Journal of Mathematical Chemistry 18 (1995) 121–132

About a certain dynamical system in molecular configuration space

Myriam M.S. de Giambiagi [a], Mario Giambiagi [a] and André Ricardo Fonseca *[b]

[a] *Centro Brasileiro de Pesquisas Físicas, Rua Dr. Xavier Sigaud 150, 22290-180 Rio de Janeiro, RJ, Brazil*

[b] *Depto. Matemática Aplicada, Universidade Federal do Rio de Janeiro, Ilha do Fundão, 21945-970 Rio de Janeiro, RJ, Brazil*

Received 9 May 1995; revised 26 July 1995

As an example of a dynamical system, glycine has been calculated through the molecular mechanics approach. Of the three planar conformations studied, only one equilibrium points region could be considered as a catchment region. On the other hand, an intriguing relation appears between the gradient norm and the spectral radius of the matrix representing the metric tensor.

1. Introduction

We started this work having in mind the study of a dynamical system [1]. As an example of an energy-minimized, geometry-optimized molecular system we chose glycine, for which we also made a search for catchment regions [2]. The corresponding calculations have been performed using the molecular mechanics (MM) method [3]. We have studied the three most usual planar conformations; we show in more detail the one exhibiting the greatest diversity in the features of the dynamical system solutions.

We have followed the spectral radius of the metric tensor, along several of the solutions. We have then gathered in a table the spectral radii obtained in the final points as well as the corresponding gradient norms. Unexpectedly, a relation seems to link these two quantities.

2. Manifolds, dynamical systems and catchment regions

We denote by R^n the set of all n-tuples of real numbers $(x_1, x_2, \ldots, x_n)$, i.e. the usual n-dimensional space of vector algebra. A set (of "points") M is defined to be a

* Part of a thesis for a master's degree in applied mathematics.

manifold if each point P of M has an open neighbourhood which has a homeomorphism f onto an open set of R^n for some n (see fig. 1). A homeomorphism is a continuous 1–1 map from one space to another, having a continuous inverse. Thus, M is locally "like" R^n [1]. The map is not required to preserve lengths or angles.

Let us now see dynamical phenomena which may be described as taking place within a manifold [4]. If the function E: $U \to R$ is defined in the open set $U \subset R^n$, its gradient ∇E: $U \to R^n$ is a vector of components

$$(\partial E/\partial x_1, \ldots, \partial E/\partial x_n), \quad \text{where } \mathbf{x} = (x_1, \ldots, x_n) \in U\,. \tag{1}$$

The gradient points to the direction along which the function E increases most steeply.

A dynamical system in an open set U is defined by the differential equation

$$\mathbf{x}' = f(\mathbf{x}); \quad f\colon U \to R^n\,. \tag{2}$$

For a gradient system,

$$f(\mathbf{x}) = -\nabla E(\mathbf{x})\,, \tag{3}$$

where E is the system's energy.

The points where $\nabla E = 0$ are called equilibrium points. The equilibrium points of (2) are classified following the behaviour of the solutions in its neighbourhood and the notion of stability given here is that attributed to Liapunov [4].

We shall say that an equilibrium point $\mathbf{x}_s$ is stable, if for every neighbourhood V of $\mathbf{x}_s$ there exists a neighbourhood W of $\mathbf{x}_s$, $W \subset V$, where any solution of (2) $\mathbf{x}(\lambda)$ with $\mathbf{x}(0)$ in W remains within V for any $\lambda > 0$ (fig. 2(a)).

An equilibrium point $\mathbf{x}_a$ is asymptotically stable if, besides being stable, any solution $\mathbf{x}(\lambda)$ of (2) with $\mathbf{x}(0)$ within W converges to $\mathbf{x}_a$ (fig. 2(b)). Finally, an equilibrium point $\mathbf{x}_u$ is unstable if there exists a neighbourhood V of it such that, for any neighbourhood W of $\mathbf{x}_u$, $W \subset V$, at least one solution $\mathbf{x}(\lambda)$ of (2) having $\mathbf{x}(0)$ within W does not remain within V for any $\lambda > 0$ (fig. 2(c)). Furthermore, if $\mathbf{x}$ is an isolated minimum of E, $\mathbf{x}$ is an asymptotically stable point of (2) [4].

