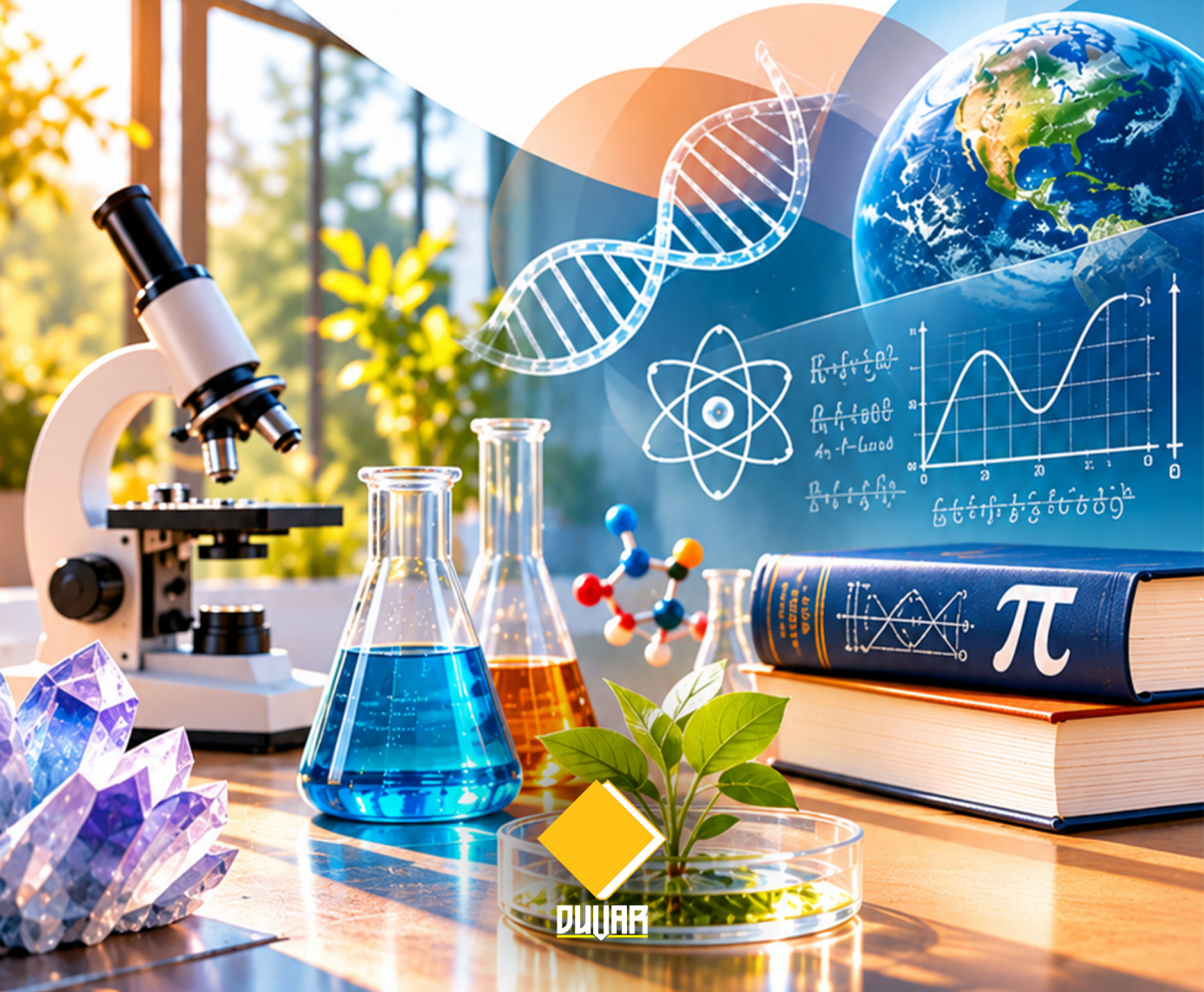
INNOVATIVE APPROACHES AND SCIENTIFIC STUDIES IN NATURAL SCIENCES AND MATHEMATICS

Editor:
Prof. Dr. Eylem DENİZ



DUVAR

INNOVATIVE APPROACHES AND SCIENTIFIC STUDIES IN NATURAL SCIENCES AND MATHEMATICS

Editor

Prof. Dr. Eylem DENİZ*

* https://orcid.org/0000-0001-8865-2086

Nnovative Approaches and Scientific Studies in Natural Sciences and Mathematics

Editor
Prof. Dr. Eylem DENİZ

General Director: Berkan Balpetek
Cover and Page Design: Duvar Design
Printing : October-2026
Publisher Certificate No: 49837
E-ISBN: 978-625-5471-58-1

© Duvar Yayınları
853 Sokak No:13 P.10 Kemeraltı-Konak/İzmir
Tel: 0 232 484 88 68
www.duvaryayinlari.com
duvarkitabevi@gmail.com

"The authors bear full responsibility for the sources, opinions, findings, results, tables, figures, images, and all other content presented in the chapters of this book. They are solely accountable for any financial or legal obligations that may arise in connection with national or international copyright regulations.The publisher and editors shall not be held liable under any circumstances"

TABLE OF CONTENTS

Chapter 1 1
Hyper-Zagreb and Forgotten Topological Indices of Triangular Cycle Graphs
Ahmet EMİN

Chapter 2 19
Quantile Regression: Theoretical Foundations,a Controlled Simulation, and an Empirical Study of the Engel Data
Berrin ERKAN

Chapter 3 40
In Vitro Behavior of Well-Crystallized Nanocrystalline Boron Nitride Synthesized by Dopant Metal Carbonates in the View of Biocompatibility
Büşra MORAN, Muhammed ÖZ

Chapter 4 58
The Effect of Distance Geometry on Spectral Clustering: An Analysis through the Minkowski Parameter
Gülay İlona TELSİZ KAYAOĞLU, Mustafa EROĞLU

Chapter 5 70
Current Systematic Botanical Nomenclature of Turkish Plant Names in Tervîhü'l-Ervâh
Harun ŞAHİN, Mevlüt ALATAŞ

Chapter 6 82
Degradation Mechanisms of Mordant Red 19, Carminic Acid, and Curcumin Molecules Using the DFT Method.
Bahar EREN

Chapter 7 102
Osculating Helices in the Multiplicative Euclidean Space. $\mathbb{E}_*^3$ and Their Duality with Normal Helices
Derya SAĞLAM, Umut SELVİ

Chapter 1

Hyper-Zagreb and Forgotten Topological Indices of Triangular Cycle Graphs

Ahmet EMİN[1]

Abstract

In this book chapter, we investigate the algebraic, combinatorial, and structural topological properties of triangular cycle graphs, denoted by $TC(n)$, which are formally constructed from figurate triangular numbers (T_n). We focus on two prominent degree-based topological descriptors widely applied in chemical graph theory, molecular structure modeling, and quantitative structure-property relationships (QSPR/QSAR): the Hyper-Zagreb index ($HM(G)$) and the Forgotten topological index ($F(G)$). By decomposing the edge set of $TC(n)$ according to endpoint vertex degrees into three mutually disjoint classes (Type A: $(2,2)$, Type B: $(2,3)$, and Type C: $(3,3)$), we establish exact combinatorial counts for each edge partition via mathematical induction that explicitly tracks both newly appended edges and the reclassification of existing boundary edges. Utilizing these partitions, we derive explicit closed-form expressions for both $HM(TC(n))$ and $F(TC(n))$ for all dimensions $n \geq 3$. In addition, alternative algebraic formulations in terms of triangular numbers T_n, discrete recurrence relations connecting consecutive dimensions, and the closed-form expression for the first Zagreb index $M_1(TC(n))$ are established. We prove analytically that the strict hierarchical chain $M_1(TC(n)) < M_2(TC(n)) < F(TC(n)) < HM(TC(n))$ holds for all $n \geq 3$, and demonstrate that the universal structural identity $HM(G) = F(G) + 2M_2(G)$ holds identically over the entire $TC(n)$ family. Finally, the theoretical closed-form formulas and the structural identity are computationally verified across all dimensions $3 \leq n \leq 80$ using an automated Python verification routine within the NetworkX framework, while the boundary cases $n = 1$ and $n = 2$ are evaluated directly from the corresponding graph structures.

Keywords: Triangular cycle graphs, polygonal cycle graphs, figurate numbers, hyper-Zagreb index, forgotten topological index, second Zagreb index, chemical graph theory.

[1] Department of Mathematics, Faculty of Engineering and Natural Sciences, Karabuk University, Karabük, Türkiye

1. Introduction and Theoretical Background

Graph theory provides an indispensable mathematical foundation for modeling discrete relational systems, molecular connectivity, and structural network properties [1-2]. The mathematical study of figurate and polygonal numbers, originating in classical Pythagorean geometry, provides an elegant bridge connecting discrete geometry, number theory, and modern combinatorics [3]. In recent years, algebraic structures, semigroup properties, and continued fraction expansions associated with polygonal and figurate numbers have been systematically characterized [4-6]. Furthermore, the realization of number-theoretic sequences as structured graph families has gained considerable momentum, linking combinatorial number sequences with chemical graph theory and complex network architectures [7].

