

Assigning Wave Functions to Graphs: A Way to Introduce Novel Topological Indices

Jorge Gálvez^{*a}, Ramón García-Domenech^a, J.Vicente de Julián-Ortiz^{a,b}

^aUnidad de Diseño de Fármacos y Conectividad Molecular, Departamento de Química Física,

^bRed de Investigación de Enfermedades Tropicales. Dept. de Biología Celular y Parasitología,

^{a,b}Facultat de Farmàcia, Universitat de València,

Av. V. Andrés Estellés, 0, 46100 Burjassot, València, Spain.

^{*}Corresponding author: jorge.galvez@uv.es

(Received January 19, 2006)

Abstract

One of the most significant differences between the topological and the physicochemical description of molecules is that the first usually makes use of discrete mathematics while the second handles continuous functions. Thus, molecular topology introduce discrete numbers called topological indices as the key descriptors, whereas quantum chemistry introduces the notion of wavefunction and its square (function density) as the mathematical objects able to encode all the information on the molecule. In this work we aim to set up a straightforward link between both formalisms by assigning wave or quasi-wave functions to the graphs. Two ways are followed: The first, on the basis of the Fourier theorem, describes the graph as a sum of sinusoidal functions. The second is a Hermite-like formalism including the graph characteristic polynomials corrected by exponential factor. As illustration of the potential interest of the method, Randić index is interpreted as a specific point of a continuous function and new topological indices are introduced as well. Their usefulness to predict boiling temperatures is also showed.

Introduction

Mathematical chemistry, and particularly molecular topology (MT), has become one of the most interesting areas of chemical research during the later years. Particularly interesting is its ability to predict many different properties, ranging from physical [1], electrochemical [2] or photochemical (chemiluminescence) [3] up to biological [4-14]. However, while significant efforts have been done on *how* MT works, much less is known about *why* it works so well. In a previous report [15] one of us has suggested that MT may be not only an alternative but also and independent approach to quantum chemistry. Perhaps the most critical difference between both formalisms is that any branch of physics (not only quantum physics) employs *dynamic* variables (such as pressure, volume, temperature) to define the state of the system, and *constitutive* magnitudes (such as formation enthalpies, lattice geometry, etc.) to describe the system structure. However, MT is not dealing with the state of the system but about the constitution of the system. Of course, the constitution is influenced by the state, but along with the continuous nature of dynamic variables it is always possible to reduce these continuous values into discrete ones what determine the structural nature. For instance, temperature changes in a continuous way, but what is really important from the structural point of view are such critical discrete values, such as melting or boiling points, conditioning whether the substance is going to hold into a gas, liquid or solid state. Indeed, these specific points are depending, in turn, on what is called as molecular structure, so closing the cycle of self-endorsement notions.

Thus, despite molecular structure plays a key role in the prediction of any experimental property of compounds, a gap between the quantum and topological descriptions still remains as an open issue.

Describing graphs through wave functions

One of the possible ways to bridge the gap between physicochemical structural descriptors and topological ones is by assigning continuous functions to any graph, in the same way as quantum chemistry does with molecules. Under this approach, these specific continuous functions would play the same role as wave functions in quantum chemistry. Furthermore, within this framework, topological indices can be considered as notable or specific points of these functions. Indeed, there are many possible ways of doing that, but we chose two

algorithms: The first is based on the Fourier theorem and the second on the graph characteristic polynomial. Let us develop each one in detail.

The trigonometric description of graphs

Starting from the classical time independent equation of waves:

$$(d^2 \Phi / dx^2) + (2\pi / \lambda)^2 \Phi = 0$$

where Φ stands for the wave function, λ is the wave length, and x is the 1-D spatial coordinate.

Assuming the de Broglie postulate, i.e. each particle has an associated wave, even if we do not introduce the mass into the wave equation ($\lambda = h / mv$), the wavefunction is depending on the quantum level as far as we confine the particle in a monodimensional box. So, the state functions for the real values of the stationary states of the particle in the box are given by:

$$\phi = (2/a)^{1/2} \sin(n \pi x/a) \quad (1)$$

where “a” is the bond length and “n” is an integer number from one on up.

Let us consider the one-edge hydrogen depleted graph, that is, the ethane graph. Its connectivity term C_{ij} , is $C_{11} = (1 \cdot 1)^{1/2} = 1.000$

It may be realized that this value is equal to the maximum Φ value for $n=1$ and $a=2$ in equation (1).