The energy along the solutions of (2) is non-increasing. These solutions are known as steepest descent paths [5,6]. The internal coordinates of an N-atoms molecular system define a $(3N\text{-}6)$ configuration space, which is a manifold of class

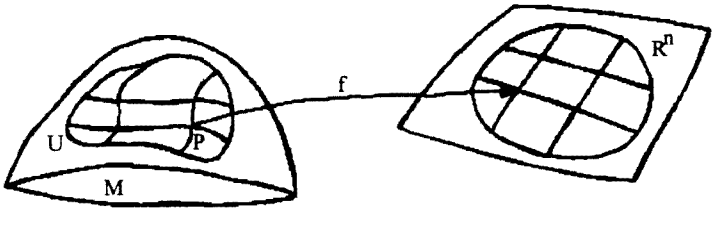

Fig. 1. Definition of a manifold.

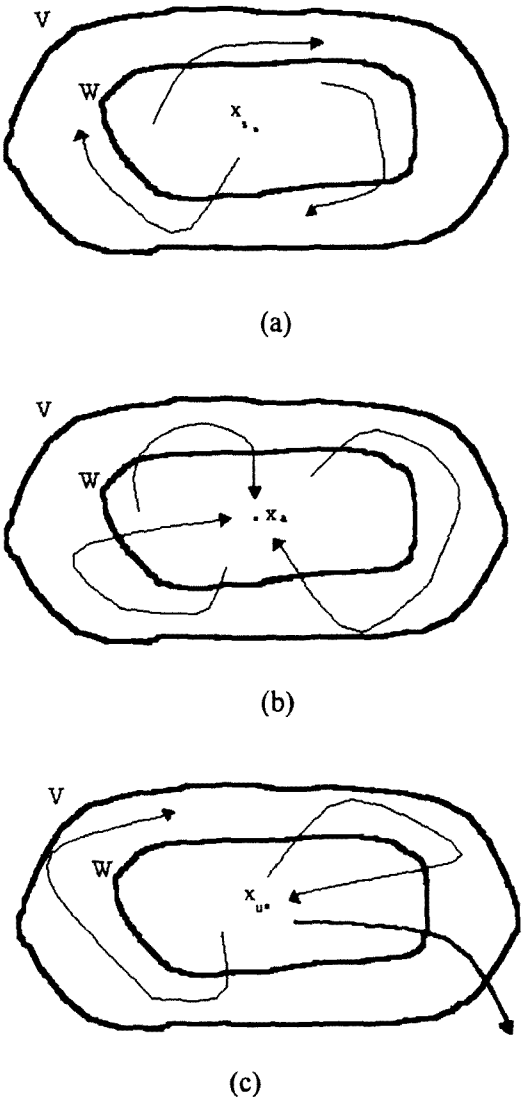

Fig. 2. (a) Stable, (b) asymptotically stable and (c) unstable equilibrium points.

C^{∞}. The gradient system, of course, may be described in terms of configuration space.

The well-known works of Fukui and Tachibana [7] apply some of the above-mentioned concepts from the theory of dynamical systems to chemical reactions.

Stable and transition structures are, respectively, asymptotically stable equilibrium points and unstable equilibrium points of the system. For asymptotically stable equilibrium points $\mathbf{x}_a$ we shall take all the solutions of (2) tending to $\mathbf{x}_a$ and for unstable equilibrium points $\mathbf{x}_u$ we take the solutions of the system tending to $\mathbf{x}_u$,

if they exist (fig. 3). The range of possible distortions of a molecular conformation which preserve chemical identity is known under the name of catchment region [1,8].

