

To appear in:

Rendiconti dell'Istituto Lombardo Accademia di Scienze e Lettere, Classe di Scienze.

ON THE REPRESENTATION OF MAPS BY LIE TRANSFORMS

ANTONIO GIORGILLI

*Dipartimento di Matematica, Via Saldini 50,
20133 — Milano, Italy.*

Sunto. Si riconsidera il problema della rappresentazione di una mappa in una forma adatta all'applicazione dei metodi di forma normale. Si mostra che ricorrendo ai metodi delle serie di Lie e delle trasformate di Lie si può costruire in modo diretto un algoritmo di normalizzazione. Si discute brevemente l'applicazione alla mappa di Schröder–Siegel e alla mappa standard di Chirikov, estendendole al caso di dimensione generica.

Abstract. The problem of representing a class of maps in a form suited for application of normal form methods is revisited. It is shown that using the methods of Lie series and of Lie transform a normal form algorithm is constructed in a straightforward manner. The examples of the Schröder–Siegel map and of the Chirikov standard map are included, with extension to arbitrary dimension.

1. Introduction

Surface transformation as a tool for studying the flow of a system of differential equations have been introduced by Poincaré ([20], Vol. III, ch. XXXIII) and deeply investigated by Birkhoff [2]. The corresponding method of *Poincaré sections* has become classical, and has been widely used also for numerical explorations. A special interesting case is that of a periodic flow, for which the corresponding surface transformation is the time- T shift, T being the period. A natural question is whether a given map may be represented by the time- T flow of a differential system. Positive answers to this question exist for symplectic maps that are perturbations of integrable ones (see, e.g., [17],[18],[22],[13] and the references therein).

In the present paper I will reconsider the problem of giving a suitable representation of a class of maps with a method that is somehow connected to, but does not coincide with, the interpolation by a periodic flow. I will rather use the formalism of

Lie transforms, which allows us to extend to maps the techniques (e.g., calculation of normal forms) that are available for differential equations.

The basic tools exploited here are not new: a short historical account is given at the beginning of sect. 2. However, most practical applications of Lie series and Lie transforms methods are related to differential equations. Typical subjects are numerical integration and coordinate transformations in perturbation theory. In the latter framework, in particular, these methods have proven to be very effective also in investigating the convergence or the asymptotic properties of perturbation series involving small divisors (see, e.g., [7], [8], [9] and references therein) and in devising effective methods for perturbation expansions via algebraic manipulation on computers (see, e.g., [10] and references therein).

It is a well known fact, however, that transporting the analytical methods of normal form theory from differential equations to maps is not straightforward. A common remark is that the case of differential equations is easier to deal with, which justifies the attempts to interpolate a map with the Poincaré section of a flow.

A natural question is whether one can write a map in such a form that transporting the normal form methods that work fine for flows becomes straightforward. Answering this question in general is a major task, of course. However, if one considers a class of maps which are perturbation of integrable ones, then the question can be positively answered.

One may consider as a basic example the Schröder–Siegel problem of iteration of analytic functions, extending it to many dimensions. However, the method developed here applies also to other cases, e.g., perturbations of integrable symplectic maps as considered in the papers quoted at the beginning of this section. As interesting models one may consider: the quadratic map investigated by Hénon [14] and its generalization in higher dimension; the twist map of an invariant annulus investigated by Poincaré [21], Birkhoff [1] and Moser [19]; the well know standard map.

In all these cases the map may be represented as a composition of two maps: an integrable one (e.g., a linear one) and a near the identity perturbation. This is a trivial well known fact, of course. However, representing the integrable map as a Lie series and the perturbation as a Lie transform allows one to implement the normal form theory as a straightforward extension of the methods used for differential equations, provided a suitable formula for the composition of Lie transforms is available. This is what I'm going to illustrate.

The paper is organized as follows. In sect. 2 a short account of the methods based on Lie series and Lie transforms is given, including the representation of a near the identity map and a composition formula. In sect. 3 the main proposition on the representation of a perturbation of an integrable map is proven. In sect. 4 it is shown how to construct a normal form algorithm for the map. The actual construction is worked out for two typical examples. A technical appendix follows.

The theory is developed at a formal level. Some hints on quantitative applications to particular models are included at the end of sect. 4.

2. Basic tools

The concept of Lie serie goes back to Sopus Lie. The use of Lie series in various problems has been widely investigated by Gröbner in a series of papers after 1957 [11]. An accurate exposition with particular emphasis on applications to numerical integration can be found in Gröbner's book [12]. The starting point, already found in Newton's work, is to express the solution of an holomorphic system of differential equations as a series expansion in time. The basic idea is to use the representation via power series as a one parameter near the identity map written in an explicit and useful form.

Lie transform may be considered as a generalization of Lie series, in a sense that will be clarified later. Its usefulness as a tool in perturbation theory has been emphasized by Hori [16] and Deprit [3], who however paid attention in particular to Hamiltonian systems. The underlying idea of Deprit's work is to generate a one parameter family of near the identity coordinate transformation using the flow of a *non autonomous* system. It is quite common in the milieu of Celestial Mechanics to call *Lie Transform* the algorithms proposed by Hori and Deprit, reserving the name *Lie series* to the algorithm based on the flow of an *autonomous* system; thus I follow the tradition.

Actually, several explicit algorithms for Lie transform have been proposed by many authors. A list of references may be found in Henrard's paper [15]. A similar algorithm for Hamiltonian systems has been introduced on a purely algebraic basis in [5]. In this paper I will follow the latter exposition, reformulating it for vector fields. A rigorous treatment may be found in [6], which has been prompted by the need of extending the contents of Gröbner's book to a more general context.

I recall here the definitions and the properties of Lie series and Lie transform working at a formal level, including what is needed in order to develop Lie methods for maps. I will omit most of the proofs, that can be found elsewhere. Furthermore, I will forget about the origin of Lie series as solutions of a system of differential equations, thus paying particular attention to the algebraic aspect.

2.1 Lie derivatives

Let $\mathcal{D} \subset \mathbf{C}^n$ be an open domain endowed with coordinates $x = (x_1, \dots, x_n)$, and let $X(x) = (X_1, \dots, X_n)$ be a complex holomorphic vector field on \mathcal{D} . Let ϕ_X^t denote the time- t flow generated by X .

The Lie derivative of a holomorphic function $f(x)$ at the point x under the flow ϕ_X^t is the new function

$$(1) \quad L_X f = \left. \frac{d}{dt} (\phi_X^t f) \right|_{t=0}$$

where $(\phi_X^t f)(x) = (f \circ \phi_X^t)(x)$.

Similarly, the Lie derivative of a holomorphic vector field v is the new vector field

$$(2) \quad L_X v = \left. \frac{d}{dt} (\phi_X^t v) \right|_{t=0}$$

where $(\phi_X^t v)(x) = (d\phi_X^t)^{-1} (v \circ \phi_X^t)(x)$. It is well known that one has

$$(3) \quad L_X v = \{X, v\} ,$$

where $\{X, v\}$ is the commutator between the vector fields X and v .

