Algebraic method for the evolution operator approximation

1998 J. Phys. A: Math. Gen. 31 4263

(http://iopscience.iop.org/0305-4470/31/18/015)

J. Phys. A: Math. Gen. **31** (1998) 4263–4275. Printed in the UK PII: S0305-4470(98)86216-1

Algebraic method for the evolution operator approximation

T Yu Mikhailova† and V I Pupyshev‡§

† N S Kurnakov Institute of General and Inorganic Chemistry, 117907, Moscow, Russia
‡ Department of Chemistry, Moscow State University, 119899, Moscow, Russia

Received 23 July 1997, in final form 5 February 1998

Abstract. The approximation of a short-time evolution operator for a molecular system with the Hamiltonian $H_0 + V$ is constructed in terms of the evolution operator of an unperturbed system, a perturbative potential V and commutators of H_0 and V. A description of a powerful method for such a construction based on the Lie-algebra technique is given. It is demonstrated that efficient representation for the evolution operator can be obtained using approximations like $\exp\left(-\frac{\mathrm{i}}{\hbar}(H_0+V)t\right)=\exp\left(-\frac{\mathrm{i}}{3\hbar}H_0t\right)\exp\left(-\frac{3\mathrm{i}}{4\hbar}Vt\right)\exp\left(-\frac{2\mathrm{i}}{3\hbar}H_0t\right)\exp\left(-\frac{\mathrm{i}}{4\hbar}Vt\right)\exp(S)$ where $S=-\frac{\mathrm{i}t^3}{48\hbar^3}[[H_0,V],V]$. This relation is accurate up to the third order in time. The simplest approximations of this type are listed.

1. Introduction

Describing the evolution of a molecular system is one of the central problems of quantum chemistry which has been studied very actively in recent years. It is examined in an enormous number of papers within very different contexts, such as unimolecular reaction dynamics, the theory of an elementary event of chemical reaction, the analysis of vibrational energy dissipation in the condensed phase etc. Without any claim to a comprehensive review we would like to mention that references [1–9] were the most helpful when creating this paper. Treating the evolution problem usually implies that the full Hamiltonian of a molecular system can be represented as a sum of the Hamiltonian H_0 and some perturbation term V. In our paper H_0 and V are assumed to be time independent and the evolution operator of the unperturbed system is known. We recall the fact that since H_0 and V are the non-commuting operators, the evolution operator of the full system is not equal to the product of the H_0 and V propagators. As a rule, the perturbation term is some potential and any function of V can be calculated easily. Hence, one usually needs to express the evolution operator of a system as the combination of the evolution operator for the unperturbed system and some functions of V. Using the notation $X = -\mathrm{i}H_0t/\hbar, Y = -\mathrm{i}Vt/\hbar$, we come to the problem of evaluating the operator $\exp(X+Y)$ in terms of $\exp(X), \exp(Y)$, Y, and, if necessary, suitable combinations of these operators such as $[X,Y] = XY - YX$ etc. The natural approach to the problem is as follows

$$\exp(X+Y) = \exp(X)\exp(Y)\exp(S) \tag{1.1}$$

where S is the correction operator to be found with a certain degree of accuracy. Further, for the sake of brevity we refer to any kind of approximation for the operator $\exp(X+Y)$

§ Author to whom correspondence should be addressed. E-mail address: vip@moleq.chem.msu.su

0305-4470/98/184263+13$19.50

as the EA (the exponent approximation). Relations of the type (1.1) are well known in Lie-algebra theory, where explicit expressions for S may be derived in terms of the commutator series. These expressions are of the most interest for us in this paper. There are several types of approximation for $\exp(S)$:

(A) One can search for an expansion of S as a linear combination of Lie elements S_k of different degree $k = 1, 2, \ldots$, i.e. expressions of the form $[\ldots[[X, Y], X], \ldots, Y]$ that are $(k-1)$-order commutators, where k denotes the number of the symbols X and Y:

$$S = S_1 + S_2 + S_3 + \cdots. \tag{1.2}$$

(B) By analogy with Zassenhaus' method [10, 11] one can search the Lie elements of degree $k - \Im_k$ or $\Re_k$ for $k = 1, 2, 3 \ldots$, determining both the left-hand and right-hand side expansions for $\exp(S)$ as

$$\exp(S) = \exp(\Re_1)\exp(\Re_2)\exp(\Re_3)\ldots \tag{1.3}$$

$$\exp(S) = \ldots\exp(\Im_3)\exp(\Im_2)\exp(\Im_1). \tag{1.4}$$

(C) Following Kumar's approach [12], one can be interested in the terms that are linear in Y, quadratic in Y etc. This approach is equivalent to perturbation theory [11]. We confine ourselves only to the terms that are linear in X or in Y.

