On the Relation between W'/W Index, Hyper-Wiener Index, and Wiener Number

Dejan Plavšić,*,† Nella Lerš,† and Katica Sertić-Bionda‡

The Rugjer Bošković Institute, P.O.B. 1016, HR-10001 Zagreb, The Republic of Croatia, and Faculty of Chemical Engineering and Technology, University of Zagreb, HR-10000 Zagreb, The Republic of Croatia

Received March 11, 1999

It is shown analytically that the W'/W index, the hyper-Wiener index, and the Wiener number are closely related graph-theoretical invariants for acyclic structures. A general analytical expression for the hyper-Wiener index of a tree is derived too.

1. INTRODUCTION

Properties of a molecule are a consequence of a complicated interplay of its topology (atomic connectivity), metric characteristics (bond lengths, valence, and torsion angles), and the detailed dynamics of electrons and nuclei.[1] Within many classes of compounds, the variations of molecular metric and electronic structure are small. Hence, changes of many of molecular properties in these classes may be considered as only topology conditioned.[2−9] Molecular topology can be represented by a (molecular) graph that is abstract, essentially a nonnumerical mathematical object.[10−15] To perform quantitative topology−property/activity studies of molecules it is necessary to quantify the structural information contained in the corresponding graphs. The characterization of a graph is usually carried out by means of graph invariants[10] (topological indices[16]).

The representation of a molecule by a topological index entails a considerable loss of information concerning the molecular structure. Hence, chemists are permanently in pursuit of novel topological indices which would improve the graph-theoretical characterization of molecular structure and enable one to make better and easily interpretable regression models of topology−property/activity relationships. In the framework of this effort Randić has recently put forward a novel bond additive molecular descriptor, the W'/W index.[17,18] The index is the sum of graphical bond orders[17,18] of all edges in a graph calculated by means of the Wiener number.[19] He also tested the W'/W index in the framework of the single-variable linear regression model by examining the van der Waals areas of heptanes[17,20] and some 20 molecular properties of octane isomers.[21] The close resemblance in quality between the regressions with the W'/W index as predictor variable and regressions based on the hyper-Wiener index[20] as well as on the Wiener number indicates that these indices encode very similar information on topology of acyclic structures. We have investigated the intercorrelation of these three indices on heptane, octane, and nonane isomers and found in all cases a high value of the coefficient of determination ($r^2 > 0.996$) and strong linear intercorrelation. The plot of the W'/W index versus the hyper-Wiener index R for nonanes is illustrated in Figure 1. Such a behavior of these three indices hints that a formal relationship might exist between them.

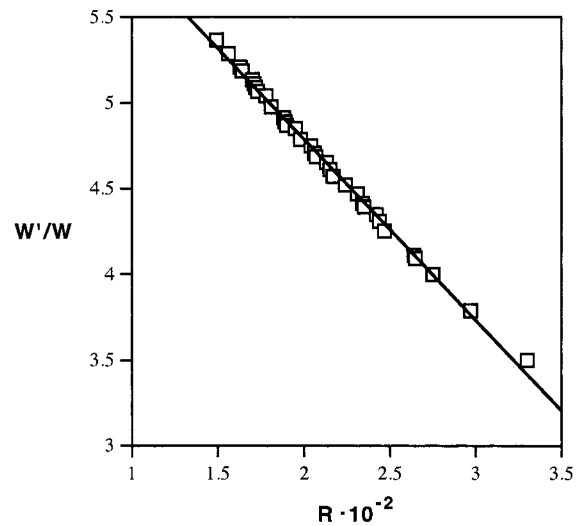


Figure 1. Plot of W'/W vs $R \times 10^{-2}$ for nonanes. The regression equation and statistical parameters are $W'/W = -1.055(\pm 0.011)\cdot(R \times 10^{-2}) + 6.899(\pm 0.023)$; $n = 35$; $r^2 = 0.997$; $s = 0.026$; $F^{1,33} = 9507$.

In this article we will discuss the relationship between the W'/W index, the hyper-Wiener index, and the Wiener number for acyclic structures.