Now, a theorem discussed by Pechukas [5] states that every point belonging to a solution of a gradient system preserves the point symmetry group of the starting point. As a corollary of this theorem, nowhere along the solution may arise a new symmetry element not present in the starting point; this can only occur in equilibrium points. Therefore, an equilibrium point corresponds to a conformation having the highest order point symmetry group within its catchment region.

If we analyse any solution within a catchment region, the order of the point symmetry groups of its elements is constant; it could increase only at an equilibrium point. This possibility may direct the search of equilibrium points. Although the study of catchment regions would require mass-scaled coordinates, they are impractical, so conventional internal coordinates will do [8].

Most familiar vector algebras involve an inner product between vectors (scalar product). The metric tensor on an n-dimensional basis $\{\boldsymbol{e}_i\}$ has components

$$g_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j \,. \tag{4}$$

These components form an $n \times n$ symmetric matrix. This matrix is required to have an inverse, of components denoted by g^{ij} [3]. In this work we shall not be concerned with the most important role of the metric tensor, namely that it maps vectors into one-forms in a 1–1 manner [1] (stated otherwise, it allows to raise or lower, respectively, covariant and contravariant tensor indices [9]). If the basis is orthonormalized, the metric tensor reduces to identity, the unit matrix.

3. Application to glycine

Glycine, the smallest amino acid, has been the subject of extensive theoretical and experimental studies [10–16]. Although we have also calculated some of the non-planar possible conformations, we restrict ourselves here to the three most usual planar conformations of glycine under a neutral form, shown in fig. 4. The most sophisticated theoretical results agree in appointing conformation I as the

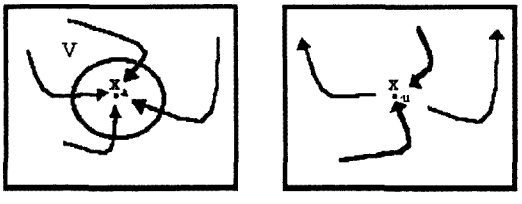

Fig. 3. Sets of solutions tending to a certain equilibrium point: $\mathbf{x}_a$, asymptotically stable; $\mathbf{x}_u$, unstable.

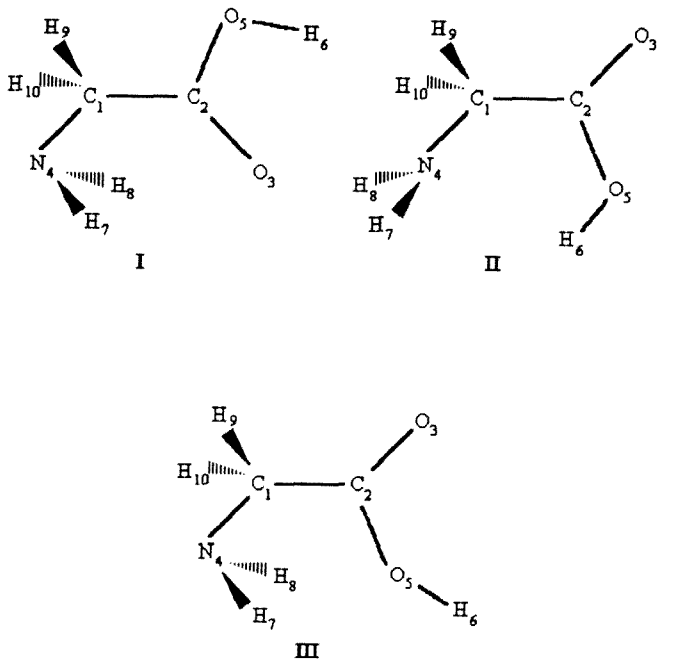

Fig. 4. The three most usual planar conformations of neutral glycine.

most stable one; however, experimental difficulties have prevented verifying this prediction [10,13] and the relative stability of other conformers is still a highly controversial issue [14–16].