There are at least two reasons that motivate the study of these approximations. First, as follows from the definition of X and Y, any Lie element of degree k corresponds to the t^k-like time dependence. This makes it interesting to consider a finite number of Lie elements with small k for the description of short-time evolution. Secondly, in general the molecular Hamiltonian H_0 has the form

$$H_0 = \sum_a \left(\frac{p_a^2}{2m_a} + U_a\right) \tag{1.5}$$

where m_a are the effective masses; U_a are the potentials that depend on the variables $q_1, q_2, \ldots, q_N$; $p_a = -\mathrm{i}\bar{h}\partial_a$ (∂_a symbol denotes the derivative with respect to q_a, $\partial_{ab} = \partial_a\partial_b$ etc). In this case all Lie elements may be calculated directly using the commutation relation $[q_a, p_b] = \mathrm{i}\bar{h}\delta_{a,b}$. For example, one obtains for V and H_0

$$[V, H_0] = \sum_a \frac{\mathrm{i}\bar{h}}{2m_a}(p_a\partial_a V + \partial_a V p_a) \tag{1.6}$$

$$[[V, H_0], H_0] = \sum_{ab} \frac{\bar{h}^2}{4m_a m_b}\{4m_b\partial_a U_b\partial_a V - \bar{h}^2\partial_{aabb}V - 2(p_a p_b\partial_{ab}V + \partial_{ab}V p_a p_b)\} \tag{1.7}$$

$$[[V, H_0], V] = \sum_a \frac{\bar{h}^2}{m_a}(\partial_a V)^2. \tag{1.8}$$

Sometimes, for instance when treating the dynamics of a weakly-coupled molecular complex, it is reasonable to assume that the potential V is not only small enough (in some sense) but also smooth enough to consider $\partial_a V$ as the terms of the smaller order of magnitude. In other words it should be advantageous to take into account only the commutators of the lowest orders at least after a relatively short time of evolution. Unfortunately, as follows from (1.6) and (1.7) the calculation of $\exp([X, Y])$ and $\exp([[X, Y], X])$ and the calculation of $\exp(X + Y)$ itself do not differ in the effort required, unless the potential V is not a linear function of all the variables. On the contrary, computing the values of expressions like $\exp([[X, Y], Y])$ is trivial for any V by virtue of (1.8). Thus, in the present work we develop a method of the EAs construction, treating only the lowest orders of S_k and paying special attention to the minimization of the numerical

coefficients at the terms of the sort (1.6), (1.7). We derive the general EA expression in the form

$$\exp(X+Y) \approx \exp(a_p X)\exp(b_p Y)\ldots\exp(a_1 X)\exp(b_1 Y) \tag{1.9}$$

where a_i, b_i are rational numbers and the correction term is presented as $\exp(S)$ (like (1.1)) in terms of the Lie elements in accordance with representations (1.2)–(1.4).

The structure of this paper is as follows. Section 2 presents a brief review of the algebraic assertions needed later in section 3 for the construction of the Lie elements $S_k, \Im_k, \Re_k$ for small k in the EAs of the sort (1.9). The approach used is a variant of the Magnus–Dynkin technique [10, 13]. Section 4 gives an analysis of the EA (1.1) as an illustration. The general problem (1.9) is considered in section 5, where some compact and effective EAs are derived and section 6 contains remarks on the use of the EAs for long-time evolution operator calculations in the spirit of the famous Lie–Trotter formulae.

To avoid any possible misunderstanding we would like to highlight that all power series or products like (1.2)–(1.4) are considered from the pure formal algebraic point of view and all problems connected with the radii of convergence, operator norms and domains are omitted. (An analysis of this range of problems is presented, for example, by Kumar [12].) We focus our attention on the algebraic structure of the expressions in use.

2. Algebraic tools

This section lists the principal definitions, formulae and statements from Lie-algebra theory employed in our further constructions. All the details and modern terminology may be found, e.g. in [14, 15]. All the operations are performed on the complex numerical field, symbols a, b, c, d are used for complex numbers.

Let us consider X and Y to be the non-commuting symbols generating a free associative algebra of polynomials $\aleph = \aleph(X, Y)$, i.e. a set of formal linear combinations of monomials

$$\chi = x_1 x_2 \ldots x_m \tag{2.1}$$

where $x_i = X$ or Y with customary multiplication defined. The degree m of the monomial (2.1) is denoted as $\deg\chi$. Functions on $\aleph$ are considered as the formal power series.

It is also possible to construct a Lie algebra on generators X and Y with the anticommutative operation $[\cdot,\cdot]$ as multiplication. Then a Lie element $\sigma(\chi)$ is an analogue for the monomial χ:

$$\sigma(\chi) = [\ldots[[x_1, x_2], x_3], \ldots, x_m] \tag{2.2}$$

where $\sigma(X) = X, \sigma(Y) = Y, \sigma(\text{constant}) = 0$. The non-zero Lie elements $\sigma(\chi)$ and their formal linear combinations form a subspace L in $\aleph$. The homogeneous Lie polynomials with degree $m > 0$ generate a linear subspace L_m in L. We consider σ to be a symbol of linear mapping $\sigma\colon \aleph \to L$ defined according to (2.2) for the monomials (2.1) and continued linearly to function from $\aleph$ onto L. For example

$$\sigma(Y\exp(X)) = \sum_{k=0}^{\infty}\frac{\sigma(YX^k)}{k!} = Y + [Y, X] + \tfrac{1}{2}[[Y, X], X] + \cdots \tag{2.3}$$

$$\sigma(Y(1+aX)\exp(aX)) = \sum_{k=0}^{\infty}\frac{\sigma(YX^k)a^k(k+1)}{k!}. \tag{2.4}$$

The following key statements [10, 13–15] are used further for any element $u \in L_n$ (we suppose $n > 0$).

The Dynkin–Specht theorem.

$$\sigma(u) = nu. \tag{2.5}$$

The Wever lemma. For any $\chi \in \aleph$

$$\sigma(u^2\chi) = 0. \tag{2.6}$$

It is critical that the degree of u is defined. Relation (2.6) is not valid for unspecified $u \in L$.