2. DEFINITIONS

Wiener Number. The Wiener number, $W = W(G)$, of a connected undirected graph G with N vertices is defined as[16]

$$W = \sum_{i=1}^{N-1} \sum_{j>i}^{N} (\mathbf{D})_{ij} \tag{1}$$

where $(\mathbf{D})_{ij}$ denotes the element in the i-th row and j-th column of the distance matrix $\mathbf{D}$ of the graph G. The summation goes over all the entries above the main diagonal of $\mathbf{D}$. If G is a connected undirected acyclic graph (tree), T, with N vertices, the Wiener number can also be expressed as[19]

$$W = \sum_{e_{ij}} C_{e_{ij}} \tag{2}$$

† The Rugjer Bošković Institute.
‡ Faculty of Chemical Engineering and Technology.

10.1021/ci9900229 CCC: $19.00
Published on Web 10/30/1999

where

$$C_{e_{ij}} = {}^iX_{e_{ij}}(N - {}^iX_{e_{ij}});\quad 1 \le {}^iX_{e_{ij}} \le N-1$$
$$= {}^iX_{e_{ij}}\,{}^jX_{e_{ij}} \qquad (3)$$

e_{ij} denotes the edge connecting vertices i and j of T. The summation in eq 2 goes over all edges in T. ${}^iX_{e_{ij}}$ and ${}^jX_{e_{ij}}$ in eq 3 denote the number of vertices of T on the side of the vertex i and on the side of the vertex j of the edge e_{ij}, respectively.

Graphical Bond Order $W_{e_{ij}}/W$. The graphical bond order $W_{e_{ij}}/W$ of an edge e_{ij} of a connected undirected graph G is defined by[17,18]

$$W_{e_{ij}}/W = \frac{W(G - e_{ij})}{W(G)} \qquad (4)$$

where $W(G)$ is the Wiener number of G and $W(G - e_{ij})$ denotes the Wiener number of the spanning subgraph $G - e_{ij}$ obtained from G by deleting the edge e_{ij}. $G - e_{ij}$ is connected if and only if G contains at least one ring and the edge e_{ij} is one of the edges making up the ring(s). A disconnected $G - e_{ij}$ has two components, say G_1 and G_2, and the Wiener index in this case is by definition given by the expression[17,18]

$$W(G - e) = W(G_1) + W(G_2) \qquad (5)$$

***W′/W* Index.** The W'/W index of a connected undirected graph G is defined as[17,18]

$$W'/W = \sum_{e_{ij}} W_{e_{ij}}/W = \frac{1}{W(G)}\sum_{e_{ij}} W(G - e_{ij}) \qquad (6)$$

where the summation goes over all edges of G.

Hyper-Wiener Index. The hyper-Wiener index R was recently introduced by Randić for an acyclic structure.[20] The R index, $R = R(T)$, of a tree T is defined as

$$R(T) = \sum_{p_{rs}} {}^rX_{p_{rs}}\,{}^sX_{p_{rs}} \qquad (7)$$

where p_{rs} represents the path connecting vertices r and s of T. ${}^rX_{p_{rs}}$ and ${}^sX_{p_{rs}}$ denote the number of vertices of T on each side of the path p_{rs}, including r and s, respectively. The summation runs over all paths in T. Note, if paths of length 1 (edges) are the only paths taken into consideration, than eq 7 is reduced to eq 2. The original definition was extended so as to be applicable for all connected graphs.[22] The hyper-Wiener index, $R = R(G)$, of a connected graph G with N vertices is defined as

$$R = {}^1\!/_2 \sum_{i=1}^{N-1}\sum_{j>i}^{N} [(\mathbf{D})_{ij} + (\mathbf{D})_{ij}^2] \qquad (8)$$