The Lie derivative is a linear operator mapping the space of holomorphic functions (respectively holomorphic vector fields) into itself. It is also immediate to check also that the property

$$L_{\alpha X + \beta Y} = \alpha L_X + \beta L_Y$$

holds true, where X, Y are vector fields and α, β are complex numbers. Further useful properties are the following. For two functions f, g the Leibniz rule applies, namely

$$L_X(fg) = fL_Xg + gL_Xf, \quad L_X^s(fg) = \sum_{j=0}^s \binom{s}{j} (L_X^j f) (L_X^{s-j} g), \quad s \geq 1.$$

For two vector fields v, w one has

$$L_X\{v, w\} = \{L_Xv, w\} + \{v, L_Xw\}.$$

Finally, denoting by $[L_X, L_Y] = L_XL_Y - L_YL_X$ the commutator between the Lie derivatives with respect to the vector fields X, Y one has

$$[L_X, L_Y] = L_{\{X, Y\}}.$$

The latter two properties are just different writings of Jacobi's identity for the commutator between vector fields, namely $\{X, \{v, w\}\} + \{v, \{w, X\}\} + \{w, \{X, v\}\} = 0$. In particular, if the vector fields X, Y do commute, namely if $\{X, Y\} = 0$ then we have $[L_X, L_Y] = 0$.

It will also be useful to write the explicit expression of the Lie derivatives in coordinates. For a function f one has

$$(4) \quad L_X f = \sum_{j=1}^n X_j \frac{\partial}{\partial x_j} f.$$

For a vector field one gets the expression of the commutator, namely

$$(L_X v)_j = \sum_{l=1}^n \left(X_l \frac{\partial v_j}{\partial x_l} - v_l \frac{\partial X_j}{\partial x_l} \right),$$

where the l.h.s. is the j -th component of the vector field $L_X v$.

2.2 Lie series and Lie transform

Let again $\mathcal{D} \subset \mathbf{C}^n$ be an open domain, and let $X = (X_1, \dots, X_n)$ be a holomorphic vector field. The Lie series operator is defined as

$$(5) \quad \exp(L_X) = \sum_{s \geq 0} \frac{1}{s!} L_X^s$$

A family of near the identity transformations depending on a parameter ε may be constructed as

$$(6) \quad y = \exp(\varepsilon L_X) x$$

or, in explicit form for the coordinates,

$$(7) \quad y_j = \exp(\varepsilon L_X)x_j = x_j + \varepsilon X_j(x) + \frac{\varepsilon^2}{2}L_X X_j(x) + \dots, \quad j = 1, \dots, n.$$

The Lie transform is introduced as follows. Let $X = \{X_1, X_2, \dots\}$ be a sequence of holomorphic vector fields, that I will call the *generating sequence*; here the lower index labels the element of the sequence, not the component of the field in coordinates, that will be denoted, e.g., by $X_{1,j}$. The Lie transform operator is defined as

$$(8) \quad T_X = \sum_{s \geq 0} E_s^X,$$

where the sequence E_s^X of linear operators is recursively defined as

$$(9) \quad E_0^X = 1, \quad E_s^X = \sum_{j=1}^s \frac{j}{s} L_{X_j} E_{s-j}^X.$$

The superscript in E^X is introduced in order to specify which sequence of vector fields is intended. However, I will remove it when unnecessary. By letting the sequence to have only one vector field different from zero, e.g., $X = \{0, \dots, 0, X_k, 0, \dots\}$ it is easily seen that one gets $T_X = \exp(L_{X_k})$.

Writing $\varepsilon^s X_s$ in place of X_s a one parameter family of near the identity transformations may be defined as

$$(10) \quad y = T_X x,$$

i.e., in coordinates,

$$(11) \quad y_j = x_j + \varepsilon X_{1,j}(x) + \varepsilon^2 \left[\frac{1}{2} L_{X_1} X_{1,j}(x) + X_{2,j}(x) \right] + \dots, \quad j = 1, \dots, n.$$

Here I used the ε expansion in order to make clear the connection with classical methods based on power expansion in a small parameter. However, in many cases it is convenient to just consider the vector field X_s to be “small of order s ” in some appropriate sense (e.g., using a norm) the order being implicit in the label of the field. Adding the powers of ε the reader will easily check that every term in the definition of the operator E_s^X carries a factor ε^s , so that E_s^X is of order s . An equivalent procedure is to determine the order of E_s^X as the sum of the indices of $L_{X_j} E_{s-j}^X$, which is s indeed. In the rest of the paper I will remove the parameter, unless it has a particular meaning. If the reader gets confused, he or she may just rewrite a formula by adding the powers of ε , check that everything is put in the correct order, and then set $\varepsilon = 1$.

The Lie series and Lie transform are linear operators acting on the space of holomorphic functions and of holomorphic vector fields on the domain \mathcal{D} . They preserve products between functions and commutators between vector fields, i.e., if f, g are functions and v, w are vector fields then one has

$$(12) \quad T_X(fg) = T_X f \cdot T_X g, \quad T_X\{v, w\} = \{T_X v, T_X w\}.$$

Here, replacing T_X with $\exp(L_X)$ gives the corresponding property for Lie series. Moreover both operators are invertible. The inverse of $\exp(L_X)$ is $\exp(-L_X)$, which is a natural fact if one recalls the origin of Lie series as a solution of an autonomous system of differential equations. The inverse of T_X takes the form

$$(13) \quad \begin{aligned} (T_X)^{-1} &= \sum_{s \geq 0} G_s^X, \\ G_0^X &= 1, \quad G_s^X = - \sum_{j=1}^s \frac{j}{s} G_{s-j}^X L_{X_j}. \end{aligned}$$

I come now to a remarkable property which justifies the usefulness of Lie methods in perturbation theory. I will adopt the name *exchange theorem* introduced by Gröbner. Let f be a function and v be a vector field. Consider the near the identity transformation (10) (or (6) for Lie series) and denote by J the differential of the transformation (10) (or (6)), namely, in coordinates, the jacobian matrix with elements $J_{j,k} = \frac{\partial y_j}{\partial x_k}$. Then one has

$$(14) \quad f(y)|_{y=T_X x} = (T_X f)(x), \quad J^{-1}v(y)|_{y=T_X x} = (T_X v)(x).$$

This result should be interpreted as follows. On the l.h.s. of the equalities there is the transformed function (resp. vector field) calculated via the usual method of substitution of variables. The r.h.s. is the transformed function (resp. vector field) via the Lie transform, where the variables are renamed as x . The claim is that both operations give the same result. The remarkable fact is that if one uses expansions order by order, then the r.h.s. gives it in a straightforward way for the transformed function (resp. vector field) in terms of Lie derivatives, thus requiring only operations that are easily performed, e.g., via algebraic manipulation on computers. Obtaining the same result via substitution of variables is a definitely longer process, unless one stops the expansion at very low order. This claim may appear a little obscure, but it will turn out to be immediately evident if one writes the expansions as power series in a parameter.