If we do the same calculation for a 1-2 edge, as for instance either of the two edges in propane's graph, then its connectivity term is $C_{12} = (1 \cdot 2)^{1/2} = 0.7071$. Likewise, this value is the maximum wave function value for $n=2$, $a=4$. The same stands for the C_{13} and C_{14} terms, for which the equivalencies are $n=3$, $a=6$ and $n=4$, $a=8$, respectively. The connectivity terms are so equivalent to the maximum probability values for particles (electrons) placed at different quantum levels within boxes of different lengths.

The values for each edge wave functions are:

$$\phi_{11} = (2/2)^{1/2} \sin(\pi x/2)$$

$$\phi_{12} = (2/4)^{1/2} \sin(\pi x/2)$$

$$\phi_{13} = (2/6)^{1/2} \sin(\pi x/2)$$

$$\phi_{14} = (2/8)^{1/2} \sin(\pi x/2)$$

An alternative way to reach the same result is taking into account the topological valence values for each edge, while maintaining constant the box length, as follows:

$$\phi_{ij} = 1/(\delta_{ij})^{1/2} \sin(\delta_{ij} \pi x/2)$$

That expression is pledged because the probability of locating an electron within a given ij edge (or bond into a chemical translation) is inverse to δ_{ij} . [15]

The wavefunctions associated to the (1,1) (1,2) (1,3) and (1,4) edges are:

$$\phi_{11} = (1/1)^{1/2} \sin (\pi x/2)$$

$$\phi_{12} = (1/2)^{1/2} \sin (\pi x)$$

$$\phi_{13} = (1/3)^{1/2} \sin (3\pi x/2)$$

$$\phi_{14} = (1/4)^{1/2} \sin (2\pi x)$$

The corresponding wavelengths (in arbitrary units) are 4, 2, 1.333 and 1.000 respectively and the corresponding topological levels as defined for their valence values are 1, 2, 3, 4, respectively.

It is easy to demonstrate that any other possible edge into a graph has an associate wave function that is derived from the four edges outlined, which can be denominated as *fundamental edges*.

From these results, it is noteworthy that topological level values are the same as quantum levels in the particle-in-a-box approach. This is a one another link between both formalisms, i.e. between molecular topology and quantum chemistry.

Furthermore, it is easy to realize that the maximum of the edges-wave functions (corresponding to sinus values =1) is equal to the connectivity terms C_{ij} .

Since any graph is composed of the four fundamental edges (1,1), (1,2) (1,3) and (1,4), what are the edges associated to ethane, propane, isobutane and neopentane, respectively, the discrete character of the connectivity indices can be associated to different wave function quantum levels. This also agrees with the Fourier theorem, which stands that any wave equation, despite its complexity, may be expressed as a sum of sinusoidal functions.

Thus, if we consider isopentane, its wave function would be:

$$\Psi_{\text{isopentane}} = \phi_{13} \cdot \phi_{12} + 2 \phi_{13} + \phi_{12} = 0.408 \sin (3\pi x/2) \sin (\pi x) + 1.1547 \sin (3\pi x/2) + 0.707 \sin (\pi x).$$

It is interesting to realize that, under this formalism, the Randić index, or first order connectivity term, χ , is just the maximum value of the graph-wave function.

Table 1 shows the corresponding wave function equations for a set of alkanes, as well as the values of the differences between maxima and minima for each graph wave function (maximum-minimum gap, MMG).

Table 1. Wave function equations and the values of the differences between maxima and minima into each graph wave function for a set of alkanes.

