

“Comment” on the Reply to the paper “On the Elber–Karplus Reaction Path-Following Method and Related Procedures”

L.L. Stachó^a, Gy. Dömötör^b and M.I. Bán^{b,*}

^a Bolyai Institute, JATE University of Szeged, H-6720 Szeged, Aradi Vértanúk tere 1, Hungary
E-mail: stacho@math.u-szeged.hu

^b Institute of Physical Chemistry, JATE University of Szeged, H-6720 Szeged, POB 105, Hungary
E-mail: {domotor, m.i.ban}@chem.u-szeged.hu

Received 10 May 2000

The flaws in the “Reply” [1] to our paper [2] have been pointed out. Elber and Karplus (EK) have not disproved our irrefutable global statement that the energy average cannot be minimized which rebuts the theoretical background of EK-type calculations. Another statement of ours has shown that even a curve for which the average energy is locally minimal for all directional perturbations in the sense of classical variational calculus cannot be identical with the reaction path (RP) defined as a steepest descent path (SDP). EK found an error in the early preprint of our theoretical paper [3] and because of this error they qualified our correct variational statement as false for all the SDPs consisting of a straight line each. Mixing global and variational arguments, EK refuted our criticism in a logically incorrect manner. In this Comment we prove that both of our earlier statements invariably remain in force and the criticism included in those has been as well-established and solid as was before.

KEY WORDS: Elber–Karplus procedures, reaction path-following, DDRP method

1. Introduction

Based on both theoretical considerations and numerical experiences, we published a paper [2] which pointed out serious conceptual errors in the popular Elber–Karplus (EK) algorithm [5] and in its sequels of path-following procedures [6–9]. In a polemic paper [1] the authors replied to our arguments, insisting on the validity of their strategy. However, as we are going to point out in this “Comment”, EK do not touch the main issue at all. They absolutely disregard our statement asserting [2,3] that the average functional

$$A = \frac{\int_R^P U(\vec{\ell}) |d\vec{\ell}|}{\int_R^P |d\vec{\ell}|} \quad (1)$$

* Corresponding author.

has no globally minimizing curve unless the two minima belonging to reactants and products are joined by a curve with constant energy values (what is a chemically irrelevant and absurd case). Instead, EK exhibits a computational error in the Euler–Lagrange equation in an early preprint of the theoretical paper [3] cited in [2] in which the minimum principle of the EK method has been examined from a variational point of view. The paper [3] has been corrected before publication, with the conclusion that a reaction path (RP in the usual Fukui sense) is no local minimum of the energy average functional on some straight line containing it in the space of the curves joining the configurations of reactants and products. Knowing the correct form of the Euler–Lagrange equation, EK have not proceeded toward the (easy) modification of our later considerations in the preprint of [3] leading to the above mentioned negative outcome for the valuation of their method. Moreover, EK have insisted [1] that minimization of the energy average functional leads to some curve which perhaps is no Fukui-type RP, however, it passes near the Fukui-type RP if this latter has small curvature. To illustrate their ideas, they take the function

$$U(x, y) = kx^2 + y^4 - y^2 \quad (2)$$

as a counter-example to our arguments. They even suggested that the SDP and the approximate RP defined as the energy average minimizing curve should be regarded same in this case. At this stage we recall the incontrovertible mathematical fact [2] that none of the curves joining one local minimum to another can minimize the energy average functional. However, a discretized version of any EK-type method with Powell minimization, without doubt, produces easily the straight line segment RP of the function U when using a fixed number of points for curve representations. In order to clarify the situation, we are to examine the case of this model function in detail. By precise MAPLE calculations [10] we prove the existence of arbitrarily small admissible modifications of the RP of U with smaller energy average than that of the RP. With this observation we mean that the minimization algorithms implemented by EK do not work properly. We point out the reason: discretized approximate line integrals along with finite dimensional minimization procedures (like the Powell algorithm) with a fixed number of points are unreliable tools in minimization of functionals of line integrals. This is a serious warning for designers of numerical algorithms based on variational principles.

For their defense, EK quote another variational principle due to Olender–Elber (OE) [8] stating that RPs minimize the line integral of the gradient norm. We show by a counter-example that, unfortunately, this claim is also false in general. The original sketched proof [8] is completely wrong. However, this latter principle works under some mathematically heavy but chemically not unrealistic conditions if the number of critical points on the potential energy surface is small, but for completely different reasons [11] as are the arguments in [8].

Clearing up the main theoretical reasons for the numerical misgivings of EK-type methods, we designed numerical experiments to establish that coincidences of paths with the RP as the results of such algorithms are incidental. In the meantime we have published such results [4,12].