We have used Császár's [15] internal coordinates as starting parameters in the MM-THOR package used [17]; the program allows options generating energy surfaces and energy contour lines. In order to obtain the results shown in this work, we were led to decrease the grid step of the original program from 10^{-2} to 10^{-10}. The gradient norm, accordingly, is considered zero when attaining 10^{-10}.

We report in table 1 the relative energies obtained for each conformation. It is seen that before geometry optimization, conformation I appears as the most stable one, while optimization favours conformation III. Nevertheless, the difference between I and III being then only 0.14 kcal/mol, it can hardly be considered meaningful. Glycine appears thus to be an essentially flexible molecule [18], this feature accounting for the unsettled discussion about conformers stability.

Table 1
Relative energies E of conformations I, II and III (in −kcal/mol) before and after geometry optimization.

Conf.	I	II	III
E (before opt.)	3.7367	0.000	3.2021
E (after opt.)	14.0893	12.3869	14.2282

Figure 5 shows the energy surface of conformation II as a function of the torsion angles of N_4 and O_5 around the plane defined by the atoms C_1, C_2 and O_3. We denote them respectively as $N \cong \tau(N_4C_1C_2O_3)$ and $O \cong \tau(O_5C_1C_2O_3)$. Figure 6 shows the corresponding contour energy lines of this smooth and shallow surface. In the illustrations of conformation I (not shown here), we also find a huge number of equilibrium points. As it happens in much more sophisticated calculations, the sensitivity to small energy variations is low. In ref. [19] different calculation methods are compared when studying glycine energy surfaces; the conclusion is that reproducing such small energy differences is a challenge for computational chemistry. More recent works [14,15] add uncertainty rather than settle the question.

Turning back to fig. 6, we call N_2 the region of equilibrium points around conformation II. For conformation I there are solutions converging and others diverging from the corresponding region N_1 (not shown here), where N and/or O differ from 180°. N_1 is hence an instability region; since conformation I belongs to it, shares this feature. Császár [15] suggests that the repulsion of lone electron pairs on the nitrogen and oxygen atoms may destabilize some glycine planar forms, leading to saddle points. The instability of region N_2 is more accentuated than that of N_1, due to the two wide regions with lowest energy corresponding to non-planar conformations (see figs. 5 and 6). As in N_1, we find in N_2 the symmetry group of planar conformations, of higher order symmetry.

Similarly, fig. 7 shows the equilibrium points region N_3. All solutions near N_3 not only remain near it, but within the region after a certain time, characterizing thus asymptotically stable equilibrium (conformation III may hence be considered

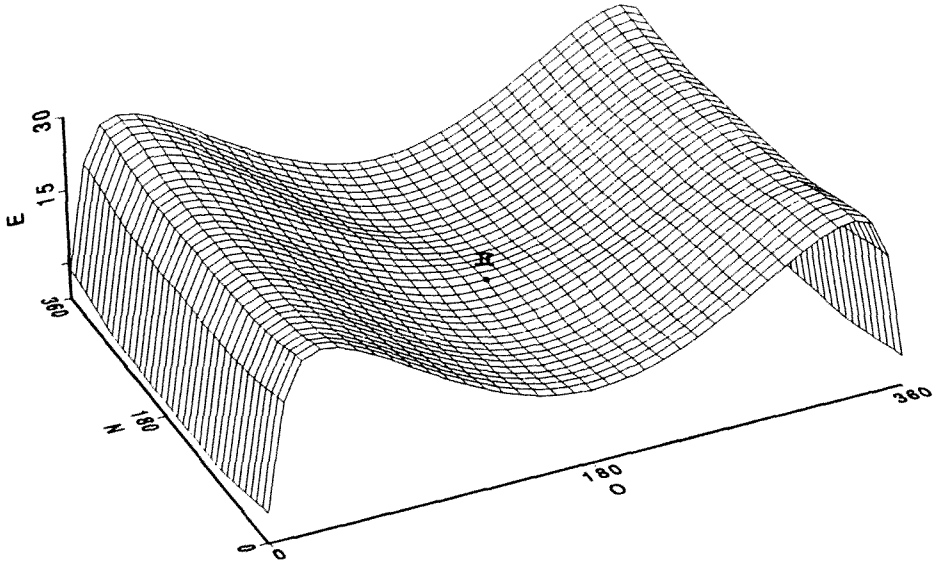

Fig. 5. Energy surface of conformation II as a function of the torsion angles $\tau(N_4C_1C_2O_3)$ (denoted N) and $\tau(O_5C_1C_2O_3)$ (denoted O). The units of energy E are kcal/mol.