Among the family of regular polygonal numbers, triangular numbers defined by $T_n = \frac{n(n+1)}{2}$ form the canonical prototype, expressing the sum of the first n positive integers. When points in these geometric arrangements are treated as vertices and adjacent spatial points are connected by edges according to natural geometric boundary rules, one obtains an intriguing class of network topologies known as polygonal cycle graphs [8, 9]. In particular, the triangular cycle graph family, denoted by $TC(n)$, provides a rich structural benchmark exhibiting rotational symmetry, periodic boundary layers, and a systematic partition between boundary vertices and internal network cycles [10, 11].

In chemical graph theory, mathematical chemistry, and complex network analysis, numerical invariants that characterize the underlying topology of a molecular graph and remain invariant under graph automorphisms are designated as topological indices or molecular descriptors [12-14]. Topological indices serve as vital quantitative descriptors in Quantitative Structure-Property Relationship (QSPR) and Quantitative Structure-Activity Relationship (QSAR) studies, enabling researchers to correlate physicochemical properties-such as boiling point, strain energy, and thermodynamic stability-directly with molecular connectivity patterns [13, 14].

Among degree-based topological descriptors, the classical first and second Zagreb indices, introduced by Gutman and Trinajstić [12] and surveyed comprehensively in [13, 14], represent the foundational metrics:

$$M_1(G) = \sum_{u\in V(G)} d(u)^2 = \sum_{uv\in E(G)} (d(u) + d(v)),$$
$$M_2(G) = \sum_{d(u)d(v)} d(u)d(v), \tag{1.1}$$

where $d(u)$ denotes the degree of vertex $u \in V(G)$. While the classical Zagreb indices provide baseline information regarding network branching and overall connectivity, they occasionally fail to distinguish fine-grained structural differences in polycyclic topologies. To address these limitations, several higher-order degree-based descriptors have been developed.

In particular, the Hyper-Zagreb index, denoted by $HM(G)$, was introduced by Shirdel, Rezapour, and Sayadi [15] and studied for various graph classes by Asok and Kureethara [16]:

$$HM(G) = \sum_{uv \in E(G)} (d(u) + d(v))^2. \tag{1.2}$$

By imposing a quadratic penalty on the sum of neighbor degrees, the Hyper-Zagreb index exhibits heightened sensitivity to high-degree adjacencies, proving useful in QSPR modeling of thermodynamic parameters and molecular stability [15, 16].

Parallel to this development, Furtula and Gutman in 2015 formally established the Forgotten topological index (also widely known as the F-index) [17]:

$$F(G) = \sum_{u \in V(G)} d(u)^3 = \sum_{uv \in E(G)} (d(u)^2 + d(v)^2). \tag{1.3}$$

Extensive empirical benchmarking across chemical datasets has shown that the F-index provides strong predictive capability, demonstrating comparable or improved correlation relative to M_1 for specific properties such as acentric factors, vaporization enthalpies, and entropy parameters [13, 14, 17].

Recent investigations have explored topological properties across diverse graph structures and network families, including spacer dendrimer architectures [18], figurate number graphs [7], and algorithmic graph layout analysis [9]. For triangular cycle graphs, preliminary studies established the first Zagreb index $M_1(TC(n))$ [10] and the second Zagreb index $M_2(TC(n))$ [11]. However, to the best of the authors' knowledge, explicit closed-form formulations for higher-order descriptors-specifically the Hyper-Zagreb index and the Forgotten topological index-have not previously been reported for the $TC(n)$ family. The principal objectives of this book chapter are:

1. To establish a rigorous combinatorial classification of the edge set of $TC(n)$ into distinct endpoint degree partitions via an inductive transition argument tracking edge reclassifications.

2. To prove explicit closed-form expressions for the Hyper-Zagreb index $HM(TC(n))$ and Forgotten index $F(TC(n))$ for all $n \geq 3$.

3. To formulate exact expressions in terms of figurate triangular numbers T_n and derive discrete recurrence relations connecting consecutive dimensions.

4. To state and verify the closed-form expression for $M_1(TC(n))$ and establish the strict hierarchy $M_1 < M_2 < F < HM$ for all $n \geq 3$.

5. To demonstrate the exact analytical validity of the universal structural identity $HM(G) = F(G) + 2M_2(G)$ across the entire $TC(n)$ architecture.

6. To present a reproducible Python implementation verifying the theoretical closed-form expressions and structural identity for all dimensions $3 \leq n \leq 80$, with boundary dimensions $n = 1$ and $n = 2$ evaluated directly from their topological structures.

2. Preliminaries and Graph Construction

Throughout this chapter, all graphs considered are finite, connected, undirected, and simple (containing neither self-loops nor multiple edges). For a graph $G = (V, E)$, $|V(G)|$ denotes the vertex cardinality (order) and $|E(G)|$ denotes the edge cardinality (size). The neighborhood of a vertex $u \in V(G)$ is denoted by $N(u) = \{v \in V(G): uv \in E(G)\}$, and its degree is $d(u) = |N(u)|$.

Definition 1 (Triangular Numbers [3]). For each positive integer $n \in \mathbb{Z}^+$, the n-th triangular number T_n is defined by the discrete arithmetic sum:

$$T_n = \sum_{k=1}^{n} k = \frac{n(n+1)}{2}. \tag{2.1}$$

The sequence begins with $T_1 = 1$, $T_2 = 3$, $T_3 = 6$, $T_4 = 10$, $T_5 = 15$, $T_6 = 21$, ... By standard convention, we set $T_0 = 0$.

Definition 2 (Triangular Cycle Graph TC(n)). Let $n \geq 1$ be an integer. The n-th triangular cycle graph, denoted by $TC(n) = (V, E)$, is a planar connected graph formed by arranging T_n vertices in triangular lattice layers and establishing cyclic boundary and inner adjacencies. The vertex set is given by $V(TC(n)) = \{1,2,3, \dots, T_n\}$. The edge set $E(TC(n))$ is governed by three deterministic geometric connectivity rules:

1. Horizontal Inner Adjacency: For every vertex $i \in \{1, \dots, T_n - 1\}$, $i \sim i + 1$ whenever i is not a triangular number ($i \notin \{T_1, T_2, \dots, T_n\}$, where $T_1 = 1$).

2. Right Exterior Boundary Adjacency: For each $k \in \{1,2, \dots, n-1\}$, $T_k \sim T_{k+1}$ (connecting consecutive triangular number terminals along the right boundary).

3. Left Exterior Boundary Adjacency: For each $k \in \{1,2, \dots, n-1\}$, $(T_{k-1} + 1) \sim (T_k + 1)$ (connecting consecutive level starts along the left boundary).

The designation triangular cycle graph originates from the structural realization introduced in [10, 11], wherein each consecutive level $k \in \{2, \dots, n\}$ introduces a

cycle boundary that encloses internal 3-cycles. For $n \geq 3$, $TC(n)$ is a 2-connected planar graph containing $n-1$ concentric cycle layers.

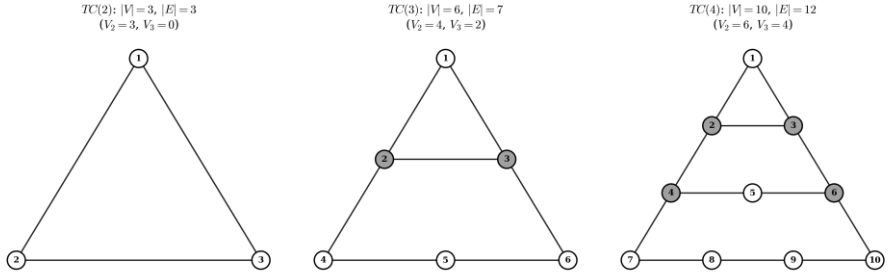


Figure 1: Geometric Configurations and Natural Vertex Labeling of TC(2), TC(3), and TC(4) Triangular Cycle Graphs (White nodes: degree 2, Gray nodes: degree 3).

As visualized in Figure 1, for $n = 2$, $TC(2)$ is isomorphic to the complete 3-cycle $C_3 \cong K_3$, wherein every vertex possesses degree 2. For $n \geq 3$, the network exhibits a degree distribution where vertices possess degrees exclusively in the set $\{2,3\}$.

