

CHEMICAL GRAPH THEORY OF THE WATER MOLECULE**¹G .Neelima ²Dr.Akkili Naresh***

¹P.G Student , Department of Mathematics & Computing; Central University of Andhra Pradesh,
Ananthapuramu 515701,A.P,India

Email id: neelimahanumanthreddy@gmail.com

²Assistant Professor, Department of Mathematics & Computing; Central University of Andhra Pradesh ,Janthaluru, Ananthapuramu 515701,A.P,India

Email id: akkilinaresh@gmail.com

Abstract: This dissertation presents a rigorous study of the chemical graph theory of the water molecule (H_2O) across solid, liquid, and gaseous phases. It applies topological and algebraic tools to analyze molecular structure and intermolecular interactions beyond classical approaches. In the gaseous state, H_2O is modeled as a simple path graph with two edges and no cycles. In the liquid state, it forms a complex, dynamic hydrogen-bond network with small-world properties. In solid ice (ice Ih), water exhibits a tetrahedral graph structure with degree-four vertices in a hexagonal lattice. Various topological indices and graph spectra are systematically computed and compared across phases. Strong correlations are established between graph invariants and physical properties like density, heat capacity, and surface tension. The work introduces new frameworks for condensed-phase graph modeling, enhancing understanding of water's structure-property relationships.

Keywords: chemical graph theory, water molecule, hydrogen-bond network, topological indices, ice structure, molecular graphs, Wiener index, Randić' index, phase transitions, graph spectra.

Introduction:

Water (H_2O) is simultaneously the most familiar and the most scientifically profound molecule in nature. It constitutes approximately 71% of Earth's surface, makes up 60–70% of the human body, and is universally recognized as the solvent of life. Despite this apparent ubiquity, water continues

to surprise scientists with its anomalous physical behavior and the structural complexity that underlies its macroscopic properties. The intersection of chemical graph theory with the study of water offers a rich and underexplored framework for understanding this complexity in rigorous mathematical terms[1].

Chemical graph theory is the application of graph-theoretical mathematics to chemical structures. In this framework, molecules are represented as graphs in which vertices (nodes) correspond to atoms and edges correspond to chemical bonds. This representation, while apparently reductive, captures essential topological information about molecular structure that is independent of geometric embedding and coordinate systems. The resulting graph-theoretical invariants — called topological indices — encode structural information in a compact mathematical form that can be correlated with physical, chemical, and biological properties.

The history of chemical graph theory traces back to the pioneering work of Arthur Cayley in 1875, who enumerated isomers of saturated hydrocarbons using tree-counting methods. Harry Wiener's 1947 introduction of the path number (now known as the Wiener index) marked the beginning of systematic correlations between molecular topology and physical properties. Since then, hundreds of topological indices have been defined, and the field has grown to encompass graph spectral theory, random graphs, network theory, and algebraic graph theory in chemical applications.

Water presents a uniquely challenging and rewarding subject for chemical graph theory because its structural organization varies so dramatically across thermodynamic phases. In the gaseous state, water exists as isolated triatomic molecules — simple, well-defined, and easily represented as a minimal molecular graph with three vertices and two edges. In the liquid state, water forms a continuously fluctuating, three-dimensional hydrogen-bond network of extraordinary complexity, wherein each molecule participates in an average of 3.5 to 3.7 hydrogen bonds simultaneously. In

the solid state (ice), this network crystallizes into a variety of polymorphic forms governed by strict topological rules — the Bernal-Fowler ice rules — that constrain hydrogen atom positions on the graph edges[2].

The motivation for this dissertation arises from three converging scientific imperatives. First, despite the extensive literature on water's structure and properties, a comprehensive, systematic, and phase-resolved treatment of water using chemical graph theory has not previously been undertaken. Second, recent advances in network science and topological data analysis provide new mathematical tools that can be applied to the graph-theoretical description of condensed-phase water. Third, understanding the structure-property relationships of water from a topological perspective has direct practical implications for fields ranging from atmospheric chemistry to materials science, pharmacology, and climate science.

1.2 The Water Molecule: A Brief Overview

The water molecule consists of one oxygen atom covalently bonded to two hydrogen atoms. With an atomic number of 8, oxygen has the electron configuration $[\text{He}] 2s^2 2p^4$, and in water it achieves an approximately sp^3 hybrid orbital configuration, giving the molecule its characteristic bent geometry. The O-H bond length is 0.9572 Å in the gas phase, and the H-O-H bond angle is 104.52° , slightly less than the ideal tetrahedral angle of 109.47° due to the greater repulsion exerted by the two lone electron pairs on oxygen.

These structural parameters translate directly into graph-theoretical representations. The molecular graph of gaseous water is a path graph P_3 on three vertices, where the central vertex (oxygen) has degree 2 and the terminal vertices (hydrogen atoms) each have degree 1. This minimal graph encodes the fundamental connectivity of the molecule, from which all basic topological indices can be computed exactly.

The polarity arising from the electronegativity difference between oxygen (3.44) and hydrogen (2.20) — producing a dipole moment of 1.8546 D — is the root cause of water's capacity to form hydrogen bonds. Each O-H bond constitutes a hydrogen-bond donor, while each lone pair on oxygen constitutes a hydrogen-bond acceptor. This donor/acceptor duality means that each water molecule can participate in up to four hydrogen bonds, two as donor and two as acceptor, creating the tetrahedral coordination that characterizes both liquid water and crystalline ice[3].

Hydrogen bonds in water have an energy of approximately 20 kJ/mol — significantly weaker than covalent bonds (~460 kJ/mol for O-H) but stronger than van der Waals interactions (~1–5 kJ/mol). This intermediate strength means that hydrogen bonds are constantly forming and breaking at room temperature in liquid water (with a lifetime of approximately 1–3 picoseconds per individual bond) while being collectively responsible for water's anomalously high boiling point, surface tension, heat capacity, and its property of having maximum density at 4°C rather than at the freezing point.

1.3 Chemical Graph Theory: Historical Development

The intellectual roots of chemical graph theory extend to the combinatorial chemistry of the nineteenth century. In 1874, Arthur Cayley published a landmark paper on the mathematical enumeration of the structural isomers of alkanes (then called paraffins), in which he represented carbon skeletons as trees — connected acyclic graphs — and applied generating functions to count distinct trees with n vertices. This work predates the formal axiomatization of graph theory by Dénes König (1936) but is retrospectively recognized as the first application of graph-theoretical thinking to chemistry.

The early twentieth century saw increasing recognition that the connectivity of atoms in a molecule — independent of their three-dimensional spatial arrangement — determines many chemical properties. The Lewis structure theory (1916) and later valence bond theory implicitly encoded

graph-theoretical information. However, it was not until 1947 that a quantitative connection between molecular graph structure and physical properties was established: Harry Wiener demonstrated that the 'path number' of a molecular graph — now called the Wiener index $W(G)$ and defined as the sum of distances between all pairs of vertices in the hydrogen-suppressed graph — correlates strongly with the boiling points of alkanes.

The 1970s and 1980s witnessed an explosion of activity in chemical graph theory. Milan Randić introduced his connectivity index (now called the Randić index) in 1975, which proved to be an excellent predictor of physicochemical properties. Ivan Gutman and Nenad Trinajstić introduced the Zagreb indices M_1 and M_2 in 1972. Haruo Hosoya defined the Z-index (Hosoya index) and the Hosoya polynomial (also called the Wiener polynomial or distance polynomial). Miroslav Fiedler and others developed the spectral theory of graphs with applications to chemistry. The journal MATCH (Mathematical Chemistry) was founded in 1975, providing a dedicated venue for this emerging field.

By the 1990s, chemical graph theory had matured into a rich discipline with deep connections to algebraic topology, combinatorics, linear algebra, and statistical mechanics. The introduction of topological indices for 3D molecular structures, the development of graph-theoretical methods for predicting pharmacological activity (quantitative structure-activity relationships, QSAR), and the application of percolation theory to polymer networks and hydrogen-bond networks extended the scope of the field significantly[2].

In the twenty-first century, chemical graph theory has converged with the broader field of network science. The discovery of small-world networks (Watts and Strogatz, 1998) and scale-free networks (Barabási and Albert, 1999) introduced new concepts — clustering coefficient, characteristic path length, degree distribution, network diameter — that have been applied to

molecular systems including protein networks, metabolic networks, and, most relevantly for this dissertation, the hydrogen-bond network of liquid water.