The proof of the identities (14) may be worked out using the algebraic properties of T_X . For, a holomorphic function (or vector field) may be expanded in power series of the variables, and by exploiting the linearity and the preservation of product the operator T_X can be moved from the variables (the substitution) to the whole function. This justifies the name “exchange theorem”, since the symbol of the Lie transform operator is exchanged with the symbol of the function.

2.3 Representation of a near the identity transformation

Let now a near the identity transformation be given in the form

$$(15) \quad y_j = x_j + \varepsilon \varphi_{1,j}(x) + \varepsilon^2 \varphi_{2,j}(x) + \dots, \quad j = 1, \dots, n,$$

where $\varphi_{1,j}(x), \varphi_{2,j}(x), \dots$ are holomorphic functions.

Proposition 1: Any transformation of the form (15) may be written as a Lie transform of the coordinates $y = T_X x$ with the generating sequence

$$(16) \quad X_{1,j} = \varphi_{1,j}, \quad X_{r,j} = \varphi_{r,j} - \sum_{k=1}^{r-1} \frac{k}{r} L_{X_k} E_{r-k} x_j, \quad j = 1, \dots, n, \quad r > 2.$$

In view of this proposition it may appear that Lie transform is more general and more attractive than Lie series, since the latter can not represent any near the identity transformation. However, a similar result may be obtained if one makes use of the *composition of Lie series*.

Let again $\{X_1, X_2, \dots\}$ be a sequence of holomorphic vector fields. Consider the sequence of transformations $\{S_X^{(0)}, S_X^{(1)}, S_X^{(2)}, \dots\}$ recursively defined as

$$(17) \quad S_X^{(0)} = 1, \quad S_X^{(r)} = \exp(L_{X_r}) \circ S_X^{(r-1)}.$$

We may well consider in formal sense the limit

$$(18) \quad S_X = \dots \circ \exp(L_{X_r}) \circ \dots \circ \exp(L_{X_2}) \circ \exp(L_{X_1})$$

as an operator obtained by composition of Lie series. A noticeable fact, pointed out by Fassò [4], is that one has

$$(19) \quad S_X = \sum_{s \geq 0} : E_s^X :$$

where $: E_s^X :$ denotes a reordering of E_s^X in the following sense: a composition of r Lie derivatives $L_{X_{k_1}} \circ L_{X_{k_2}} \circ \dots \circ L_{X_{k_r}}$ is reordered as

$$: L_{X_{k_1}} \circ L_{X_{k_2}} \circ \dots \circ L_{X_{k_r}} : = L_{X_{\sigma_1(k)}} \circ L_{X_{\sigma_2(k)}} \circ \dots \circ L_{X_{\sigma_r(k)}}$$

where $\sigma(k) = \{\sigma_1(k), \dots, \sigma_r(k)\}$ is any permutation of $k = \{k_1, \dots, k_r\}$ such that $\sigma_1(k) \geq \sigma_2(k) \geq \dots \geq \sigma_r(k)$. E.g., $: L_{X_1} L_{X_2} : = L_{X_2} L_{X_1}$.

Proposition 2: Any transformation of the form (15) may be represented via a composition of Lie series of the form $y = S_X x$ with a generating sequence $\{X_1, X_2, \dots\}$ that can be explicitly determined with a recursive procedure.

As the reader will notice, an explicit expression for the vector fields is missing in the statement. Such an expression may be produced exploiting (19), but it turns out to be quite useless, because it has a non recursive form. However, a recursive procedure for determining X_1, X_2, \dots may be easily constructed by trying the first steps. Setting $X_{1,j} = \varphi_{1,j}$ one has

$$y_j - \exp(L_{X_1}) x_j = \varphi_{2,j} - \frac{1}{2} L_{X_1} X_{1,j} + \dots$$

where the dots denote terms at least of third order. Setting $X_{2,j} = \varphi_{2,j} - \frac{1}{2} L_{X_1} X_{1,j}$ one gets that $y_j - \exp(L_{X_2}) \circ \exp(L_{X_1}) x_j$ starts with terms at least of third order, which are used in order to determine X_3 , and so on. Such a procedure is easily implemented, e.g., via computer algebra, and it is in fact also the scheme of proof of the proposition. I add just a remark concerning the actual use of propositions 1 and 2. Apparently, it

seems unreasonable to do extra work in order to calculate the generating sequence of the transformation. However, if one wants to transform either a function or a vector field, pushing the calculation at high orders, then using the exchange theorem turns out to be definitely more effective than performing a substitution.

2.4 Composition formulæ

I begin with a formula for the commutation between two Lie transforms. *Let X, Y be generating sequences. Then one has*

$$(20) \quad T_X \circ T_Y = T_W \circ T_X ,$$

with the generating sequence $W = \{T_X Y_1, T_X Y_2, \dots\}$. Here, the Lie transform T_X may be replaced by $\exp(L_X)$ in case X is a vector field, and similarly for Y .

Generally speaking the formula above does not seem very interesting because the vector field W turns out to be itself a series. However, it is useful in some case, typically when one is able to put W in a manageable form by explicitly calculating the sum of the series. It will be used later, in sect. 3.2.

It is instead more interesting to observe that since $y = T_X \circ T_Y x$ is a near the identity transformation then in view of proposition 1 there exists a generating sequence Z such that $y = T_Z x$. The following proposition gives an explicit expression for Z .

Proposition 3: *Let X, Y be generating sequences. Then one has $T_X \circ T_Y = T_Z$ where Z is the generating sequence recursively defined as*

$$(21) \quad Z_1 = X_1 + Y_1 , \quad Z_s = X_s + Y_s + \sum_{j=1}^{s-1} \frac{j}{s} E_{s-j}^X Y_j .$$

The proof requires a long sequence of tedious calculations that can hardly be found in previous papers. Thus I include it in appendix A. Composition formulæ for any combination of Lie series and Lie transforms are easily obtained by suitably elaborating formula (21). E.g., if X, Y are vector fields then one has $\exp(L_X) \circ \exp(L_Y) = T_W$, where $W = \{W_1, W_2, \dots\}$ is the generating sequence

$$(22) \quad W_1 = X + Y , \quad W_s = \frac{1}{s!} L_X^{s-1} Y .$$

The latter formula reminds the well known Baker–Campbell–Hausdorff composition of exponentials. The difference is that the result is expressed as a Lie transform instead of an exponential.

3. Representation of a map by a composition of Lie Transforms

I shall consider two cases. The first one is a map in a neighbourhood of an equilibrium, which may be expanded in Taylor series as

$$(23) \quad z' = \Lambda z + v_1(z) + v_2(z) + \dots , \quad z \in \mathbf{C}^n$$

where Λ is a $n \times n$ complex matrix and $v_s(z)$ is a homogeneous polynomial of degree $s + 1$. The second example is a real analytic map

$$(24) \quad \begin{aligned} \varphi' &= \varphi + \omega(I) + \varepsilon f_1(\varphi, I) + \varepsilon^2 f_2(\varphi, I) + \dots \\ I' &= I + \varepsilon g_1(\varphi, I) + \varepsilon^2 g_2(\varphi, I) + \dots \end{aligned}$$

where $(\varphi, I) \in \mathbf{T}^n \times \mathcal{G}$, with $\mathcal{G} \subset \mathbf{R}^m$ open, and ε is a small perturbation parameter.

