

# The Energy of a Graph and its Size Dependence.

## An Improved Monte Carlo Approach

Harald Friepertinger, Ivan Gutman<sup>a</sup>, Adalbert Kerber<sup>b</sup>, Axel Kohnert<sup>b</sup>,  
and Dušica Vidović<sup>a</sup>

Institut für Mathematik, Karl Franzens Universität Graz, Heinrichstraße 36, A-8010 Graz, Austria

<sup>a</sup> Faculty of Science, University of Kragujevac, P. O. Box 60, YU-34000 Kragujevac, Yugoslavia

<sup>b</sup> Lehrstuhl II für Mathematik, Universität Bayreuth, D-95440 Bayreuth, Germany

Reprint requests to Prof. I. G.; Fax: +381 34 335040; E-mail: gutman@knez.uis.kg.ac.yu

Z. Naturforsch. **56 a**, 342–346 (2001); received April 2, 2001

In an earlier work [Gutman et al., Chem. Phys. Lett. **297**, 428 (1998)] the average energy  $\langle E \rangle$  of graphs with  $n$  vertices and  $m$  edges was examined, in particular its dependence on  $n$  and  $m$ . The quantity  $\langle E \rangle$  was computed from a set of randomly, but not uniformly, constructed  $(n, m)$ -graphs. We have now improved our method by constructing the  $(n, m)$ -graphs uniformly, so that every  $(n, m)$ -graph has equal probability to be generated. Differences between the old and new approaches are significant only in the case of graphs with a small number of edges.

**Key words:** Energy (of Graph); Total  $\pi$ -electron Energy; Random Graphs; Monte Carlo Methods.

### Introduction

If  $G$  is a graph and  $\lambda_1, \lambda_2, \dots, \lambda_n$  are its eigenvalues [1] then the *energy* of  $G$  is defined as

$$E = E(G) = \sum_{i=1}^n |\lambda_i|.$$

In the case when  $G$  is the graph representation of a conjugated molecule,  $E(G)$  is closely related to the total  $\pi$ -electron energy of that molecule. This fact has been outlined on numerous occasions (see, e.g., [2 - 4] and the references cited therein) and will not be elaborated here again.

In order to be consistent with our earlier work [5] as well as with the terminology of Hückel molecular orbital (HMO) theory [6] (from where  $E(G)$  originates), we will say that  $E(G)$  is expressed in units of the HMO resonance integral  $\beta$ .

The numbers of vertices and edges of the graph  $G$  are denoted by  $n$  and  $m$ , respectively. It is customary to say that  $G$  is an  $(n, m)$ -graph.

In the case when  $G$  represents a conjugated molecule [2, 7], the condition  $n - 1 \leq m \leq 3n/2$  must be obeyed. For such molecular graphs it was established long time ago [8 - 11] that the gross part of  $E(G)$  is determined by the parameters  $n$  and  $m$ ,

implying that within any set of  $(n, m)$ -graphs the variation of the  $E$ -values is relatively small. In order to find out the dependence of  $E$  on the size parameters  $n$  and  $m$  it is advantageous to deal with the average value of  $E$ , denoted by  $\langle E \rangle$ , where the averaging is done over some reasonably chosen set of  $(n, m)$ -graphs. Indeed, for molecular graphs  $\langle E \rangle$  is known to be a monotonically increasing function of both  $n$  and  $m$ , whose analytical form is, to a good approximation, equal to  $a\sqrt{mn}$ , where  $a$  is a constant [10, 12 - 15].

If, however, the number of edges increases beyond the “chemical” limit  $3n/2$ , then the average graph energy cannot remain a monotonically increasing function of  $m$ , but must attain a maximum and start to decrease. In order to learn more about this phenomenon a series of computer experiments were recently performed [5], in which the  $(n, m)$ -graphs needed for the calculation of  $\langle E \rangle$  were constructed by the following Monte Carlo procedure: For a chosen fixed value of  $n$  one starts with the graph without edges. Then edges are added to it one-by-one, so that in each step any pair of non-adjacent vertices has equal probability to become connected. The procedure ends when a total of  $n(n-1)/2$  edges is added, resulting in the complete  $n$ -vertex graph.