Lemma 1 (Vertex Degree Distribution [10]). *Let $n \geq 2$ be an integer. In the triangular cycle graph $TC(n)$, the vertex set partitions into degree 2 and degree 3 subsets with cardinalities:*

$$|V_2(TC(n))| = \frac{n^2-3n+8}{2}, \tag{2.2}$$

$$|V_3(TC(n))| = 2n-4. \tag{2.3}$$

Proof. We proceed by mathematical induction on n. For the base case $n = 2$, $|V(TC(2))| = T_2 = 3$, and all 3 vertices have degree 2. Equations (2.2) and (2.3) yield $|V_2| = (4-6+8)/2 = 3$ and $|V_3| = 4-4 = 0$, holding true. Assume the hypothesis holds for dimension $n-1$: $|V_2(TC(n-1))| = ((n-1)^2 - 3(n-1)+8)/2$ and $|V_3(TC(n-1))| = 2n-6$. The graph $TC(n)$ is obtained from $TC(n-1)$ by appending the n vertices of the n-th level: $\{T_{n-1}+1, \dots, T_n\}$. These new vertices form a base path where all n nodes possess degree 2 in $TC(n)$. Concurrently, the former boundary vertices $T_{n-2}+1$ and T_{n-1} from the $(n-1)$-th level, which had degree 2 in $TC(n-1)$, are connected to $T_{n-1}+1$ and T_n, respectively, increasing their degrees from 2 to 3. Consequently, the net count of degree 3 vertices increases by exactly 2: $|V_3(TC(n))| = (2n-6)+2 = 2n-4$. Subtracting this from the total order T_n yields

$$|V_2(TC(n))| = \frac{n(n+1)}{2} - (2n-4) = \frac{n^2-3n+8}{2}. \blacksquare$$

Corollary 1 (Edge Cardinality). *For $n \geq 2$, the total number of edges in $TC(n)$ is:*

$$|E(TC(n))| = \frac{n^2+3n-4}{2}. \tag{2.4}$$

Proof. By the Handshaking Lemma,

$$2\big|E\big(TC(n)\big)\big| = \sum_{v \in V} d(v) = 2|V_2| + 3|V_3|.$$

Substituting the expressions from Lemma 1:

$$\begin{aligned} 2|E(TC(n))| &= 2\left(\frac{n^2-3n+8}{2}\right) + 3(2n-4) = n^2 - 3n + 8 + 6n - 12 \\ &= n^2 + 3n - 4. \end{aligned}$$

Dividing by 2 establishes Equation (2.4). ■

3. Edge Partitions and Structural Induction

Because all vertices in $TC(n)$ possess degrees in $\{2,3\}$ for $n \geq 3$, the edge set $E(TC(n))$ decomposes naturally into three mutually disjoint classes based on the unordered pair of endpoint degrees:

- **Type A Edges ($E_{(2,2)}$):** Edges connecting two vertices of degree 2.
- **Type B Edges ($E_{(2,3)}$):** Edges connecting a vertex of degree 2 with a vertex of degree 3.
- **Type C Edges ($E_{(3,3)}$):** Edges connecting two vertices of degree 3.

We denote the respective partition cardinalities by
$N_A(n) = |E_{(2,2)}|$, $N_B(n) = |E_{(2,3)}|$, and $N_C(n) = |E_{(3,3)}|$.

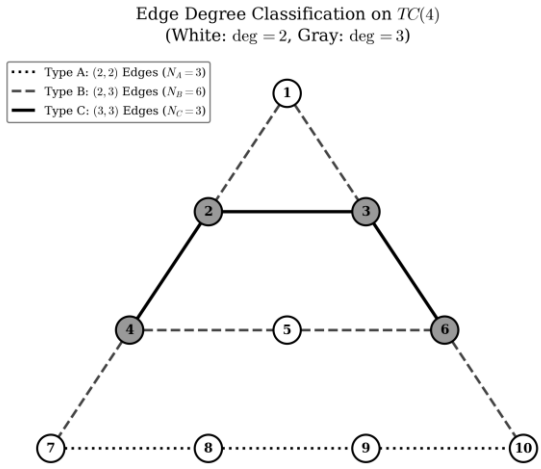


Figure 2: Edge Partition Decomposition on TC(4): Type A (2,2) edges (dotted), Type B (2,3) edges (dashed), and Type C (3,3) edges (solid).

Lemma 2 (Edge Partition Cardinalities [11]). *Let* $n \geq 3$ *be an integer. The edge partition cardinalities of the triangular cycle graph* $TC(n)$ *are given by:*

$$N_A(n) = |E_{(2,2)}| = \frac{n^2-5n+10}{2} \tag{3.1}$$

$$N_B(n) = |E_{(2,3)}| = 2n - 2 \tag{3.2}$$

$$N_C(n) = |E_{(3,3)}| = 2n - 5 \tag{3.3}$$

Proof. We proceed by mathematical induction on n. For base case $n = 3$, $TC(3)$ consists of 6 vertices and 7 edges (see Figure 1). Direct inspection confirms $N_A(3) = 2$, $N_B(3) = 4$, $N_C(3) = 1$, and their sum is $2 + 4 + 1 = 7 = |E(TC(3))|$, establishing the base step.

Assume the partition formulas hold for $TC(n-1)$ for some $n \geq 4$:

$N_A(n-1) = (n^2 - 7n + 16)/2$, $N_B(n-1) = 2n - 4$, $N_C(n-1) = 2n - 7$.

When appending level n, n new vertices and $n + 1$ new edges are added. All n new vertices have degree 2. The 2 former base-corners receive an additional incident edge, increasing their degrees from 2 to 3. Tracking transitions:

(1) Newly added: $n - 1$ horizontal edges connect degree 2 pairs ($+(n-1)$ to N_A); 2 boundary edges connect degree 3 former corners to degree 2 new corners ($+2$ to N_B).

(2) Reclassifications: 2 horizontal edges incident to former corners transition from (2,2) to (2,3) (-2 to N_A, $+2$ to N_B); 2 former exterior boundary edges transition from (2,3) to (3,3) (-2 to N_B, $+2$ to N_C). Net increments:

$\Delta N_A = (n-1) - 2 = n - 3$; $\Delta N_B = 2 + 2 - 2 = +2$; $\Delta N_C = +2$.

Applying these increments yields

$$N_A(n) = \frac{n^2-5n+10}{2},\ N_B(n) = 2n - 2,\ N_C(n) = 2n - 5.$$

Their sum satisfies $N_A + N_B + N_C = \frac{n^2+3n-4}{2} = \left|E\left(TC(n)\right)\right|$, completing the induction. ■

4. The Hyper-Zagreb Index of Triangular Cycle Graphs

We now derive the explicit closed-form formulation for the Hyper-Zagreb index $HM(TC(n))$, followed by its representation in terms of triangular numbers and its discrete recurrence relation.

Theorem 1 (Closed Form of Hyper-Zagreb Index). *For any integer $n \geq 3$, the Hyper-Zagreb index of the triangular cycle graph $TC(n)$ is:*

$$HM(TC(n)) = 8n^2 + 82n - 150 \tag{4.1}$$

Proof. By definition, the sum over all edges $uv \in E(TC(n))$ is decomposed across the three edge subsets: $HM(TC(n)) = 16N_A(n) + 25N_B(n) + 36N_C(n)$. Substituting Equations (3.1)-(3.3):

$$\begin{aligned} HM(TC(n)) &= 16\left(\frac{n^2-5n+10}{2}\right) + 25(2n-2) + 36(2n-5) \\ &= 8(n^2 - 5n + 10) + (50n - 50) + (72n - 180) \\ &= 8n^2 + 82n - 150. \blacksquare \end{aligned}$$

Corollary 2 (Hyper-Zagreb Index via Triangular Numbers). *For $n \geq 3$, the Hyper-Zagreb index of $TC(n)$ can be expressed in terms of the n-th triangular number T_n as:*

$$HM(TC(n)) = 16T_n + 74n - 150 \tag{4.2}$$

Proof. Recalling that $T_n = \frac{n(n+1)}{2}$, we have $n^2 = 2T_n - n$. Substituting this into Equation (4.1): $HM(TC(n)) = 8(2T_n - n) + 82n - 150 = 16T_n + 74n - 150$. ■

Corollary 3 (Discrete Recurrence for Hyper-Zagreb Index). *For $n \geq 4$, the Hyper-Zagreb index satisfies the first-order linear recurrence relation:*

$$HM(TC(n)) = HM(TC(n-1)) + 16n + 74 \tag{4.3}$$

Proof. Computing the forward difference:

$$\begin{aligned} HM(TC(n)) - HM(TC(n-1)) &= (8n^2 + 82n - 150) \\ &\quad -(8(n-1)^2 + 82(n-1) - 150) \\ &= 16n + 74. \blacksquare \end{aligned}$$

Remark 1 (Boundary Cases $n = 1$ and $n = 2$): For $n = 1$, $TC(1) \cong K_1$ contains no edges, hence $HM(TC(1)) = 0$. For $n = 2$, $TC(2) \cong C_3$ contains 3 edges of type (2,2), yielding $HM(TC(2)) = 3 \times 4^2 = 48$. Note that the general polynomial evaluated at $n = 2$ yields $8(4) + 82(2) - 150 = 46 \neq 48$; the closed form is valid strictly for $n \geq 3$.

