H. Gamkrelidze

Exponential representation of solutions of ordinary differential equations

In: Jiří Fábera (ed.): Equadiff IV, Czechoslovak Conference on Differential Equations and Their Applications. Proceedings, Prague, August 22-26, 1977. Springer-Verlag, Berlin, 1979. Lecture Notes in Mathematics, 703. pp. [118]--129.

Persistent URL: http://dml.cz/dmlcz/702211

EXPONENTIAL REPRESENTATION OF SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS

R. Gamkrelidze, Moscow

I shall describe here a kind of calculus for solutions of ordinary differential equations developed jointly with my collaborator A.Agrachev. This calculus is based on the exponential representation of the solutions and reflects their most general group-theoretic properties. In deriving the calculus we were strongly influenced by problems of control and optimization and it is shaped according to the needs of these theories. Nevertheless it might be considered, as I believe, not merely as a technical tool for dealing with control problems only but could also be of more general interest. This may justify my choice of the topic for the Equadiff conference.

1. Differential equations considered

Let us consider a differential equation in $\mathbb{R}^n$

(1) $$\dot{z} = X_t(z)$$

where $X_t(z)$ is a C^∞-function in $z \in \mathbb{R}^n$ for $\forall t \in \mathbb{R}$, measurable in t for $\forall z \in \mathbb{R}^n$ and satisfying the condition

(2) $$\|X_t\|_k \le \mu_k(t) , \quad \int_{\mathbb{R}} \mu_k(t)dt < \infty , \quad k=0,1,\dots$$

where $\|\cdot\|_k$ denotes the norm of the uniform convergence in $\mathbb{R}^n$ up to the k-th derivative.

Our first goal is to find a suitable representation of the flow defined by (1), that is, of a family of C^∞-diffeomorphisms F_t, $t \in \mathbb{R}$, of $\mathbb{R}^n$ satisfying the equation

(3) $$\frac{d}{dt} F_t x = X_t(F_t x) , \quad F_o = Id \quad \forall x \in \mathbb{R}^n .$$

The existence of F_t is guaranteed by (2).

2. Transforming (3) into a linear "operator equation"

There is a procedure transforming the nonlinear equation (3) into a certain linear "operator equation" for F_t. To describe it let me introduce some standard notions.

Φ will denote the algebra of all C^∞-scalar functions $f,g,\dots$ on $\mathbb{R}^n$ with the topology of the uniform convergence on compact sets for every derivative. $\mathcal{A}$ stands for the associative algebra of all continuous linear transformations of Φ. The composition of two elements A_1, A_2 in $\mathcal{A}$ will be denoted by $A_1 \circ A_2$. The operators from $\mathcal{A}$ can be applied also to vector-valued functions on $\mathbb{R}^n$. Denote by Θ the identity mapping of $\mathbb{R}^n$: $\Theta(x) \equiv x$. We shall say that a sequence of operators $A_1, A_2, \ldots$ from $\mathcal{A}$ is convergent to A iff the sequence of functions $A_1\Theta, A_2\Theta, \ldots$ converges in Φ to $A\Theta$. Every diffeomorphism F of $\mathbb{R}^n$ will be considered as an element of $\mathcal{A}$: $Ff(x) = f(Fx)$, $\forall x \in \mathbb{R}^n$, and the set of all C^∞-diffeomorphisms of $\mathbb{R}^n$ will be denoted by $\mathcal{D}$. By $\mathcal{L}$ we shall denote the Lie algebra of all C^∞-vector fields on $\mathbb{R}^n$, which is a subspace of $\mathcal{A}$ characterized by the differentiation rule $X(fg) = (Xf)g + f(Xg)$ $\forall X \in \mathcal{L}$, $\forall f,g \in \Phi$. The Lie bracket of two fields will be denoted as usual by $[X,Y] = X \circ Y - Y \circ X =$ $= (\mathrm{ad}\, X)Y$. The following important relation holds:

$$(4) \qquad F \circ X \circ F^{-1} \stackrel{\text{def}}{=} (\mathrm{Ad}\, F)X \in \mathcal{L} \ \forall X \in \mathcal{L}, \ \forall F \in \mathcal{D}.$$