Alkane	Sinusoidal wave function	MMG ^a
Ethane	$\Psi_{\text{ethane}} = \phi_{11} = \sin(\pi x / 2)$	2.00
Propane	$\Psi_{\text{propane}} = 2\phi_{12} = 1.4142 \sin(\pi x)$	2.83
Isobutane	$\Psi_{\text{isobutane}} = 3\phi_{13} = 1.7320 \sin(3\pi x / 2)$	3.46
Butane	$\Psi_{\text{butane}} = 2\phi_{12} + \phi_{12}\phi_{12} = 1.4142\sin(\pi x) + 0.5 \sin^2(\pi x)$	2.91
Neopentane	$\Psi_{\text{neopentane}} = 4\phi_{14} = 2 \sin(2\pi x)$	4.00
2M4	$\Psi_{2M4} = \phi_{13}\phi_{12} + 2\phi_{13} + \phi_{12} = 0.4082\sin(3\pi x / 2)\sin(\pi x) + 1.1547\sin(3\pi x / 2) + 0.7071\sin(\pi x)$	3.72
n-pentane	$\Psi_{\text{n-pentane}} = 2\phi_{12} + 2\phi_{12}\phi_{12} = 1.4142\sin(\pi x) + 1.0 \sin^2(\pi x)$	2.91
22MM4	$\Psi_{22M4} = \phi_{14}\phi_{12} + 3\phi_{14} + \phi_{12} = 0.3535\sin(2\pi x)\sin(\pi x) + 1.5\sin(2\pi x) + 0.7071\sin(\pi x)$	4.41
23MM4	$\Psi_{23MM4} = \phi_{13}\phi_{12} + 2\phi_{13} + \phi_{12} = 0.4082\sin(3\pi x / 2)\sin(\pi x) + 1.1547 \sin(3\pi x / 2) + 0.7071\sin(\pi x)$	4.61
2M5	$\Psi_{2M5} = 2\phi_{13} + \phi_{13}\phi_{12} + \phi_{12}\phi_{12} + \phi_{12} = 1.1547 \sin(3\pi x / 2) + 0.4082\sin(3\pi x / 2)\sin(\pi x) + 0.5 \sin^2(\pi x) + 0.7071\sin(\pi x)$	3.72
3M5	$\Psi_{3M5} = 2\phi_{12} + 2\phi_{13}\phi_{12} + \phi_{13} = 1.4142 \sin(\pi x) + 0.8164\sin(3\pi x / 2)\sin(\pi x) + 0.5773\sin(3\pi x / 2)$	4.00
n-hexane	$\Psi_{\text{n-hexane}} = 2\phi_{12} + 3\phi_{12}\phi_{12} = 1.4142\sin(\pi x) + 1.5 \sin^2(\pi x)$	3.25
23MM5	$\Psi_{23MM5} = 3\phi_{13} + \phi_{13}\phi_{12} + \phi_{13}\phi_{13} + \phi_{12} = 1.7320 \sin(3\pi x / 2) + 0.4082\sin(3\pi x / 2)\sin(\pi x) + 0.3333 \sin^2(3\pi x / 2) + 0.7071\sin(\pi x)$	4.88

^aMMG= Maximun-minimum gap.

Other algorithms allowing manipulation of wave functions are possible as well. For example the (1,1), (1,2) (1,3) and (1,4) edges can be described as:

$$\begin{aligned}
 \phi_{11} &= \sin(\pi x/2 + \pi/2); & \phi_{11}(0) &= 1.0000 = \chi_{11} \\
 \phi_{12} &= \sin(\pi x/2 + \pi/4) & \phi_{12}(0) &= 0.7071 = \chi_{12} \\
 \phi_{13} &= 2/3 \sin(\pi x/2 + \pi/3) & \phi_{13}(0) &= 0.5773 = \chi_{13} \\
 \phi_{14} &= \sin(\pi x/2 + \pi/6) & \phi_{14}(0) &= 0.5000 = \chi_{14}
 \end{aligned}$$

As in the former description, any other graph edge may be derived from these four fundamental edges. In this case, the differences in the associated wave functions are not in the amplitude but in the phase angle (except for the (1,3) edge that includes both).

In this case, the connectivity index χ is just the intercept for each graph equation ($\chi = \Psi(0)$).

The polynomial approach

Many previous attempts are reported on the graph characterization through their characteristic polynomials. [16, 17] However, the use of the single polynomials leads to asymptotic curves, whose characteristic points are not very representative as graph descriptors. It is interesting to realize that the eigenvalues for the four fundamental graphs above defined (corresponding to ethane, propane, isobutane and neopentane) are identical to the χ values for each one.

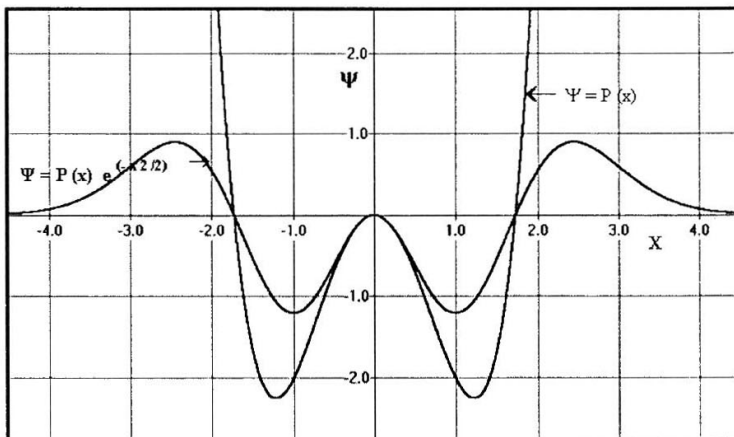


Figure 1.- Comparison between the single polynomial and the Hermite-like function for the n-butane graph. Observe that the AUC makes finite for the exponential quasi-wave function.