2. Discussion

We have already proved [2–4,12,13] that the average (1) of the line integral of the energy function serving as a base for the EK method and its sequels has no minimum. In our proof we have presented a procedure by which to any curve L we can construct another curve L' with the same starting and endpoints for which the average of the line integral is strictly smaller than that for L unless the energy function assumes its minimum at every point of L . EK did not even try to disprove this our main statement because it was irrefutable and indisputably faultless. Occasional numerical errors in the calculations of our given examples may occur because of the lack of the original authentic algorithms and the use of a general version of the Powell method [14] which does not require the knowledge of the analytic derivative of the function to be minimized. However, it should now be stressed: the gravest insufficiency of the EK method is that based on the minimum of the line integral average *which really does not exist*. Our other objection against the strategy of the EK-type methods is that – though the minimum for the discretized form of the line integral does exist – the concrete calculations based on the discretized line integral are still strongly parameter dependent. EK argue that the penalties that avoid rigid body translations and rotations do not influence the value of the functional and the equidistant constraint does not affect significantly the shape of the RP. Nevertheless, using fixed numbers of points (as they do) this statement is not true as we have already shown by examples [2,4,12,13]. Therefore, the hollow pretext by EK, simply saying that they do not use rigid body constraint in the Müller–Brown curve calculations since it is unnecessary in that case, cannot be accidental. Their true reason for not using the penalty function to avoid rigid body movements is that they would get much worse results if they employed the choice of $\lambda = \lambda' = 4000, 8000$ or 16000 as they had done in former examples.

Another – even more general – statement has been taken in [3] where we have shown that by using variational analysis (i.e., weaker constraints) even local minima of the line integral average can not be regarded as a steepest descent path (SDP). There we have also acknowledged EK for reminding us to a calculational error. However, after correcting this error our statements [2] are invariably valid. In their "Reply" [1] EK hoped to find a counter-example by constructing the 2D potential energy surface of the energy function (2). They argue that the y -axis joining the two minima is an SDP ($x = 0$) and this is also the minimum of the line integral average. Out of this is only true that the y -axis is a SDP and may also be true that the curve calculated by the EK procedure is coinciding with it. Nevertheless we can easily demonstrate that line integral averages smaller than those calculated for curves running along the y -axis do really exist.

By using MAPLE V calculations [10] ($k = 1$) with the control sequence

```
a:=sqrt(1/2); assume(t>0); assume(u>=-Pi/2,u<=Pi/2);
C:=s+t*sqrt(a^2-s^2); Cu:=simplify(subs(s=a*sin(u),C));
ut:=solve(diff(Cu,u)=0,u);
Lt:=int(diff(Cu,u),u=-Pi/2..ut)-int(diff(Cu,u),u=ut..Pi/2);
U:=k*x^2+y^4-y^2; Fu:=subs(x=0,y=Cu,U)*diff(Cu,u);
```

$\text{It} := \text{int}(\text{Fu}, u = -\text{Pi}/2..ut) - \text{int}(\text{Fu}, u = ut, .\text{Pi}/2);$ $\text{At} := \text{It}/\text{Lt};$
 $\text{DDA0} := \text{limit}(\text{diff}(\text{At}, t, t), t = 0, \text{right});$
 we prove that the curves C_t ($t \geq 0$) with the coordinate components

$$(C_t)_x = t\sqrt{a^2 - s^2}, \quad (C_t)_y = s + t\sqrt{a^2 - s^2} \quad (3)$$

(here $a = 1/\sqrt{2}$ and $-a \leq s \leq a$) at any sufficiently small positive t parameter values give smaller line integral average value than the SDP curve ($x = 0$). Using a simpler assignment

$$C_t : [-a, a] \ni s \mapsto (0, s + t\sqrt{a^2 - s^2}) \quad (4)$$

for arc-length proportionally parameterized curves MAPLE calculations [10] give the second derivative

$$\left. \frac{d^2}{dt^2} A(C_t) \right|_{t=0+} = -\frac{2}{15} < 0, \quad (5)$$

whose negative value shows that the SDP is not a local minimum of the set of curves. Consequently, $A(C_t) < A(C_0)$ if $0 < t < t_0$ for some $t_0 > 0$ value and thus the functional A is not a local minimum in the sense of variational calculus. The inequality $A(C_t) < A(C_0)$ can especially easy be verified for $t = 0.01, 0.02$ and 0.03 .