5. The Forgotten Topological Index of Triangular Cycle Graphs

We now examine the Forgotten topological index $F(TC(n))$. We provide two distinct derivations: one based on vertex degree cubes and the other based on edge degree squares.

Theorem 2 (Closed Form of Forgotten Topological Index). *For any integer $n \geq 3$, the Forgotten topological index of the triangular cycle graph $TC(n)$ is:*

$$F(TC(n)) = 4n^2 + 42n - 76 \tag{5.1}$$

Proof. We establish the theorem using two distinct combinatorial formulations.
Method 1 (Vertex Degree Cubes):

$$\begin{aligned} F(TC(n)) &= 8|V_2| + 27|V_3| = 8\left(\tfrac{n^2-3n+8}{2}\right) + 27(2n-4) \\ &= 4(n^2 - 3n + 8) + 54n - 108 = 4n^2 + 42n - 76. \end{aligned}$$

Method 2 (Edge Degree Squares):

$$\begin{aligned} F(TC(n)) &= 8N_A + 13N_B + 18N_C = 8\left(\tfrac{n^2-5n+10}{2}\right) + 13(2n-2) + 18(2n-5) \\ &= 4(n^2 - 5n + 10) + 26n - 26 + 36n - 90 = 4n^2 + 42n - 76. \end{aligned}$$

Both derivations coincide identically. ■

Corollary 4 (Forgotten Index via Triangular Numbers). *For $n \geq 3$, the Forgotten topological index of $TC(n)$ satisfies:*

$$F(TC(n)) = 8T_n + 38n - 76 \tag{5.2}$$

Proof. Substituting $n^2 = 2T_n - n$ into Equation (5.1):

$$F(TC(n)) = 4(2T_n - n) + 42n - 76 = 8T_n + 38n - 76. \blacksquare$$

Corollary 5 (Discrete Recurrence for Forgotten Index). *For $n \geq 4$, the Forgotten topological index satisfies the recurrence relation:*

$$F(TC(n)) = F(TC(n-1)) + 8n + 38 \tag{5.3}$$

Proof. Taking the difference between consecutive dimensions:

$$F(TC(n)) - F(TC(n-1)) = (4n^2 + 42n - 76) - (4(n-1)^2 + 42(n-1) - 76) = 8n + 38. \blacksquare$$

6. Structural Identities, Index Hierarchy, and Comparison

A well-known structural identity in degree-based topological index theory connects the Hyper-Zagreb, Forgotten, and second Zagreb indices directly from algebraic expansion [15, 16].

Theorem 3 (Universal Structural Identity). *For every simple graph G, the following structural identity holds identically:*

$$HM(G) = F(G) + 2M_2(G) \tag{6.1}$$

Proof. For any edge $uv \in E(G)$, expanding the binomial square yields

$$(d(u) + d(v))^2 = d(u)^2 + d(v)^2 + 2d(u)d(v).$$

Summing over all edges $uv \in E(G)$ yields Equation (6.1). ■

To provide a complete picture of the Zagreb descriptor family on $TC(n)$, we also recall the first and second Zagreb indices.

Theorem 4 (First Zagreb Index [10]). *For any integer $n \geq 3$, the first Zagreb index of the triangular cycle graph $TC(n)$ is:*

$$M_1(TC(n)) = 2n^2 + 12n - 20 \tag{6.2}$$

Proof. By vertex summation,

$M_1(TC(n)) = 4|V_2| + 9|V_3| = 4(\frac{n^2-3n+8}{2}) + 9(2n-4) = 2n^2 + 12n - 20.$

Alternatively, summing over edge partitions:

$4N_A + 5N_B + 6N_C = 2n^2 + 12n - 20.$ ■

Proposition 1 (Analytical Identity Verification on TC(n)). *For $n \geq 3$, the expressions $F(TC(n)) = 4n^2 + 42n - 76$ and $M_2(TC(n)) = 2n^2 + 20n - 37$ [11] identically satisfy Theorem 3, exactly reproducing the Hyper-Zagreb formula $HM(TC(n)) = 8n^2 + 82n - 150$.*

$$\begin{aligned} F(TC(n)) + 2M_2(TC(n)) &= (4n^2 + 42n - 76) + 2(2n^2 + 20n - 37) \\ &= 8n^2 + 82n - 150 = HM(TC(n)). \end{aligned}$$

Furthermore, the recurrence differences satisfy

$\Delta F + 2\Delta M_2 = (8n + 38) + 2(4n + 18) = 16n + 74 = \Delta HM,$

validating internal dimensional consistency.

Theorem 5 (Strict Index Hierarchy). *For all dimensions $n \geq 3$, the four topological descriptors satisfy the strict inequality ordering:*

$$M_1(TC(n)) < M_2(TC(n)) < F(TC(n)) < HM(TC(n)) \tag{6.3}$$

Proof. We evaluate the differences between successive descriptors in the proposed hierarchy:

1. For $M_2 - M_1$:
   $M_2(TC(n)) - M_1(TC(n)) = (2n^2 + 20n - 37) - (2n^2 + 12n - 20) = 8n - 17$
   For all $n \geq 3$, $8n - 17 \geq 8(3) - 17 = 7 > 0$. Hence
   $M_1(TC(n)) < M_2(TC(n))$.
2. For $F - M_2$:
   $$\begin{aligned} F(TC(n)) - M_2(TC(n)) &= (4n^2 + 42n - 76) - (2n^2 + 20n - 37) \\ &= 2n^2 + 22n - 39. \end{aligned}$$
   For all $n \geq 3$, $2n^2 + 22n - 39 \geq 2(9) + 22(3) - 39 = 45 > 0$. Hence $M_2(TC(n)) < F(TC(n))$.
3. For $HM - F$:
   $HM(TC(n)) - F(TC(n)) = (8n^2 + 82n - 150) - (4n^2 + 42n - 76$

$$= 4n^2 + 40n - 74 = 2M_2\big(TC(n)\big).$$

Because all vertex degrees in $TC(n)$ satisfy $d(u) \geq 2$, every edge weight $d(u)d(v) \geq 4 > 0$, implying $M_2(TC(n)) > 0$ for all $n \geq 3$. Thus $HM(TC(n)) > F(TC(n))$. ■

Remark 2 (Boundary Behavior at $n = 2$): For $n = 2$, $TC(2) \cong C_3$, each vertex has degree 2, giving $M_1(TC(2)) = 3 \times 2^2 = 12$ and $M_2(TC(2)) = 3 \times (2 \times 2) = 12$. Thus, M_1 and M_2 coincide at $n = 2$, after which the trajectory of M_2 strictly diverges above M_1 for all $n \geq 3$, as shown in Figure 3.

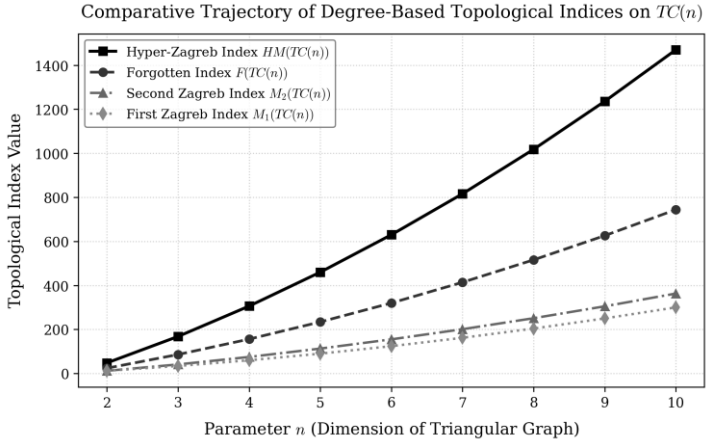


Figure 3: Comparative Quadratic Trajectories of Degree-Based Topological Descriptors (M_1, M_2, F, HM) on Triangular Cycle Graphs TC(n) for n ∈ [2, 10].

As illustrated in Figure 3, all four degree-based indices exhibit quadratic growth rates $O(n^2)$. The dominant coefficient $(8n^2)$ of the Hyper-Zagreb index reflects its heightened sensitivity to network scale in triangular cycle networks.

7. Algorithmic Modeling and Computational Verification

To ensure complete reproducibility and computational validation, an algorithmic implementation in Python utilizing the NetworkX library was developed to construct $TC(n)$ and evaluate all topological metrics directly from adjacency matrices.