3. RELATIONSHIP BETWEEN W'/W INDEX, HYPER-WIENER INDEX, AND WIENER NUMBER

Let T be a connected undirected acyclic graph with N vertices, and let e_{ij} be an edge of T, see Figure 2. The graph $T - e_{ij}$ obtained from T by deleting the edge e_{ij} has two components, T_1 and T_2, with ${}^iX_{e_{ij}}$ and ${}^jX_{e_{ij}}$ vertices, respectively. Clearly, ${}^iX_{e_{ij}} + {}^jX_{e_{ij}}$ is equal to N. The Wiener number of the spanning subgraph $T - e_{ij}$ of T is smaller than the Wiener number of T due to the absence of the contribution of the edge e_{ij} being equal to ${}^iX_{e_{ij}}{}^jX_{e_{ij}}$ and the decrease in contributions of the remaining edges. An edge of $T - e_{ij}$, say e_{ab} (see Figure 2), makes a contribution to $W(T - e_{ij})$ equal to the difference between the contribution of the corresponding edge in T to $W(T)$, ${}^aX_{e_{ab}}{}^bX_{e_{ab}}$, and the product ${}^aX_{e_{ab}}{}^jX_{e_{ij}}$. Hence, the difference between $W(T)$ and $W(T - e_{ij})$ is given by

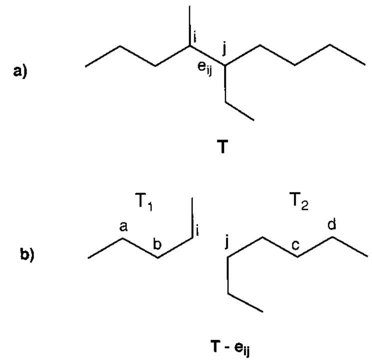


Figure 2. (a) A connected undirected acyclic graph T. An edge of T is denoted by e_{ij}. (b) Spanning subgraph $T - e_{ij}$ of T with components T_1 and T_2.

$$W(T) - W(T - e_{ij}) = {}^iX_{e_{ij}}{}^jX_{e_{ij}} + {}^jX_{e_{ij}} \sum_{e_{ab}\in E_1} {}^aX_{e_{ab}} + {}^iX_{e_{ij}} \sum_{e_{cd}\in E_2} {}^dX_{e_{cd}} \qquad (9)$$

on the condition that

$$(\mathbf{D})_{ia} > (\mathbf{D})_{ib} \qquad (10)$$

and

$$(\mathbf{D})_{jd} > (\mathbf{D})_{jc} \qquad (11)$$

${}^aX_{e_{ab}}$ (${}^dX_{e_{cd}}$) is the number of vertices of T_1 (T_2) on the side of the vertex a (d) of the edge e_{ab} (e_{cd}). E_1 and E_2 denote the sets of edges of the components T_1 and T_2, respectively. The summation in the second (third) term on the right hand side of eq 9 runs over all edges of the component T_1 (T_2) of the graph $T - e_{ij}$.

The graphical bond order $W_{e_{ij}}/W$ of the edge e_{ij} of T can be obtained by dividing eq 9 by $W(T)$:

$$W_{e_{ij}}/W = 1 - \frac{1}{W(T)}\left({}^iX_{e_{ij}}{}^jX_{e_{ij}} + {}^jX_{e_{ij}} \sum_{e_{ab}\in E_1} {}^aX_{e_{ab}} + {}^iX_{e_{ij}} \sum_{e_{cd}\in E_2} {}^dX_{e_{cd}}\right) \qquad (12)$$

This quantity represents the “importance” of the edge e_{ij} in T.

The W'/W index of T is the sum of graphical bond orders of all $N - 1$ edges of T:

$$W'/W = N - 1\,\frac{1}{W(T)}\sum_{e_{ij}} \left({}^{i}X_{e_{ij}}\,{}^{j}X_{e_{ij}} + {}^{j}X_{e_{ij}} \sum_{e_{ab}\in E_1} {}^{a}X_{e_{ab}} + {}^{i}X_{e_{ij}} \sum_{e_{cd}\in E_2} {}^{d}X_{e_{cd}}\right) \quad (13)$$