Consider X_t, $t \in \mathbb{R}$, in (1) as a nonstationary (time-dependent) vector field on $\mathbb{R}^n$. It is not difficult to show that (3) is equivalent to the linear "operator equation" for the flow F_t

$$(5) \qquad \frac{d}{dt} F_t = F_t \circ X_t, \quad F_0 = \mathrm{Id} \Leftrightarrow F_t = \mathrm{Id} + \int_0^t F_\tau \circ X_\tau \, d\tau,$$

where the operations of differentiation and integration in t should be understood in the "weak" sense: first apply the operator to an arbitrary function from Φ and then differentiate or integrate. The equivalence between (3) and (5) should be understood literally - the existence of a unique solution of (3) implies the existence of a unique solution F_t, $t \in \mathbb{R}$, for (5) which at the same time necessarily turns out to be a flow and vice versa. Certainly we can always consider the flow F_t only for values of t sufficiently close to zero since the equation $F_t = F_{t_0} + \int_{t_0}^t F_\tau \circ X_\tau \, d\tau$, t_0-arbitrary fixed, has exactly the same properties as (5), which permits to restore the whole flow F_t, $t \in \mathbb{R}$.

Call the formal series

$$(6) \qquad \mathrm{Id} + \sum_{i=1}^{\infty} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{i-1}} d\tau_i X_{\tau_i} \circ X_{\tau_{i-1}} \circ \ldots \circ X_{\tau_1},$$

arising when solving the linear equation (5) formally, the Volterra series corresponding to (5).

3. Exponential representation of the flow

Suppose the field X_t analytic on $\mathbb{C}^n$ and subject to the condition (2), where the norms $\|\cdot\|_k$ should be understood (in this case) as norms of the uniform convergence in a certain complex σ_k--neighbourhood ($\sigma_k > 0$) of $\mathbb{R}^n \subset \mathbb{C}^n$. Then the Volterra series (6) converges (in the above defined sense) for every t rendering the integral $\int_0^t \mu_0(\tau)d\tau$ a sufficiently small value to an analytic diffeomorphism $x \mapsto F_t x$, and the obtained flow is the unique solution of the equation (5) (proof by the method of majorants).

The fields X_t generally do not commute for different values of t, hence the order of the factors in the term $\int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{i-1}} d\tau_i X_{\tau_i} \cdots X_{\tau_1}$ could not be changed and the corresponding times τ_j increase from left to right: $0 \le \tau_i \le \ldots \le \tau_1 \le t$. Adopting the terminology used by physicists we call the flow F_t to which the series (6) converges the right chronological exponent of X_t and denote it by

$$F_t = \overrightarrow{\exp} \int_0^t X_\tau \, d\tau \quad , \tag{7}$$

the arrow indicating the direction of growth of the τ_j-s in the successive terms of the "right" Volterra series (6).

In the general C^∞-case the series (6) is not convergent, however, we can call the unique solution of (5) (which exists and is a flow according to the standard existence theorem for (3)) the right chronological exponent of X_t and denote it with the same symbol (7). The following basic asymptotics may justify this convention:

$$\left\| \left\{ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau - \left(\mathrm{Id} + \sum_{i=1}^{k} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{i-1}} d\tau_i X_{\tau_i} \circ \cdots \circ X_{\tau_1}\right)\right\} f \right\|_{Q,0} \le C_{Q,k} \left(\int_0^t \|X_\tau\|_{\hat{Q},k+1} d\tau \right)^{k+1} \|f\|_{\hat{Q},k+1} \quad \forall f \in \Phi ,$$

$$k = 1, 2, \ldots ,$$

where $\|\cdot\|_{Q,j}$, $j=0,1,2,\ldots$ denotes here and in the sequel the norm of the uniform convergence of all derivatives up to the order j on an arbitrary compact set $Q\in\mathbb{R}^n$, $\hat{Q}$ - a compact neighbourhood of Q of radius $\int_0^t \mu_0(\tau)d\tau$.

However, we can go even further in interpreting the symbol (7) and describe a sort of "summation procedure" which enables us to "sum up" the Volterra series (6) in the general C^∞-case to a family of operators F_t which turns out to be a flow satisfying the equation (5), and thus give an existence proof for (5). Uniqueness is an easy consequence of the fact that F_t is a flow.