Listing 1: Python Verification and Topological Index Computation on TC(n)

```python
import networkx as nx

def build_triangular_cycle_graph(n):
    """Constructs the triangular cycle graph TC(n) based on triangular number coordinates."""
    if n == 1:
        G = nx.Graph()
        G.add_node(1)
        return G

    T = lambda k: k * (k + 1) // 2
    total_vertices = T(n)
    G = nx.Graph()
    G.add_nodes_from(range(1, total_vertices + 1))

    # Triangular numbers up to T_n (including T_1 = 1)
    triangular_set = {T(k) for k in range(1, n + 1)}

    # 1. Horizontal inner adjacency (within each level)
    for i in range(1, total_vertices):
        if i not in triangular_set:
            G.add_edge(i, i + 1)

    # 2. Right exterior boundary adjacency (connecting T_i ~ T_{i+1})
    for i in range(1, n):
        G.add_edge(T(i), T(i + 1))

    # 3. Left exterior boundary adjacency (connecting (T_{i-1}+1) ~ (T_i+1))
    for i in range(1, n):
        G.add_edge(T(i - 1) + 1, T(i) + 1)

    return G
```

```python

def compute_topological_indices(G):
    """Computes M1, M2, F, and HM topological indices directly from graph topology."""
    deg = dict(G.degree())
    m1 = sum(d**2 for d in deg.values())
    m2 = sum(deg[u] * deg[v] for u, v in G.edges())
    f_idx = sum(d**3 for d in deg.values())
    hm_idx = sum((deg[u] + deg[v])**2 for u, v in G.edges())
    return m1, m2, f_idx, hm_idx

def verify_theoretical_formulas(max_n=80):
    """Verifies boundary cases directly and closed-form expressions for n in [3, max_n]."""
    # Direct evaluation of boundary cases n=1 and n=2
    G1 = build_triangular_cycle_graph(1)
    assert G1.number_of_nodes() == 1 and G1.number_of_edges() == 0

    G2 = build_triangular_cycle_graph(2)
    m1_2, m2_2, f_2, hm_2 = compute_topological_indices(G2)
    assert m1_2 == 12 and m2_2 == 12 and f_2 == 24 and hm_2 == 48
    assert hm_2 == f_2 + 2 * m2_2  # Structural identity holds at n=2

    # Verification of closed-form formulas for n = 3 to max_n
    for n in range(3, max_n + 1):
        G = build_triangular_cycle_graph(n)
        m1, m2, f_val, hm_val = compute_topological_indices(G)

        # Theoretical closed-form formulas (valid for n >= 3)
        m1_th = 2 * n**2 + 12 * n - 20
        m2_th = 2 * n**2 + 20 * n - 37
        f_th  = 4 * n**2 + 42 * n - 76
        hm_th = 8 * n**2 + 82 * n - 150

        assert m1 == m1_th, f"M1 mismatch at n={n}"
        assert m2 == m2_th, f"M2 mismatch at n={n}"
        assert f_val == f_th, f"F mismatch at n={n}"
        assert hm_val == hm_th, f"HM mismatch at n={n}"
        assert hm_val == f_val + 2 * m2, f"Identity mismatch at n={n}"
    print(f"Boundary cases (n=1,2) and formulas for n in [3, {max_n}] verified."
    )
```

Table 1: Structural Invariants, Edge Partitions, and Topological Index Values of $TC(n)(n = 1, \cdots, 12)$

n	T_n	$\|E\|$	$\|V_2\|$	$\|V_3\|$	N_A	N_B	N_C	$M_1(TC(n))$	$M_2(TC(n))$	$F(TC(n))$	$HM(TC(n))$
1	1	0	0	0	0	0	0	0	0	0	0
2	3	3	3	0	3	0	0	12	12	24	48
3	6	7	4	2	2	4	1	34	41	86	168
4	10	12	6	4	3	6	3	60	75	156	306
5	15	18	9	6	5	8	5	90	113	234	460
6	21	25	13	8	8	10	7	124	155	320	630
7	28	33	18	10	12	12	9	162	201	414	816
8	36	42	24	12	17	14	11	204	251	516	1018
9	45	52	31	14	23	16	13	250	305	626	1236
10	55	63	39	16	30	18	15	300	363	744	1470
11	66	75	48	18	38	20	17	354	425	870	1720
12	78	88	58	20	47	22	19	412	491	1004	1986

Note: Closed-form formulas are valid for $n \geq 3$. *Values for* $n = 1,2$ *are determined directly from graph topology.*

The theoretical closed-form formulas for the four degree-based indices and their structural identity are verified computationally for all dimensions $3 \leq n \leq 80$. The boundary cases $n = 1$ and $n = 2$ are evaluated directly from the corresponding graph structures. Specifically, for $n = 1, TC(1) \cong K_1$ has no edges, yielding zero for all edge-based descriptors. For $n = 2, TC(2) \cong C_3$ is a 2-regular cycle where $M_1(TC(2)) = 12$, $M_2(TC(2)) = 12$, $F(TC(2)) = 24$, and $HM(TC(2)) = 48$, confirming that $HM = F + 2M_2 = 24 + 2(12) = 48$ and $M_1 = M_2 = 12$. For all $n \geq 3$, the closed-form polynomials match the graph calculations with zero deviation. For example, at dimension $n = 6$:

$$M_1(TC(6)) = 2(36) + 12(6) - 20 = 124$$
$$M_2(TC(6)) = 2(36) + 20(6) - 37 = 155$$
$$F(TC(6)) = 4(36) + 42(6) - 76 = 320$$
$$HM(TC(6)) = 8(36) + 82(6) - 150 = 630$$

and

$$F(TC(6)) + 2M_2(TC(6)) = 320 + 2(155) = 320 + 310 = 630 = HM(TC(6)),$$

confirming complete numerical consistency.

8. Conclusions and Future Perspectives

In this chapter, we investigated the algebraic, combinatorial, and topological characteristics of the triangular cycle graph family $TC(n)$ constructed from figurate

triangular numbers. Through systematic combinatorial analysis, the vertex set was partitioned into degrees $\{2,3\}$, and the edge set was decomposed into three disjoint endpoint classes ($E_{(2,2)}, E_{(2,3)}, E_{(3,3)}$) using an inductive argument tracking edge reclassifications.

The principal contributions established herein include:

1. Rigorous mathematical proof of explicit closed-form expressions for the Hyper-Zagreb index $HM(TC(n)) = 8n^2 + 82n - 150$ and Forgotten index $F(TC(n)) = 4n^2 + 42n - 76$ for all $n \geq 3$.

2. Formulation of alternative representations in terms of triangular numbers T_n and first-order linear recurrence relations governing dimensional transitions.

3. Explicit closed-form determination of the first Zagreb index $M_1(TC(n)) = 2n^2 + 12n - 20$ and analytic proof of the strict hierarchy $M_1 < M_2 < F < HM$ for all $n \geq 3$.

4. Analytical and computational verification of the fundamental structural identity $HM(G) = F(G) + 2M_2(G)$ across all graph dimensions.

5. Development of a reproducible Python validation workflow confirming complete agreement between theoretical closed-form theorems and computational network models across $3 \leq n \leq 80$, with boundary dimensions $n = 1,2$ evaluated directly.

Future Research Perspectives: The structural partitioning methodology formulated in this study provides a versatile mathematical template for investigating other figurate architectures, including square, pentagonal, hexagonal [8], and generalized polygonal cycle graphs [7]. Promising avenues for subsequent research encompass evaluating distance-based and degree-exponential descriptors-such as the Sombor indices [19]-as well as investigating spectral graph properties and metric dimension on generalized polygonal cycle networks.

References

[1] Sarp, Ü. (2023). Visualising connections between types of polygonal number. *The Mathematical Gazette*, 107(568), 56–64. https://doi.org/10.1017/mag.2023.7

[2] Emin, A. (2022). Some algebraic structure on figurate numbers. *Bitlis Eren University Journal of Science*, 11(2), 604-612. https://doi.org/10.17798/bitlisfen.1068022

[3] Deza, E. & Deza, M.M. (2012). *Figurate Numbers*. Singapore: World Scientific Publishing. https://doi.org/10.1142/8188

[4] Gebeş, E., Emin, A., & Sarp, Ü. (2024). *On triangular pentagonal and square pentagonal numbers*. In M. Çevik (Ed.), Scientific Research Reports (ch. 4, pp. 47–56). İzmir, Türkiye: Duvar Publications.

[5] Emin, A. & Sarp, Ü. (2022). Simple continued fractions for centered polygonal numbers, *Transactions of A. Razmadze Mathematical Institute*, 176(2), 197-205.

[6] Emin, A. & Sarp, Ü. (2024). Groupoid and semigroup construction on isosceles triangular numbers, *BSEU Journal of Science*, 11(2), 325-334. https://doi.org/10.35193/bseufbd.1313160

[7] Demir, B., Sarp, Ü., & Emin, A. (2026). On the topological indices of a generalized graph family used in number-theoretic graph realizations, *Gulf Journal of Mathematics*, 22(2), 1–14. https://doi.org/10.56947/gjom.v22i2.4197

[8] Emin, A., Demir, B., & Sarp, Ü. (2025). *Hexagonal cycle graphs: A graph-theoretic realization of hexagonal numbers*. In M. Çevik (Ed.), Advances in Scientific Research (ch. 17, pp. 200-208). İzmir, Türkiye: Duvar Publications.