By noting that

$$\sum_{e_{ij}} \left({}^{i}X_{e_{ij}}\,{}^{j}X_{e_{ij}} + {}^{j}X_{e_{ij}} \sum_{e_{ab}\in E_1} {}^{a}X_{e_{ab}} + {}^{i}X_{e_{ij}} \sum_{e_{cd}\in E_2} {}^{d}X_{e_{cd}}\right) = 2R(T) - W(T) \quad (14)$$

the relationship between the W'/W index, the hyper-Wiener index, and the Wiener number of T immediately follows:

$$W'/W = N - \frac{2R(T)}{W(T)} \quad (15)$$

Since $2R$ is equal to the sum of the (unnormalized) second moment of distance D_2 and the Wiener number,[22] the W'/W index can also be expressed as

$$W'/W = N - 1 - \frac{D_2(T)}{W(T)} \quad (16)$$

4. DERIVATION OF GENERAL EXPLICIT FORMULA FOR HYPER-WIENER INDEX OF TREE

A closer examination of eqs 6 and 15 reveals that it is possible by means of them to arrive at the general explicit formula for calculation of the hyper-Wiener index of T. To wit, combining eqs 6 and 15 one can write

$$R(T) = (NW - W')/2 \quad (17)$$

where W' is the sum of the Wiener numbers of all the spanning subgraphs $T - e$ of T. It is well-known that the Wiener number can be calculated in a number of ways.[16,19,23−34] If one selects the route[23]

$$W = \sum_{n=1}^{n\le N-1} n\,{}^{n}p \quad (18)$$

where ${}^{n}p$ denotes the number of paths of length n in T, then the general explicit formula for calculation of $R(T)$ can be derived in a rather simple way.

The total number of paths of a given length n in the spanning subgraphs $T - e$ of T, ${}^{n}P$, is given by the expression

$${}^{n}P = \sum_{T-e} {}^{n}p_{T-e} = (N - n - 1)\,{}^{n}p \quad (19)$$

where ${}^{n}p_{T-e}$ is the number of paths of length n in $T - e$, and the summation runs over all the spanning subgraphs $T - e$ of T.

W' of T is given by

$$W' = \sum_{n=1}^{n\le N-1} n\,{}^{n}P = \sum_{n=1}^{n\le N-1} n(N - n - 1)\,{}^{n}p \quad (20)$$

Combining eqs 17, 18, and 20, one obtains

$$R(T) = {}^{1}\!/_{2} \sum_{n=1}^{n\le N-1} (n^2 + n)\,{}^{n}p \quad (21)$$

To calculate the R index of T it is necessary and sufficient to have knowledge of ${}^{n}p$'s of T. The number of paths of a given length n in T can be calculated either by means of the recursive relation or using the general analytical expression for ${}^{n}p$ in T.[35] The recursive relation reads as

$${}^{n}p = \sum_{(\mathbf{D})_{ij}=n-2} v_i v_j - 2({}^{n-1}p) - {}^{n-2}p; \quad n \ge 3 \quad (22)$$

where v_i is the valence of the vertex i, and $(\mathbf{D})_{ij} = n - 2$ denotes that the topological distance between the vertices i and j is equal to $n - 2$. The summation runs over all the pairs of vertices of T separated by the paths of length $n - 2$. The initial conditions are

$${}^{1}p = N - 1 \quad (23)$$

$${}^{2}p = {}^{1}\!/_{2} \sum_i v_i^2 - N + 1 \quad (24)$$

The general analytical expression for ${}^{n}p$ in T reads as follows:

$${}^{n}p = \sum_{m=1}^{n-2+k} (-1)^{m+s}(n - m - 1 + t) \sum_{\substack{(\mathbf{D})_{ij}=m \\ i<j}} v_i v_j + (-1)^{n}\left(\frac{n-1}{2} \sum_{(\mathbf{D})_{ij}=0} v_i v_i - N + 1\right) \quad (25)$$

Note that $\sum_{(\mathbf{D})ij=0} v_i v_j = \sum_i v_i^2$. The parameters k, s, and t take the following values:

$$k = \begin{cases} 2 & \text{for } n = 1 \\ 1 & \text{for } n = 2 \\ 0 & \text{for } n \ge 3 \end{cases} \quad (26)$$

$$s = \begin{cases} 1 & \text{for } n = \text{odd} \\ 0 & \text{for } n = \text{even} \end{cases} \quad (27)$$