To describe the "summation procedure" take the "δ-type" analytic mollifier

$$\omega_\varepsilon = \frac{1}{(\sqrt{\pi}\,\varepsilon)^n}\, e^{-(\frac{z}{\varepsilon})^2} \qquad (\varepsilon \to 0)$$

and consider the convolution

$$X_t^{(\varepsilon)}(z) = \omega_\varepsilon * X_t = \frac{1}{(\sqrt{\pi}\,\varepsilon)^n}\int_{\mathbb{R}^n} e^{-(\frac{z-w}{\varepsilon})^2} X_t(w)\,dw \quad .$$

The obtained field $X_t^{(\varepsilon)}$ is an entire-analytic field on $\mathbb{C}^n$ for every $\varepsilon > 0$ subject to (2) (up to a constant factor for $\mu_k(t)$), thus the corresponding Volterra series is convergent to a flow $F_t^{(\varepsilon)}$ which satisfies the equation (5). It turns out that $F_t^{(\varepsilon)}$ $(\varepsilon \to 0)$ is a Cauchy family of flows (in the topology of the uniform convergence on compact sets of $\mathbb{R}^n$ for every derivative) and converges to a flow F_t which is the unique solution of (5). We consider the flow F_t , $t\in\mathbb{R}$ as the "generalized sum" of the right Volterra series (6) and call it the right chronological exponent of X_t :

$$\frac{d}{dt}\,\overrightarrow{\exp}\int_0^t X_\tau\, d\tau = \overrightarrow{\exp}\int_0^t X_\tau\, d\tau \circ X_t \ .$$

The left Volterra series and the corresponding left chronological exponent could be considered in a completely symmetrical way

$$G_t = \overleftarrow{\exp}\int_0^t X_\tau\, d\tau = \mathrm{Id} + \sum_{i=1}^{\infty}\int_0^t d\tau_1\int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{i-1}} d\tau_i X_{\tau_1}\circ \ldots \circ X_{\tau_i} \ .$$

The flow $\overleftarrow{\exp}\int_0^t -X_\tau\, d\tau$ satisfies the "adjoint operator equation" for (5)

$$\frac{d}{dt} G_t = -X_t \circ G_t \ , \qquad G_0 = \mathrm{Id} \ ,$$

which is equivalent to the linear partial differential equation of the first order in $\mathbb{R}^n$

$$\frac{\partial \omega(t,x)}{\partial t} + \sum_{i=1}^{n} X_t^i(x) \frac{\partial \omega(t,x)}{\partial x^i} = \frac{\partial \omega}{\partial t} + X_t \omega = 0 \ ,$$

$\omega(t,x) = G_t f(x)$, $f(x) = G_0 f(x) = \omega(0,x)$ - the initial function. Evidently

$$\overrightarrow{\exp}\int_0^t X_\tau\, d\tau \circ \overleftarrow{\exp}\int_0^t -X_\tau\, d\tau = \overleftarrow{\exp}\int_0^t -X_\tau\, d\tau \circ \overrightarrow{\exp}\int_0^t X_\tau d\tau = \mathrm{Id} \ .$$

In the "commutative case" that is if $\left[X_t \ , \int_0^t X_\tau\, d\tau\right] = 0$ $\forall\, t \in \mathbb{R}$ we have

$$\overrightarrow{\exp}\int_0^t X_\tau\, d\tau = \overleftarrow{\exp}\int_0^t X_\tau\, d\tau = \mathrm{Id} + \sum_{i=1}^{\infty} \frac{1}{i!} \left(\int_0^t X_\tau\, d\tau\right)^i =$$
$$= e^{\int_0^t X_\tau d\tau} \ .$$

For example, if $X_t \equiv X$ then $\overrightarrow{\exp}\int_0^t X_\tau\, d\tau = e^{tX}$.

To demonstrate the flexibility of the obtained representation I shall derive formulas expressing two basic objects in the theory of ordinary differential equations - the perturbing flow of a given flow F_t and the variation of F_t .