The advantage of the above described procedure is the ease (and speed) by which the required  $(n, m)$ -

0932-0784 / 01 / 0500-0342 \$ 06.00 © Verlag der Zeitschrift für Naturforschung, Tübingen · www.znaturforsch.com



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitalized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition “no derivative works”). This is to allow reuse in the area of future scientific usage.

graphs are generated. It, however, generates *labeled*  $(n, m)$ -graphs uniformly at random, i.e., every labeled  $(n, m)$ -graph has equal probability to be generated. This seems to be a serious drawback, because not all graphs can be labeled in the same number of ways, implying that different graphs are generated with different probabilities. In other words, by the above described procedure non-symmetric species are produced more frequently than the symmetric ones. This point is described in more detail in the subsequent section.

### Labeled vs. Unlabeled Graphs

Consider a graph  $G$  and label its vertices by  $1, 2, \dots, n$ . Let  $\mathcal{P}$  be a permutation of the numbers  $1, 2, \dots, n$ . If the condition: “the vertices labeled by  $i$  and  $j$  are adjacent in  $G$  if and only if the vertices labeled by  $\mathcal{P}(i)$  and  $\mathcal{P}(j)$  are adjacent in  $G$ ” holds for all  $i, j = 1, 2, \dots, n$ , then  $\mathcal{P}$  is said to be an *automorphism* of the graph  $G$ . The set of all automorphisms of  $G$  forms the *automorphism group*  $\mathcal{A}(G)$  of  $G$  [2]. Let  $|\mathcal{A}(G)|$  be the size (= number of elements) of  $\mathcal{A}(G)$ .

Intuitively, automorphisms are often conceived as symmetry operations on the graph. Graphs for which  $|\mathcal{A}(G)|$  is large are usually described as possessing a high symmetry [2].

An unlabeled  $n$ -vertex graph can be considered as an orbit of the symmetric group  $S_n$  on the set of labeled  $n$ -vertex graphs. It is clear from the theory of group actions (since the length of an orbit is the index of the automorphism group of each of its elements in the acting group) that an  $n$ -vertex graph can be labeled in  $n!/|\mathcal{A}(G)|$  distinct ways. Consequently, a non-symmetric graph (with trivial automorphism group) can be labeled in  $n!$  ways, whereas the number of labelings of symmetric graphs is less than  $n!$ .

Consider, as a simple example, the set of all  $(4, 3)$ -graphs, i.e., unlabeled graphs with 4 vertices and 3 edges. There are three such graphs,  $G_A, G_B, G_C$ , depicted in Figure 1.

$G_A$  can be labeled in 12 distinct ways whereas  $G_B$  and  $G_C$  can be labeled in only 4 distinct ways, see Figure 1.

Evidently, the average energy of the unlabeled  $(4, 3)$ -graphs is

$$\begin{aligned}\langle E \rangle_{\text{unlab}} &= \frac{1}{3} [E(G_A) + E(G_B) + E(G_C)] \\ &= \frac{1}{3} [4.4721 + 3.4641 + 4.0000] = 3.9787,\end{aligned}$$

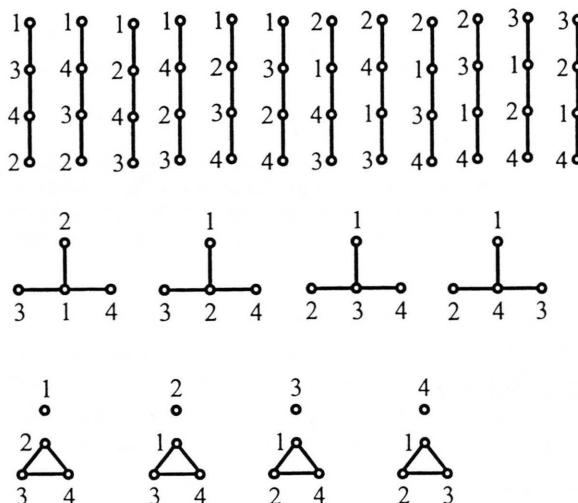
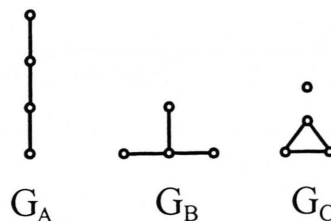


Fig. 1. Unlabeled and labeled graphs with 4 vertices and 3 edges.

whereas if the averaging is done over labeled  $(4, 3)$ -graphs one obtains a somewhat different value:

$$\begin{aligned}\langle E \rangle_{\text{lab}} &= \frac{1}{12+4+4} [12 E(G_A) + 4 E(G_B) + 4 E(G_C)] \\ &= \frac{1}{20} [12 \cdot 4.4721 + 4 \cdot 3.4641 + 4 \cdot 4.0000] \\ &= 4.1761.\end{aligned}$$

In [5] the averaging of the graph energy was done over labeled graphs, which – for reasons explained above – cannot be considered as fully satisfactory. We now report a novel averaging procedure that is aimed at overcoming this problem.