[9] Sarp, Ü., & Demir, B. (2025). *Kamada-Kawai algorithm: A comprehensive analysis from global graph layout to modern challenges*. In Ş. Türkoğlu (Ed.), Theoretical and Applied Research in Natural Sciences and Mathematics (ch. 5, pp. 63-75). İzmir, Türkiye: Duvar Publications.

[10] Demir, B., Sarp, Ü., & Emin, A. (2025). On the first Zagreb index of triangular cycle graphs, in *Proceedings of the 5th International Symposium of Scientific Research and Innovative Studies (ISSRIS'25)*, 157–161.

[11] Sarp, Ü., Demir, B., & Emin, A. (2025). On the second Zagreb index of triangular cycle graphs, in *Proceedings of the 9th International Students Science Congress (ISSC 2025)*, İzmir, Türkiye, 331–339. https://doi.org/10.52460/issc.2025.043

[12] Gutman, I., & Trinajstić, N. (1972). Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons, *Chemical Physics Letters*, 17(4), 535-538. https://doi.org/10.1016/0009-2614(72)85099-1

[13] Borovićanin, B., Das,K.C., Furtula, B., & Gutman, I. (2017). Bounds for Zagreb indices, *MATCH Communications in Mathematical and in Computer Chemistry*, 78(1), 17-100. https://match.pmf.kg.ac.rs/electronic_versions/Match78/n1/match78n1_17-100.pdf

[14] Ali, A., Gutman, I., Milovanović, E., & Milovanović, I. (2018). Sum of powers of the degrees of graphs: Extremal results and bounds, *MATCH Communications in Mathematical and in Computer Chemistry*, 80(1), 5–84. https://match.pmf.kg.ac.rs/electronic_versions/Match80/n1/match80n1_5-84.pdf

[15] Shirdel, G.H., Rezapour, H., & Sayadi, A.M. (2013). The hyper-Zagreb index of graph operations, *Iranian Journal of Mathematical Chemistry*, 4(2), 213–220. https://doi.org/10.22052/ijmc.2013.5294

[16] Asok, A., & Kureethara, J.V. (2020). Characterization of the forgotten topological index and the hyper–Zagreb index for the unicyclic graphs, *MATCH Communications in Mathematical and in Computer Chemistry*, 84(3), 739–751. https://match.pmf.kg.ac.rs/electronic_versions/Match84/n3/match84n3_739-751.pdf

[17] Furtula, B., & Gutman, I. (2015). A forgotten topological index, *Journal of Mathematical Chemistry*, 53(4), 1184-1190. https://doi.org/10.1007/s10910-015-0480-z

[18] Demir, B., & Sarp, Ü. (2025). *A graph-theoretic study of polygonal core spacer dendrimers*. In N. Gümrükçüoğlu (Ed.), Modern Research and Investigations in Natural and Mathematical Sciences III (ch. 1, pp. 1-12). Lyon, France: Livre de Lyon publications.

[19] Gutman, I. (2021). Geometric approach to degree–based topological indices: Sombor indices, *MATCH Communications in Mathematical and in Computer Chemistry*, 86(1), 11-16.https://match.pmf.kg.ac.rs/electronic_versions/Match86/n1/match86n1_11-16.pdf

Chapter 2

Quantile Regression: Theoretical Foundations, a Controlled Simulation, and an Empirical Study of the Engel Data

Berrin ERKAN[1]

ABSTRACT

This study examines the theoretical foundations and practical interpretation of quantile regression at two levels. First, it distinguishes conditional-mean regression from conditional-quantile regression using asymmetric absolute loss, median regression, linear programming, bootstrap inference, and out-of-sample quantile evaluation. Second, it investigates the income–food expenditure relationship in the Engel dataset, comprising 235 Belgian working-class households observed in 1857. The empirical analysis estimates conditional quantiles at τ=0.10, 0.25, 0.50, 0.75, and 0.90; compares ordinary least squares (OLS); uses 600 resamples generated by the pairs bootstrap, 100 repeated train–test splits, a log–log sensitivity model, and an influential-observation check. Expressed per franc of income, the income slope rises from 0.402 at the lower conditional quantile to 0.686 at the upper quantile, and the 95% bootstrap interval for their difference excludes zero. An additional controlled Monte Carlo experiment with 180 replications illustrates how slopes can differ across quantiles under a known data-generating process. The findings demonstrate that quantile regression reveals conditional distributional heterogeneity, while cross-sectional associations do not identify causal effects or individual transitions across quantiles. Additional checks examine income-stratified calibration, duplicate income–expenditure value pairs, and paired out-of-fold model comparisons.

Keywords: Quantile Regression; Conditional Quantiles; Asymmetric Absolute Loss; Pinball Loss; Engel Data.

[1] Assist. Prof. Dr., Çanakkale Onsekiz Mart University, Department of Statistics, berrin.erkan@comu.edu.tr, Orcid: 0000-0001-5889-7461

1. INTRODUCTION

Standard linear regression typically asks how the conditional mean of an outcome changes with explanatory variables. However, economic variables such as income, wages, and health expenditure can be skewed, and predictors may affect not only the conditional mean but also dispersion and tails. A single OLS slope may therefore fail to represent relationships across the conditional distribution.

Quantile regression, systematically developed by Koenker and Bassett (1978), models a selected conditional quantile directly. Median regression is a special case. Comparing coefficients over several levels of τ reveals whether a predictor has different conditional associations at low and high outcome levels (Koenker and Hallock, 2001).

The aim is not merely to draw several quantile lines. The theoretical objective is to define what a quantile function represents, identify the optimization problem behind estimation, explain coefficient interpretation, and quantify uncertainty. The empirical objective is to implement a coherent workflow of estimation, comparison, evaluation, and sensitivity analysis on real data.

The Engel dataset is particularly suitable: income and food expenditure are strongly positively associated, yet expenditure dispersion is not constant over income. Quantile regression can describe this variation. Nevertheless, the dataset is historical and observational; it is not representative of current consumption patterns and cannot by itself establish a causal effect of income.

2. LITERATURE REVIEW AND STUDY POSITIONING

2.1. Theoretical development of quantile regression

Koenker and Bassett (1978) connected conditional-quantile estimation to minimizing asymmetrically weighted absolute errors. The framework allows coefficients to vary across quantiles. Koenker and Bassett (1982) further established how regression quantiles can be used to investigate heteroscedasticity.

Koenker and Hallock (2001) use applications including food expenditure and wages to show why the quantile approach complements ordinary least squares. Koenker (2005) provides a unified treatment of computation, inference, and extensions. A central starting point here is that a conditional quantile slope need not equal a conditional mean slope, even under an otherwise meaningful regression model.

Buchinsky (1998) systematically examines covariance estimation, the joint assessment of coefficients across quantiles, and tests concerning the error distribution in cross-sectional applications. This work underscores that quantile

regression is more than drawing fitted lines for the median or two tail quantiles: defensible inference requires an appropriate account of uncertainty and a coherent interpretation of the quantile process. The present study does not claim to construct simultaneous confidence bands for the entire process. Instead, it undertakes a limited, prespecified contrast between two slopes while resampling complete household records together.

2.2. Goodness of fit and predictive evaluation

Koenker and Machado (1999) developed a goodness-of-fit perspective specific to quantile regression, distinct from conventional R^2. Gneiting and colleagues (2023) distinguish calibration diagnostics from out-of-sample loss evaluation for quantile models. The present study therefore centers predictive comparisons on pinball loss, a score aligned with the target quantile.

Predictive evaluation requires two distinct questions: How much does a model reduce pinball loss for its target quantile, and how well do its stated quantiles or intervals agree with realized outcomes? Gneiting, Balabdaoui, and Raftery (2007) emphasize joint attention to calibration and sharpness in probabilistic forecasts. The quantile-focused review by Gneiting et al. (2023) extends this distinction to conditional quantiles and prediction intervals. Consequently, a lower average loss is not treated here as proof of correct coverage in every income stratum.

The Engel dataset is a well-established teaching example. The present contribution is not a claim of novel data discovery but an integrated exposition combining theory, controlled simulation, pairs-bootstrap inference, repeated out-of-sample evaluation, sensitivity analysis, and reproducible bilingual reporting.

2.3. Quantile inference, joint contrasts, and bootstrap limitations

Statistical significance of a slope at one quantile does not establish a significant difference between slopes at two quantiles. The appropriate unit of inference is the contrast itself; because both slopes are fitted to the same households, their sampling dependence matters. Koenker (2005) and Buchinsky (1998) discuss the importance of conditional density estimation and covariance structure in quantile regression inference. Resampling complete income–food-expenditure pairs in the Engel application is intended to preserve this within-record dependence in a practical pairs-bootstrap procedure.