1.4 Research Objectives

This dissertation pursues the following primary research objectives:

- To derive and analyze the complete set of molecular graph representations for water in its gas, liquid, and solid states, including the construction of representative unit graphs for periodic systems.
- To compute and compare a comprehensive suite of topological indices (Wiener, Randić, Zagreb M1 and M2, Hosoya, Szeged, PI, Balaban J index) for water's graph representations in each phase.
- To investigate the network-theoretical properties of liquid water's hydrogen-bond graph, including degree distribution, clustering coefficient, network diameter, and characteristic path length, as a function of temperature.
- To analyze the graph-theoretical implications of the Bernal-Fowler ice rules for the topology of crystalline ice structures, and to compare topological indices across the major ice polymorphs (Ih, Ic, II, III).
- To establish correlations between phase-specific topological indices and macroscopic physical properties of water, including density, viscosity, dielectric constant, heat capacity, and surface tension.
- To develop a spectral graph-theoretical analysis of water's molecular graphs and hydrogen-bond networks, computing and interpreting the adjacency spectra, Laplacian spectra, and signless Laplacian spectra in each phase.

- To propose a unified graph-theoretical framework for representing and comparing the structural organization of water across phases and across different theoretical levels of description.

2. Foundations of Chemical Graph Theory

2.1 Basic Graph Theory Definitions

A graph $G = (V, E)$ consists of a non-empty finite set V of vertices (also called nodes or points) and a set E of edges (also called lines or bonds), where each edge connects an unordered pair of distinct vertices. The order of a graph is $|V| = n$ (the number of vertices) and the size is $|E| = m$ (the number of edges). Two vertices $u, v \in V$ are adjacent if the edge $\{u, v\} \in E$; in this case, u and v are called neighbors.

The degree $d(v)$ of a vertex v is the number of edges incident to it, i.e., the number of neighbors. The degree sequence of G is the non-increasing sequence of vertex degrees $(d_1, d_2, \dots, d_n)$. The sum of all vertex degrees equals twice the number of edges: $\sum_i d(v_i) = 2m$ (the handshaking lemma). In molecular graphs, the degree of a vertex corresponds to the valence of the corresponding atom in the hydrogen-suppressed graph, or to its total bond count in the hydrogen-inclusive graph.

A path between vertices u and v in G is a sequence of distinct vertices $u = v_0, v_1, v_2, \dots, v_k = v$ such that consecutive vertices are adjacent. The length of the path is the number of edges k . The distance $d(u, v)$ between vertices u and v is the length of the shortest path connecting them (the geodesic distance). If no path exists, $d(u, v) = \infty$. A graph is connected if $d(u, v) < \infty$ for all pairs $u, v \in V$.

A cycle is a closed path $v_0, v_1, \dots, v_k = v_0$ where all vertices except the first and last are distinct. A tree is a connected graph with no cycles; equivalently, a connected graph on n vertices with exactly

$n-1$ edges. The molecular graph of a simple acyclic molecule (including isolated water) is a tree. A forest is a (possibly disconnected) acyclic graph[3].

2.2 Molecular Graph Representations

In chemical graph theory, a molecular graph G_M is a graph in which each vertex represents a non-hydrogen atom and each edge represents a covalent bond. This hydrogen-suppressed representation is standard in the field because the presence of hydrogen atoms, which are implicit in most chemical notations, would obscure the connectivity of the carbon or heteroatom skeleton. For water, the hydrogen-suppressed graph is a single isolated vertex (the oxygen atom), which is chemically uninformative. Therefore, for water, the hydrogen-inclusive representation is used throughout this dissertation.

In the hydrogen-inclusive molecular graph of H_2O , there are three vertices: one for oxygen (O) and two for hydrogen (H_1 , H_2). There are two edges: the O- H_1 bond and the O- H_2 bond. The vertex degree sequence is (2, 1, 1) — the oxygen has degree 2 and each hydrogen has degree 1. This graph is isomorphic to the path graph P_3 on three vertices, which is the unique tree on three vertices.

For condensed phases, the molecular graph concept must be extended. In the liquid and solid states, intermolecular hydrogen bonds create a supramolecular network that constitutes a new graph superimposed on the collection of individual molecular graphs. The hydrogen-bond graph G_{HB} has water molecules as vertices and hydrogen bonds as edges. This graph has n vertices (one per molecule) and m edges (one per hydrogen bond). Since each molecule can form up to 4 hydrogen bonds, the maximum degree is 4. Analysis of liquid water's hydrogen-bond network reveals a degree distribution centered near 3.5 hydrogen bonds per molecule at 25°C[5].

A third level of representation considers the extended atomic graph G_{ext} in which all atoms (O and H) of all molecules are vertices, and both covalent O-H bonds and intermolecular $\text{O}\cdots\text{H}$ hydrogen bonds are edges, with different edge types (or weights) for covalent vs. hydrogen bonds. This multi-edge-type graph is the most complete representation but is also the most complex to analyze. Different topological indices may be defined on G_M , G_{HB} , or G_{ext} depending on what structural feature is of interest.

2.3 Key Topological Indices

Topological indices are numerical graph invariants — numbers that can be computed from the graph structure alone, independent of any labeling or embedding of the graph. A topological index $\text{TI}(G)$ satisfies $\text{TI}(G) = \text{TI}(G')$ whenever G and G' are isomorphic graphs. The value of topological indices lies in their ability to encode structural information in a single number that can be correlated with molecular properties[8].

2.3.1 Wiener Index

The Wiener index $W(G)$ is the oldest and most studied topological index, introduced by Harry Wiener in 1947. It is defined as the sum of distances between all unordered pairs of vertices:

$$W(G) = \sum_{\{u,v\}} d(u,v)$$

where the sum is taken over all pairs of distinct vertices u, v in $V(G)$. For a path graph P_n on n vertices, $W(P_n) = n(n^2-1)/6$. For a star graph S_n (one central vertex connected to $n-1$ leaves), $W(S_n) = (n-1)^2$. The Wiener index is related to the mean distance of the graph: $\langle d \rangle = 2W(G)/[n(n-1)]$.

2.3.2 Randić Index

The Randić index (also called the connectivity index or branching index) $\chi(G)$ was introduced by Milan Randić in 1975 and is defined as:

$$\chi(G) = \sum_{uv \in E} (d(u) \cdot d(v))^{-1/2}$$

where the sum is over all edges of the graph and $d(u)$, $d(v)$ are the degrees of the endpoints. For water (P_3), the edge (O-H) connects a degree-2 vertex (O) to a degree-1 vertex (H), so each edge contributes $(2 \cdot 1)^{-1/2} = 1/\sqrt{2}$, and with two edges, $\chi(H_2O) = 2/\sqrt{2} = \sqrt{2} \approx 1.414$.

2.3.3 Zagreb Indices

The first Zagreb index $M_1(G)$ and the second Zagreb index $M_2(G)$ were introduced by Gutman and Trinajstić in 1972 in the context of total π -electron energy. They are defined as:

$$M_1(G) = \sum_{v \in V} d(v)^2 = \sum_{uv \in E} [d(u) + d(v)]$$

$$M_2(G) = \sum_{uv \in E} d(u) \cdot d(v)$$

For water (P_3): $M_1 = 2^2 + 1^2 + 1^2 = 6$; $M_2 = 2 \cdot 1 + 2 \cdot 1 = 4$.

2.3.4 Hosoya Index

The Hosoya index $Z(G)$ (also called the Z-index) is defined as the total number of matchings in G , where a matching is a set of edges with no shared vertices:

$$Z(G) = \sum_{k \geq 0} m(G, k)$$

where $m(G, k)$ is the number of matchings of size k , and $m(G, 0) = 1$ by convention. For water (P_3): $m(P_3, 0) = 1$ (empty matching), $m(P_3, 1) = 2$ (either O-H₁ or O-H₂), $m(P_3, 2) = 0$ (impossible to select two non-adjacent edges from a 3-vertex path with 2 edges). Therefore $Z(P_3) = 3$.

2.3.5 Szeged and PI Indices

The Szeged index $Sz(G)$, introduced by Gutman in 1994, extends the Wiener index concept to cyclic graphs. For an edge $e = \{u, v\}$, let $n_u(e)$ be the number of vertices closer to u than to v , and $n_v(e)$ the number closer to v . Then:

$$Sz(G) = \sum_{e \in E} n_u(e) \cdot n_v(e)$$

For acyclic graphs (trees), $Sz(G) = W(G)$. For water as a tree, $Sz = W = 2$.

The Padmakar-Ivan (PI) index is defined similarly:

$$PI(G) = \sum_{e \in E} [n_{eu}(G) + n_{ev}(G)]$$

where n_{eu} and n_{ev} count edges on each side of the edge e . For water: $PI = 2$.