4. <u>The perturbing flow</u>

Suppose the field X_t and the corresponding flow $F_t = \overrightarrow{\exp}\int_0^t X_\tau\, d\tau$ fixed. Call an arbitrary field Y_t a perturbing field for X_t , the flow $\overrightarrow{\exp}\int_0^t (X_\tau + Y_\tau)d\tau$ - the corresponding perturbed flow.

<u>Problem.</u> Find a flow $C_t = C_t(Y_\tau)$ satisfying the equation

$$(8) \qquad \overrightarrow{\exp}\int_0^t (X_\tau + Y_\tau)d\tau = C_t(Y_\tau) \circ \overrightarrow{\exp}\int_0^t X_\tau\, d\tau = C_t(Y_\tau) \circ F_t .$$

We call $C_t(Y_\tau)$ the perturbing flow for F_t corresponding to Y_t.

The proposed solution coincides with the method of variation of constants and could be carried out as follows. According to (4) we can consider $\mathrm{Ad}\, F_t$ in the formula

(9) $$(\mathrm{Ad}\, F_t)Z = F_t \circ Z \circ F_t^{-1}\,, \qquad Z \in \mathscr{L}\,,$$

as a time-dependent linear transformation of $\mathscr{L}$. Differentiating we obtain the equation

$$\frac{d}{dt}\,\mathrm{Ad}\, F_t = \mathrm{Ad}\, F_t \circ \mathrm{ad}\, X_t\,,$$

which suggests the notation

(10) $$\mathrm{Ad}\, F_t = \overrightarrow{\exp} \int_0^t \mathrm{ad}\, X_\tau\, d\tau\,.$$

Differentiation of (8) yields

$$\frac{d}{dt}\, C_t(Y_\tau) = C_t(Y_\tau) \circ (\mathrm{Ad}\, F_t)Y_t\,,$$

whence combining (8) and (10) we come to formulas

(11) $$C_t(Y_\tau) = \overrightarrow{\exp} \int_0^t (\mathrm{Ad}\, F_\tau)Y_\tau\, d\tau = \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \mathrm{ad}\, X_s ds\right) Y_\tau\, d\tau\,,$$

$$\overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau)\, d\tau = \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \mathrm{ad}\, X_s ds\right) Y_\tau\, d\tau \circ \overrightarrow{\exp} \int_0^t X_\tau\, d\tau\,,$$

asserting two basic facts. 1) If X_t, Y_t are analytic (in t and z) and satisfy (2) then evaluating all chronological exponents on the right sides as the corresponding formal right Volterra series and performing the indicated operations we come to convergent (in appropriate regions) series defining the flows standing on the left sides. 2) For the general C^∞-case the equalities (10)-(11) should be understood in the following asymptotic sense:

$$\left\| \left\{ \overrightarrow{\exp} \int_0^t \mathrm{ad}\, X_\tau\, d\tau - \left(\mathrm{Id} + \sum_{i=1}^{k} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{i-1}} d\tau_i\; \mathrm{ad}\, X_{\tau_i} \circ \ldots \circ \mathrm{ad}\, X_{\tau_1}\right)\right\} Z \right\|_{Q,0} \le C_{Q,k} \left(\int_0^t \|X_\tau\|_{\hat{Q},k+1}\, d\tau \right)^{k+1} \|Z\|_{\hat{Q},k+1} \quad \forall Z \in \mathscr{L}\,,$$

$$k = 1, 2, \ldots\;;$$

$$\Big\| \overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau)\, d\tau - \Big(Id + \sum_{i=1}^{k} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{i-1}} d\tau_i (Ad\, F_{\tau_i}) \cdot Y_{\tau_i} \circ \ldots \circ (Ad\, F_{\tau_1}) Y_{\tau_1}\Big) \circ F_t \Big\|_{Q,0} \le C_{Q,k} \Big(\int_0^t \| F_\tau^{-1} \|_{\hat{Q},k+1} \| Y_\tau \|_{\hat{Q},k}\, d\tau \Big)^{k+1}$$

$\| F_t \|_{\hat{Q},k+1}$. In case of time independent fields $X_t \equiv X$, $Y_t \equiv Y$ we have $e^{tX} \circ Z \circ e^{-tX} = e^{t\, ad\, X} Z$, $e^{t(X+Y)} = \overrightarrow{\exp} \int_0^t e^{\tau\, ad\, X} Y d\tau \circ e^{tX}$.