Corollary. If $u \in L_n$, $v \in L_n$ for arbitrary n and any $\chi \in \aleph$ then

$$\sigma(uv\chi) = -\sigma(vu\chi). \tag{2.7}$$

The last three statements follow from an important property of σ-mapping [15]: for $v \in L$, $\chi \in \aleph$

$$\sigma(\chi v) = [\sigma(\chi), v]. \tag{2.8}$$

This can be proved easily for $v \in L_m$ by induction on m. Due to (2.6) for any number a, any $\chi \in \aleph$ and $u \in L_n$, it is evident that the following relations arise for σ:

$$\sigma(\exp(au)) = a\sigma(u) \tag{2.9a}$$

$$\sigma(\exp(au)\chi) = \sigma(\chi) + a\sigma(u\chi) \tag{2.9b}$$

$$\sigma(u\exp(au)\chi) = \sigma(u\chi). \tag{2.9c}$$

Relation (2.5) demonstrates that σ maps from L_n into L_n and is reversible on L_n. As L is a direct sum of subsets $L_n (n > 0)$, the contraction of σ on L is also reversible. For example,

$$\text{if } u \in L \text{ and } \sigma(u) = 0 \text{ then } u = 0. \tag{2.10}$$

Together with σ, another linear mapping $\pi : \aleph \to L$ may be introduced by the following relations:

$$\pi(\text{constant}) = 0 \qquad \pi(\chi) = \frac{\sigma(\chi)}{\deg \chi} \tag{2.11}$$

for χ of the form (2.1) and continued linearly to function from $\aleph$ onto L. It is obvious (see (2.5)) that π is an identical operation on L. The main properties of this mapping are the following. First, by virtue of (2.9a),

$$\text{for } u \in L_m \qquad \pi(\exp(u)) = u. \tag{2.12}$$

Secondly, one can easily see that

$$\pi(YX^n) = \frac{\sigma(YX^n)}{n+1} = \frac{[\ldots[[Y, X], X], \ldots, X]}{n+1} \tag{2.13a}$$

$$\pi(Y\exp(X)) = \sum_{k=0}^{\infty} \frac{\sigma(YX^k)}{(k+1)!} \tag{2.13b}$$

$$\pi(Y(1+aX)\exp(aX)) = \sum_{k=0}^{\infty} \frac{a^k}{k!}\sigma(YX^k) = \sigma(Y\exp(aX)). \tag{2.13c}$$

This mapping is of interest to our goals for the following reason. As demonstrated later, the properties (2.9) of σ allow us to estimate the value $\sigma(S)$ for unknown $S \in L$ for some chosen EA. Then one may recover the value of S by applying the following lemma.

Lemma 1. If $S \in L$, $\chi \in \aleph$ then it follows from $\sigma(S) = \sigma(\chi)$ that

$$S = \pi(\chi). \tag{2.14}$$

Proof. Let $T = \pi(\chi)$. By virtue of (2.5), (2.11) for the monomials (2.1) with $m > 0$ and hence for any $\chi \in \aleph$

$$\sigma(\pi(\chi)) = \sigma(\chi). \tag{2.15}$$

That is why

$$\sigma(S - T) = \sigma(S) - \sigma(T) = \sigma(S) - \sigma(\pi(\chi)) = \sigma(S) - \sigma(\chi) = 0$$

and the assertion follows from (2.10). □

The method of Magnus [10] and Dynkin [13] employed in the present paper for EA construction is based on this lemma and the properties of the σ and π functions listed earlier. The principal point in applying lemma 1 is to demonstrate the fact that the values of function $S = S(X, Y)$ lie in L. For a simple check it is convenient to utilize the Friedrichs' criterion [10, 15].

Let $S(X, Y)$ be an element of $\aleph(X, Y)$. Let us introduce two additional generators X', Y', supposing $[X, Y] \neq 0$, $[X', Y'] \neq 0$, $[X, X'] = [X, Y'] = [Y, X'] = [Y, Y'] = 0$. Then we construct a free associative algebra of polynomials $\tilde{\aleph}(X, Y, X', Y')$ and define an element $S(X + X', Y + Y') \in \tilde{\aleph}$ that has the same particular form as the function $S(X, Y)$.

Friedrichs' criterion. The element $S(X, Y)$ belongs to L if and only if

$$S(X + X', Y + Y') = S(X, Y) + S(X', Y'). \tag{2.16}$$

Applying Friedrichs' criterion is straightforward. For instance, set $S(X, Y) = \ln(\exp(X)\exp(Y))$, then

$$\begin{aligned} S(X + X', Y + Y') &= \ln(\exp(X + X')\exp(Y + Y')) \\ &= \ln(\exp(X)\exp(Y)\exp(X')\exp(Y')) = \ln(\exp(S(X, Y))\exp S(X', Y')) \\ &= S(X, Y) + S(X', Y'). \end{aligned}$$

We have used only the properties of the formal power series for the exponent and logarithm for commuting values. Hence

$$\exp(X)\exp(Y) = \exp(S) \tag{2.17}$$

where $S \in L$. Here S is represented by the well known Baker–Campbell–Hausdorff (BCH) formula [10, 13, 15, 16]:

$$S = X + Y + \tfrac{1}{2}[X, Y] + \tfrac{1}{12}([[X, Y], Y] + [[Y, X], X]) + \cdots. \tag{2.18}$$

We present one other example: according to Friedrichs' criterion $\exp(-X)Y\exp(X)$ is a Lie element. On the other hand, owing to (2.7), (2.9*b*)

$$\sigma(\exp(-X)Y\exp(X)) = \sigma(Y\exp(X)) - \sigma(XY\exp(X)) = \sigma(Y(1 + X)\exp(X)).$$

In accordance with lemma 1 and (2.13*c*), we come to another well known BCH formula:

$$\exp(-X)Y\exp(X) = \pi(Y(1 + X)\exp(X)) = Y + [Y, X] + \tfrac{1}{2}[[Y, X], X] + \cdots.$$

3. The Magnus–Dynkin technique

Within the scope of our approach to the problem of the EA one may obtain correction terms in the form $\exp(S)$, where S is a Lie element. Thus, due to the properties of σ the problem of S-determination is reduced to the equation

$$\sigma(\exp(S)) = \sigma(q) \tag{3.1}$$

where q is an element of $\aleph$, calculated with the help of relations (2.9). Applying lemma 1 it is possible to estimate S in the manner described in points (A), (B) and (C) of the introduction.