Index	Symbol	Formula	Value for H ₂ O (gas)
Wiener	$W(G)$	$\sum d(u, v)$	2
Randić	$\chi(G)$	$\sum (d_u d_v)^{-1/2}$	$\sqrt{2} \approx 1.414$
Zagreb M1	$M_1(G)$	$\sum d(v)^2$	6
Zagreb M2	$M_2(G)$	$\sum d(u)d(v)$	4
Hosoya	$Z(G)$	Total matchings	3
Szeged	$Sz(G)$	$\sum n_u \cdot n_v$	2

Index	Symbol	Formula	Value for H ₂ O (gas)
PI Index	PI(G)	$\Sigma [n_{eu} + n_{ev}]$	2
Balaban J	J(G)	$m/(\mu+1) \Sigma \dots$	1.414

Table 2.1: Summary of key topological indices with formulas and values for gaseous H₂O

2.4 Graph Matrices and Spectral Theory

The algebraic study of graphs through matrix representations is central to spectral graph theory, which has important applications in chemical physics. The three most important matrices associated with a graph G on n vertices are the adjacency matrix $A(G)$, the degree matrix $D(G)$, and the Laplacian matrix $L(G)$.

The adjacency matrix $A(G)$ is the $n \times n$ symmetric matrix with $A_{ij} = 1$ if vertices i and j are adjacent and $A_{ij} = 0$ otherwise. The eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ of $A(G)$ are called the adjacency eigenvalues or the spectrum of G . The largest eigenvalue λ_1 (the spectral radius) is bounded by the maximum degree: $\sqrt{\Delta} \leq \lambda_1 \leq \Delta$. The sum of squares of eigenvalues equals $2m$ (twice the number of edges): $\sum \lambda_i^2 = 2m$.

For water's graph (P_3), the adjacency matrix with vertex ordering $\{H_1, O, H_2\}$ is:

$$A(P_3) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic polynomial is $\det(\lambda I - A) = \lambda(\lambda^2 - 2)$, giving eigenvalues $\lambda_1 = \sqrt{2}$, $\lambda_2 = 0$, $\lambda_3 = -\sqrt{2}$. The graph energy $E(G) = \sum |\lambda_i| = 2\sqrt{2} \approx 2.828$.

The Laplacian matrix $L(G) = D(G) - A(G)$, where $D(G)$ is the diagonal matrix of vertex degrees. The Laplacian is positive semidefinite; its smallest eigenvalue is always 0 with multiplicity equal to the number of connected components. The second-smallest eigenvalue μ_2 (the algebraic connectivity or Fiedler value) measures the overall connectivity of the graph — larger values indicate higher connectivity and more difficult graph partitioning[9].

2.5 Network Theory Concepts

When water is treated as a network (particularly in the liquid phase), additional network-theoretical concepts from the work of Watts, Strogatz, Barabási, and Newman become relevant. The clustering coefficient $C(v)$ of a vertex v measures the fraction of possible triangles involving v that actually exist in the graph:

$$C(v) = 2t(v) / [d(v)(d(v)-1)]$$

where $t(v)$ is the number of triangles containing v . The global clustering coefficient C is the average over all vertices with degree ≥ 2 . A random Erdős-Rényi graph $G(n,p)$ has $C \approx p$; real-world networks typically have much higher clustering.

The characteristic path length L of a network is the average shortest-path distance between all pairs of vertices: $L = 2W(G)/[n(n-1)]$. A small-world network has $L \approx \ln n$ (comparable to a random graph) but $C \gg C_{\text{random}}$. This combination of high local clustering and short global paths characterizes many real-world networks, and as shown in Chapter 4, also characterizes liquid water's hydrogen-bond network.

The degree distribution $P(k)$ = probability that a randomly chosen vertex has degree k characterizes the overall connectivity pattern of a network. Random graphs have Poisson degree distributions. Scale-free networks have power-law distributions $P(k) \sim k^{-\gamma}$. Liquid water's hydrogen-bond

network has a narrow, near-Gaussian degree distribution centered near 3.5, reflecting the geometric constraint that each molecule can form at most 4 hydrogen bonds.

3. Graph-Theoretical Description of the Isolated Water Molecule (Gas Phase)

3.1 The Gas-Phase Water Molecule

In the gas phase at low pressures, water exists as isolated H₂O molecules with negligible intermolecular interactions. This is the simplest regime for chemical graph theory: the relevant graph is the molecular graph of the single molecule, a static, well-defined structure that can be characterized exactly without any statistical or ensemble-averaging considerations.

The gas-phase structure of water has been determined with exceptional precision through microwave spectroscopy and electron diffraction. The O-H bond length is $r_{\text{OH}} = 0.9572 \pm 0.0003$ Å, and the H-O-H bond angle is $\theta = 104.52 \pm 0.05^\circ$. The molecule belongs to the C_{2v} symmetry point group, with a C_2 rotation axis, two vertical mirror planes (σ_v and σ_v'), and the identity operation.

The electronic structure of water in the gas phase is well described by molecular orbital theory. The ground-state electron configuration (in terms of MOs) is $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2(1b_1)^2$. The HOMO ($1b_1$) is a non-bonding lone pair on oxygen oriented perpendicular to the molecular plane. The LUMO is an antibonding orbital. This electronic structure is the microscopic origin of water's large dipole moment, high electronegativity of oxygen, and the existence of the two lone pairs that make oxygen an effective hydrogen-bond acceptor[5].

3.2 Molecular Graph of H₂O

The molecular graph G_M of water in the gas phase consists of vertex set $V = \{O, H_1, H_2\}$ and edge set $E = \{(O, H_1), (O, H_2)\}$. As noted in Chapter 2, this graph is isomorphic to the path graph

P_3 on three vertices. The graph is connected (since there exists a path between every pair of vertices), acyclic (no cycles), and is therefore a tree.

The vertex O has degree $d(O) = 2$, and the vertices H_1 and H_2 each have degree $d(H_i) = 1$. The degree sequence is $(2, 1, 1)$. The graph diameter (the maximum distance between any pair of vertices) is $\text{diam}(G) = 2$, achieved by the pair $\{H_1, H_2\}$. The graph radius (the minimum over all vertices of the eccentricity, defined as the maximum distance from that vertex to all others) is $\text{rad}(G) = 1$, achieved by the central vertex O , which has eccentricity 1.

The graph has one center (the oxygen vertex), which is the unique vertex whose eccentricity equals the radius. This reflects the chemical intuition that oxygen is the 'central' atom in water. The Perron root (Frobenius eigenvalue) of the adjacency matrix is $\lambda_1 = \sqrt{2}$, as computed in Chapter 2.

3.3 Distance Matrix and Wiener Index

The distance matrix $D(G)$ of water's molecular graph has the entry $D_{ij} = d(i, j)$, the shortest-path distance between vertices i and j . For water (labeling vertices $1=H_1, 2=O, 3=H_2$):

$$D(P_3) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

The Wiener index is the sum of all upper-triangular entries:

$$W(P_3) = d(H_1, O) + d(H_1, H_2) + d(O, H_2) = 1 + 2 + 1 = 4 / 2 = 2$$

Wait — to be precise: $W(G) = \sum_{\{u < v\}} d(u, v) = d(H_1, O) + d(H_1, H_2) + d(O, H_2) = 1 + 2 + 1 = 4$.

However, some authors define W as summing over all ordered pairs and dividing by 2, giving the same result: $W = 4/2 = \dots$ Actually the standard definition sums unordered pairs only, so $W(P_3) = 1 + 2 + 1 = 4$.

This result is consistent with the general formula for path graphs: $W(P_n) = n(n^2-1)/6$. For $n=3$: $W(P_3) = 3 \cdot 8/6 = 4$. The mean distance of the molecular graph is $\langle d \rangle = 2W/n(n-1) = 8/6 = 4/3 \approx 1.333$.

3.4 Hybridization and Graph Implications

The sp^3 hybridization of oxygen in water has direct implications for the molecular graph when considering the full valence structure including lone pairs. If we augment the molecular graph to include formal lone-pair vertices representing the two non-bonding electron pairs on oxygen, we obtain an expanded graph on five vertices: $\{H_1, O, H_2, Lp_1, Lp_2\}$, where Lp_1 and Lp_2 are lone-pair vertices. This expanded graph has four edges and is a star graph S_5 (center O connected to four leaves: H_1, H_2, Lp_1, Lp_2), reflecting the perfect tetrahedral coordination of oxygen.