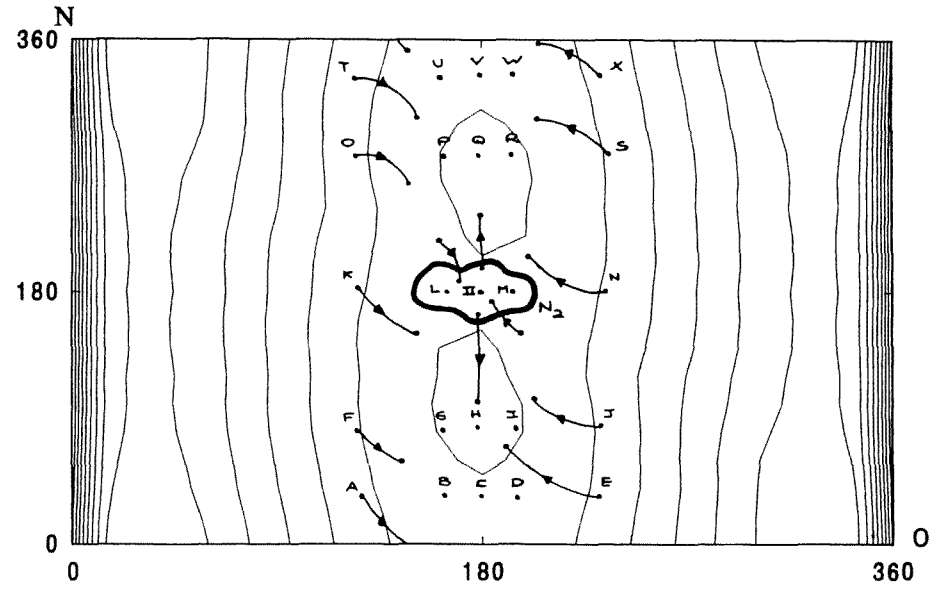

Fig. 6. Contour energy lines of conformation II.

as such, agreeing with refs. [12] and [13]). Now it is clearer that the equilibrium point is one of the conformations having the highest order symmetry group within its catchment region; as conformation III is an equilibrium point within N_3 and it