3.1 The unperturbed map

By unperturbed map I mean here (as usual) either the linear part of (23) or the map (24) with $\varepsilon = 0$, which is a Kronecker map on a family of invariant tori parameterized by the actions I .

Let me start with the linear part $z' = \Lambda z$ of the map (23). Let $\Lambda = e^{\mathbf{A}}$ with a complex $n \times n$ matrix \mathbf{A} . For the vector field $\mathbf{A}z$ one has $L_{\mathbf{A}z}z = \mathbf{A}z, \dots, L_{\mathbf{A}z}^s z = \mathbf{A}^s z$, and so also $\exp(L_{\mathbf{A}z})z = \sum_{s \geq 0} \frac{1}{s!} \mathbf{A}^s z = e^{\mathbf{A}}z$. Using the exchange theorem we may transform a function $f(z')$ and a vector field $v(z')$ as

$$(25) \quad \begin{aligned} \exp(L_{\mathbf{A}z})f(z) &= f(z') \Big|_{z'=e^{\mathbf{A}}z} = f(\Lambda z) , \\ \exp(L_{\mathbf{A}z})v(z) &= e^{-\mathbf{A}}v(z') \Big|_{z'=\exp(L_{\mathbf{A}z})z} = \Lambda^{-1}v(\Lambda z) . \end{aligned}$$

The simplest case occurs when $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ is a diagonal matrix.

I point out that here the exchange theorem is used only in order to represent a transformation of a function and of a vector field as the action of a Lie series operator. Of course, in this case everybody would perform the transformations via a direct substitution as in the last member of the formulæ above, since this is actually the sum of the Lie series in closed form and requires no further expansion.

I come now to considering the map (24), that for $\varepsilon = 0$ writes

$$(26) \quad \varphi' = \varphi + \omega(I) , \quad I' = I .$$

Introducing the $(n + m)$ -dimensional vector field $\Omega = (\omega(I), 0)$ one immediately gets

$$L_{\Omega}(\varphi, I) = (\omega(I), 0) , \quad L_{\Omega}^s(\omega(I), 0) = (0, 0) \quad \text{for } s > 1 .$$

Thus, the map may be rewritten as $(\varphi', I') = \exp(L_{\Omega})(\varphi, I)$. Using again the exchange theorem as above we may transform a function $f(\varphi, I)$ or a $(n + m)$ -dimensional vector field $v(\varphi, I)$ as

$$(27) \quad \begin{aligned} \exp(L_{\Omega})f(\varphi, I) &= f(\varphi', I') \Big|_{(\varphi', I')=\exp(L_{\Omega})(\varphi, I)} = f(\varphi + \omega(I), I) , \\ \exp(L_{\Omega})v(\varphi, I) &= \mathbf{J}^{-1}v(\varphi', I') \Big|_{(\varphi', I')=\exp(L_{\Omega})(\varphi, I)} = \mathbf{J}^{-1}v(\varphi + \omega(I), I) , \end{aligned}$$

where \mathbf{J} and its inverse \mathbf{J}^{-1} are the jacobian block matrices

$$\mathbf{J} = \begin{pmatrix} \mathbf{I}_n & \mathbf{B} \\ 0 & \mathbf{I}_m \end{pmatrix} , \quad \mathbf{J}^{-1} = \begin{pmatrix} \mathbf{I}_n & -\mathbf{B} \\ 0 & \mathbf{I}_m \end{pmatrix} , \quad \mathbf{B} = \left\{ \frac{\partial \omega_j}{\partial I_l} \right\}_{1 \leq j \leq n, 1 \leq l \leq m} ,$$

\mathbf{I}_s denoting the $s \times s$ identity matrix.

3.2 The full map

The maps (23) or (24) considered in the last section are just interesting examples. In more general terms one can consider an unperturbed map $x' = f_0(x)$ which has some nice properties and can be represented as a Lie transform $x' = \exp(L_X)$ with an appropriately defined vector field X . In view of the particular role of the latter map I will denote $R = \exp(L_X)$.

Then one considers a perturbed map

$$(28) \quad x' = Rx + f_1(x) + f_2(x) + \dots$$

where $f_s(x)$ is of order s in some reasonable sense. E.g., $f_s(x)$ should be replaced by $v_s(z)$ in (23), and by $\varepsilon f_s(\varphi, I)$ in (24), with the obvious change of the symbols for the coordinates.

Proposition 4: *Consider the map (28) where R is a Lie series operator. Then there exist generating sequences of vector fields $V(x) = \{V_1(x), V_2(x), \dots\}$ and $W(x) = \{W_1(x), W_2(x), \dots\}$ with $W_s = RV_s$ such that one has both*

$$(29) \quad x' = R \circ T_V x \quad \text{and} \quad x' = T_W \circ R x$$

Proof. Using the linearity of the Lie series operator R rewrite the map (29) as

$$x' = R(x + R^{-1}(f_1 + f_2 + \dots)) .$$

This is a representation of the map as the composition of two operations, namely

$$\tilde{\varphi}(x) = x + R^{-1}(f_1 + f_2 + \dots) , \quad x' = R\tilde{\varphi}(x) .$$

In view of proposition 1 one may determine the generating sequence $V(x) = \{V_1(x), V_2(x), \dots\}$ such that $\tilde{\varphi}(x) = T_V x$. Applying R to both members and using linearity one readily gets $x' = R \circ T_V x$, namely the first of (29). Using the identity (20) of proposition 3, which clearly applies also to the Lie series operator R , one gets $R \circ T_V = T_W \circ R$ with W as in the statement, which gives the second of (29). *Q.E.D.*

3.3 Conjugating maps

I come now to the following question. *Let two maps*

$$(30) \quad x' = T_W \circ R x , \quad y' = T_Z \circ R y$$

be given, where R is an invertible Lie series operator and $W = \{W_1, W_2, \dots\}$, $Z = \{Z_1, Z_2, \dots\}$ are generating sequences. To find whether the maps are conjugated by a holomorphic near the identity transformation

$$(31) \quad y = x + \varphi_1(x) + \varphi_2(x) + \dots .$$

Using the same operator R in both maps means only that the unperturbed maps are trivially conjugated. In view of propositions 1 and 2 it is natural to consider the transformation (31) as generated by either a Lie transform or a Lie series.