From now on we label with a bottom index k the degree of the component S_k and with the indices x or y the terms which are linear in X or Y, respectively. The operations $(\cdot)_k$, $(\cdot)_{x,y}$ are mutually interchangeable and are also interchangeable with σ and π for the elements of the form (2.1) and thus for all elements of $\aleph$.

To construct the type A and B EAs it is practical to use the following.

Lemma 2. The solution of equation (3.1) may be expressed in terms of approximations (1.2)–(1.4) as

$$\begin{aligned} &S_1 = \Im_1 = \Re_1 = \pi(q_1) \qquad S_2 = \Im_2 = \Re_2 = \pi(q_2) \\ &S_3 = (\Re_3 + \Im_3)/2 \\ &\Re_3 = \pi(q_3 - \pi(q_1)\pi(q_2)) \qquad \Im_3 = \pi(q_3 - \pi(q_2)\pi(q_1)). \end{aligned} \tag{3.2}$$

For $k > 3$, the relevant expressions for $\Im_k, \Re_k$ have the same form as the following one for S_k:

$$S_k = \pi(S_k) = \pi(q_k - \eta_k) \tag{3.3}$$

where η_k is some combination of S_j for $j < k$.

Proof. In accordance with (2.15) one can reduce problem (3.1) to the system of equations

$$\sigma(\pi(\exp(S)))_k = \sigma(\pi(q))_k \tag{3.4}$$

which leads (by applying lemma 1) to the system:

$$\pi(\exp(S))_k = \pi(q)_k. \tag{3.5}$$

For an EA correction of the type (1.2)

$$(\exp(S))_k = \left(1 + S + \frac{1}{2}S^2 + \cdots + \frac{1}{k!}S^k\right)_k \tag{3.6}$$

since we only need to take into account terms of degree less than or equal to k. Several components from the right-hand side of (3.6) vanish owing to the Wever lemma (2.6). For instance for $k > 1$

$$\left(\pi\left(\frac{1}{k!}S^k\right)\right)_k = \pi\left(\frac{1}{k!}S_1^k\right) = 0.$$

As a result (3.5) is transformed into (3.3). Realization of this technique proposed by Magnus [10] leads to (3.2). □

Studying the terms of S which are linear in X or in Y is difficult enough in the general case, but may be simplified significantly with the help of several additional restrictions.

Lemma 3. Let $(S_x)_1 = 0$. Then for equation (3.1)

$$S_y = \pi(q_y). \tag{3.7}$$

Proof. As if $(S_x)_1 = 0$, we set $S = S_y + \tilde{S}$, where $\tilde{S}$ should contain the terms that are higher than linear in Y. In this case

$$(\exp(S))_y = (1 + S + \tfrac{1}{2}S^2 + \ldots)_y = (1 + S)_y = S_y$$

as S^k has a degree greater or equal to k with respect to Y. Further conclusions are the same as in the proof of (3.3). One should mention that the Wever lemma is not valid here.

Note that a trivial modification of the proof of lemma 3 demonstrates that $((\exp(S))_x)_1 = (S_x)_1 = (q_x)_1$ and hence lemma 3 can be applied to solve (3.1) if $(q_x)_1 = 0$.

In general it is rather difficult to calculate q_i in order to apply lemma 2 to solve (3.1). The essence of our technique is to employ lemma 3 for the construction of S_x and S_y. In both cases it is possible to utilize standard algebraic tools to compute $\pi(q)$ in the Y-linear (or X-linear) approximation. As mentioned in the introduction, the corrections of small degree are of the most interest for practical goals. Generally, we use only those EAs that coincide with the exact expressions in the limit $t \to 0$. It means that in all those EAs $S_1 = 0$ and by virtue of lemma 2 $\Re_3 = \Im_3 = S_3$ for the case at hand. Under this assumption the S-terms with degree $k \leqslant 3$ are

$$S_2 + S_3 = c_2[X, Y] + c_3[[X, Y], Y] + d_3[[Y, X], X]. \tag{3.8}$$

It is clear that c_2 and c_3 may be found as the coefficients at $(S_x)_2$ and $(S_x)_3$, while c_2, d_3 as the coefficients at $(S_y)_2, (S_y)_3$ similarly. Hence, it is sufficient to evaluate only S_x and S_y and to apply (3.2), (3.7) to settle problem (3.1) for approximations (A), (B) and (C) of several lowest orders. □

4. Examples of exponent approximation

Let us consider, for instance, a simple approximation of the form:

$$\exp(X + Y) \approx \exp(X)\exp(Y). \tag{4.1}$$

There are several EAs of this type, obtained by rearranging the cofactors:

$$\exp(X + Y) = \exp(X)\exp(Y)\exp(S) \tag{4.2}$$

$$\exp(X + Y) = \exp(X)\exp(S)\exp(Y) \tag{4.3}$$

$$\exp(X + Y) = \exp(S)\exp(X)\exp(Y) \tag{4.4}$$

together with those resulting from X, Y permutation. For all these six cases the function S differ. The method described in section 3 allows us to find S for any one of the problems (4.2)–(4.4) by means of trivial algebraic manipulations.