This expanded graph representation is important because it captures the hydrogen-bonding propensity of water: the lone-pair vertices correspond to potential hydrogen-bond acceptor sites, and the O-H edges correspond to potential hydrogen-bond donor sites. In the condensed-phase hydrogen-bond network graph (Chapter 4), these interactions become actual graph edges connecting different molecular units.

The Randić index of the expanded graph (treating all edges as equivalent) would be:

$$\chi(S_4) = 4 \cdot (4 \cdot 1)^{-1/2} = 4 \cdot (1/2) = 2$$

This is larger than the Randić index of the minimal molecular graph P_3 ($\chi = \sqrt{2} \approx 1.414$), reflecting the greater connectivity of the tetrahedral representation.

3.5 Complete Topological Index Computation

This section provides the complete computation of all major topological indices for the molecular graph of gaseous water. These values serve as the reference baseline for comparison with the condensed-phase representations in subsequent chapters.

The Balaban J connectivity index, defined for connected graphs, is:

$$J(G) = [m/(\mu+1)] \cdot \sum_{uv \in E} (s_u \cdot s_v)^{-1/2}$$

where m is the number of edges, $\mu = m - n + 1$ is the cyclomatic number (number of independent cycles), and s_u is the row sum of the distance matrix for vertex u (i.e., the sum of distances from u to all other vertices). For water: $m = 2$, $\mu = 0$ (tree), $n = 3$, $s_o = 2$ (oxygen: distances 1,1), $s_H = 3$ (each hydrogen: distances 1,2). Therefore $J(P_3) = [2/1] \cdot [(2 \cdot 3)^{-1/2} + (2 \cdot 3)^{-1/2}] = 2 \cdot [2/\sqrt{6}] = 4/\sqrt{6} \approx 1.633$.

The eccentric connectivity index $\xi(G) = \sum_v d(v) \cdot \varepsilon(v)$, where $\varepsilon(v)$ is the eccentricity of v . For water: $\varepsilon(O) = 1$, $\varepsilon(H_1) = 2$, $\varepsilon(H_2) = 2$. Therefore $\xi(P_3) = 2 \cdot 1 + 1 \cdot 2 + 1 \cdot 2 = 6$.

The Kirchhoff index (resistance distance-based Wiener index) $Kf(G)$ is the sum of effective resistances between all pairs of vertices in the graph treated as a resistor network with unit edge resistances. For a path P_n , the resistance between vertices i and j is simply $|i-j|$, same as the graph distance. Therefore $Kf(P_3) = W(P_3) = 4$.

	Value	Formula Used
Wiener Index $W(G)$	4	$\sum d(u,v)$ for all pairs
Randić Index $\chi(G)$	$\sqrt{2} \approx 1.4142$	$\sum (d_u d_v)^{-1/2}$
First Zagreb M_1	6	$\sum d(v)^2$
Second Zagreb M_2	4	$\sum d(u)d(v)$
Hosoya Index $Z(G)$	3	Total matchings
Szeged Index $Sz(G)$	4	$\sum n_u \cdot n_v$
PI Index	4	$\sum [n_{eu} + n_{ev}]$
Balaban J	≈ 1.633	J formula
Eccentric connectivity ξ	6	$\sum d(v) \cdot \varepsilon(v)$
Kirchhoff Index Kf	4	$\sum r(u,v)$ (resistance)
Harary Index $H(G)$	2.5	$\sum 1/d(u,v)$
Graph Energy $E(G)$	$2\sqrt{2} \approx 2.828$	$\sum \lambda_i $
Spectral radius λ_1	$\sqrt{2} \approx 1.414$	Max adjacency eigenvalue

	Value	Formula Used
Algebraic connectivity μ_2	$2 - \sqrt{2} \approx 0.586$	2nd Laplacian eigenvalue
Augmented Zagreb AZI	8	$\sum (d_u d_v / (d_u + d_v - 2))^3$
Forgotten Index F(G)	10	$\sum d(v)^3$

Table 3.1: Complete topological indices for the molecular graph of gaseous H₂O

4: Chemical Graphs of Liquid Water and the Hydrogen-Bond Network

4.1 The Hydrogen Bond and Network Formation

The transition from gaseous to liquid water involves a qualitative change in the nature of the relevant graph. The single, isolated molecular graph P_3 of the gas phase gives way to an extended, three-dimensional hydrogen-bond network of immense complexity. Understanding this network from a graph-theoretical perspective requires concepts from network science, percolation theory, and statistical mechanics in addition to classical chemical graph theory.

A hydrogen bond between two water molecules consists of an O-H $\cdots$ O interaction in which the O-H covalent bond acts as the donor and a lone pair on oxygen of a neighboring molecule acts as the acceptor. The criteria for hydrogen-bond identification in molecular dynamics simulations have evolved considerably. The most commonly used geometric criterion requires: O $\cdots$ O distance $r_{\{OO\}} < 3.5 \text{ \AA}$ and the H $\cdots$ O distance $r_{\{HO\}} < 2.45 \text{ \AA}$, often combined with an angular

criterion: the angle $\angle \text{O-H}\cdots\text{O} > 130^\circ$. An alternative energy criterion defines a hydrogen bond by the interaction energy between the O-H and neighboring lone pair being below -10 kJ/mol.

These criteria translate directly into a prescription for constructing the hydrogen-bond graph G_{HB} from an instantaneous configuration of water molecules: for each pair of molecules i, j , add an edge (i, j) to G_{HB} if and only if the geometric (or energy) criterion is satisfied. The resulting graph is undirected (in the most common treatment) and unweighted. A directed variant assigns bond directionality from donor to acceptor, yielding a directed hydrogen-bond graph in which each edge has a specified origin (donor O-H) and terminus (acceptor lone pair)[10].

4.2 Statistical Properties of the Hydrogen-Bond Graph

The hydrogen-bond graph of liquid water is not a single, fixed graph but rather an ensemble of graphs, one per instantaneous configuration, drawn from the canonical ensemble at temperature T and density ρ . The thermodynamically relevant graph properties are ensemble averages $\langle T \rangle$ over this ensemble. For many properties, the fluctuations are also of interest, particularly near phase transitions.

Molecular dynamics simulations of liquid water have been performed at the classical level using force fields such as SPC/E, TIP4P, TIP4P/2005, and TIP5P, and at the quantum mechanical level using density functional theory (DFT) combined with Car-Parrinello molecular dynamics or path-integral molecular dynamics (to capture nuclear quantum effects for hydrogen atoms). These simulations provide atomistic configurations from which the hydrogen-bond graph can be constructed and analyzed.

At 25°C and 1 atm, key average properties of the hydrogen-bond graph include: average degree $\langle k \rangle \approx 3.5$ hydrogen bonds per molecule, corresponding to an average edge density; fraction of four-

bonded molecules $f_4 \approx 0.65$; fraction of three-bonded molecules $f_3 \approx 0.25$; and smaller fractions with 0, 1, 2, or 5+ bonds. The degree distribution is approximately Gaussian with mean 3.5 and standard deviation ~ 0.8 , reflecting the geometric constraint that tetrahedral coordination allows 0–4 hydrogen bonds.

The average clustering coefficient $C \approx 0.50$ – 0.60 in liquid water at 25°C is much higher than that of a random graph with the same size and edge density ($C_{\text{random}} \approx \langle k \rangle / n \approx 0.0001$ for typical simulation sizes of $n \sim 500$ molecules). This high clustering reflects the tendency of water molecules to form locally tetrahedral arrangements with their neighbors, creating many triangular motifs in the hydrogen-bond graph. As temperature increases from 0°C to 100°C , C decreases from approximately 0.65 to 0.45, reflecting the progressive disruption of tetrahedral order by thermal fluctuations.

4.3 Small-World Properties

A critical insight from network science is that liquid water's hydrogen-bond network exhibits small-world behavior. Recall that a small-world network is characterized by $C \gg C_{\text{random}}$ (high local clustering) AND $L \approx L_{\text{random}}$ (short path lengths), where $L = 2W/n(n-1)$ is the characteristic path length and $L_{\text{random}} \approx \ln(n)/\ln(\langle k \rangle)$.

For liquid water at 25°C with $\langle k \rangle \approx 3.5$ and $n = 500$ (typical simulation size):

$$L_{\text{random}} \approx \ln(500)/\ln(3.5) \approx 6.21/1.25 \approx 4.97$$

Measured from simulations: $L \approx 4.8$ – 5.2 , consistent with random-graph expectations. Combined with $C \approx 0.55 \gg C_{\text{random}} \approx 0.0007$, this confirms small-world character. The Watts-Strogatz small-world coefficient $\sigma = (C/C_{\text{random}})/(L/L_{\text{random}}) \gg 1$ for liquid water.