The second formula shows that even if the fields X, Y are time-independent the corresponding perturbing flow is expressed through a chronological exponent.

Consider a flow $F_t = \overrightarrow{\exp} \int_0^t Y_\tau^{(1)} d\tau \circ \ldots \circ \overrightarrow{\exp} \int_0^t Y_\tau^{(m)} d\tau$ and suppose we have to define the field Z_t that generates F_t : $F_t = \overrightarrow{\exp} \int_0^t Z_\tau\, d\tau$. We call Z_t the right chronological logarithm of F_t and denote $Z_t = \overrightarrow{\log}\, F_t$. It is expressed by (see (9)-(10))

(12)
$$\overrightarrow{\log}\Big(\overrightarrow{\exp} \int_0^t Y_\tau^{(1)} d\tau \circ \ldots \circ \overrightarrow{\exp} \int_0^t Y_\tau^{(m)} d\tau\Big) \overset{def}{=}$$

$$\overset{def}{=} \lambda(Y_\tau^{(1)}, \ldots, Y_\tau^{(m)}) = F_t^{-1} \circ \frac{d}{dt} F_t =$$

$$= \Big(\overleftarrow{\exp} \int_0^t -ad\, Y_\tau^{(m)} d\tau \circ \ldots \circ \overleftarrow{\exp} \int_0^t -ad\, Y_\tau^{(2)} d\tau\Big) Y_t^{(1)} +$$

$$+ \Big(\overleftarrow{\exp} \int_0^t -ad\, Y_\tau^{(m)} d\tau \circ \ldots \circ \overleftarrow{\exp} \int_0^t -ad\, Y_\tau^{(3)} d\tau\Big) Y_t^{(2)} +$$

$$+ \ldots + \Big(\overleftarrow{\exp} \int_0^t -ad\, Y_\tau^{(m)} d\tau\Big) Y_t^{(m-1)} + Y_t^{(m)} .$$

5. The variation of a flow

We start with the following problem. For a given flow $\hat{F}_t = \overrightarrow{\exp} \int_0^t Y_\tau\, d\tau$ determine a field $V_t(Y_\tau)$ which satisfies in an appropriate asymptotic sense (to be formulated precisely) the relation

$$\overrightarrow{\exp}\int_0^t Y_\tau\, d\tau \cong e^{V_t(Y_\tau)} = \mathrm{Id} + \sum_{i=1}^{\infty} \frac{1}{i!}\,(V_t(Y_\tau))^i \ .$$

It is natural to call $V_t(Y_\tau)$ the usual (not chronological) logarithm of the flow considered and to denote

$$V_t(Y_\tau) = \ln \overrightarrow{\exp}\int_0^t Y_\tau\, d\tau \ .$$

For a precise formulation we have to consider noncommutative nonassociative polynomials over $\mathbb{R}$, $P(Y_1,\dots,Y_k)$ in $k=1,2,\dots$ $\mathcal{L}$ - valued variables Y_i with the Lie bracket multiplication; the P-s consequently will also be $\mathcal{L}$ - valued. A simple and explicit algorithm (see n°6) prescribes a universal sequence of such polynomials

(13) $$P_1(Y_1),\ P_2(Y_1,Y_2),\dots,\ P_k(Y_1,\dots,Y_k),\dots\ ,$$

P_k - homogeneous of degree k in its variables, each variable having degree 1, for which the following theorem is valid.