Nevertheless, percentile confidence intervals from 600 bootstrap resamples do not guarantee nominal coverage under every possible data-generating mechanism. Household heterogeneity, sparse information in tail quantiles, and misspecification of linear conditional quantiles may affect validity. Searching across many slope contrasts also raises multiple-testing concerns. Accordingly,

the prespecified upper-versus-lower slope difference is treated as the primary contrast, while other differences are described as secondary or exploratory findings.

2.4. Quantile crossing, calibration, and model choice

Conditional quantile functions estimated separately need not be ordered monotonically in the quantile index. Chernozhukov, Fernández-Val, and Galichon (2010) develop monotone rearrangement to address such crossing. The absence of crossings among five fitted linear quantiles over a 400-point grid spanning observed Engel incomes indicates that rearrangement is not required for these particular curves on that grid. It does not guarantee ordering for other quantiles or extrapolated income levels.

Near-nominal aggregate coverage of an 80% prediction band likewise need not imply adequate coverage at every income level. The distinction between unconditional or marginal calibration and conditional calibration in Gneiting et al. (2023) is useful for interpreting the linear model's approximately 72.5% coverage among low-income households. Income-stratified diagnostics here are instructive, not definitive population-level calibration tests, because the held-out records recur across splits and subgroup uncertainty remains.

2.5. Contribution and scope relative to the literature

Koenker and Hallock (2001) already use the same historical Engel data to illustrate what quantile regression reveals beyond the conditional mean. The present research therefore does not claim discovery of a previously unknown food-expenditure relationship. Its contribution is pedagogical and reproducible integration: it explains the Koenker–Bassett (1978) estimator using a controlled simulation; compares real-data slopes using a household-paired bootstrap; evaluates pinball loss alongside marginal and income-stratified coverage; and checks repeated observations and the influence of extreme incomes.

The findings are descriptive and arise from a cross-sectional observational sample. A larger slope at a higher conditional food-expenditure quantile does not establish that an individual household moves between expenditure strata as its income changes, or that income has a causal effect on its food spending. Those questions require different identification strategies. This article is positioned as a verifiable teaching and analytical application, not as a new estimator or proof of policy effects with external validity.

3. THEORETICAL FRAMEWORK

3.1. Conditional mean and conditional quantile

A conventional OLS model targets the conditional mean E(Y | X=x). By contrast, the τ-quantile is the outcome threshold below which the conditional distribution places probability τ when X=x. For a continuous conditional distribution, τ=0.10 denotes the lower decile, τ=0.50 the median, and τ=0.90 the upper decile.

$$Q_\tau(Y \mid X = x) = \inf\{y: F_{Y|X=x}(y) \geq \tau\}, \quad 0 < \tau < 1$$

A linear conditional-quantile model estimates a separate coefficient vector at each τ and does not require the vector to remain constant. Importantly, τ=0.90 does not refer to the highest-income 10% of households: the quantile belongs to the conditional distribution of food expenditure at a fixed income level.

$$Q_\tau(Y \mid X = x) = x^\top \beta(\tau) = \beta_0(\tau) + \beta_1(\tau)x$$

3.2. Asymmetric absolute loss and estimation

Quantile regression replaces squared errors with piecewise-linear penalties that differ for positive and negative residuals. This pinball, or check, loss penalizes underprediction more heavily when τ is large. At τ=0.50 both sides receive equal weight, yielding median regression.

$$\rho_\tau(u) = u[\tau - \mathbf{1}(u < 0)]$$

$$\hat{\beta}(\tau) = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{n} \rho_\tau \left(y_i - x_i^\top \beta\right)$$

The objective is convex but not differentiable at zero residual. It can therefore be formulated as a linear program. In this application each residual is split into positive and negative components, and the HiGHS linear-programming solver is used. This also provides an independent numerical check against iterative approximate fitting algorithms.

Objective (β, u^+, u^-): $\tau\sum_i u_i^+ + (1-\tau)\sum_i u_i^-$
Subject to: $X\beta + u^+ - u^- = y$; $u^+ \geq 0$ and $u^- \geq 0$

3.3. Coefficient interpretation and important limits

$\beta_1(\tau)$ describes the model-based change in the τth conditional quantile of Y for a one-unit change in X. It does not describe a specific household moving from the 10th to the 90th percentile. In cross-sectional data, quantile levels describe

different conditional positions, so neither within-person transitions nor causal income effects follow from these coefficients.

Quantile regression does not require Gaussian errors or constant error variance for its definition, but this does not remove all dependence and sampling assumptions. Independent or appropriately sampled observations, a meaningful conditional-quantile specification, and sufficient design information remain important for inference. High-leverage predictor observations can affect median regression as well.

3.4. Uncertainty and bootstrap inference

Coefficients estimated at different quantiles are statistically dependent because they use the same sample. It is inappropriate to construct an interval for $\beta_1(0.90) - \beta_1(0.10)$ by treating two separate standard errors as independent. This study uses the pairs bootstrap: households are resampled as income–expenditure pairs, and all five quantiles are re-estimated on each resample.

$$\hat{\Delta} = \beta_1(0.90) - \beta_1(0.10); \quad 95\%\ \mathrm{CI} = [q_{0.025}(\Delta^*), q_{0.975}(\Delta^*)]$$

Percentile bootstrap intervals are transparent computationally but do not guarantee exact coverage in small samples or extreme quantiles. The 600 resamples here serve to assess the empirical slope contrast; more advanced inference could examine alternative bootstrap designs and a larger number of replications.

3.5. Fit, pinball loss, and calibration

Conventional R^2 is not the primary fit measure for quantile regression. An in-sample pseudo-R^2 summarizes the reduction in check loss relative to an intercept-only quantile model. Here $L\tau$ denotes the total pinball loss at the target quantile. Its numerical magnitude should not be compared directly with the usual OLS R^2.

$$R_\tau^2 = 1 - \frac{L_\tau(\hat{\beta}_\tau)}{L_\tau(\hat{q}_\tau)}$$

Out-of-sample evaluation uses pinball loss at each τ; smaller values indicate better estimation of the target quantile. Evaluation criteria must match the forecast target, such as absolute error for a median and squared error for a conditional mean. Treating an OLS mean prediction as though it were a valid quantile prediction at every τ is a target mismatch; such values are included only as an instructive contrast.

$$PB_\tau = \frac{1}{m}\sum_{j=1}^{m} \rho_\tau\left(y_j - \hat{Q}_\tau(x_j)\right)$$

Calibration concerns the expectation that roughly 90% of outcomes lie below a valid τ=0.90 forecast. Nevertheless, correct average coverage over all observations does not establish conditional calibration at every income level. A 10th–90th quantile band is an interval for the conditional outcome distribution, not a confidence interval for regression coefficients.

4. DATA, RESEARCH DESIGN, AND REPRODUCIBILITY

4.1. Engel dataset and variables

The primary dataset contains annual income (income) and annual food expenditure (foodexp) for 235 Belgian working-class households in 1857. Both variables are measured in Belgian francs. The data share their source with the Engel dataset in the R quantreg package and are distributed directly with statsmodels. This is a historical cross-section; neither the currency nor the observation period has been converted into contemporary expenditure levels.

Data validation found no missing observations, and both variables were positive in all records. A log–log sensitivity model can therefore be estimated. In the OLS and linear-quantile models, income is divided by 1,000 for numerical scaling; reported slopes describe the modeled expenditure difference for a 1,000-franc income difference.

Table 1. Engel Data: Descriptive Statistics

Variable	Mean	Std. dev.	Q1	Median	Q3	Minimum	Maximum
Income	982.5	519.2	638.9	884.0	1,164.0	377.1	4,957.8
Food spending	624.2	276.5	429.7	582.5	743.9	242.3	2,032.7
Spending/income	0.656	0.094	0.592	0.662	0.719	0.339	0.912

Source: calculated from the Engel dataset.

4.1.1. Data quality and repeated values

Among the 235 source observations, 232 income–expenditure pairs are unique. Two pairs occur repeatedly: one appears three times and another twice, giving three additional rows with identical measured values. These may be duplicate records, but they may equally represent distinct households with coincidentally identical amounts. Without household identifiers, deleting them as

errors is not justified. The primary analysis retains all 235 rows and uses unique-value pairs only as a sensitivity scenario.

Table 2. Sensitivity To Duplicate Income–Expenditure Value Pairs

Quantile	235-row slope	232-pair slope	Change
0.10	401.77	398.35	-3.41
0.25	474.10	474.10	0.00
0.50	560.18	557.87	-2.31
0.75	644.01	643.24	-0.78
0.90	686.30	686.30	-0.00

When identical-value rows are temporarily collapsed, the lower and median slopes change from 401.766 to 398.354 and from 560.181 to 557.871, respectively; the upper slope remains unchanged. This limited sensitivity check shows that the main ordering of slopes is not driven by repeated measured values; it is not a recommendation to delete source records.