Without doubt, several numerical EK type algorithms find the proper RP (the curve C_0) in this case. However, this fact is not, least, a good message for the users of such programs. Namely, if their minimization works properly, they should have found curves better than the RP arbitrarily close to it. The explanation of this paradox is as follows. If we use a fixed number, e.g. n , of nearly equidistant points to represent curves and if we apply numerical approximation with a given order in terms of the points of representation for line integrals, then the parts of the curves C_t outside the segment C_0 do not contain any points of their representation for small values of the parameter t such that $\text{length}(C_t) - \text{length}(C_0) < a/n = 1/n\sqrt{2}$. Thus the small, nevertheless – for the integral average crucial – part $C_t \setminus C_0$ (i.e., the piece of the segment C_t outside C_0) of the curve C_t has no influence at all to the numerical calculation of the energy average along the curve C_t for sufficiently small values of t . This consideration can be modified in a straightforward manner to the case of any curve C_0^* joining the global minimum with strictly positive definite Hessian of an energy function with another local minimum with higher energy value. Therefore, in general, *discretized numerical integration with a fixed number of points along the curves in such calculations cannot be combined with finite dimensional minimization algorithms (e.g., Powell's minimization) to determine minimizing curves of functionals of line integrals.*

We do not deny, of course, the existence of the line integral or its average (EK give the matter the appearance as if we would do) but we do not accept the existence of the minimum of the line integral average. Parts of our paper cited carelessly or in an intentionally wrong way can be suited to mislead the reader uninformed in the given topic. EK state that "other algorithms developed by Elber and coworkers for determining RPs for larger systems, such as that of Ulitsky and Elber (UE) [9] and of OE [8] compute

the exact SDP". The proof of this statement given in [8] is false. OE disregard that by varying the curve its length will also be changed. Assuming two minima and one saddle point (SP) and accepting some chemically not irreal conditions, an absolutely correct proof [11] entirely different from that described in [8] can be given. To verify this we give a simple 2D artificial counterexample the RP of which determined by the OE method [8] is, definitely, not a SDP. Let us take the function

$$U = (1 - r^2)^8 - r^{14}(2 - r^2)^7 \cos 6\varphi, \quad (6)$$

where r and φ are polar coordinates. Fukui's RPs are the segments of the unit circle with its centre in the origin and those radii of the circle which join the centre of the circle with the local minima defined by the relation

$$P_k = (\cos k\pi/m, \sin k\pi/m), \quad k = 0, \dots, 2m - 1. \quad (7)$$

Denote the curve between the points $(-1, 0)$ and $(1, 0)$ by C_t which is composed from the semicircle lying on the upper semiplane and having a radius t and its center in the origin, and from two straight line segments of the section of length $1 - t$. Let $G(C)$ be the functional defined by equation (26) in [8]:

$$G(C) = \int_C \|\nabla U\| d\ell. \quad (8)$$

Direct MAPLE calculations [10] show that

$$G(C_{1/2}) < G(C_0), \quad G(C_{1/2}) < G(C_1). \quad (9)$$

As G is not minimized by C_0 and C_1 therefore from these two inequalities follows that the curve which joins the points $(0, 1) = P_0$ and $(0, -1) = P_3$ and minimize G cannot be a SDP. This our counter-example verifies that the proof given in OE [8] is incorrect. Nevertheless, though the curves minimizing G can be used as an optional definition for a new RP concept they cannot be used as a new determination method of SDPs. In the meantime we gained negative experiences by using the authentic Czerminski–Elber (CzE) [6] algorithm implemented in the program package TINKER [7]. These results were presented at the WATOC'99 congress [4,12].

3. Conclusions

1. EK have never mentioned in their papers that the minimum of the average of the line integral does not exist, therefore it is unfair that they suppress this fact which we have proved by exact mathematical arguments. In the same time all the RP-following algorithms operating by the EK strategy have been based on the existence of the minimum of the energy functional average.

2. We have also proved that it is not correct to use a new concept introduced by EK which defines the RP as the minimization curve of the energy functional average. We reject the way by which EK are attacking our strictly mathematical proof and its

results without exact mathematical contra-argumentation. The "Reply" mostly contains mathematically unproven statements, therefore the reader who reads this and our paper in succession will not be able to judge the two works objectively with equal measures and to make distinction between the qualities and truth-contents of the two papers. We are also arguing against the statement according to which the piling up of the points at the end positions is well known from the literature and therefore, in the first part of our proof there is no novelty. EK and other authors using the EK strategy have never investigated *the reason* for this congestion of points. Had they investigated it, they should have come to realize that this congestion is nothing else but just *the consequence of the non-existence of the energy average minimum*. Thus, those who have been using the EK strategy automatically and without critics have always been rejoicing whenever they get sections of curves between the congestional places of clustered points corresponding, approximately, to some experimental data while just this congestion should have raised suspicions and doubts against the method in the authors using the EK strategy. Another question – based on experiences of computer experiments with the EK-type algorithms – raised by us from the beginning: whether the accidental agreements between curve sections joining the congestion places and the experimental data can or cannot be ascribed to the side-effects inherent in the algorithms or in the numerical realizations of the methods applied (cf. the RP calculations on the Müller–Brown potential in [1,2]). To our opinion, EK (and followers) should have only stated fairly that by using their algorithms or numerical methods they obtained numerous different chemically relevant and explainable results in some cases and with given constraints (parameters and/or penalty functions). Such an "experimental" method – even without adequate mathematical foundation – could be acceptable, especially when applied to very large chemical systems. It is, nevertheless, absolutely unacceptable if someone tries to justify such an "experimental" method by a deficient concept and faulty mathematical argumentations. This way would then be a pseudo-scientific and not a true scientific approach to the real problem and, therefore, must not be accepted in exact natural sciences. It is important to call attention to the potentially dangerous tendency by promising a possibility to theoretically unqualified users to apply an algorithm and/or numerical method automatically and therefore without critics for characterizing large organic and biological systems, which, instead of using a theoretically well-established concept, is only pretended to give a correct definition of the RP.