Proposition 5: Let $X = \{X_1, X_2, \dots\}$ be a generating sequence of the near the identity transformation $y = T_X x$. Then the maps (30) are conjugated if

$$(32) \quad T_W \circ T_{RX} = T_X \circ T_Z .$$

More explicitly, the following relations must be satisfied:

$$(33) \quad DX_1 = Z_1 - W_1 , \quad D = R - 1$$

$$(34) \quad DX_s = Z_s - W_s + \sum_{j=1}^{s-1} \frac{j}{s} (E_{s-j}^X Z_j - E_{s-j}^W RX_j) , \quad s > 1 .$$

Proof. In the l.h.s. of the map $y' = T_Z \circ R y$ set $y' = T_X x'$ and then substitute $x' = T_W \circ R x$. This must be the same as substituting $y = T_X x$ in the r.h.s. of the map. Thus the identity

$$T_X x' \Big|_{x'=T_W \circ R x} = T_Z \circ R y \Big|_{y=T_X x}$$

must be true, and by the exchange theorem this gives $T_W \circ R \circ T_X = T_X \circ T_Z \circ R$. In view of (20) one has $R \circ T_X = T_{RX} \circ R$, so that (32) is readily found in view of the invertibility of R . Apply now proposition 3 to both members of the latter equality. By the first of (21) we get $W_1 + RX_1 = X_1 + Z_1$, namely (33). For $s > 1$ we get

$$W_s + RX_s + \sum_{j=1}^{s-1} \frac{j}{s} E_{s-j}^W RX_j = X_s + Z_s + \sum_{j=1}^{s-1} \frac{j}{s} E_{s-j}^X Z_j ,$$

from which (34) readily follows. Q.E.D.

A similar result for Lie series holds true if in (32) one replaces the Lie transform T_X with the infinite composition of Lie series S_X as given by (18). However, (34) must be restated in a more elaborated manner, proceeding step by step. Let me say that the maps (32) are conjugated up to order r in case there exists a finite generating sequence $X = \{X_1, \dots, X_r\}$ such that the transformation $y = S^{(r)} x$ makes the difference between the maps to be of order higher than r , i.e.,

$$S_X^{(r)} x' \Big|_{x'=T_W \circ R x} - T_Z \circ R y \Big|_{y=S_X^{(r)} x} = \mathcal{O}(r+1) ,$$

where $S^{(r)}$ is defined by (17). The maps are trivially conjugated up to order r if the generating sequences W, Z coincide up to order r , i.e., $W_1 = Z_1, \dots, W_r = Z_r$. For this implies $T_W x - T_Z x = \mathcal{O}(r+1)$.

Proposition 6: Let the generating sequences of the maps (30) coincide up to order $r-1$ and let X_r be a vector field of order r generating the near the identity transformation $y = \exp(L_{X_r} x)$. Then the maps are conjugated up to order r if

$$(35) \quad T_W \circ \exp(L_{RX_r}) = \exp(L_{X_r}) \circ T_Z .$$

More explicitly, the following relations must be satisfied:

$$(36) \quad DX_r = Z_r - W_r, \quad D = R - 1;$$

$$(37) \quad Z_s - W_s = \frac{r}{s} E_{s-r}^W R X_r - \sum_{j=1}^{\lfloor (s-1)/r \rfloor} \frac{s-jr}{s \cdot j!} L_{X_r}^j Z_{s-jr} \quad \text{for } s > r.$$

The proof of (35) is a straightforward adaptation of that proposition 5 using a generating sequence X with all elements zero except X_r . For, in this case one has $T_X = \exp(L_{X_r})$, so that $E_s^X = 0$ if s is not a multiple of r . Removing from (34) all vanishing term, with some patience one obtains (37).

4. Normal form algorithm

A standard and useful tool in perturbation theory is the construction of a normal form for either a map or a system of differential equations. This is a classical and widely investigated problem, so I will limit the discussion to indicating how the known methods may be revisited within the scheme of representation of maps presented in this paper. Constructing a normalization algorithm for maps is indeed an easy matter in view of the results of sect. 3.3. Actually, two different algorithms may be devised, the first one based on Lie transform, the second one based on composition of Lie series.

4.1 A general formulation

Let me start with the Lie transform. The key point is that formula (32) must be considered as an equation for the generating sequence X , which must be so determined that the transformed sequence Z has some nice property that characterizes it as being in normal form. Look now at equations (34), which are just a rewriting of (32) order by order. The reader will immediately see that for every $s > 1$ one has to solve recursively an equation of the form

$$(38) \quad DX_s + Z_s = \Psi_s$$

where Ψ_s is known, since it is determined by X_1, \dots, X_{s-1} and W_1, \dots, W_{s-1} , which are known. Thus, the problem is only that the prescription that Z has a normal form should be imposed so that eq. (38) may be solved for X_s and Z_s . This is the standard problem in normal form theory for both maps and differential equations. If the process can be worked out, at least formally, then the generating sequence X produces a coordinate transformation $y = T_X x$, with inverse $x = T_X^{-1} y$, such that the map in the new coordinates writes $y' = T_Z \circ R y$.

The algorithm based on composition of Lie series appears to be more elaborated, since it requires using proposition 6 as an iteration step. Precisely, one constructs an infinite sequence $\{W^{(r)}\}_{r \geq 0}$ of generating sequences, with $W^{(0)} = W$, and a generating sequence $X = \{X_1, X_2, \dots\}$ such that for every $r > 0$ the generating sequences $W^{(r-1)}$ and $W^{(r)}$ coincide up to order $r - 1$, and $Z_1 = W_1^{(r)}, \dots, Z_r = W_r^{(r)}$ are in normal form and do not change with the next iteration. Proposition 6 shows how to determine

X_r and $W_r^{(r)} = Z_r$ by solving eq. (36) with $W_r^{(r-1)}$ in place of W_r . Then the whole generating sequence $W^{(r)}$ is constructed as given by (37), putting $W^{(r-1)}$ in place of W and $W^{(r)}$ in place of Z . The normal form is thus determined step by step.

At first sight, the reader may think that this is a too complicated process. However, by implementing the algorithm using algebraic manipulation he or she will realize that there is no substantial increase of complexity with respect to the algorithm based on a single Lie transform, and that in some cases the composition of Lie series may even be more effective.

4.2 Back to examples

Let me illustrate how a normal form may be constructed for the examples (23) and (24). Actually, this means that I should explain how to characterize the normal form and how to solve eq. (38).

Assume that the matrix Λ in (23) has a diagonal form, namely $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, and define $\lambda_j = e^{\mu_j + i\omega_j}$. This is a n -dimensional version of the problem of iteration of analytic maps investigated by Schröder [23], who gave the formal solution for the case $n = 1$. Clearly one has

$$R = \exp(L_{\Omega x}) , \quad \Omega = \text{diag}(\mu_1 + i\omega_1, \dots, \mu_n + i\omega_n) ,$$

and applying the method of proposition 4 the generating sequence W may be determined so that W_s is a homogeneous polynomial of degree $s + 1$. Following Schröder one tries to conjugate the map to its linear part. This means that the normal form of the map should be $\zeta' = \Lambda\zeta$, which in the Lie transform representation means that one wants $Z = \{0, 0, \dots\}$, the null sequence. Thus, according to (38), the generating sequence is determined by solving for X_s the equation

$$(39) \quad DX_s = \Psi_s , \quad D = \exp(L_{\Omega x}) - 1 ,$$

where $\Psi_s(x_1, \dots, x_n)$ is a homogeneous polynomial of degree $s+1$. The relevant property is that D is diagonal on the basis of monomials $x^k \mathbf{e}_j = x_1^{k_1} \dots x_n^{k_n} \mathbf{e}_j$, where $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ is the canonical basis of \mathbf{C}^n . For, in view of the second of (25) one has

$$D x^k \mathbf{e}_j = (e^{\langle k, \mu + i\omega \rangle - \mu_j - i\omega_j} - 1) x^k \mathbf{e}_j .$$

