Pfaffian in any open subset U of W where $|\lambda| < 1$ for all eigenvalues of B in U. Now we apply this to $SO(n)$ to get the following theorem.

Theorem 6.3. On the open subset W of $SO(n)$ consisting of the C whose eigenvalues are of the form $\exp(i\theta)$ for $-\frac{\pi}{3} < \theta < \frac{\pi}{3}$, the usual series for $\log(z) = \log(1+(z-1))$ converges to a Pfaffian function of the entries of the matrix C. Moreover, the real and imaginary parts of the entries of $\log(C)$ are Pfaffian in the real and imaginary parts of the entries of C.

I have not tried to extend this 'logarithmic domain' yet, and expect to tackle this problem in a subsequent, lengthier paper. In any case, the above is a natural domain for a logarithm.

I now state a satisfactory decidability result for the orthogonal groups with this logarithm.

Theorem 6.4. Fix n, and consider the structure $\mathscr{M}$ on $\mathbb{R}$ given by the semi-algebraic structure together with the partial logarithm on $SO(n)$ given in Theorem 6.3. Then, assuming Schanuel's Conjecture, $\mathscr{M}$ is decidable. Moreover, there is a constructive model-completeness in terms of the preceding Pfaffian primitives.

Proof (sketch). We sketch a proof – it is a special case of a very general result whose proof will be given elsewhere (but see Macintyre (2005a)). That proof depends on the work in Macintyre and Wilkie (1996) and/or Gabrielov and Vorobjov (2004). The real Pfaffian property of log is crucial.

The main idea of the proof is to show that the structure is constructively model-complete (this does not need Schanuel) and then to use the idea of Macintyre and Wilkie (1996) to give an algorithm for deciding existential formulas. One should pass constructively to the connected components (these are semi-algebraic) of the domain of the above partial logarithm, and work in the system having the Pfaffian primitives on each of the components. □

Note: In later work I will show how to adapt the above to prove the decidability of suitable logarithms on arbitrary compact linear groups.

7. Concluding remarks

I consider the evidence from the above, and from my work on elliptic functions, as very strongly in favour of a conjecture that all Lie groups have some kind of decidability result for a natural logarithm, modulo a plausible conjecture in transcendence theory. On the other hand, the evidence for the corresponding exponentials on the Lie algebras points to analogues of Zilber's Conjectures on the complex exponential.

References

Barut, A. O., Zeni, J. R. and Laufer, A. (1994) The exponential map for the conformal group O(2,4). arXiv:hep-th/9408105v3.

https://doi.org/10.1017/S0960129508006622 Published online by Cambridge University Press

Bertolin, C. (2002) Périodes de 1-motifs et transcendance (in French). *J. Number Theory* **97** (2) 204–221.

Bianconi, R. (1990) *Some Results in the Model Theory of Analytic Functions*, Ph.D. Thesis, Oxford.

Bianconi, R. (1997) Nondefinability Results for Expansions of the Field of Real Numbers by the Exponential Function and by the Restricted Sine Function. *J. Symb. Log.* **62** (4) 1173–1178.

Gabrielov, A. and Vorobjov, N. (2004) Complexity of computations with Pfaffian and Noetherian functions. In: Normal forms, bifurcations and finiteness problems in differential equations. *NATO Sci. Ser. II Math. Phys. Chem.* **137** 211–250.

Macintyre, A. J. (2005a) Elementary Theory of Elliptic Functions 1; The Formalism and a Special Case. In: Edmundo, M., Richardson, D. and Wilkie, A. (eds.) *O-minimal Structures, Proceedings of the RAAG Summer School Lisbon 2003*, Lecture Notes in Real Algebraic and Analytic Geometry, Cuvillier Verlag 104–131.

Macintyre, A. J. (2005b) Some Observations about the Real and Imaginary Parts of Complex Pfaffian Functions. To appear in: Macpherson, D. *et al.* (eds.) Proceedings of INI 2005 Model Theory meeting.

Macintyre, A. J. and Wilkie, A. (1996) On the decidability of the real exponential field. In: Odifreddi, P. (ed.) *Kreiseliana. About and Around Georg Kreisel*, A. K. Peters 441–467.

Serre, J. P. (1965) *Lie Algebras and Lie groups*, W. A. Benjamin.

Wilkie, A. (1996) Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function. *J Amer Math Soc* **9** (4) 1051–1094.

Zilber, B. I. (2004) Pseudo-exponentiation on algebraically closed fields of characteristic zero. *Ann. Pure Appl. Logic* **132** (1) 67–95.