A possible alternative is combining the polynomial expression with the squared exponential in a similar way as the Hermite polynomials are introduced into the wave function of the harmonic oscillator.

Thus, we define the quasi-wave function for a graph as:

$$\Psi = P(x) e^{(-x^2/2)}$$

being $P(x)$ the graph characteristic polynomial. The difference between the two representations is displayed in Figure 1. The most significant difference is that the area under the curve becomes finite in the exponential approach, which makes sense to the concept of quantum probability as the overall integral between 0 and ∞ .

It is noteworthy the appearing of a maximum at about $x=2.4$. That maximum also stands for the rest of alkanes in the range between $x=1$ and $x=5$. Because of the similarity to the Hermite approach for harmonic oscillator wave functions we call our approach as Hermite-like wave functions.

Table 2 illustrates the characteristic polynomial, as well as the χ connectivity index, maximum eigenvalues, ordinates of the maxima placed between $x=1$ and $x=5$ ($\Psi_{\max 1-5}$), area under the curve and boiling temperatures for a set of alkanes.

Table 2.- Connectivity index χ ; characteristic polynomial; maximum eigenvalues; maximum Ψ between $x=1$ and $x=5$; area under the curve AUC and boiling temperature, T_b , for a set of alkanes.

Alkane	χ	Characteristic polynomial	Maximum eigenvalues	$\Psi_{\max, \text{exp.}}$ $x=0 \text{ and } x=5$	AUC	$T_b(K)$
ethane	1.0000	$x^2 - 1$	1.000	0.4463	0.6065	184.4
propane	1.4142	$x^3 - 2x$	1.414	0.5592	0.7358	230.9
isobutane	1.7320	$x^4 - 3x^2$	1.732	0.8962	1.1594	261.3
butane	1.9142	$x^4 - 3x^2 + 1$	1.618	0.9486	1.2765	272.5
neopentane	2.0000	$x^5 - 4x^3$	2.000	1.6922	2.1630	282.5
2M4	2.2700	$x^5 - 4x^3 + 2x$	1.845	1.8379	2.4772	300.9
n-pentane	2.4142	$x^5 - 4x^3 + 3x$	1.732	1.9188	2.6775	309.1
22MM4	2.5610	$x^6 - 5x^4 + 3x^2$	2.074	3.9767	5.3290	322.7
23MM4	2.6430	$x^6 - 5x^4 + 4x^2$	2.000	4.1095	5.6415	331.0
2M5	2.7700	$x^6 - 5x^4 + 5x^2$	1.902	4.2497	5.9898	333.0
3M5	2.8080	$x^6 - 5x^4 + 5x^2 - 1$	1.932	4.2318	5.9877	336.6
n-hexane	2.9140	$x^6 - 5x^4 + 6x^2 - 1$	1.802	4.3784	6.3219	341.7
23MM5	3.1807	$x^7 - 6x^5 + 8x^3 - 2x$	2.053	10.4415	14.894	362.8

Results and discussion

The formalism we introduce here, i.e. assigning wave or quasi-wave functions to molecular graphs, has a widespread potential applicability. One of the most interesting usefulness is opening a new way to introduce novel topological indices, as specific points of such functions.

A detailed description on how this approach can be achieved is out of the goal of this paper and it will be hopefully dealt in the near future, but let's illustrate, just as an example, a simple case: The good fitting between the maximum amplitude (MA) of the sinusoidal wave functions and the boiling points for groups of alkane isomers. The same happens with the ordinate values of the maxima as well as with AUC for the Hermite-like approach.

The following regression equations were the best (highest F-ratio), obtained correlating boiling points with a set of topological indices including the new ones introduced by first time here, as well as the connectivity up to the fourth order:

$$\begin{aligned}
 &BP = 59.80 \chi + 167.92 \\
 &N=58 \quad R^2 = 0.9591 \quad SE= 8.23 \quad F = 128.4 \\
 &BP = 1.444 (MA) + 54.23 \chi + 179.66 \\
 &N=58 \quad R^2 = 0.9688 \quad SE= 6.78 \quad F = 191.2 \\
 &BP = -4.93 (\Psi_{\max 1.5}) + 87.28 \chi + 112.85 \quad (2) \\
 &N=58 \quad R^2 = 0.9878 \quad SE= 4.36 \quad F = 366.4
 \end{aligned}$$

Thus, although the χ index continues to be the best single descriptor for boiling points prediction for the selected set of alkanes, the inclusion of some of the novel topological indices led to a significant improvement in the prediction as compared with the rest of connectivity indices.