3. Another statement of EK, without proof again, that the calculation results provided by their numeric method (and those by other methods based on the EK strategy) are independent of the parameters and/or penalty functions chosen. We have experienced quite the contrary and could not find positive arguments supporting the statement of EK (cf. Müller–Brown curve calculations [1,2]). EK have criticized our numerical experiments saying that we did not use the authentic algorithms. An algorithm or numerical method should be reproducible by the information given in a theoretical paper. We accomplished the reproduction as much as it was made possible by the description given in the theoretical papers. We are, of course, aware of all the deficiencies and know that the algorithms used by us cannot be perfectly the same as the original ones (just

because the EK paper [5] did not give sufficient basis for the correct reproduction). Nevertheless, we have made several variants [4,12,13] following the principles described in the EK paper [5] (and in papers of other authors [6,8,9,15]) so they may not differ substantially from the authentic versions and, therefore, they should produce the same results with minor deviations. We have already offered and now repeat the possibility to show and give our authentic DDRP algorithm on our homepage asking to exchange it with the authentic other algorithms. In this way the numeric calculation results could be reproduced and checked by the authors working on the same field, immediately. Many misunderstandings could then be cleared up directly and easily.

4. We were ready to admit that we made an error in the derivation of the mathematical proof described at an earlier stage of the preprint form of our paper. This means that equation (6) in the "Reply" [1] is the right expression. For the detection of this error we are very much obliged to EK and this fact has been acknowledged in the published version [3]. However, it must be stressed that this error (made *in the proof* of the arguments in the preprint of [3]) does affect in no way the arguments themselves and their consequences to the EK strategy. Therefore EK have not enough base and right to come to the conclusion that because of the error in the proof our fundamental mathematical arguments are also faulty.

Acknowledgements

Supported by the Hungarian Scientific Research Fund OTKA, Grant No. T 020743.

References

- [1] R. Elber and M. Karplus, Chem. Phys. Lett. 311 (1999) 335.
- [2] L.L. Stachó, Gy. Dömötör and M.I. Bán, Chem. Phys. Lett. 311 (1999) 328.
- [3] L.L. Stachó, Gy. Dömötör and M.I. Bán, J. Math. Chem. 26 (1999) 87.
- [4] Gy. Dömötör, L.L. Stachó and M.I. Bán, Comparative study of global path-following methods, in: *WATOC'99 (5th World Congress of Theoretically Oriented Chemists), Book of Abstracts*, London, UK, 1–6 August 1999, p. 335; to be published in J. Comput. Chem. (in preparation).
- [5] R. Elber and M. Karplus, Chem. Phys. Lett. 139 (1987) 3755.
- [6] R. Czerminski and R. Elber, J. Chem. Phys. 92 (1990) 5580.
- [7] S. Huston and J.W. Ponder, TINKER: Software tools for molecular design, Version 3.5 (October 1997) (Copyright Jay William Ponder, 1990–1997).
- [8] R. Olender and R. Elber, J. Mol. Struct. (Theochem) 398–399 (1997) 63.
- [9] A. Ulitsky and R. Elber, J. Chem. Phys. 96 (1990) 1510.
- [10] Maple V Release 4, Version 4.00b (1981–1996).
- [11] L.L. Stachó, Gy. Dömötör and M.I. Bán, J. Math. Chem. 28 (2000) 241.
- [12] L.L. Stachó, M.I. Bán and Gy. Dömötör, Reaction paths defined by global variational principles, in: *WATOC'99 (5th World Congress of Theoretically Oriented Chemists), Book of Abstracts*, London, UK, 1–6 August 1999, p. 333; published in [11].
- [13] Gy. Dömötör, L.L. Stachó and M.I. Bán, J. Mol. Struct. (Theochem) 501–502 (2000) 509.
- [14] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, Numerical recipes, in: *The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986) p. 294.
- [15] S.S.L. Chiu, J.J.W. Douall and I.H. Hillier, J. Chem. Soc. Farad. Trans. 90 (1994) 1575.