Let us examine (4.2) as an example. In this case

$$\exp(S) = \exp(-Y)\exp(-X)\exp(X + Y). \tag{4.5}$$

As follows from Friedrichs' criterion, S is a Lie element. Applying (2.9) one obtains for $\sigma(\exp(S))$:

$$\begin{aligned}\sigma(\exp(S)) &= \sigma(\exp(-X)\exp(X + Y)) - \sigma(Y\exp(-X)\exp(X + Y)) \\ &= \sigma(\exp(X + Y)) - \sigma(X\exp(X + Y)) - \sigma(Y\exp(-X)\exp(X + Y)) \\ &= (X + Y) - \sigma((X + Y)\exp(X + Y)) + \sigma(Y(1 - \exp(-X))\exp(X + Y)) \\ &= \sigma(Y(1 - \exp(-X))\exp(X + Y)).\end{aligned} \tag{4.6}$$

Since $1-\exp(-X)=X+\cdots$, $\exp(S)$ does not contain an S_1-type component. It is also clear that

$$(\sigma(\exp(S))_y=\sigma(Y(1-\exp(-X))\exp(X))=\sigma(Y(\exp(X)-1)). \tag{4.7}$$

By virtue of (2.13*b*), (3.7) and (4.7)

$$S_y=\pi(Y(\exp(X)-1))=\sum_{k=1}^{\infty}\frac{\sigma(YX^k)}{(k+1)!} \tag{4.8}$$

and similarly

$$S_x=-\pi(XY\exp(Y))=-\sum_{k=1}^{\infty}\frac{\sigma(XY^k)}{(k+1)!}k. \tag{4.9}$$

Hence, it is obvious that in (4.2)

$$c_2=-\tfrac{1}{2}\qquad c_3=-\tfrac{1}{3}\qquad d_3=\tfrac{1}{6}$$

and we come to the relation

$$\exp(X+Y)=\exp(X)\exp(Y)\exp(-\tfrac{1}{2}[X,Y]-\tfrac{1}{3}[[X,Y],Y]+\tfrac{1}{6}[[Y,X],X]+\ldots) \tag{4.10}$$

which may be compared to (2.17), (2.18). Different relations of this kind may be found in the literature [11, 12, 16] and many of them may be obtained by means of elementary tools especially in the case when some additional assumptions about X, Y and $[X,Y]$ are implied. In contrast our approach does not need any additional assumption of that sort. The simplest relations, that were found useful for practical goals (see e.g. [8] where the problem of molecular evolution was studied within the frameworks of adiabatic approximation), are presented in tables 1 and 2.

Table 1 contains the values of the symbols S_x, S_y and c_k, d_k coefficients for the EAs (4.2)–(4.4). Table 2 contains the values of S_k, $\Im_k$, $\Re_k$ for the simplest formulae of the kind

$$\exp(X+Y)=\exp(X)\exp(S). \tag{4.11}$$

Table 1. The linear components of the correction operator and the weights of the lowest-order commutators for different representations of $\exp(X+Y)$.

Exponent approximation	S_X-component	S_Y-component	c_2 $[X,Y]$	c_3 $[[X,Y],Y]$	d_3 $[[Y,X],X]$
$e^Xe^Xe^S$	$\pi(-XYe^Y)$	$\pi(Y(e^X-1))$	$-\frac{1}{2}$	$-\frac{1}{3}$	$\frac{1}{6}$
$e^Ye^Xe^S$	$\pi(X(e^Y-1))$	$\pi(-YXe^X)$	$\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{3}$
$e^Xe^Se^Y$	$\pi(X(e^{-Y}-1))$	$\pi(Y(e^X-1))$	$-\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{6}$
$e^Ye^Se^X$	$\pi(X(e^Y-1))$	$\pi(Y(e^{-X}-1))$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{6}$
$e^Se^Xe^Y$	$\pi(X(e^{-Y}-1))$	$\pi(YXe^{-X})$	$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{3}$
$e^Se^Ye^X$	$\pi(XYe^{-Y})$	$\pi(Y(e^{-X}-1))$	$\frac{1}{2}$	$-\frac{1}{3}$	$\frac{1}{6}$
$e^{X/2}e^Ye^{X/2}e^S$	$\pi(X/2((1-Y)e^Y-1))$	$\pi(Y(e^X-(1+X/2)e^{X/2}))$	0	$-\frac{1}{12}$	$\frac{1}{24}$
$e^{X/2}e^Ye^Se^{X/2}$	$\pi(X/2((1-Y)e^Y-1))$	$\pi(Y(e^{X/2}/2+e^{-X/2}/2-1))$	0	$-\frac{1}{12}$	$\frac{1}{24}$
$e^{X/2}e^Se^Ye^{X/2}$	$\pi(X/2((1+Y)e^{-Y}-1))$	$\pi(Y(e^{X/2}/2+e^{-X/2}/2-1))$	0	$-\frac{1}{12}$	$\frac{1}{24}$
$e^Se^{X/2}e^Ye^{X/2}$	$\pi(X/2((1+Y)e^{-Y}-1))$	$\pi(Y(e^{-X}-(1-X/2)e^{-X/2}))$	0	$-\frac{1}{12}$	$\frac{1}{24}$
$e^{Y/2}e^Xe^{Y/2}e^S$	$\pi(X(e^Y-(1+Y/2)e^{Y/2}))$	$\pi(Y/2((1-X)e^X-1))$	0	$\frac{1}{24}$	$-\frac{1}{12}$

Table 2. The lowest components for the simplest approximations of the types A, B and C.