In closing, we may conclude that the use of graph wave functions allows, within the framework of the independent and fundamental nature of topological indices, to consider molecular energies (at least electronic and vibrational) as particular cases of topological indices. In fact, whilst energy is the eigenvalue of the Hamiltonian operator over a given wave function, topological indices would result on applying different operators over the selected wave functions. Under such a view, energy would be just a particular case of topological index, what could explain why electronic and vibrational energies [7, 15] correlate so well with some topological indices. This is interesting because topology, in the sense we use this word here, would be an alternative and independent approach to energy, or at least, it may suggest new ways of dealing with the well-known concepts of molecular energies and wave functions.

Conclusions

The possibility of describing graph by continuous functions is a very useful approach for several reasons:

- First, it makes possible introducing new topological indices as specific points of continuous functions describing each graph.
- Second, it allows setting up a parallelism between the quantization of energy and the discrete character of topological descriptors.
- Finally, this formalism involves a straightforward relationship between molecular energy and topological descriptors. Under this view, molecular energies, particularly electronic and vibrational ones, would be particular cases of topological indexes derived from the application of the Hamiltonian operator, just one of the many possible operators to apply over the wave functions.

Acknowledgments

The authors acknowledge financial support from the Spanish *Red de Investigación de Centros de Enfermedades Tropicales, RICET, (C03/04)*, the project FIS-2005 PI052128 of the *Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Ministerio de Sanidad*, Spain and the grant number BQU2003-07420-C05 of the former *Ministerio de Ciencia y Tecnología* within the Spanish *Plan Nacional I+D*.

References

1. R. García-Domenech, J.V. de Julián-Ortiz, J. Phys. Chem. B. 106 (2002) 1501.
2. J. Gálvez, R. García-Domenech, R. Soler, P.V. Olcina, An. Quím. 89 (1993) 475.
3. L. Lahuerta Zamora, Y. Fuster Mestre, M.J. Duarte, G.M. Antón-Fos, R. García-Domenech, J. Gálvez, J. M-Calatayud, Anal. Chem. 73 (2001) 4301.
4. R. García-Domenech, J.V. de Julián-Ortiz, M.J. Duarte, J.M. García-Torrecillas, G.M. Antón-Fos, I. Ríos Santamarina, C. de Gregorio Alapont, J. Gálvez, SAR & QSAR Environ. Res. 12 (2001) 237.
5. J. Gálvez, R. García-Domenech, J.V. de Julián-Ortiz, R. Soler, J. Chem Inf. Comput Sci. 35 (1995) 272.
6. E. Estrada & E. Uriarte, Curr. Med. Chem. 8 (2001) 1699.
7. R.K. Mishra, R. García-Domenech, J. Gálvez, J. Chem. Inf. Comput. Sci. 41 (2001) 387.

8. C. de Gregorio-Alapont, L.B. Kier, L.H. Hall, *J. Comput. Aided Mol. Des.* 12 (1998) 557.
9. S.C. Basak, D.R. Mills, A.T. Balaban, B.D. Gute, *J. Chem. Inf. Comput. Sci.* 41 (2001) 671.
10. O. Ivanciuc, *J. Chem. Inf. Comput. Sci.* 40 (2000) 1412.
11. R.D. Enriz, H.A. Baldoni, M.A. Zamora, E.A. Jauregui, M.E. Sosa, C.E. Tonn, J.M. Luco, M. Gordaliza, *J. Agric. Food Chem.* 48 (2000) 1384.
12. J.V. de Julián-Ortiz, J. Gálvez, C. Muñoz-Collado, R. García-Domenech, C. Gimeno-Cardona, *J. Med. Chem.* 42 (1999) 3308.
13. J. V. de Julián-Ortiz, C. de Gregorio Alapont, I. Ríos-Santamarina, R. García-Domenech, J. Gálvez, *J. Mol. Graph. Model.* 20 (2001) 84.
14. A. García-García, J. Gálvez, J. V. de Julián-Ortiz, R. García-Domenech, C. Muñoz-Collado, R. Guna, R. Borrás, *J. Biomol. Screen.* 10 (2005) 206.
15. J. Gálvez, *J. Mol. Struct. (Theochem)* 429 (1998) 255.
16. K. Balasubramanian *J. Comput. Chem.* 5 (1984) 387.
17. R. Ramaraj, K. Balasubramanian *J. Comput. Chem.* 6 (1985) 122.