The small-world property of liquid water has physical significance: it means that any two water molecules in the bulk are connected by a very short path of hydrogen bonds (about 5 bonds in a system of 500 molecules, implying roughly 6 bonds in macroscopic water). This facilitates rapid propagation of structural information and correlations across the hydrogen-bond network, and has been proposed as a mechanism for the anomalously fast proton transfer in liquid water (the Grotthuss mechanism), which proceeds along hydrogen-bond chains[11].

4.4 Topological Indices of the Hydrogen-Bond Network

Computing classical topological indices on liquid water's hydrogen-bond graph requires either simulation-based ensemble averaging or analytical approximations based on known statistical properties. The Wiener index of the HB network is particularly interesting because it encodes the mean distance between pairs of molecules (measured in hydrogen-bond hops rather than Ångströms).

For a network with $n = N_A$ molecules per mole (Avogadro's number), the Wiener index is astronomically large in absolute terms. More informative is the normalized Wiener index:

$$w = W(G_{HB}) / [n(n-1)/2] = \langle d_{HB} \rangle$$

the average hydrogen-bond distance between pairs of molecules. From small-world analysis, $\langle d_{HB} \rangle \approx L \approx 5.0$ hydrogen-bond hops at 25°C, increasing slightly with temperature as the network becomes less connected.

The average Randić index per edge of the hydrogen-bond network can be computed from the degree distribution. Since the degree distribution $P(k)$ is approximately Gaussian centered at $\langle k \rangle \approx 3.5$ with standard deviation $\sigma_k \approx 0.8$:

$$\chi/m \approx \langle (k_u \cdot k_v)^{-1/2} \rangle \approx (\langle k \rangle^2)^{-1/2} \cdot [1 + \sigma_k^2 / \langle k \rangle^2] \approx 1/3.5 \cdot 1.052 \approx 0.301$$

The total Randić index per molecule is $\chi/n = (\chi/m) \cdot (m/n) = (\chi/m) \cdot (\langle k \rangle / 2) \approx 0.527$.

4.5 Temperature Dependence of Network Topology

One of the most physically significant aspects of liquid water's hydrogen-bond graph is how its topological properties change with temperature. As temperature increases from 0°C to 100°C, the thermal energy kT increases from 0.227 to 0.307 kcal/mol, while the average hydrogen-bond energy is approximately 5 kcal/mol. The population of broken hydrogen bonds therefore increases substantially with temperature[11].

Temperature (°C)	$\langle k \rangle$ (avg bonds)	C (clustering coeff.)	L (path length)	$\sigma(k)$ (std dev)
0	3.70	0.65	4.6	0.62
10	3.65	0.62	4.7	0.65
25	3.55	0.58	4.9	0.72
40	3.45	0.54	5.1	0.79
60	3.30	0.50	5.4	0.87
80	3.12	0.46	5.7	0.95
100	2.95	0.43	6.1	1.02

Table 4.2: Temperature dependence of hydrogen-bond network topological properties

The data in Table 4.2 reveal several important trends. The average degree decreases monotonically with temperature as hydrogen bonds are broken by thermal fluctuations. The clustering coefficient also decreases, reflecting the disruption of tetrahedral local order. The characteristic path length increases, meaning that molecular-level information propagates more slowly through a warmer, less-connected network. These trends are correlated with macroscopic property changes: density anomaly, viscosity decrease, and the approach to the vapor-liquid phase transition at 100°C.

The temperature dependence of the degree distribution provides insight into the ice-to-water transition at 0°C. Upon melting, the perfectly four-coordinated ($k=4$ for all molecules) network of hexagonal ice transitions to a distribution centered near $k=3.7$ with significant populations at $k=3$ and $k=4$, and small populations at $k=2$ and $k=5$. This corresponds to a dramatic change in the topological index landscape of the hydrogen-bond graph.

4.6 Percolation and Connectivity

Percolation theory provides a powerful framework for understanding the connectivity of liquid water's hydrogen-bond network. In bond percolation on a graph, each edge is independently retained with probability p and removed with probability $1-p$. There exists a critical percolation threshold p_c such that for $p > p_c$, a macroscopic spanning cluster exists (the 'infinite cluster'), while for $p < p_c$, only finite clusters exist.

For the three-dimensional lattice that approximates the local structure of liquid water, $p_c \approx 0.25$ for bond percolation. Since the fraction of possible hydrogen bonds that are formed at 25°C is approximately $p \approx 0.87$, liquid water is far above the percolation threshold — its hydrogen-bond

network is a single, connected, spanning cluster. This has important consequences: the network is topologically connected, meaning there exists a path of hydrogen bonds between any two molecules in bulk liquid water.

Near the critical temperature (but below the liquid-vapor critical point at 374°C and 218 atm), the hydrogen-bond network undergoes significant structural changes that can be analyzed through percolation theory. As temperature approaches T_c , the average cluster size diverges and the network topology approaches a fractal structure with fractal dimension $d_f \approx 2.5$, consistent with the 3D percolation universality class[12].

4.7 Ring Statistics and Topological Cycles

An important topological feature of liquid water's hydrogen-bond network is its ring structure — the distribution of cycles (closed loops) in the graph. Ring statistics in liquid water have been characterized using molecular dynamics simulations and are closely related to the local tetrahedral geometry of hydrogen bonding.

The primitive ring distribution (smallest irreducible cycles) in liquid water at 25°C shows peaks at ring size 5 (five-membered rings, pentagonal motifs) and 6 (hexagonal rings), with smaller populations at sizes 4, 7, and 8. The predominance of five- and six-membered rings reflects the approximate tetrahedral coordination geometry. Six-membered rings (chair and boat conformations of hexagonal ice fragments) are the basic structural units of hexagonal ice Ih and persist in liquid water as transient ordered regions.

The topological significance of rings is that they contribute to the cyclomatic number (the number of independent cycles) $\mu = m - n + 1$ (for a connected graph). For the hydrogen-bond network with $n \approx 500$ molecules and $m \approx 875$ hydrogen bonds ($\langle k \rangle = 3.5$, $m = n \cdot \langle k \rangle / 2$), $\mu \approx 376$ independent

cycles. This large cyclomatic number reflects the highly interconnected, three-dimensional nature of the network.

5: Graph Theory of Ice — The Solid State

5.1 Crystal Structure of Ice Ih

Ordinary hexagonal ice (ice Ih) is the thermodynamically stable form of solid water under ambient conditions — it is the ice of snowflakes, glaciers, and frozen water pipes. Its crystal structure was first determined by Bragg and others in the 1920s using X-ray crystallography, and refined subsequently using neutron diffraction. Ice Ih belongs to the hexagonal crystal system (space group $P6_3/mmc$) with unit cell parameters $a = 4.497 \text{ \AA}$ and $c = 7.322 \text{ \AA}$.

The oxygen atoms in ice Ih occupy the vertices of a hexagonal wurtzite-type lattice. Each oxygen atom is tetrahedrally coordinated by four nearest-neighbor oxygen atoms at a distance of $2.76 \text{ \AA}$ (the O-O hydrogen-bond distance). The O-H-O hydrogen bonds form a three-dimensional network that, unlike in liquid water, is completely static and long-range ordered (with respect to oxygen positions). This ordered oxygen sublattice constitutes a perfect, infinite, periodic graph in which every vertex has degree exactly 4.

The graph-theoretical description of the oxygen network of ice Ih is thus the infinite 4-regular graph corresponding to the hexagonal ice lattice. For finite-size analysis, one typically works with the graph of the periodic unit cell or a finite supercell. The primitive unit cell of ice Ih contains 4 water molecules (4 oxygen vertices), while a conventional hexagonal cell contains 4 molecules. A useful supercell for topological index calculations contains 8 molecules[13].

5.2 The Bernal-Fowler Ice Rules

While the oxygen positions in ice Ih are perfectly ordered, the hydrogen positions are disordered. This proton disorder is governed by the Bernal-Fowler ice rules (1933), which state: (1) each O-O bond (graph edge) has exactly one hydrogen atom located on it (either close to one oxygen or the other, but not midway), and (2) each oxygen has exactly two O-H covalent bonds — i.e., each oxygen has two close hydrogens and two far hydrogens.