Exponent approximation	$S_1=\Im_1=\Re_1$		$S_2=\Im_2=\Re_2$	S_3		$\Im_3$		$\Re_3$	
	X	Y	$[X,Y]$	$[[X,Y],Y]$	$[[Y,X],X]$	$[[X,Y],Y]$	$[[Y,X],X]$	$[[X,Y],Y]$	$[[Y,X],X]$
$e^X e^S$	0	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$-\frac{1}{3}$	$\frac{1}{6}$
$e^S e^X$	0	1	$\frac{1}{2}$	$-\frac{1}{12}$	$\frac{1}{6}$	$-\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$e^{X/2}e^S e^{X/2}$	0	1	0	$-\frac{1}{12}$	$\frac{1}{24}$	$-\frac{1}{12}$	$\frac{1}{24}$	$-\frac{1}{12}$	$\frac{1}{24}$

In this latter case one readily finds that

$$S_y = \pi(Y\exp(X)) = \sum_{k=0}^{\infty} \frac{\sigma(YX^k)}{(k+1)!}. \tag{4.12}$$

This relationship may also be obtained in a more standard form [11, 12]. It is easy to prove this by the induction of a trivial assertion:

Lemma 4. For the eigenvector $|0\rangle$ of the operator X with the eigenvalue x_0

$$\sigma(YX^n)|0>=(x_0-X)^nY|0\rangle. \tag{4.13}$$

In the Y-linear approximation this means that

$$\exp(X+Y)|0\rangle = \exp(X)\exp(S)|0\rangle = \exp(X)(1+S_y)|0\rangle$$
$$= \exp(X)\left(1+\sum_{k=0}^{\infty}\frac{(x_0-X)^n}{(k+1)!}Y\right)|0\rangle. \tag{4.14}$$

For the spectral projector Q corresponding to the eigenvalue x_0, $P=1-Q$, relation (4.14) transforms to (note that the operator $(x_0-X)^{-1}P$ should be defined correctly)

$$\exp(X+Y)|0\rangle \approx \exp(x_0)(1+QY+(x_0-X)^{-1}P(1-\exp(X-x_0)Y))|0\rangle. \tag{4.15}$$

This formula is standard for first-order perturbation theory.

5. The exponential approximations of a general case

Let us treat the EA which may be written as the following expansion with a finite number of cofactors:

$$\exp(X+Y) = \exp(a_pX)\exp(b_pY)\ldots\exp(a_1X)\exp(b_1Y)\exp(S) \tag{5.1}$$

where a_i, b_i are the numerical coefficients. Let us denote

$$\alpha_0=\beta_0=0 \qquad \alpha_j=\sum_{i=1}^{j}a_i \qquad \beta_j=\sum_{i=1}^{j}b_i. \tag{5.2}$$

Then the relation $S_1=0$ means

$$\alpha_p=\beta_p=1. \tag{5.3}$$

Theorem 1. For an EA of the type (5.1) with conditions (5.3)

$$S_y=\sum_{k=0}^{\infty}\frac{\sigma(YX^k)}{k!}\left(\frac{1}{k+1}-\sum_{j=1}^{p}b_j\alpha_{j-1}^k\right) \qquad S_x=\sum_{k=0}^{\infty}\frac{\sigma(XY^k)}{k!}\left(\frac{1}{k+1}-\sum_{j=1}^{p}a_j\beta_j^k\right). \tag{5.4}$$

Proof. It is evident that

$$\exp(S) = \exp(-b_1 Y)\exp(-a_1 X)\ldots\exp(-b_p Y)\exp(-a_p X)\exp(X+Y)$$

and S satisfies Friedrichs' criterion. The use of an approximation, linear with respect to Y, under assumption (5.3), yields

$$(\exp(S))_y = (\exp(-X)\exp(X+Y))_y - \sum_{j=1}^{p} b_j \exp(-\alpha_{j-1}X)Y\exp(\alpha_{j-1}X).$$

According to lemma 3 and due to the properties (2.9) of the σ- mapping (in just the same manner as that used for (4.6)) we obtain

$$S_y = \pi(Y\exp(X) - \sum_{j=1}^{p} b_j \exp(-\alpha_{j-1}X)Y\exp(\alpha_{j-1}X)) = \pi(Y F_p(X)) \tag{5.5}$$

where

$$F_p(X) = \exp(X) - \sum_{j=1}^{p} b_j(1+\alpha_{j-1}X)\exp(\alpha_{j-1}X).$$

With the help of (2.13*b*), (2.13*c*) we find (5.4). The value of S_x is calculated analogously. □

Note that, as it in fact must be, the second-order components of S_x and S_y are equal in magnitude by virtue of (5.3).