Thus determining the vector field X_s is an easy matter if none of the eigenvalues of D is zero. For, writing the homogeneous polynomial vector field as

$$\Psi_s = \sum_{j=1}^n \mathbf{e}_j \sum_{|k|=s+1} \psi_{j,k} x^k$$

the solution of (39) is readily found to be

$$X_s = \sum_{j=1}^n \mathbf{e}_j \sum_k \frac{\psi_{j,k}}{e^{\langle k, \mu + i\omega \rangle - \mu_j - i\omega_j} - 1} x^k .$$

Thus the map may be formally linearized if the *nonresonance condition*

$$(40) \quad e^{\langle k, \mu + i\omega \rangle - \mu_j - i\omega_j} \neq 1 \text{ for } k \in \mathbf{Z}_+^n , |k| > 1 \text{ and } j = 1, \dots, n$$

is fulfilled. In the case $n = 1$ considered by Schröder a resonance may occur only if λ is a root of the unity.

If the resonance condition is violated then a different definition of normal form must be introduced. Precisely, the space \mathcal{P}^s of homogeneous polynomials of any degree s splits into two complementary subspaces

$$\mathcal{N}^s = \mathbf{D}^{-1}\{0\} , \quad \mathcal{R}^s = \mathbf{D}(\mathcal{P}^s) ,$$

namely the kernel of \mathbf{D} and the image of \mathcal{P}^s through \mathbf{D} . For, \mathbf{D} maps \mathcal{P}^s onto itself, so that both \mathcal{N}^s and \mathcal{R}^s are subspaces of \mathcal{P}^s , and \mathbf{D} is diagonal. Then the operator \mathbf{D} may be uniquely inverted on \mathcal{R}^s . Thus eq. (38) may be solved by splitting $\Psi_s = \Psi_s^{\mathcal{N}} + \Psi_s^{\mathcal{R}}$, with obvious meaning of the superscripts, and setting

$$Z_s = \Psi_s^{\mathcal{N}} , \quad X_s = \mathbf{D}^{-1}\Psi_s^{\mathcal{R}} , \quad X_s \in \mathcal{R}^s .$$

The latter condition makes the solution unique. The procedure thus described is a standard one in normal form theory. Different solutions may be considered, of course, depending on what one is looking for.

The convergence of the transformation to normal form in the non resonant case has been proved by Siegel [24] for the case $n = 1$ under the additional hypothesis that λ satisfies a diophantine condition. The work of Siegel represents a milestone for the problem of convergence of perturbation series with small divisors.

Let me now come to the model (24). In view of the particular form of the map it is convenient to represent the vector fields separating, so to say, the φ component from the I component by writing $\begin{pmatrix} X \\ Y \end{pmatrix}$ in place of X , where $X(\varphi, I)$ and $Y(\varphi, I)$ are a n -dimensional and a m -dimensional vector function, respectively. The explicit form of the commutator is written as

$$(41) \quad \left\{ \begin{pmatrix} X \\ Y \end{pmatrix}, \begin{pmatrix} V \\ W \end{pmatrix} \right\} = \left(\sum_{l=1}^n \left(X_l \frac{\partial V}{\partial \varphi_l} - V_l \frac{\partial X}{\partial \varphi_l} \right) + \sum_{l=1}^m \left(Y_l \frac{\partial V}{\partial I_l} - W_l \frac{\partial X}{\partial I_l} \right) \right) \\ \left(\sum_{l=1}^n \left(X_l \frac{\partial W}{\partial \varphi_l} - W_l \frac{\partial Y}{\partial \varphi_l} \right) + \sum_{l=1}^m \left(Y_l \frac{\partial W}{\partial I_l} - W_l \frac{\partial Y}{\partial I_l} \right) \right)$$

Recall also that in this case with the notation above one has

$$\mathbf{D} = \exp(L_\Omega) - 1 , \quad \Omega = \begin{pmatrix} \omega(I) \\ 0 \end{pmatrix} .$$

In the r.h.s. of (38) we may expand $\Psi_s(\varphi, I)$ in Fourier series as

$$\Psi_s = \begin{pmatrix} \sum_{k \in \mathbf{Z}^m} \alpha_k(I) e^{i\langle k, \varphi \rangle} \\ \sum_{k \in \mathbf{Z}^m} \beta_k(I) e^{i\langle k, \varphi \rangle} \end{pmatrix}$$

with known coefficients $\alpha_k(I)$ and $\beta_k(I)$. Using a similar expansion for

$$X_s = \begin{pmatrix} \sum_{k \in \mathbf{Z}^m} c_k(I) e^{i\langle k, \varphi \rangle} \\ \sum_{k \in \mathbf{Z}^m} d_k(I) e^{i\langle k, \varphi \rangle} \end{pmatrix} ,$$

in view of (27) the action of the operator \mathbf{D} is given by

$$\mathbf{D}X_s = \begin{pmatrix} \sum_{k \in \mathbf{Z}^n} (e^{i\langle k, \omega(I) \rangle} - 1) c_k(I) e^{i\langle k, \varphi \rangle} - \sum_{k \in \mathbf{Z}^n} e^{i\langle k, \omega(I) \rangle} \mathbf{B}d_k(I) e^{i\langle k, \varphi \rangle} \\ \sum_{k \in \mathbf{Z}^n} (e^{i\langle k, \omega(I) \rangle} - 1) d_k(I) e^{i\langle k, \varphi \rangle} \end{pmatrix} .$$

Thus, one would be tempted to solve eq. (38) by setting

$$(42) \quad Z_s = \begin{pmatrix} \alpha_0(I) \\ \beta_0(I) \end{pmatrix}, \quad X_s = \begin{pmatrix} \sum_{k \in \mathbf{Z}^n \setminus \{0\}} \left[\frac{\alpha_k(I)}{e^{i\langle k, \omega(I) \rangle} - 1} + \frac{e^{i\langle k, \omega(I) \rangle} \mathbf{B} \beta_k(I)}{(e^{i\langle k, \omega(I) \rangle} - 1)^2} \right] e^{i\langle k, \varphi \rangle} \\ \sum_{k \in \mathbf{Z}^n \setminus \{0\}} \frac{\beta_k(I)}{e^{i\langle k, \omega(I) \rangle} - 1} e^{i\langle k, \varphi \rangle} \end{pmatrix}.$$

However, some denominator could vanish at some point $I \in \mathcal{G}$, the action's domain, or at least become very small. This is indeed the classical problem of small divisors in Celestial Mechanics, which was well known to, e.g., Lagrange and Laplace.