### Uniformly Generated Unlabeled $(n, m)$ -Graphs and their Average Energies

The theory on which the uniform generation of random  $(n, m)$ -graphs is based is described in the book

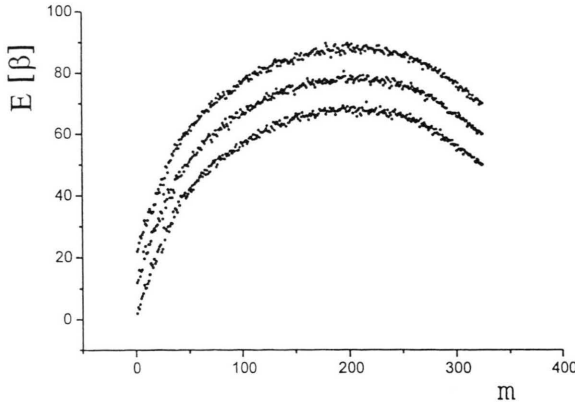


Fig. 2. Energy (in units of the HMO resonance integral  $\beta$ ) of graphs with  $n = 26$  vertices and  $m$  edges, generated uniformly at random; results of three Monte Carlo experiments are shown; the middle and top “curves” are shifted upwards respectively by 10 and 20  $\beta$  units.

[16], especially in its Section 9.6. The actual procedure employed is a pertinent modification and refinement of the Dixon-Wilf algorithm [17, 18] which allows to generate elements at random that are distributed uniformly over the orbits of an acting finite group.

**The Dixon-Wilf algorithm.** Assume that a finite group  $G$  is acting on a finite set  $X$ . We can generate elements  $x \in X$  that are uniformly distributed over the orbits of  $G$  on  $X$  in the following way:

- Choose a conjugacy  $C$  class of elements of the group  $G$  with the probability

$$p(C) := \frac{|C| \cdot |X_g|}{|G| \cdot |G \backslash X|},$$

where  $g$  denotes an element in  $C$ ,  $X_g$  the set of fixed points of  $g$ , and  $G \backslash X$  the set of orbits of  $G$  on  $X$ .

- Pick an element  $g$  from  $C$  and construct a fixed point  $x \in X_g$  of it, uniformly at random.

Then, for each  $\omega \in G \backslash X$ , we have

$$p(x \in \omega) = \frac{1}{|G \backslash X|},$$

i. e.,  $x$  is distributed over the orbits of  $G$  uniformly at random.

In the case of generating labeled  $n$ -vertex graphs, we take for  $X$  the set of all mappings  $f$  from the set of  $\binom{n}{2}$  pairs of vertices into the set  $\{0, 1\}$ , interpreting

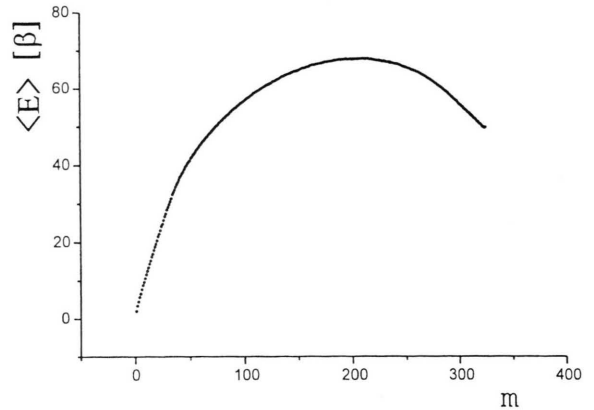


Fig. 3. The average of 250 Monte Carlo experiments specified in Fig. 2; the shape of the  $\langle E \rangle_{\text{unlab}}-m$  line is same for all  $n \geq 10$ , whereas for  $n \leq 8$  it monotonically increases; the case  $n = 9$  is borderline; exactly the same properties were found for the  $\langle E \rangle_{\text{lab}}-m$  lines [5].

$f(\{i, j\}) = 1$  as an edge between the vertices with labels  $i$  and  $j$ . The group  $G$  is the symmetric group acting on the set of vertices, and hence also on the set of pairs of vertices, and finally on the set of these mappings, too, in a canonic way (see [16]).