<u>Theorem.</u> For every field Y_t consider the formal series

(14) $$V_t(Y_\tau) = \sum_{i=1}^{\infty}\int_0^t d\tau_1\int_0^{\tau_1} d\tau_2 \dots \int_0^{\tau_{i-1}} d\tau_i P_i(Y_{\tau_1},\dots,Y_{\tau_i}) = \sum_{i=1}^{\infty} V_t^{(i)}(Y_\tau)$$

and call it the formal vector field associated with Y_t . Then the following asymptotics holds:

(15) $$\left\| \overrightarrow{\exp}\int_0^t Y_\tau\, d\tau - e^{\sum_{i=1}^{k} V_t^{(i)}(Y_\tau)} \right\|_{Q,0} \le C_{Q,k}\left(\int_0^t \|Y_\tau\|_{\hat{Q},k+1}\, d\tau\right)^{k+1} , \quad k=1,2,\dots ,$$

also expressed by either of the relations

(16) $$\ln \overrightarrow{\exp}\int_0^t Y_\tau\, d\tau \cong \sum_{i=1}^{\infty} V_t^{(i)}(Y_\tau) = V_t(Y_\tau) ,$$

$$\overrightarrow{\exp}\int_0^t Y_\tau\, d\tau \cong e^{\sum_{i=1}^{\infty} V_t^{(i)}(Y_\tau)} = e^{V_t(Y_\tau)} \ .$$

As an immediate consequence of (15) we obtain

$$(17) \qquad (V_t^{(i)}(Y_\tau))x = 0,\ i=1,\dots,k-1 \Longrightarrow (\overrightarrow{\exp}\int_0^t Y_\tau\, d\tau)x = x + (V_t^{(k)}(Y_\tau))x + \mathcal{O}(\int_0^t \|Y_\tau\|_{Q,k+1} d\tau)^{k+1} \quad \forall x\in\mathbb{R}^n,\ k=1,2,\dots .$$

Formulas (15), (17) justify the forthcoming terminology. Call the field $V_t^{(k)}(Y_\tau)$ the k-th variation of the identity flow Id_t corresponding to Y_t - the perturbing field of the zero field (which generates the identity flow), and denote $V_t^{(k)}(Y_\tau) = \delta^{(k)}Id_t(Y_\tau)$. Similarly, call the formal field (14) the (full) variation $\delta Id_t(Y_\tau)$ of the identity flow

$$(18) \qquad \delta Id_t(Y_\tau) = \sum_{i=1}^{\infty} \delta^{(i)} Id_t(Y_\tau),$$

and the formal series

$$(19) \qquad e^{V_t(Y_\tau)} = Id + \sum_{i=1}^{\infty} \frac{1}{i!}(V_t(Y_\tau))^i$$

- the formal flow corresponding to the formal field $V_t(Y_\tau)$. According to (14), (16) the following basic asymptotic expansion is valid:

$$(20) \qquad \overrightarrow{\exp}\int_0^t Y_\tau\, d\tau \cong e^{V_t(Y_\tau)} = e^{\delta Id_t(Y_\tau)} = Id + \sum_{i=1}^{\infty}\frac{1}{i!}(\delta Id_t(Y_\tau))^i,$$

$$\delta Id_t(Y_\tau) = \sum_{i=1}^{\infty}\delta^{(i)}Id_t(Y_\tau) = \sum_{i=1}^{\infty}\int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \dots \int_0^{\tau_{i-1}} d\tau_i P_i(Y_{\tau_1},\dots,Y_{\tau_i}).$$

We call it the "Maclaurin series expansion" (around the zero field) of the flow $\overrightarrow{\exp}\int_0^t Y_\tau\, d\tau$ - the perturbing flow for Id_t under Y_t, which can be also regarded as the corresponding perturbed flow. A composition rule in the set of formal flows (19) defined by (see (12))

$$e^{V_t(Y_\tau^{(1)})} \circ e^{V_t(Y_\tau^{(2)})} = e^{V_t(\lambda(Y_\tau^{(1)}, Y_\tau^{(2)}))}$$

turns it into a multiplicative group.