Assume for a moment that no divisor actually vanishes. This is true, e.g., if $\omega \in \mathbf{R}^n$ is a constant vector and the non resonance condition $e^{i\langle k, \omega \rangle} \neq 1$ for $k \in \mathbf{Z}^m \setminus \{0\}$ is satisfied. Then the generating sequence $Z(I)$ is independent of the angle variables φ . Thus one has

$$L_{Z_s(I)} \varphi = Z_s(I), \quad L_{Z_s}^r \varphi = 0 \text{ for } r > 1,$$

and the map in normal form is written as

$$\varphi' = \varphi + \omega'(I), \quad I' = I + g'(I)$$

where $\omega'(I)$ and $g'(I)$ are determined via the normalization process. Thus the dynamics of the actions I is separated from that of the angles φ .

A simpler form of the normalized map is found in case the map possesses some interesting symmetries. Let me give an example. Say that a vector field $\begin{pmatrix} X \\ Y \end{pmatrix}$ is of type $(+, -)$ or of type $(-, +)$, respectively, if it satisfies

$$\begin{pmatrix} X(-\varphi, I) \\ Y(-\varphi, I) \end{pmatrix} = \begin{pmatrix} X(\varphi, I) \\ -Y(\varphi, I) \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} X(-\varphi, I) \\ Y(-\varphi, I) \end{pmatrix} = \begin{pmatrix} -X(\varphi, I) \\ Y(\varphi, I) \end{pmatrix}.$$

With a little patience, looking at the explicit expression (41) of the commutator, one checks that the commutators obey the rules symbolically expressed by the table

$\{\cdot, \cdot\}$	$(+, -)$	$(-, +)$.
$(+, -)$	$(-, +)$	$(+, -)$	
$(-, +)$	$(+, -)$	$(-, +)$	

Assume now that the map (24) satisfies the symmetry

$$f_s(-\varphi, I) = f_s(\varphi, I), \quad g_s(-\varphi, I) = -g_s(\varphi, I).$$

This means that the system is reversible. It is not difficult to check that in this case the generating sequence $W(\varphi, I)$ is of type $(+, -)$. Then, with a little more patience and using induction, one checks also that at every step of the normalization procedure one has that Ψ_s is of type $(+, -)$, and so by solving eq. (38) one gets Z_s of type $(+, -)$ and X_s of type $(-, +)$. This implies that in (42) one has $Z_s = \begin{pmatrix} \alpha_0(I) \\ 0 \end{pmatrix}$. In turn this

implies that the normalized map is written as

$$\varphi' = \varphi + \omega'(I), \quad I' = I,$$

representing a Kronecker map on a family of invariant tori with angles $\omega'(I)$.

All this is formal, of course. Making a rigorous statement, as is known, is a definitely more complicated matter. If the angles $\omega(I)$ of the unperturbed map do depend on the actions I then the formal construction above is expected to fail due to the presence of zero divisors, and an analog of Poincaré's theorem on non existence of holomorphic first integrals for Hamiltonian systems of differential equations applies. However one can prove, possibly with some extra condition on $\omega(I)$, that KAM theory applies, thus showing the existence of a big set on invariant tori carrying a Kronecker map with strongly non resonant angles. On the other hand, it should also be possible to prove a theorem of Nekhoroshev's type on exponential stability. If the angles ω are constant, then some normal form may be constructed, possibly taking into account the resonances, but the series so constructed are expected to be divergent, unless one looks for a Kolmogorov's normal form on an invariant torus. However, all this matter goes behind the limits of the present note, which deals only with formal aspects.

A. Proof of proposition 3

Let me state a preliminary identity. If X is a generating sequence and V a vector field then one has

$$(43) \quad E_s^X L_V = \sum_{j=0}^s L_{E_j^X V} E_{s-j}^X.$$

The proof is worked out by induction, since the equality is trivial for $s = 0$ and moreover for $s = 1$ it is just Jacobi's identity for commutators. Here is the complete calculation for $s > 1$. The notation is made simpler by writing E_j in place of E_j^X , since there is no confusion.

$$\begin{aligned} E_s L_V &= \sum_{m=1}^s \frac{m}{s} L_{X_m} E_{s-m} L_V \\ &= \sum_{m=1}^s \frac{m}{s} L_{X_m} \sum_{j=0}^{s-m} L_{E_j V} E_{s-m-j} \\ &= \sum_{m=1}^s \sum_{j=0}^{s-m} \frac{m}{s} \left(L_{L_{X_m} E_j V} + L_{E_j V} L_{X_m} \right) E_{s-m-j} \\ &= \sum_{m=1}^s \sum_{j=0}^{s-m} \frac{m}{s} \left(L_{L_{X_m} E_{s-j-m} V} E_j + L_{E_j V} L_{X_m} E_{s-j-m} \right) \\ &= \sum_{j=0}^{s-1} \frac{s-j}{s} \sum_{m=1}^{s-j} \frac{m}{s-j} \left(L_{L_{X_m} E_{s-j-m} V} E_j + L_{E_j V} L_{X_m} E_{s-j-m} \right) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{j=0}^{s-1} \frac{s-j}{s} \left(L_{E_{s-j}V} E_j + L_{E_jV} E_{s-j} \right) \\
 &= \sum_{j=1}^s \frac{j}{s} L_{E_jV} E_{s-j} + \sum_{j=0}^{s-1} \frac{s-j}{s} L_{E_jV} E_{s-j} = \sum_{j=0}^s L_{E_jV} E_{s-j} .
 \end{aligned}$$

Jacobi's identity for commutators written as $L_X L_w - L_w L_X = L_{L_X w}$ is used in order to obtain the third equality.

Coming to the proof of (21) of proposition 3, by definition of Lie transform one has

$$(44) \quad T_X \circ T_Y = \left(\sum_{l \geq 0} E_l^{(X)} \right) \left(\sum_{k \geq 0} E_k^{(Y)} \right) = \sum_{s \geq 0} \sum_{m=0}^s E_m^{(X)} E_{s-m}^{(Y)} .$$

On the other hand for the generating sequence Z defined as in (21) one has

$$(45) \quad E_s^Z = \sum_{l=1}^s \frac{l}{s} L_{X_l+Y_l} E_{s-l}^Z + \sum_{l=2}^s \frac{l}{s} \sum_{m=1}^{l-1} \frac{m}{l} L_{E_{l-m}^X Y_m} E_{s-l}^Z .$$