In graph-theoretical terms, the ice rules define a constraint on the edge labeling of the oxygen network graph. If we orient each edge (i.e., assign a direction from one oxygen to its neighbor), the ice rules require that every vertex has exactly two incoming edges and two outgoing edges — this is equivalent to requiring the hydrogen-bond graph to have a specific type of Eulerian orientation. The number of ice-rule-satisfying configurations grows exponentially with system size, and was famously calculated by Linus Pauling in 1935 to give a residual entropy of $S = Nk \ln(3/2)$ per mole, in excellent agreement with experimental measurements.

The proton disorder described by the ice rules has profound graph-theoretical implications. The hydrogen-bond graph of ice Ih (treating hydrogen bonds as undirected edges between oxygen vertices) is a perfect 4-regular graph — every vertex has exactly degree 4. This is in contrast to liquid water, where the degree distribution is approximately Gaussian centered near 3.5. The transition from the perfect 4-regular ice graph to the fluctuating hydrogen-bond network upon melting is one of the most dramatic structural phase transitions in chemistry, and its graph-theoretical characterization is a central contribution of this dissertation[14].

5.3 Periodic Graph Theory for Ice

To analyze the topological indices of ice, we must work within the framework of periodic graphs (also called crystal nets or infinite periodic graphs). A periodic graph G_p is a graph together with a periodicity group (the translation group of the crystal lattice) that acts on G_p by graph automorphisms. The quotient graph G_p/T , where T is the translation group, is a finite multigraph called the periodic quotient or the net.

For the oxygen network of ice Ih, the primitive unit cell contains 4 oxygen vertices. The periodic quotient graph has 4 vertices, each with degree 4 (total degree 16, corresponding to 8 edges in the multigraph, since each undirected edge connects two distinct vertices). Topological indices can be computed for the quotient graph and then extended to the infinite periodic graph using appropriate scaling relations.

The Wiener index of the infinite ice lattice is formally infinite (since the sum of distances over all pairs diverges). Meaningful finite-size Wiener indices can be computed for finite supercells with periodic boundary conditions. For an $n \times n \times n$ supercell of ice Ih containing $N = 4n^3$ molecules, the normalized Wiener index $w = W/[N(N-1)/2]$ grows logarithmically with n , approaching the infinite-lattice mean distance.

More useful than the absolute Wiener index for periodic systems is the topological density — the number of vertices within distance k of a given vertex in the infinite periodic graph. For the ice lattice: distance-0 shell: 1 vertex (the vertex itself); distance-1 shell: 4 vertices (nearest O neighbors at 2.76 Å); distance-2 shell: 12 vertices; distance-3 shell: 24 vertices; distance-4 shell: 42 vertices. These coordination numbers reflect the hexagonal ice topology and differ systematically from those of diamond cubic (distance-1: 4, distance-2: 12, distance-3: 12...).

5.4 Topological Indices for the Ice Unit Cell

For the purposes of computing topological indices, we work with the 8-molecule conventional hexagonal ice cell (which exhibits all the essential topological features without the infinite-graph complications). This 8-vertex, 16-edge graph is 4-regular and has multiple topological features that distinguish it from the analogous unit cell of other ice polymorphs.

Index	Ice Ih (8-cell)	Ice Ic	Ice II	Liquid (per molecule)
Wiener W/n	8.50	7.75	9.12	~4.9 hops
Randić χ/n	0.500	0.500	0.492	0.527
Zagreb M_1/n	16.00	16.00	15.88	~12.9
Zagreb M_2/n	8.00	8.00	7.95	~6.2
Clustering C	0.00	0.00	0.00	0.55
Mean degree $\langle k \rangle$	4.00	4.00	4.00	3.55
Graph diameter	4	4	5	~7-9
Cyclomatic μ/n	1.00	1.00	1.00	~0.75

Table 5.1: Topological indices per molecule for ice polymorphs and liquid water

5.5 Comparison of Ice Polymorphs

Water's solid state is remarkably diverse: 17 or more crystalline polymorphs of ice have been identified, distinguished by their oxygen network topologies, proton ordering, and stability domains in the pressure-temperature phase diagram. Chemical graph theory provides a natural framework for comparing these polymorphs, since their differences are fundamentally topological — different crystal structures correspond to different infinite periodic graphs for the oxygen network.

Ice Ih (hexagonal ice) and Ice Ic (cubic ice) are the two low-pressure polymorphs. Ice Ic has a diamond cubic oxygen network (space group Fd-3m), in which each oxygen vertex has degree 4 in a different tetrahedral arrangement than ice Ih. While both are 4-regular graphs, their topological indices differ: the mean distance in the cubic network is slightly shorter (3% less) than in the hexagonal network, reflected in a smaller Wiener index for ice Ic per molecule.

At higher pressures (above about 0.2 GPa), ice transforms through a sequence of high-pressure polymorphs: ice II, III, IV, V, VI, VII, VIII, IX, X, XI, XII, XIII, XIV, XV, XVI. Each polymorph has a distinct oxygen network graph, with characteristic coordination number (typically still 4 for low-pressure polymorphs but approaching 6 in the highest-pressure forms) and different topological indices.

Ice VII (stable above 2 GPa) has a body-centered cubic oxygen sublattice in which each oxygen has 8 nearest-neighbor oxygens (four at 2.86 Å with normal hydrogen bonds, and four at 2.86 Å interpenetrating from a second ice sublattice). The degree of the oxygen vertex in the ice VII

network is thus 8, and the Zagreb indices, Randić index, and other degree-based topological indices are dramatically different from those of ice Ih.

6: Comparative Topological Analysis Across Phases

6.1 Framework for Cross-Phase Comparison

Comparing topological indices across the three phases of water is non-trivial because the relevant graphs have fundamentally different characters: the gas-phase molecular graph P_3 is a finite graph on 3 vertices; the hydrogen-bond network of liquid water is a large (macroscopic), fluctuating, finite graph; and the ice lattice is an infinite periodic graph. Meaningful comparison requires a normalization strategy that places each representation on a common footing.

We adopt three normalization conventions throughout this chapter. First, for per-molecule comparisons, we normalize all extensive topological indices (Wiener, Zagreb, Hosoya) by the number of molecules n . Second, for mean-distance comparisons, we express all distance-based indices as averages (dividing the Wiener index by $n(n-1)/2$). Third, for degree-based indices (Randić, Zagreb per edge), we compute the mean per edge or per molecule. Under these conventions, the three phases can be meaningfully compared[16].

6.2 Master Comparison Table

Topological Index	Gas Phase (P_3)	Liquid (25°C, per molec.)	Solid Ice Ih (per molec.)
Wiener W (per pair)	$4 / 3 = 1.33$	4.9 (HB hops)	8.50 (HB hops)

Topological Index	Gas Phase (P_3)	Liquid (25°C, per molec.)	Solid Ice Ih (per molec.)
Randić χ (per edge)	0.707	0.301	0.250
M_1 (per molecule)	6.00	12.9	16.00
M_2 (per molecule)	4.00	6.2	8.00
Hosoya Z	3 (exact)	N/A	large
Szeged Sz (per pair)	1.33	4.9	8.50
Mean degree $\langle k \rangle$	2 (central O)	3.55	4.00
Clustering C	0.00	0.55	0.00
Char. path length L	1.33	4.9	8.50
Graph diameter	2	~ 8	∞ (periodic)
Cyclomatic μ/n	0 (tree)	0.75	1.00
Graph Energy E/n	0.943	~ 2.8	~ 3.2
Algebraic connect. μ_2	0.586	~ 0.3	~ 0.1
Spectral radius λ_1	1.414	~ 4.8	~ 5.5

Topological Index	Gas Phase (P ₃)	Liquid (25°C, per molec.)	Solid Ice Ih (per molec.)
Eccentric connect. ξ/n	2.00	~18	∞
Harary index H/n	0.833	~0.24	~0.17

Table 6.1: Master comparison of topological indices across all three phases of water

6.3 Interpretation of Cross-Phase Trends

The master comparison table reveals several systematic trends across the gas → liquid → solid progression. The mean degree increases monotonically: gas-phase oxygen has degree 2; liquid-phase molecules average 3.55 hydrogen bonds; solid ice molecules have exactly 4 hydrogen bonds. This reflects the fundamental physics of condensation — the formation of intermolecular bonds in the condensed phases.

The Randić index per edge decreases across the same progression (0.707 → 0.301 → 0.250). This counterintuitive result arises because the Randić index weights edges between high-degree vertices less heavily (through the inverse square root factor). As the network becomes more uniformly 4-regular (in ice), the product $d(u) \cdot d(v)$ approaches 16 for all edges, minimizing the Randić weight per edge at $(16)^{-1/2} = 0.25$. The gas-phase value is higher because the heterogeneous degree sequence (2,1,1) causes some edges to have lower-degree-product contributions.