Unifying formulas (12), (20) (see also the construction of the P_i-s) we come to a generalization of the Campbell-Hausdorff formula

$$(21)\qquad \overrightarrow{\exp}\int_0^t Y_\tau^{(1)}d\tau\circ\ldots\circ\overrightarrow{\exp}\int_0^t Y_\tau^{(m)}d\tau \cong$$
$$\cong e^{\sum_{i=1}^{\infty}\int_0^t d\tau_1\ldots\int_0^{\tau_{i-1}} d\tau_i P_i(Z_{\tau_1},\ldots,Z_{\tau_i})},$$
$$Z_t = (\overleftarrow{\exp}\int_0^t -\mathrm{ad}\, Y_\tau^{(m)}d\tau\circ\ldots\circ\overleftarrow{\exp}\int_0^t -\mathrm{ad}\, Y_\tau^{(2)}d\tau)Y_t^{(1)} +$$
$$+\ldots+(\overleftarrow{\exp}\int_0^t -\mathrm{ad}\, Y_\tau^{(m)}d\tau)Y_t^{(m-1)} + Y_t^{(m)}.$$

The usual Dynkin form of this formula (when $Y_t^{(i)} \equiv Y$, $m=2$, $t=1$) seems to be unnecessarily complicated which results from the fact that it actually carries out all i-fold integrations indicated in (21). For analytic fields (both in t and z) all formal series involved in (18)-(21) are convergent provided the appropriate norms of the Y_t-s are sufficiently small (I shall not go here into the details of precise formulation).

The crucial advantage of the introduced variations consists in the validity of the asymptotic relation (20) and in their invariant form - the $\delta^{(k)}\mathrm{Id}_t$-s are vector fields and thus belong to the first tangent bundle of the underlying space (in our case of $\mathbb{R}^n$), consequently they act not only on Φ but also on $\mathbb{R}^n$ as "infinitesimal displacements" of $\mathbb{R}^n$. This permits to obtain by (18) the "formal infinitesimal displacement" of $\mathbb{R}^n$ - the full variation $\delta\mathrm{Id}_t(Y_\tau)$ and finally, using the Maclaurin expansion (20) to come to the asymptotic evaluation of the perturbing flow $\overrightarrow{\exp}\int_0^t Y_\tau\, d\tau$ - the basic goal of many problems connected with ordinary differential equations, in particular of optimization problems.

"The usual variations" of the perturbing flow - the successive terms in the Volterra expansion

$$\overrightarrow{\exp}\int_0^t Y_\tau\, d\tau = \mathrm{Id} + \int_0^t d\tau_1 Y_{\tau_1} + \int_0^t d\tau_1\int_0^{\tau_1} d\tau_2 Y_{\tau_2}\circ Y_{\tau_1}+\ldots$$

have no invariant meaning starting from the quadratic term, therefore they act only on Φ but not on $\mathbb{R}^n$. Our actual achievement consists in extracting the "invariant variations" $\delta^{(k)}\mathrm{Id}_t(Y_\tau)$ from the "usual ones". Their interrelations are established by (20), for example

$$\delta^{(2)}\mathrm{Id}_t(Y_\tau) = \int_0^t d\tau_1 \int_0^1 d\tau_2 Y_{\tau_2} \circ Y_{\tau_1} - \frac{1}{2}\delta^{(1)}\mathrm{Id}_t(Y_\tau) \circ \circ \delta_t^{(1)}\mathrm{Id}_t(Y_\tau) .$$

Suppose an arbitrary flow $F_t = \overrightarrow{\exp} \int_0^t X_\tau d\tau$ rather than the identity flow is perturbed by Y_t. Then the corresponding variations $\delta^{(k)}F_t(Y_\tau)$ are defined by

$$\delta^{(k)}F_t(Y_\tau) = \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{k-1}} d\tau_k P_k((\overrightarrow{\exp} \int_0^{\tau_1} \mathrm{ad}\, X_s ds)Y_{\tau_1}, \ldots, (\overrightarrow{\exp} \int_0^{\tau_k} \mathrm{ad}\, X_s ds)Y_{\tau_k}) ,$$

the full variation $\delta F_t(Y_\tau) = \sum_{i=1}^{\infty} \delta^{(i)}F_t(Y_\tau)$, and for the perturbing flow we obtain the "Taylor series expansion around the initial field X_t":

$$\overrightarrow{\exp} \int_0^t (\overrightarrow{\exp} \int_0^\tau \mathrm{ad}\, X_s ds) Y_\tau d\tau \cong e^{\delta F_t(Y_\tau)} .$$

As a Taylor series expansion of the perturbed flow we may consider

$$\overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau) d\tau \cong e^{\delta F_t(Y_\tau)} \circ e^{\delta \mathrm{Id}_t(X_\tau)} .$$

6. Construction of the polynomials (13)

Consider the free associative algebra Ass $(\mathrm{ad}, Y_1, Y_2, \ldots)$ over $\mathbb{R}$ with (multiplicative) generators ad, $Y_1, Y_2, \ldots$, and denote by $D(a)$, $a \in$ Ass $(\mathrm{ad}, Y_1, Y_2, \ldots)$, differentiation in the algebra defined on generators by $D(a) = a(\mathrm{ad})$, $D(a)Y_i = aY_i$, $i=1,2,\ldots$.