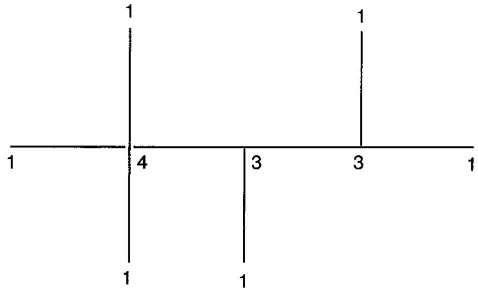
and

$$t = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n \ge 2 \end{cases} \quad (28)$$

Combining eqs 21 and 25 one obtains the general explicit formula for calculation of the R index of T

$$R(T) = {}^{1}\!/_{2} \sum_{n=1}^{n\le N-1} (n^2 + n)\left\{\sum_{m=1}^{n-2+k} (-1)^{m+s}(n - m - 1 + t) \sum_{\substack{(\mathbf{D})_{ij=m} \\ i<j}} v_i v_j + (-1)^{n}\left(\frac{n-1}{2} \sum_{(\mathbf{D})_{ij=0}} v_i v_j - N + 1\right)\right\} \quad (29)$$

where the parameters k, s, and t take the values from eqs 26, 27, and 28, respectively. The application of the formula is illustrated for the hydrogen-suppressed graph of 2,2,3,4-tetramethylpentane in Figure 3.



$$R(T) = \frac{1}{2}\Big\{2\times 8 + 6\Big[\frac{1}{2}\left(6\times 1^2 + 2\times 3^2 + 1\times 4^2\right) - 8\Big] + 12\Big[3\times 1\times 4 + 3\times 1\times 3 + 3\times 3 + 3\times 4 - \left(6\times 1^2 + 2\times 3^2 + 1\times 4^2\right) + 8\Big] + 20\Big[-2\left(3\times 1\times 4 + 3\times 1\times 3 + 3\times 3 + 3\times 4\right) + 4\times 1\times 1 + 6\times 1\times 3 + 3\times 4 + 4\times 1 + \frac{3}{2}\left(6\times 1^2 + 2\times 3^2 + 1\times 4^2\right) - 8\Big]\Big\} = 164$$

Figure 3. Calculation of the *R* index of the hydrogen-suppressed graph of 2,2,3,4-tetramethylpentane. Numbers at each site represent the corresponding graph-theoretical valences.

A special case of an acyclic graph is the path graph, P_N. Equation 29 in the case of P_N takes a rather simple form

$$R(P_N) = {}^1\!/_2 \sum_{n=1}^{n=N-1} (n^2 + n)(N - n) \qquad (30)$$

whose summation gives the formula already derived by Lukovits.[36]

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technology of the Republic of Croatia and by the Croatian–Slovenian project "Discrete Mathematics in Chemistry". We are grateful to Dr. Milan Šoškić (Zagreb) and Professor Ante Graovac (Zagreb) for useful comments.