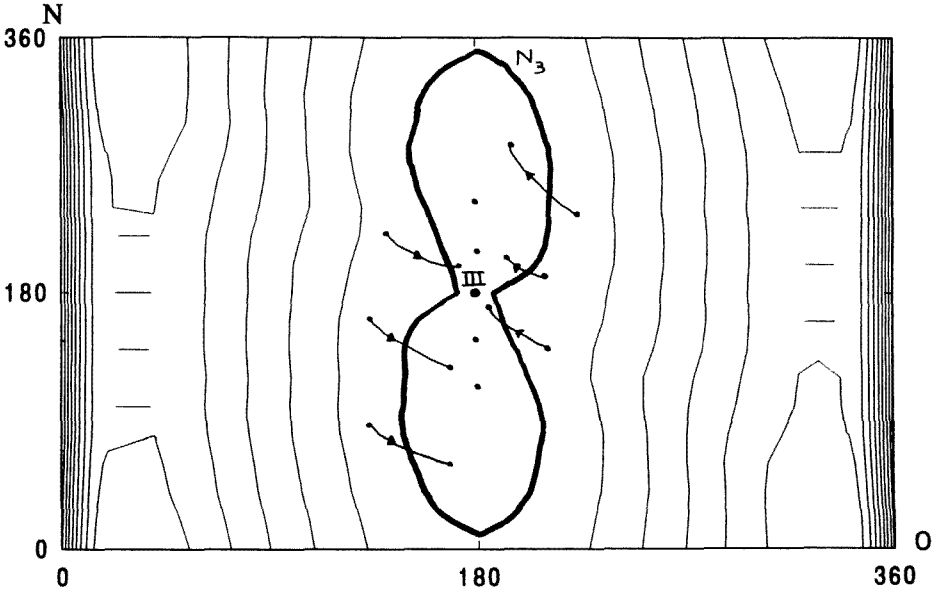

Fig. 7. Contour energy lines of conformation III.

exhibits the highest order symmetry element (planar symmetry), N_3 may be considered a catchment region.

4. Further developments

We may wonder what happens with a molecule's metric along solutions of a dynamical system such as those in fig. 6. The classical Chirgwin and Coulson work [20] mentions a noteworthy point which since then has been overlooked for many years, namely, that the well-known overlap matrix **S** of molecular calculations is the metric tensor of the Hilbert space defined by the basis of atomic orbitals. It may be considered as an operator, for it rules the raising or lowering of tensor indices [9]. So, if we desire to follow the metric behaviour along a given solution, we must choose some quantity related to the overlap matrix. We have opted for the spectral radius, i.e. the highest eigenvalue of **S**. This is a possible definition of norm [21]; we shall however speak of the spectral radius of tensor **S** so as to avoid confusion with the much more familiar definition of norm as a vector's modulus, which we shall use too.

We use the overlap matrix built from Slater-type orbitals – such as given by the CNDO/2 program [22] – of which we have calculated the eigenvalues. Although CNDO supposes *complete neglect of differential overlap*, the overlap matrix is used (and it appears in standard output) in the parameterization of the Hamiltonian matrix elements.

Of the three conformations studied, we shall at this stage focus our attention on conformation II, because it shows the greatest range of variation in the characteristics of the dynamical system solutions. We report in table 2 the behaviour of the spectral radius $\| S_{ij} \|$ along some points of solution O in fig. 6. Now, in table 2 $\| S_{ij} \|$ decreases along the solution. This is not a general trend. It tends to decrease in solutions A, J, O and X, while it tends to increase for E, F, K, N, S and T. We have found no association between these behaviours and any other quantity. Let us remind that $\| S_{ij} \|$ measures the maximum stretch which the vectors of unitary radius may undergo upon application of the **S** operator [5].

Let us look for a connection between the dynamical system solutions and the gradient norm in the final point of each solution, calculated with the PM3 Hamiltonian of the MOPAC package [23]. The norms $\| \nabla \boldsymbol{E} \|$ thus estimated have high values because they are computed with geometries optimized through a completely different method. Let us see first the $\| \nabla \boldsymbol{E} \|$ results for the three studied conformations; they are respectively 80.0562 (conf. I), 90.3208 (conf. II) and 77.7704 (conf. III). Although the differences in the relative norms is small, they indicate (in agreement with table 1) conformation III as the most stable one.

We have collected in table 3 both $\| \nabla \boldsymbol{E} \|$ and $\| S_{ij} \|$ for the solutions' final points displayed in fig. 6. Table 3 strongly suggests a link between the two quantities. The programs with which they are estimated, both make only an indirect use of

Table 2
Variation of the overlap matrix spectral radius $\| S_{ij} \|$ along some points of solution O (see fig. 6).