For a fixed value of  $n$ , ( $5 \leq n \leq 27$ ), graphs were generated uniformly at random by gradually increasing the value of  $m$  from  $m = 1$  up to  $m = n(n-1)/2$ . For each graph thus obtained the energy was calculated. Typical results of such calculations are shown in Figure 2.

The three “curves” seen in Fig. 2 have similar shapes, but are (of course!) somewhat different and full of humps. These features are in full agreement with what has been observed in [5]. In order to get an estimate of  $\langle E \rangle_{\text{unlab}}$  we have repeated the above described Monte Carlo experiments 250 times and calculated the respective average energies. The number of repetitions was chosen to be 250 in order to be able to compare the present results with those from [5]. A typical result is shown in Figure 3.

At first glance the “curve” shown on Fig. 3 is smooth, but a detailed inspection reveals that its smoothness is far from being perfect. Also this detail agrees with what has been observed in [5].

Our main goal is to analyze the difference between the  $(n, m)$ -dependence of  $\langle E \rangle_{\text{lab}}$  (as calculated in [5]) and  $\langle E \rangle_{\text{unlab}}$  (as calculated in this work). To our surprise, the two  $\langle E \rangle-m$  “curves” were found to coincide almost completely.

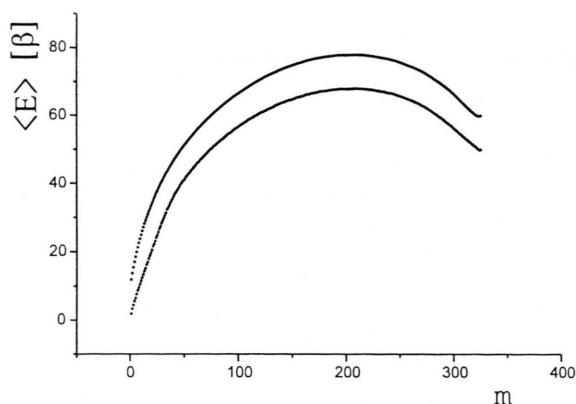


Fig. 4. The “curve” from Fig. 3 and the analogous “curve” produced by the method of [5], shifted upwards by 10  $\beta$  units; without this shift there would be no visible difference between the two lines.

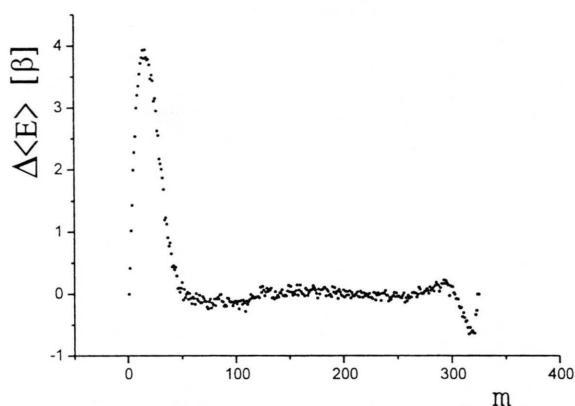


Fig. 5. The difference between the two average energies, (1), for  $n = 26$ ; except for the first few and the last few values of  $m$ , all  $\Delta$ -values are practically equal to zero.

In Fig. 4 an  $\langle E \rangle_{\text{unlab}}-m$  is compared with an  $\langle E \rangle_{\text{lab}}-m$  “curve”. In Fig. 5

$$\Delta \langle E \rangle = \langle E \rangle_{\text{lab}} - \langle E \rangle_{\text{unlab}} \quad (1)$$

is plotted versus  $m$ .

Results such as shown in Figs. 4 and 5 were obtained for all examined values of  $n$ .

## Discussion and Concluding Remarks

The main conclusion of this work is that there is no noteworthy difference between the averaging of graph energies over sets of unlabeled and labeled  $(n, m)$ -graphs, except for graphs with very few edges and

almost complete graphs. For medium large values of  $m$  (which are just those interesting for the theory of graph energy and total  $\pi$ -electron energy) the present method and the method developed in [5] are approximately the same. This finding is in harmony with the well known fact that the automorphism groups  $\mathcal{A}(G)$  of graphs with  $n$  vertices tend to  $\mathcal{A}(G) = \{1\}$ , as  $n$  goes to infinity.