As mentioned in the introduction, the EA which has zero coefficients for terms of the type $[YX^k]$ represents the most interest for practical purposes. This requirement leads to a system of equations (see(5.4)) for some M:

$$\sum_{j=1}^{p} b_j \alpha_{j-1}^k = \frac{1}{k+1} \qquad k = 0, 1, 2, \ldots, M. \tag{5.6}$$

Relations of this type are known as the moment problem which has a unique solution in the positive area (this follows [17] from the positive definity of the matrix with elements $A_{ij} = \left(\frac{1}{i+j+1}\right)$). For instance, for $p = 2$, $M = 2$ problem (5.6) has an explicit solution that leads to the following EAs:

$$\exp(X+Y) \approx \exp\left(\frac{X}{3}\right)\exp\left(\frac{3}{4}Y\right)\exp\left(\frac{2}{3}X\right)\exp\left(\frac{Y}{4}\right)\exp(S) \tag{5.7a}$$

$$\exp(X+Y) \approx \exp\left(\frac{Y}{4}\right)\exp\left(\frac{2}{3}X\right)\exp\left(\frac{3}{4}Y\right)\exp\left(\frac{X}{3}\right)\exp(S) \tag{5.7b}$$

where $S = -\frac{1}{48}[[X, Y], Y]$. For both EAs (5.7) the difference between the right-hand side and the left-hand side contains terms of the fourth and higher degree, i.e. (5.7) is accurate up to the third order. The replacement of the $\exp(S)$ cofactor in the spirit of (4.2)–(4.4) gives different expressions for S but does not change the lowest-order correction component in accordance with the following.

Theorem 2. Let us consider two EAs:

$$\exp(X+Y) = \exp(A)\exp(B)\exp(S)$$
$$\exp(X+Y) = \exp(A)\exp(\tilde{S})\exp(B)$$

for some Lie elements A, B. The lowest-order non-zero Lie components of S and $\tilde{S}$ have the same value.

Proof. In accordance with the definition of S we have

$$\exp(\tilde{S}) = \exp(-A)\exp(X+Y)\exp(-B) = \exp(B)\exp(S)\exp(-B).$$

Due to (2.19) this means that

$$\exp(\tilde{S}) = \exp(S) + [B, \exp(S)] + \tfrac{1}{2}[[\exp(S), B], B] + \cdots$$

and now it is sufficient to modify the proof of lemma 3 to demonstrate that the lowest-degree non-zero Lie components of $\exp(\tilde{S})$ and $\exp(S)$ are the same.

Note that if $S = S_n + S_{n+1} + \cdots$ and B_1 is not equal to zero, then $\tilde{S}_{n+1} = S_{n+1} + [B_1, S_n]$ and similar estimations may be found for the higher orders. □

With the help of our formulae it is possible to find different EAs of a particular specified form. Here one should note that for the case under study (see (1.8)) $\sigma(XY^k) = 0$ when $k > 2$. Hence, to provide the simultaneous vanishing of the coefficients at the terms $\sigma(XY^m)$ and $\sigma(YX^m)$ we need to resolve together with (5.6) only the following relation:

$$\sum_{j=1}^{p} a_j \beta_j^2 = \tfrac{1}{3}. \tag{5.8}$$

As relation (5.3) shows, the problem is insoluble for $p = 2, M = 2$. For $p = 3$ it is easy to express b_2, b_3 as functions of α_1 and α_2, applying (5.6) for $M = 2$. The substitution of these results in (5.8) leads to an equation for α_1 and α_2:

$$\alpha_1^2 3(4\alpha_2 - 3) - \alpha_1\alpha_2 3(4\alpha_2 - 3) + (3\alpha_2 - 2)^2 = 0 \tag{5.9}$$

and it is not difficult to demonstrate the absence of any real-valued solutions in positive numbers a_i, b_i. We have noticed that among others the continuum of solutions with some negative a_i, b_i exists for this equation. We present only one here as an example

$$\exp(X+Y) \approx \exp\left(\frac{7}{24}X\right)\exp\left(\frac{13}{51}Y\right)\exp\left(-\frac{X}{24}\right)\exp\left(-\frac{2}{3}Y\right)\exp\left(\frac{3}{4}X\right)\exp\left(\frac{24}{17}Y\right) \tag{5.10}$$

which has the same accuracy as (5.7). A similar EA constructed by Bandrauk and Shen [18] has seven cofactors and irrational coefficients—the latter include negative ones. However, the use of evolution operator expansions with variable time direction seems to be strange enough and inconvenient in practice. In our opinion EA (5.7) is optimal.

In practice one may use large p values and, in any case, equations (5.6), (5.8) allow us to exclude commutators of the sort $\sigma(XY^m)$ or $\sigma(YX^m)$. Delivering from terms such as $\sigma(XYXY)$ needs much more effort and we doubt whether this effort is reasonable.

6. Concluding remarks

Applying the EAs considered earlier for numerical modelling of a real molecular system we find they have some rather unpleasant properties. While for relatively short times t the evolution operator $U(t)$ does not differ significantly from an approximated one $W(t)$ and the difference may be estimated as

$$\|U(t) - W(t)\| \leqslant C_m t^m \tag{6.1}$$

where C_m is some constant and m is the lowest non-zero order in $(U(t))^{-1}W(t)$, at relatively long evolution times this difference may grow significantly. Whereas we usually consider the Hermitean Hamiltonians, the evolution operators $U(t)$, $W(t)$ are unitary in our technique,

thus they differ in norm by no more than 2, but this is too rough for practical calculation. The way out may be found in employing the formulae which are very popular for large t in semigroup theory and for numerical approaches to the problems of molecular dynamics i.e. the EAs based on the Lie–Trotter formulae [19, 20]

$$\exp(X+Y) = \lim_{n\to\infty}\left(\exp\left(\frac{X}{n}\right)\exp\left(\frac{Y}{n}\right)\right)^n \tag{6.2}$$

as well as its symmetrized analogue [19, 21, 22]

$$\exp(X+Y) = \lim_{n\to\infty}\left(\exp\left(\frac{X}{2n}\right)\exp\left(\frac{Y}{n}\right)\exp\left(\frac{X}{2n}\right)\right)^n. \tag{6.3}$$

([20] presents a complete analysis of the evolution operator for the case of very small evolution times.)