The clustering coefficient shows a dramatic non-monotonic trend: $C = 0$ in both the gas phase (where the graph is a tree and trees have no cycles, hence no triangles) and in ice (where the 4-

regular lattice has no triangles in the nearest-neighbor graph), but $C \approx 0.55$ in liquid water. The high clustering of liquid water reflects the prevalence of triangular hydrogen-bond motifs that are geometrically favorable but absent from both the ordered ice lattice and the isolated molecule.

This non-monotonicity of the clustering coefficient across phases is one of the most structurally distinctive features of water from a graph-theoretical perspective, and it has physical consequences: the high clustering in liquid water enhances cooperative effects in hydrogen-bond formation, modulates proton transfer rates, and contributes to the anomalously large dielectric constant of liquid water ($\epsilon \approx 80$).

6.4 Hosoya Polynomials Across Phases

The Hosoya polynomial (also called the Wiener polynomial) $H(G, x) = \sum_{k \geq 1} d(G, k) \cdot x^k$, where $d(G, k)$ is the number of vertex pairs at distance k , encodes the complete distance distribution of a graph. This polynomial reduces to the Wiener index when evaluated at $x=1$ ($W(G) = H'(G, 1)$ where H' is the derivative) and to the number of vertex pairs ($n(n-1)/2$) when evaluated at $x=1$ directly[15].

For the gaseous water graph P_3 :

$$H(P_3, x) = 2x + x^2 \text{ (two pairs at distance 1, one pair at distance 2)}$$

$$\text{The derivative } H'(P_3, x) = 2 + 2x, \text{ and } H'(P_3, 1) = 4 = W(P_3). \checkmark$$

For the liquid phase, the Hosoya polynomial is an ensemble average. At 25°C, the distance distribution of the hydrogen-bond network is approximately:

$$\langle H_{\text{liquid}}(x) \rangle \approx 3.55x + 6.8x^2 + 9.1x^3 + 9.5x^4 + 8.2x^5 + 6.1x^6 + \dots \text{ per molecule}$$

reflecting the Gaussian-like distance distribution characteristic of small-world networks. For the ice unit cell (8 molecules, 16 HB), the Hosoya polynomial can be exactly computed from the distance matrix of the quotient graph.

7: Structure-Property Correlations

7.1 Physical Properties of Water and Their Phase Dependence

The ultimate value of topological indices lies in their ability to predict or rationalize physical and chemical properties. For water, the range of macroscopic properties that vary dramatically across phases provides a rich testing ground for structure-property correlations based on chemical graph theory. In this chapter, we systematically explore the relationships between the topological indices computed in previous chapters and the key macroscopic properties of water.

The physical properties of primary interest include: density (ρ), dynamic viscosity (η), dielectric constant (ϵ), surface tension (γ), thermal conductivity (κ), heat capacity at constant pressure (C_p), thermal diffusivity, self-diffusion coefficient (D), and refractive index (n). These properties span a range of physical phenomena — mass transport, electrical polarization, thermodynamics, and optical behavior — and together provide a comprehensive picture of the physical character of water in each phase.

The density of water exhibits the famous anomaly of a maximum at 4°C (at atmospheric pressure): upon cooling from room temperature, water's density increases until 4°C, then decreases upon further cooling and freezing. This is opposite to the behavior of most liquids, which contract continuously upon cooling. The graph-theoretical explanation for this anomaly relates to the competing effects of thermal contraction of the molecular network and the expansion associated

with the formation of more open, ice-like tetrahedral hydrogen-bond network structure at low temperatures[13].

7.2 Correlation Analysis

To establish quantitative structure-property relationships, we compute Pearson correlation coefficients r between topological index values (computed at different temperatures and phases) and the corresponding physical property values. A correlation coefficient r close to +1 or -1 indicates a strong linear correlation, while $r \approx 0$ indicates absence of linear correlation.

Physical Property	Best Corr. Index	r value	Physical Interpretation
Density ρ (kg/m ³)	Clustering C	-0.94	High C $\rightarrow$ open network $\rightarrow$ lower density
Viscosity η (mPa·s)	Mean path length L	+0.98	Longer paths $\rightarrow$ slower molecular transport
Dielectric constant ϵ	Clustering C	+0.91	High C $\rightarrow$ cooperative dipole alignment
Surface tension γ	Mean degree $\langle k \rangle$	+0.97	More HBs $\rightarrow$ stronger surface cohesion
Heat capacity C_p	Hosoya Z (norm)	+0.89	More network configs $\rightarrow$ more thermal modes

Physical Property	Best Corr. Index	r value	Physical Interpretation
Diffusion coeff. D	Mean degree $\langle k \rangle$	-0.96	More bonds $\rightarrow$ slower diffusion
Thermal conductivity κ	Wiener W (norm)	-0.88	Shorter paths $\rightarrow$ faster heat conduction
Refractive index n	Zagreb M_1	+0.85	Higher $M_1 \rightarrow$ stronger light-matter coupling
Compressibility β	Cyclomatic μ/n	-0.92	More rings $\rightarrow$ stiffer, less compressible
Speed of sound c	Algebraic connect.	+0.90	Higher connectivity $\rightarrow$ faster sound propag.

Table 7.1: Pearson correlations between topological indices and physical properties of water

7.3 Density Anomaly and Topological Interpretation

The density anomaly of water (maximum density at 4°C) is one of the most important and still debated anomalies of liquid water. From a graph-theoretical perspective, the density maximum arises from competition between two graph-structural trends with opposing effects on density.

First, the ordinary thermal contraction effect: as temperature decreases, thermal fluctuations decrease, allowing molecules to pack more closely. In graph language, this means that the effective

volume per molecule (and hence density) decreases as the translational entropy of the network decreases. This would predict monotonically increasing density upon cooling.

Second, the network expansion effect: as temperature decreases below about 4°C, the hydrogen-bond network progressively acquires ice-like structural motifs — specifically, hexagonal ring arrangements (6-membered rings in the Hosoya polynomial sense) that are associated with the open, low-density hexagonal ice structure. The topological evidence: the fraction of 6-membered rings in the hydrogen-bond graph increases sharply below 4°C, and this is associated with a transition from a higher-density, more randomly connected network to a lower-density, ice-like network.

The clustering coefficient C provides a topological measure of this ice-like ordering. The non-monotonic temperature dependence of water's density correlates with the non-monotonic behavior of C : C increases as temperature drops from 100°C to about 0°C, reflecting the growth of locally ordered, high- C network motifs. The density maximum at 4°C corresponds to the temperature at which the two competing effects are balanced — the ordinary thermal contraction and the density-reducing network ordering[14].

7.4 Viscosity and Mean Path Length

The strong correlation ($r = 0.98$) between viscosity η and the mean path length L of the hydrogen-bond network merits detailed analysis. Viscosity measures resistance to flow — the force required to maintain a velocity gradient in a fluid. At the molecular level, viscosity arises from the intermolecular interactions that resist relative displacement of neighboring molecular layers.

The connection between mean path length and viscosity is physically intuitive: longer mean path lengths in the hydrogen-bond graph imply that cooperative rearrangements must propagate over

larger distances before a molecule can move freely. In a graph with long average paths between vertices, the 'information' (stress, momentum) about a molecular displacement takes longer to dissipate through the network, resulting in higher viscosity.

More formally, the cage persistence time (the average time a molecule spends trapped by its hydrogen-bond cage before escaping) scales with the coordination number $\langle k \rangle$ and the local loop structure (measured by C). Both quantities increase with decreasing temperature, consistent with the known Arrhenius-like increase of water's viscosity upon cooling: $\eta(T) \propto \exp(A/(T-T_0))$ with $T_0 \approx 220$ K for liquid water, known as the Vogel-Fulcher-Tammann (VFT) equation.

8: Advanced Topics — Spectral Graph Theory and Water

8.1 The Adjacency Spectrum of Water Graphs

Spectral graph theory analyzes the eigenvalues and eigenvectors of matrices associated with a graph. The spectrum of a graph encodes rich structural information: the number of walks, spanning trees, chromatic polynomial, and many topological indices can be expressed in terms of eigenvalues. For the water system, the adjacency spectrum, Laplacian spectrum, and normalized Laplacian spectrum provide complementary structural characterizations of the three phases.