For  $n \geq 10$  the  $\langle E \rangle-m$  line has a unique maximum, the position and height of which were investigated in [5]. The results obtained there are fully identical to what can be obtained by the present method: if  $E_{\text{max}}$  is the maximum value of  $\langle E \rangle_{\text{unlab}}$ , attained for  $m = m_{\text{max}}$ , then

$$E_{\text{max}} \approx A n^k; \quad m_{\text{max}} \approx A' m^{k'}$$

where  $A = 0.73$ ,  $k = 1.39$ ,  $A' = 0.47$  and  $k' = 1.87$ .

As seen from Fig. 5, for graphs with a small number of edges,  $\Delta \langle E \rangle$  is positive-valued, i. e.,  $\langle E \rangle_{\text{lab}}$  exceeds  $\langle E \rangle_{\text{unlab}}$ . This seems to be related to the fact that almost all components of graphs with a small number of edges are trees. Among trees (with a given number  $n$  of vertices) the path-graph  $P_n$  has maximal and the star-graph  $S_n$  minimal energy [19]. On the other hand,  $|\mathcal{A}(P_n)| = 2$  whereas  $|\mathcal{A}(S_n)| = (n-1)!$ . Consequently, the method of [5] generates path-graphs with much higher probability than star-graphs. (For an example see Fig. 1.)

For graphs with a very large number of edges (that are almost complete),  $\Delta \langle E \rangle$  is negative, implying that the most symmetric ones of such graphs have energies above average. Not much is known on the energy of such graphs. However, as seen from Fig. 5, the deviation of  $\Delta$ -values from zero are significantly smaller than in the case of graphs with a small number of edges.

From the point of view of chemical applications, the  $(n, m)$ -graphs of interest are those for which  $n-1 \leq m \leq 3n/2$ . For the theory of graph energy the case  $m \approx m_{\text{max}}$  is also important. For both such types of graphs  $\Delta \langle E \rangle$  is zero or very close to zero. Therefore, for these graphs generating labeled or unlabeled species uniformly at random is equally plausible. The algorithm for labeled graphs [5] is significantly simpler and faster. Therefore it deserves to be given preference, irrespective of the weaknesses in its theoretical foundation. In special cases and doubtful situations the theoretically fully justified generation of unlabeled graphs should be recommended.

*Acknowledgement*

I. G. thanks the Alexander von Humboldt Foundation for financial support.

- [1] D. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs – Theory and Application*, Academic Press, New York 1980.
- [2] I. Gutman and O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin 1986.
- [3] I. Gutman, *J. Math. Chem.* **1**, 123 (1987).
- [4] A. Graovac, I. Gutman, P. E. John, D. Vidović, and I. Vlah, *Z. Naturforsch.* **56a**, 307 (2001).
- [5] I. Gutman, T. Soldatović, and D. Vidović, *Chem. Phys. Lett.* **297**, 428 (1998).
- [6] C. A. Coulson, B. O’Leary, and R. B. Mallion, *Hückel Theory for Organic Chemists*, Academic Press, London 1978.
- [7] O. E. Polansky, *Z. Naturforsch.* **41a**, 560 (1986).
- [8] H. Hartmann, *Z. Naturforsch.* **2a**, 259, 263 (1947).
- [9] G. G. Hall, *Proc. Roy. Soc. London* **A229**, 251 (1955).
- [10] B. J. McClelland, *J. Chem. Phys.* **54**, 640 (1971).
- [11] W. England and K. Ruedenberg, *J. Amer. Chem. Soc.* **95**, 8769 (1973).
- [12] I. Gutman and M. Rašković, *Z. Naturforsch.* **40a**, 1059 (1985).
- [13] J. Cioslowski, *Z. Naturforsch.* **40a**, 1169 (1985).
- [14] J. Cioslowski and O. E. Polansky, *Theor. Chim. Acta* **74**, 55 (1988).
- [15] J. Cioslowski, *Topics Curr. Chem.* **153**, 85 (1990).
- [16] A. Kerber, *Applied Finite Group Actions*, Springer-Verlag, Berlin 1999.
- [17] J. D. Dixon and H. S. Wilf, *J. Algorithms* **4**, 205 (1983).
- [18] Computer programs for generation of  $(n, m)$ -graphs uniformly at random are available from the authors (H. F. and A. K.) upon request, see <http://www-ang.kfunigraz.ac.at/~friper/symmetrica.html>
- [19] I. Gutman, *Theor. Chim. Acta* **45**, 79 (1977).