N	O	$\| S_{ij} \|$
270	120	2.895
268.65	126.38	2.880
268.44	127.55	2.875
267.93	129.11	2.869
267.53	130.63	2.868
266.55	132.91	2.854
265.76	134.94	2.853
264.13	138.26	2.831
262.87	141.04	2.828
260.41	145.81	2.796
258.76	149.03	2.790
257.76	150.54	2.778
256.81	153.23	2.765
255.25	155.07	2.762
254.19	157.20	2.742
249.55	162.70	2.700
248.26	163.69	2.691
248.23	164.96	2.685

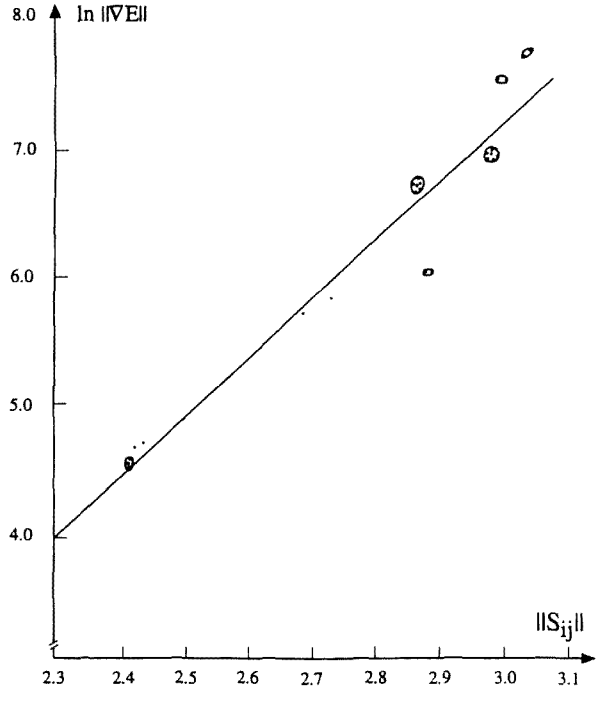

Fig. 8. Gradient norm logarithm ($\ln \| \nabla E \|$) as a function of spectral radius $\| S_{ij} \|$, both from table 3.

Table 3
Gradient norm $\| \nabla E \|$ and spectral radius $\| S_{ij} \|$ (see text) for the dynamical system solutions of fig. 6 in the final points.

Conf.	$\| \nabla E \|$	$\| S_{ij} \|$
A	401.7764	2.886
B	999.6724	2.985
C	1000.8692	2.985
D	1002.3634	2.985
E	2219.4795	3.048
F	1776.9051	3.003
G	772.1601	2.875
H	722.8660	2.876
I	723.8417	2.877
J	330.9951	2.739
K	104.8664	2.420
L	90.5965	2.412
II	**90.3208**	**2.412**
M	90.3802	2.412
N	112.3853	2.426
O	302.2469	2.685
P	726.1285	2.877
Q	725.3437	2.876
R	724.7618	2.875
S	1822.3083	3.006
T	2213.1819	3.051
U	995.0994	2.984
V	993.5896	2.984
W	992.2199	2.984
X	396.9940	2.882

overlap. The overlap integrals are not programmed in the same way in CNDO [22] and in MOPAC [23]; however, for first row atoms they cannot differ appreciably, so that the above-mentioned relation should actually apply. Figure 8 shows $\ln \| \nabla E \|$ plotted against $\| S_{ij} \|$. A least-squares calculations yields

$$\ln \| \nabla E \| = 4.447 \| S_{ij} \| - 6.238 \tag{5}$$

with a linear regression coefficient of 0.9705.

Equation (5) furnishes us a relation between $\| \nabla E \|$ computed by MM and that given by MOPAC. Let us see why. At first sight, it does not look consistent with the fact that, when $\| \nabla E \|$ is zero, its logarithm is $-\infty$. Nevertheless, it must be kept in mind that the lower bound to $\| S_{ij} \|$ is one. This is for strictly orthogonal bases, and bases are never really orthogonal; MOPAC supposes orthogonal bases in all electronic integrals, despite that S is used in the Hamiltonian parameterization. For $\| S_{ij} \| = 1$, eq. (4) gives $\ln \| \nabla E \| = -1.791$, so that $\| \nabla E \| = 0.167$, which is thus a shift in the MM-MOPAC scale.

The relation between $\| \nabla E \|$ and $\| S_{ij} \|$ given by eq. (5) is entirely unexpected and, of course, it must be taken as a preliminary result, subject to further testing.