REFERENCES AND NOTES

(1) McWeeny, R. *Methods of Molecular Quantum Mechanics*, 2nd ed.; Academic Press: London, 1989.
(2) Rouvray, D. H. Predicting Chemistry from Topology. *Sci. Am.* **1986**, *254* (3), 36–43.
(3) Turro, N. J. Geometrical and Topological Thinking in Organic Chemistry. *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 882–901.
(4) Rouvray, D. H. The Modeling of Chemical Phenomena Using Topological Indices. *J. Comput. Chem.* **1987**, *8*, 470–480.
(5) Randić, M. The Nature of Chemical Structure. *J. Math. Chem.* **1990**, *4*, 157–189.
(6) Randić, M. Chemical Structure – What is She? *J. Chem. Educ.* **1992**, *69*, 713–718.
(7) Rouvray, D. H. A Rationale for the Topological Approach to Chemistry. *J. Mol. Struct. (Theochem)* **1995**, *336*, 101–114.
(8) Randić, M. On Characterization of Chemical Structure. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 672–687.
(9) Milne, G. W. Mathematics as a Basis for Chemistry. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 639–644.
(10) Harary, F. *Graph Theory;* Addison–Wesley: Reading, MA, 1972.
(11) Kier, L. B.; Hall, L. M. *Molecular Connectivity in Structure–Activity Analysis;* Research Studies Press Ltd.: Letchworth, England, 1986.
(12) Gutman, I.; Polansky, O. *Mathematical Concepts in Organic Chemistry;* Springer–Verlag: Berlin, 1986.
(13) Bonchev, D., Rouvray, D. H., Eds. *Chemical Graph Theory. Introduction and Fundamentals;* Abacus Press, Gordon and Breach: New York, 1991.
(14) Trinajstić, N. *Chemical Graph Theory*, 2nd revised ed.; CRC Press: Boca Raton, FL, 1992.
(15) Balaban, A. T., Ed. *From Chemical Topology to Three–Dimensional Geometry;* Plenum Press: New York, 1997.
(16) Hosoya, H. Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44* , 2332–2339.
(17) Randić, M. Search for Optimal Molecular Descriptors. *Croat. Chem. Acta* **1991**, *64*, 43–54.
(18) Randić, M.; Mihalić, Z.; Nikolić, S.; Trinajstić, N. Graphical Bond Orders: Novel Structural Descriptors. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 403–409.
(19) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
(20) Randić, M. Novel Molecular Descriptor for Structure-Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478–483.
(21) Randić, M. Comparative Regression Analysis. Regressions Based on a Single Descriptor. *Croat. Chem. Acta* **1993**, *66*, 289–312.
(22) Klein, D. J.; Lukovits, I.; Gutman, I. On the Definition of the Hyper-Wiener Index for Cycle-Containing Structures. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 672–687.
(23) Bonchev, D.; Trinajstić, N. Information Theory, Distance Matrix and Molecular Branching. *J. Chem. Phys.* **1977**, *67*, 4517–4533.
(24) Bersohn, M. A Fast Algorithm for Calculation of the Distance Matrix. *J. Comput. Chem.* **1983**, *4*, 110–113.
(25) Canfield, E. R.; Robinson, R. W.; Rouvray, D.H. Determination of the Wiener Molecular Branching Index for the General Tree. *J. Comput. Chem.* **1985**, *6*, 598–609.
(26) Barysz, M.; Plavšić, D.; Trinajstić, N. A Note on Topological Indices. *Math. Chem. (Muelcheim/Ruhr)* **1986**, *19*, 89–116.
(27) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. An Algorithm for Construction of the Molecular Distance Matrix. *J. Comput. Chem.* **1987**, *8*, 170–173.
(28) Mohar, B.; Pisanski, T. How to Compute the Wiener Index of a Graph. *J. Math. Chem.* **1988**, *2*, 267–277.
(29) Senn, P. The Computation of the Distance Matrix and the Wiener Index for Graphs of Arbitrary Complexity with Weighted Vertices and Edges. *Comput. Chem.* **1988**, *12*, 219–227.
(30) Lukovits, I. General Formulas for the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 503–507.
(31) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. Some Recent Results in the Theory of the Wiener Number. *Indian J. Chem.* **1993**, *32A*, 651–661.
(32) Plavšić, D.; Nikolić, S.; Trinajstić, N.; Klein, D. J. Relation between the Wiener Index and the Schultz Index for Several Classes of Chemical Graphs. *Croat. Chem. Acta* **1993**, *66*, 345–353.
(33) Gutman, I. A New Method for the Calculation of the Wiener Number of Acyclic Molecules. *J. Mol. Struct. (Theochem)* **1993**, *285*, 137–142.
(34) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener Index: Development and Applications. *Croat. Chem. Acta* **1995**, *68*, 105–129.
(35) Plavšić, D.; Šoškić, M.; Landeka, I.; Trinajstić, N. On the Relation between the P′/P Index and the Wiener Number. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1123–1126.
(36) Lukovits, I. Formulas for the Hyper-Wiener Index of Trees. *J. Chem. Inf. Comput. Sci.* **1994**, *34* , 1079–1081.

CI9900229