Thus it is enough to check that

$$(46) \quad E_s^Z = \sum_{m=0}^s E_m^X E_{s-m}^Y , \quad s \geq 0 .$$

I proceed by induction. For $s = 0, 1$ the equality is true. For $s > 1$ calculate

$$\begin{aligned}
 \sum_{m=0}^s E_m^X E_{s-m}^Y &= \sum_{m=0}^{s-1} \frac{s-m}{s} (E_{s-m}^X E_m^Y + E_m^X E_{s-m}^Y) \\
 &= \sum_{m=0}^{s-1} \frac{s-m}{s} \sum_{l=1}^{s-m} \frac{l}{s-m} (L_{X_l} E_{s-m-l}^X E_m^Y + E_m^X L_{Y_l} E_{s-m-l}^Y) \\
 &= \sum_{m=0}^{s-1} \sum_{l=1}^{s-m} \frac{l}{s} \left(L_{X_l} E_{s-m-l}^X E_m^Y + \sum_{k=0}^m L_{E_k^X Y_l} E_{m-k}^X E_{s-m-l}^Y \right) \\
 &= \sum_{l=1}^s \sum_{m=0}^{s-l} \frac{l}{s} (L_{X_l} E_{s-m-l}^X E_m^Y + L_{Y_l} E_m^X E_{s-m-l}^Y) \\
 &\quad + \sum_{l=1}^{s-1} \sum_{m=1}^{s-l} \sum_{k=1}^m \frac{l}{s} L_{E_k^X Y_l} E_{m-k}^X E_{s-m-l}^Y .
 \end{aligned}$$

The identity (43) is used in order to obtain the third line. The first double sum in the latter expression is further elaborated as

$$\sum_{l=1}^s \frac{l}{s} \left(L_{X_l} \sum_{m=0}^{s-l} E_{s-l-m}^X E_m^Y + L_{Y_l} \sum_{m=0}^{s-l} E_m^X E_{s-l-m}^Y \right)$$

$$= \sum_{l=1}^s \frac{l}{s} L_{X_l+Y_l} \sum_{m=0}^{s-l} E_m^X E_{s-l-m}^Y = \sum_{l=1}^s \frac{l}{s} L_{X_l+Y_l} E_{s-l}^Z .$$

The latter expression coincides with the first sum in the r.h.s. of (45). The triple sum is further elaborated as

$$\begin{aligned} \sum_{l=1}^{s-1} \sum_{k=1}^{s-l} \frac{l}{s} L_{E_k^X Y_l} \sum_{h=0}^{s-l-k} E_h^X E_{s-l-k-h}^Y &= \sum_{l=1}^{s-1} \sum_{k=1}^{s-l} \frac{l}{s} L_{E_k^X Y_l} E_{s-l-k}^Z \\ &= \sum_{l=1}^{s-1} \sum_{m=l+1}^s \frac{l}{s} L_{E_{m-l}^X Y_l} E_{s-m}^Z = \sum_{m=2}^s \sum_{l=1}^{m-1} \frac{l}{s} L_{E_{m-l}^X Y_l} E_{s-m}^Z . \end{aligned}$$

Here the induction hypothesis is used in the first step. The last expression coincides with the second double sum in the r.h.s. of (45). Thus, the right member of (46) equals the last member of (45), and this concludes the proof.

References

- [1] G.D. Birkhoff: *Proof of Poincaré's Geometric Theorem*, Transactions of the American Mathematical Society **14**, 14–22 (1913)
- [2] G.D. Birkhoff: Surface transformations and their dynamical applications, Acta Mathematica **43**, 1–119 (1920).
- [3] A. Deprit: *Canonical transformations depending on a small parameter*, Cel. Mech. **1**, 12–30 (1969).
- [4] F. Fassò: *On a relation among Lie series*, Celestial Mechanics **46**, 113–118 (1989).
- [5] A. Giorgilli, L. Galgani: *Formal integrals for an autonomous Hamiltonian system near an equilibrium point*, Cel. Mech. **17** (1978), 267–280.
- [6] A. Giorgilli, L. Galgani: *Rigorous estimates for the series expansions of Hamiltonian perturbation theory*, Cel. Mech. **37**, 95–112 (1985).
- [7] A. Giorgilli, U. Locatelli: *Kolmogorov theorem and classical perturbation theory*, ZAMP **48**, 220–261 (1997).
- [8] A. Giorgilli, U. Locatelli: *A classical self-contained proof of Kolmogorov's theorem on invariant tori*, in Proceedings of the NATO ASI school "Hamiltonian systems with three or more degrees of freedom", C. Simó ed., NATO ASI series C: Math. Phys. Sci., Vol. 533, Kluwer Academic Publishers, Dordrecht–Boston–London, 72–89 (1999).
- [9] A. Giorgilli, S. Marmi: *Improved estimates for the convergence radius in the Poincaré–Siegel problem*, Discrete and Continuous Dynamical Systems series S **3**, 601–621 (2010).
- [10] A. Giorgilli, M. Sansottera: *Methods of algebraic manipulation in perturbation theory*, in *Chaos, Diffusion and Non-integrability in Hamiltonian Systems - Applications to Astronomy*, Proceedings of the 3rd La Plata International School on Astronomy and Geophysics, P.M. Cincotta, C.M. Giordano and C. Efthymiopou-

- los eds., Universidad Nacional de La Plata and Asociacin Argentina de Astronomia Publishers, La Plata, Argentina (2012).
- [11] W. Gröbner: *Nuovi contributi alla teoria dei sistemi di equazioni differenziali nel campo analitico*, Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. **23** 375-379 (1957).
 - [12] W. Gröbner: *Die Lie-Reihen und Ihre Anwendungen*, VEB Deutscher Verlag der Wissenschaften, Mathematische Monographien, **3** (1960). Italian translation: *Serie di Lie e loro applicazioni*, Ed. Cremonese, Roma, 1973.
 - [13] A. Haro: *The primitive function of an exact symplectomorphism*, Nonlinearity **13**, 1483–1500 (2000).
 - [14] M. Hénon: *Numerical Study of Quadratic Area-Preserving Mappings*, Quart. Appl. Math. **27**, 291–312 (1969).
 - [15] J. Henrard, J. Roels: *Equivalence for Lie transforms*, Celestial Mechanics **10**, 497–512 (1974).
 - [16] G. Hori: *Theory of general perturbations with unspecified canonical variables*, Publ. Astron. Soc. Japan, **18**, 287–296 (1966).
 - [17] S. Kuksin: *On the Inclusion of an Analytic Symplectomorphism Close to an Integrable One into a Hamiltonian Flow*, Russian Journal of Mathematical Physics **1**, 191–207 (1993).
 - [18] S. Kuksin, J. Pöschel: *On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications*, in: S. Kuksin, V. Lazutkin, J Pöschel (Eds.), *Seminar on dynamical systems*, Birkhäuser, Basel, 96–116 (1994)
 - [19] J.K. Moser: *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Gött., II Math. Phys. Kl. 1962, 1–20 (1962).
 - [20] Poincaré, H.: *Les méthodes nouvelles de la mécanique céleste*, Gauthier–Villars, Paris (1892).
 - [21] H. Poincaré: *Sur un théorème de géométrie*, Rendiconti del Circolo Matematico di Palermo **33**, 375–407 (1912).
 - [22] A.V. Pronin, D.V. Treschev: *On the inclusion of analytic maps into analytic flows*, Regular and Chaotic Dynamics **2**, 14–24 (1997).
 - [23] E. Schröder: *Über iterierte Functionen*, Math. Ann. **3**, 296-322 (1871).
 - [24] Siegel, C.L.: *Iterations of analytic functions*, Annals of Math. **43**, 607–612 (1942).