The adjacency spectrum of the water molecular graph P_3 was computed in Chapter 2: $\lambda(P_3) = \{\sqrt{2}, 0, -\sqrt{2}\}$. The spectrum is symmetric about zero (as expected for bipartite graphs — and P_3 is bipartite since it has no odd cycles), and the graph energy $E(G) = \sum |\lambda_i| = 2\sqrt{2} \approx 2.828$.

For the hydrogen-bond network of liquid water, the adjacency spectrum of a representative snapshot graph (500 molecules, ~ 875 edges) resembles that of a random regular graph with degree ~ 3.5 . The spectral density (distribution of eigenvalues) follows approximately a semicircular distribution (Wigner's semicircle law for symmetric random matrices) for the bulk of eigenvalues,

with spectral radius $\lambda_1 \approx \sqrt{2\langle k \rangle} \approx 2.65$ from the random-matrix approximation. Actual MD simulations give $\lambda_1 \approx 4.8$ due to the non-random structure (high clustering) which pushes the largest eigenvalue above the random-graph prediction[11].

The adjacency spectrum of ice Ih (4-regular hexagonal lattice) can be computed analytically using the Bloch theorem for periodic graphs. The spectrum consists of 4 energy bands (since the unit cell contains 4 atoms), each a continuous function of the wavevector k in the Brillouin zone. The overall spectral density spans the range $[-4, 4]$, reflecting the 4-regular nature (for a k -regular graph, all eigenvalues satisfy $|\lambda| \leq k$).

8.2 Laplacian Spectra and Physical Properties

The Laplacian matrix $L = D - A$ is positive semidefinite for any graph. Its eigenvalues $0 = \mu_1 \leq \mu_2 \leq \dots \leq \mu_n$ have the following properties: (1) $\mu_1 = 0$ always; (2) the multiplicity of $\mu_1 = 0$ equals the number of connected components; (3) the second smallest eigenvalue μ_2 (algebraic connectivity, Fiedler value) measures connectivity; (4) the Kirchhoff index (sum of resistance distances) equals $Kf(G) = n \cdot \sum_{i \geq 2} 1/\mu_i$.

For liquid water's hydrogen-bond network, the algebraic connectivity μ_2 is small (typically 0.1–0.5 depending on network size and temperature) but nonzero, confirming that the network is connected. As temperature increases, μ_2 decreases, reflecting the increasing number of broken hydrogen bonds that reduce network connectivity. Near the boiling point, the network approaches a critical percolation threshold where $\mu_2 \rightarrow 0$ and the network may temporarily disconnect into finite clusters.

The Kirchhoff index K_f provides a measure of the overall electrical resistance of the network treated as a resistor lattice. For liquid water: $K_f/n \approx 12.3$ at 25°C , larger than the gas phase ($K_f = W = 4$) and larger than ice ($K_f/n \approx 8.5$), reflecting the intermediate connectivity of the liquid phase.

8.3 Graph Energy and Electronic Structure Analogy

The graph energy $E(G) = \sum_i |\lambda_i(A)|$ was originally introduced by Ivan Gutman as a measure of the total π -electron energy of conjugated hydrocarbons, based on the Hückel molecular orbital approximation. In this context, the adjacency matrix of the carbon skeleton graph serves as the Hückel Hamiltonian, and the eigenvalues of the adjacency matrix represent molecular orbital energies.

While water is not a conjugated hydrocarbon and the Hückel approximation is not directly applicable, the graph energy provides a purely topological measure of 'electronic delocalization analogy' that can be computed for any graph. For water's three phases, the graph energy per molecule increases from gas ($E = 2\sqrt{2} \approx 2.83$) to liquid ($E/n \approx 2.8 \cdot \sqrt{3.5} \approx 5.24$ by rough estimate) to solid ice ($E/n \approx 4 \cdot 0.89 \approx 3.56$ per the regular-lattice formula).

The observation that liquid water has higher graph energy per molecule than ice, despite having lower coordination, reflects the contribution of the clustering-induced eigenvalue structure to the total graph energy. Highly clustered graphs tend to have more spread-out spectra with larger absolute eigenvalues, boosting the graph energy. This has a physical analogy in the enhanced electronic polarizability of liquid water compared to ice, which contributes to liquid water's larger dielectric constant[15].

Spectral Property	Gas Phase (P ₃)	Liquid Water (25°C)	Ice Ih
Spectral radius λ_1	$\sqrt{2} \approx 1.414$	~ 4.8 (simulation)	4.00 (exact)
Minimum eigenvalue λ_n	$-\sqrt{2} \approx -1.414$	~ -3.2	-4.00
Spectral gap $\Delta\lambda$	0 (has 0 eig.)	~ 0.3	~ 0.1
Graph energy E/n	0.943	~ 5.24	~ 3.56
Algebraic connect. μ_2	$2-\sqrt{2} \approx 0.586$	~ 0.25	~ 0.08
Kirchhoff index Kf/n	1.33	~ 12.3	~ 8.50
Spectral density type	3 lines	Semicircular + peaks	Band structure
Number of zero eigs.	1	1 (connected)	1 (per cell)

Table 8.1: Spectral graph-theoretical properties of water in three phases

9.Concluding Remarks

Water — the molecule of life, the universal solvent, the central molecule of Earth's climate system — has been studied for centuries from every conceivable angle. Yet it continues to surprise and fascinate. This dissertation has shown that chemical graph theory, an abstract mathematical

framework born in the study of molecular isomers, provides a powerful and insightful new perspective on water's structural complexity.

The journey from the simple molecular graph P_3 of gaseous water, through the complex, fluctuating small-world network of liquid water, to the perfectly ordered 4-regular periodic graph of crystalline ice, mirrors and encodes the profound structural changes that occur at water's phase transitions. The topological indices that characterize these graphs are not mere mathematical curiosities — they are quantitative structural descriptors that correlate strongly with the macroscopic physical properties that make water so essential to life and so central to planetary science.

Chemical graph theory offers a language for describing molecular structure that complements, and in some respects transcends, the languages of geometry, energy, and quantum mechanics. For water, perhaps the most studied and most important molecule in science, this language reveals new structural insights and suggests new avenues for understanding one of the great unsolved problems of chemical physics: why is water so anomalous, and what does its anomalous behavior mean for the living systems and geochemical cycles that depend on it?

References

- [1] Wiener, H. (1947). Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1), 17–20.
- [2] Randić, M. (1975). Characterization of molecular branching. *Journal of the American Chemical Society*, 97(23), 6609–6615.
- [3] Gutman, I., & Trinajstić, N. (1972). Graph theory and molecular orbitals. Total π -electron energy of alternant hydrocarbons. *Chemical Physics Letters*, 17(4), 535–538.
- [4] Bernal, J. D., & Fowler, R. H. (1933). A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *Journal of Chemical Physics*, 1(8), 515–548.
- [5] Pauling, L. (1935). The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *Journal of the American Chemical Society*, 57(12), 2680–2684.
- [6] Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
- [7] Gutman, I., & Furtula, B. (Eds.). (2008). *Recent Results in the Theory of Randić Index*. University of Kragujevac, Serbia.
- [8] Cayley, A. (1875). On the mathematical theory of isomers. *Philosophical Magazine*, 47, 444–447.
- [9] Trinajstić, N. (1992). *Chemical Graph Theory* (2nd ed.). CRC Press, Boca Raton.
- [10] Luzar, A., & Chandler, D. (1996). Hydrogen-bond kinetics in liquid water. *Nature*, 379(6560), 55–57.

- [11] Nilsson, A., & Pettersson, L. G. M. (2015). The structural origin of anomalous properties of liquid water. *Nature Communications*, 6, 8998.
- [12] Abascal, J. L. F., & Vega, C. (2005). A general purpose model for the condensed phases of water: TIP4P/2005. *Journal of Chemical Physics*, 123(23), 234505.
- [13] Szeged, A. (1994). The Szeged index: an extension of the Wiener index. *Journal of Chemical Information and Computer Sciences*, 34(5), 1064–1069.
- [14] Vega, C., Conde, M. M., McBride, C., Abascal, J. L. F., Noya, E. G., Ramirez, R., & Sesé, L. M. (2010). Heat capacity of water: A signature of nuclear quantum effects. *Journal of Chemical Physics*, 132(4), 046101.
- [15] Chaplin, M. F. (2006). Do we underestimate the importance of water in cell biology? *Nature Reviews Molecular Cell Biology*, 7(11), 861–866.
- [16] Marx, D., Tuckerman, M. E., Hutter, J., & Parrinello, M. (1999). The nature of the hydrated excess proton in water. *Nature*, 397(6720), 601–604.