Further, consider an arbitrary word from the algebra composed of generators $\mathrm{ad}, Y_1, Y_2, \ldots$. To each Y_k entering a given word w we assign a nonnegative integer - the index of Y_k in w - by the following procedure. Represent $w = w_1 Y_k w_2$, where w_1 (may be an empty word) does not contain Y_k and suppose $w_1 = v_1 \ldots . v_\ell$, where each of the v_j-s is one of the generators. Define a set $J \subset \{1, \ldots, \ell\}$ by the rule: $i \in J$ iff the following two conditions are satisfied: 1) the number of occurences of ad in the word $v_i v_{i+1} \ldots . v_\ell$ is equal to the number of occurences of the Y_j-s;

2) for every $i' > i$ the number of occurences of ad in the word $v_{i'}, v_{i'+1} \dots v_\ell$ does not exceed the number of occurences of the Y_j-s . Then the index of Y_k in w is equal to the number of elements in J (which may turn out to be empty).

Among all possible words composed of ad and the Y_j-s call regular words those which could be regarded (by appropriate distribution of parentheses) as elements of the free Lie algebra with the Y_j-s as generators and ad with its usual meaning (whenever possible this could be done only in a unique way).

Finally write down the sequence of real numbers $\beta_0 = 1$, $\beta_1 = \frac{1}{2}$, $\beta_i = \frac{1}{i!} B_i$, $i=2,3,\dots$, where B_i , $i \geq 2$, is the i-th Bernoulli number: $B_3 = B_5 = B_7 = \dots = 0$, $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$,... .

Now consider an element

$(D(\mathrm{ad}\,Y_k) \circ \dots \circ D(\mathrm{ad}\,Y_2))Y_1 \in \mathrm{Ass}\,(\mathrm{ad}, Y_1, Y_2, \dots)$, $k \geq 2$, which is obtained from Y_1 by successive applications of the differentiation operators $D(\mathrm{ad}\,Y_2), \dots, D(\mathrm{ad}\,Y_k)$ and which is the sum of $(2k-3)!!$ regular words

$(D(\mathrm{ad}\,Y_k) \circ \dots \circ D(\mathrm{ad}\,Y_2))Y_1 = w_1 + \dots + w_{(2k-3)!!}$, each of the symbols $Y_1, \dots, Y_k$ entering every word w_j exactly once. Denote the index of Y_i in w_j by N_{ij} and define $P_1(Y_1) = Y_1$,

$$P_k(Y_1, \dots, Y_k) = \beta_{N_{11}} \cdots \beta_{N_{k1}} w_1 + \beta_{N_{12}} \cdots \beta_{N_{k2}} w_2 + \dots + \beta_{N_{1(2k-3)!!}} \cdots \beta_{N_{k(2k-3)!!}} w_{(2k-3)!!} , \quad k \geq 2 .$$

Here are the first four polynomials:

$$P_1 = Y_1 , \quad P_2 = \frac{1}{2} [Y_2 , Y_1] ,$$

$$P_3 = \frac{1}{6} [Y_3 , [Y_2 , Y_1]] + \frac{1}{6}[[Y_3 , Y_2] , Y_1] ,$$

$$P_4 = \frac{1}{2}([[Y_4 , Y_3] , [Y_2 , Y_1]] + [[[Y_4 , Y_3] , Y_2] , Y_1] + [Y_4 , [[Y_3 , Y_2] , Y_1]] + [Y_3 , [[Y_4 , Y_2] , Y_1]] .$$

Author's address: Steklov Institute, 42 Vavilov Str.,
Moscow 117333, USSR