Taking into account that $U(t) = (U(t/n))^n$ and applying (6.1) one may see after ordinary transformations that [19]

$$\begin{aligned}\|U(t) - (W(t/n))^n\| &= \|(U(t/n))^n - (W(t/n))^n\| \\ &= \|\sum_{k=0}^{n-1}(U(t/n))^{n-k-1}(U(t/n) - W(t/n))(W(t/n))^k\| \\ &\leqslant n\|U(t/n) - W(t/n)\| \leqslant nC_m(t/n)^m. \end{aligned}\tag{6.4}$$

As a result

$$U(t) = \lim_{n\to\infty}(W(t/n))^n. \tag{6.5}$$

Let us consider as an illustration relations of the type (5.1) setting all the exponents $a_j = b_j = 1/p$ (this is a case of Trotter expansion (6.2)). Then $S_y = \pi(YF_p)$, where (see (5.5))

$$F_p(X) = \exp(X) - \sum_{j=1}^{p}\frac{1}{p}\left(1 + \frac{j-1}{p}X\right)\exp\left(\frac{j-1}{p}X\right). \tag{6.6}$$

The sum on the right-hand side of this expression is a Darbou sum for the integral

$$\int_0^1 (1+zx)\exp(zx)\,\mathrm{d}z = \exp(x)$$

and the tendency of F_p to vanish as $p \to \infty$ illustrates relations (6.2) and (6.3). (For (6.3) the Darbou sum is replaced with the trapezoidal formula.) The efficiency of the high-order formulae like (6.2) and (6.3) is demonstrated by a large number of numerical experiments [5–8, 18, 21, 22].

As follows from (6.4), the growth of t in calculations with a given accuracy level for $\|U(t) - (W(t/n))^n\|$ leads to the non-linear growth of n. For small m the time-step (t/n) decreases considerably while t increases, e.g. for t growth 100 times for $m = 2$ (the case of Trotter's EA, (6.2)), the time-step should be decreased 100 times for given accuracy. For the case corresponding to (6.3) ($m = 3$), the time-step should be decreased 10 times, and for our favourite EA corresponding to (5.7) with $m = 4$—only 2.2 times. The calculation efforts increase only twice here compared with (6.2).

In conclusion we would like to mention again that the constructions described here are effective for a finite time of evolution. Unfortunately, the technique derived here is not potent enough for a direct description of asymptotic $t \to \infty$. An analysis of the formulae from tables 1 and 2 demonstrates that in relations like (4.15) the terms responsible for the delta-form tendency are corrected by some non-singular functions when $t \to \infty$. Roughly

speaking, the singular term in (4.15) is 'omitted' in EAs of the type (5.1) with a finite-degree correction as an approximation for S. Nevertheless, the method described seems to be at least attractive for the analysis of the S-matrix in the scattering theory.

Acknowledgments

This work has been supported by the Russian Fund of Fundamental Research, project 95–03–09710. The authors are deeply indebted to Professor N F Stepanov whose pertinent comments have led to substantial improvements in this manuscript.

References

[1] Miller W H, Schwartz S D and Tromp J W 1983 *J. Chem. Phys.* **79** 4889
[2] Taylor H and Simons J 1986 *J. Phys. Chem.* **90** 580
[3] Yamashita K and Miller W H 1985 *J. Chem. Phys.* **82** 5475
[4] Alverellos J and Metiu H 1988 *J. Chem. Phys.* **88** 4957
[5] Schwartz S D 1989 *J. Chem. Phys.* **91** 7621
[6] Schwartz S D 1994 *J. Chem. Phys.* **100** 8795
[7] Schwartz S D 1994 *J. Chem. Phys.* **101** 10436
[8] Schwartz S D 1996 *J. Chem. Phys.* **104** 1394
[9] Meyer H D, Kucar J and Cederbaum L S 1988 *J. Math. Phys.* **29** 1417
[10] Magnus W 1954 *Commun. Pure Appl. Math* **7** 649
[11] Wilcox R M 1967 *J. Math. Phys.* **8** 962
[12] Kumar K 1965 *J. Math. Phys.* **6** 1928
[13] Dynkin E B 1947 *Dokl. Akad. Nauk* **LVII** 323 (in Russian)
[14] Serre J P 1992 *Lie-algebras and Lie-groups* (Berlin: Springer)
[15] Postnikov M M 1982 *Lie-groups and Lie-algebras* (Moscow: Nauka)
[16] Weiss G H and Maradudin A A 1962 *J. Math. Phys.* **3** 771
[17] Gantmakher F R 1988 *The Theory of Matrices* (Moscow: Nauka) (in Russian)
[18] Bandrauk A D and Shen H 1991 *Chem. Phys. Lett.* **176** 428
[19] Reed M and Simon B 1972 *Methods of Modern Mathematical Physics* vol 1 (New York: Academic)
Reed M and Simon B 1975 *Methods of Modern Mathematical Physics* vol 2 (New York: Academic)
[20] Makri N and Miller W H 1989 *J. Chem. Phys.* **90** 904
[21] Fleck J A Jr, Morris J R and Feit M D 1976 *Appl. Phys.* **10** 129
[22] Feit M D and Fleck J A Jr 1983 *J. Chem. Phys.* **78** 301