5. Concluding remarks

Mezey [8] states a theorem interrelating catchment regions and point symmetries of nuclear configurations. One aspect of this theorem is that "The distribution of catchment regions and their critical points are the properties of the potential energy surface, that is, they depend on *energy relations*, whereas the point symmetry of various nuclear configurations are purely *geometrical properties*, not directly dependent on energy. Consequently, *the theorem interrelates two very different molecular properties*" [8, p 3796].

This aspect is exactly the kind of result we have found, a link between the norm of the gradient (a quantity belonging to the energy domain) and the spectral radius of the overlap matrix (a typically geometrical quantity).

References

[1] B.F. Schutz, *Geometrical Methods of Mathematical Physics* (Cambridge University Press, 1980) chap. 2.
[2] P.G. Mezey, Theor. Chim. Acta 58 (1981) 309; 63 (1983) 9.
[3] I.N. Levine, *Quantum Chemistry* (Prentice-Hall, Englewood Cliffs, NJ, 1991) chap. 16.
[4] M.W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra* (Academic Press, N. York, 1980) chap. 9.
[5] P. Pechukas, J. Chem. Phys. 64 (1976) 1516.
[6] R.M. Minyaev and D.J. Wales, Chem. Phys. Lett. 218 (1994) 413.
[7] A. Tachibana and K. Fukui, Theor. Chim. Acta 49 (1978) 321; Theor. Chim. Acta 51 (1979) 275;
K. Fukui, Acc. Chem. Res. 14 (1981) 363.
[8] P.G. Mezey, J. Am. Chem. Soc. 112 (1990) 3791.
[9] M.S. de Giambiagi, M. Giambiagi and F.E. Jorge, Z. Naturforsch. 39(a) (1984) 1259; Theor. Chim. Acta 68 (1985) 337;
M.S. de Giambiagi, M. Giambiagi and P. Pitanga, Chem. Phys. Lett. 129 (1986) 367.
[10] H.L. Sellers and L. Schäfer, J. Am. Chem. Soc. 100 (1978) 7728.
[11] R.D. Brown, P.D. Godfrey, J.W.V. Storey and M.P. Bassez, J. Chem. Soc., Chem. Commun. (1978) 547.
[12] B.T. Luke, A.G. Gupta, G.H. Loew, J.G. Lawless and D.H. White, Int. J. Quant. Chem. 11 (1984) 117.
[13] K. Iijima, K. Tanaka and S. Onuma, J. Mol. Struct. 246 (1991) 257.
[14] M. Ramek, V.K.W. Cheng, R.F. Frey, S.Q. Newton and L. Schäfer, J. Mol. Struct. 235 (1991) 1.
[15] A.G. Császár, J. Am. Chem. Soc. 114 (1992) 9568.
[16] C.-H. Hu, M. Sen and H.F. Schaefer III, J. Am. Chem. Soc. 115 (1993) 2923.
[17] THOR, Force Field and Molecular Dynamics Simulation, is a software package developed by K.C. Mundim, P.G. Pascutti and P.M. Bisch, to be published;
E.P.G. Areas, P.G. Pascutti, S. Schreier, K.C. Mundim and P.M. Bisch, Brazilian J. Med. Biol. Res. 27 (1994) 527.

[18] J. Koca and P.H.J. Carlsen, J. Mol. Struct. 246 (1991) 165.
[19] P. Palla, C. Petrongolo and J. Tomasi, J. Phys. Chem. 84 (1980) 435.
[20] B.H. Chirgwin and C.A. Coulson, Proc. Roy. Soc. London A201 (1950) 196.
[21] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965) pp. 58–59.
[22] J.A. Pople and D.L. Beveridge, *Approximate Molecular Orbital Theory* (McGraw-Hill, New York, 1970).
[23] J.J.P. Stewart, J. Comp.-Aided Mol. Design 4 (1990) 1 (special issue).