Because some households have high incomes, mean income exceeds median income. The mean ratio of food expenditure to income is 0.656, and its median is 0.662. These are descriptive features of historical household expenditure shares, not estimates of a policy effect or contemporary consumption pattern.

4.2. Models and research hypotheses

The reference model estimates E(foodexp|income) by OLS. Main quantile regressions use linear specifications at $\tau=\{0.10, 0.25, 0.50, 0.75, 0.90\}$. An auxiliary log–log model assesses conditional-quantile elasticities and functional-form sensitivity. The primary empirical expectation is that lower- and upper-quantile income slopes may differ, accompanied by greater conditional dispersion at higher incomes.

$$Q_\tau(\text{foodexp} \mid \text{income}) = \alpha_\tau + \beta_\tau(\text{income}/1000)$$
$$Q_\tau[\log(\text{foodexp}) \mid \text{income}] = a_\tau + b_\tau \log(\text{income})$$

4.3. Computation, uncertainty, and diagnostic checks

The five quantiles are estimated by the HiGHS linear-programming algorithm. Their check-loss optima are cross-checked against independent statsmodels QuantReg solutions. Six hundred bootstrap resamples of household pairs yield 95% percentile intervals. Breusch–Pagan and White diagnostics are applied to

OLS residuals. Influential-observation sensitivity is examined by refitting models after excluding the top 1% of income observations.

Predictive evaluation uses 100 random 80% training–20% test splits. Each training sample fits a linear quantile model, a log–log quantile model, and an unconditional-quantile baseline that ignores predictors. Pinball loss is evaluated on the same held-out households. Because splits overlap, the 100 performance estimates are not treated as independent observations and no artificial p-values are calculated from them.

Table 3. Study analysis and validation protocol

Stage	Design	Purpose
Main data	n = 235; income / foodexp	Cross-sectional application
Quantile levels	0.10; 0.25; 0.50; 0.75; 0.90	Conditional distribution
Bootstrap	600 paired resamples	Coefficient uncertainty
Holdout evaluation	100 × 80% train / 20% test	Pinball loss and coverage
Monte Carlo	180 replications; n = 400	Known-parameter validation
Robustness	Exclude top 1% income	Extreme-observation sensitivity
Randomness	Seed = 20261025	Reproducibility

5. CONTROLLED MONTE CARLO DEMONSTRATION

5.1. Constructing known conditional quantiles

To illustrate estimation when the true quantile slopes are known, we first construct a synthetic data-generating process. X follows a uniform distribution on [0,1] and ε is standard normal. Error scale depends on X, so the conditional response distribution widens as X increases. The true slopes at different quantiles can be calculated analytically.

$$X \sim \text{Uniform}(0,1), \quad \epsilon \sim N(0,1), \quad Y = 1 + 2X + (0.3 + 0.8X)\epsilon$$
$$Q_\tau(Y \mid X = x) = [1 + 0.3\Phi^{-1}(\tau)] + [2 + 0.8\Phi^{-1}(\tau)]x$$

Each replication generates n=400 observations, for a total of 180 replications. For τ=0.10, 0.50, and 0.90, we report the true slope, the mean estimated slope, bias, and RMSE. This Monte Carlo design does not prove the empirical Engel results; it validates the mechanism through which differing quantile slopes arise under a known model.

Table 4. Controlled Monte Carlo results

τ	True slope	Mean estimate	Bias	RMSE
0.10	0.975	0.954	-0.021	0.223
0.50	2.000	1.996	-0.004	0.146
0.90	3.025	3.043	0.018	0.197

180 replications; n=400 per replication. True slope $\beta_1(\tau)=2+0.8\Phi^{-1}(\tau)$.

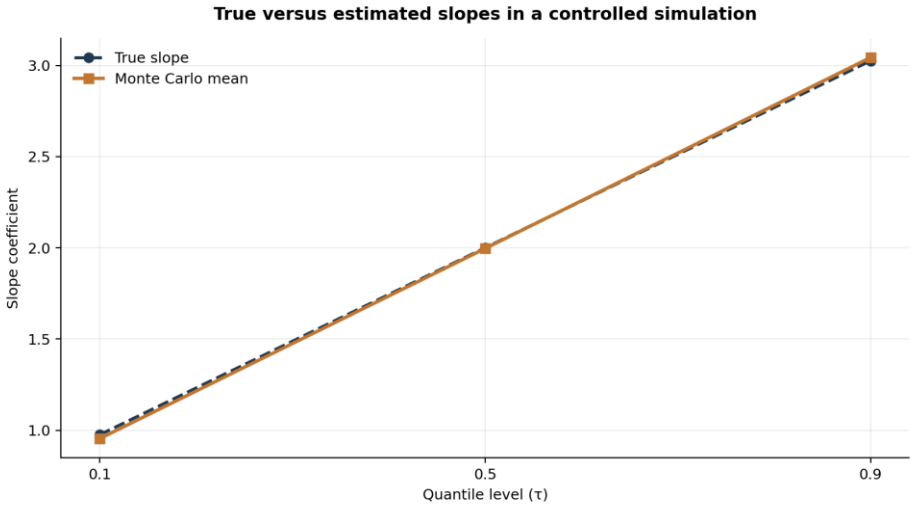


Figure 1. True and mean estimated slopes in the synthetic process.

The true slopes are approximately 0.975, 2.000, and 3.025; their estimated means, 0.954, 1.996, and 3.043, are close. The median slope being near the conditional mean slope agrees with the zero-mean Gaussian error design. Lower and upper quantile slopes differ because the error scale depends on X, not because the algorithm creates an artificial advantage.

6. EMPIRICAL RESULTS FOR THE ENGEL DATA

6.1. OLS benchmark and heteroscedasticity

OLS estimates an intercept of 147.475 francs and an income slope of 485.178 francs per 1,000 francs of income. Its in-sample R^2 is 0.830, and the Pearson correlation is 0.911. The Breusch–Pagan LM statistic is 109.263 ($p<0.001$) and the White LM statistic is 181.120 ($p<0.001$), providing strong evidence against

constant variance in OLS residuals. These diagnostics motivate examining dispersion but do not, by themselves, establish that the linear quantile specification is perfectly correct.

6.2. Conditional-quantile estimates

Estimated income coefficients are positive at all five quantiles, but slopes generally increase with τ. At a fixed income level, the conditional income association is steeper toward the upper part of the expenditure distribution. Individual 95% bootstrap intervals are reported for each slope; they describe coefficient uncertainty, not a distribution of future household expenditure.

Table 5. Quantile coefficients, bootstrap intervals, and fit

τ	Intercept	Income slope†	95% CI lower	95% CI upper	Pseudo-R^2	Pinball (train)
0.10	110.14	401.766	336.018	494.248	0.495	16.47
0.25	95.48	474.103	388.885	514.251	0.554	30.14
0.50	81.48	560.181	471.361	612.388	0.621	37.36
0.75	62.40	644.014	578.719	691.950	0.697	27.78
0.90	67.35	686.299	633.439	729.814	0.765	14.43

† Expenditure contrast for 1,000 francs of income; 600 pairs-bootstrap resamples.

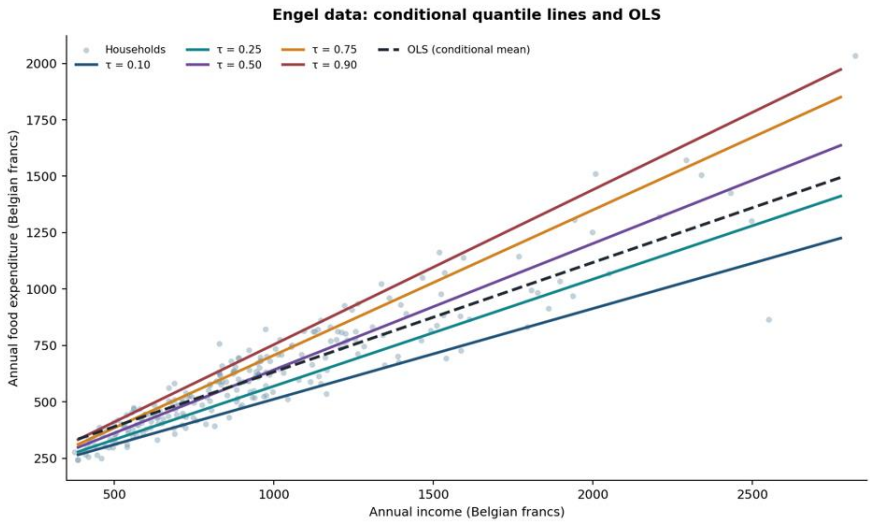


Figure 2. OLS and five quantile-regression lines over Engel observations.

For a modeled income difference of 1,000 francs, the estimated conditional food-expenditure difference is approximately 402 francs at τ=0.10, 560 francs at τ=0.50, and 686 francs at τ=0.90. These are conditional model contrasts, not causal claims that increasing a particular person's income would change their spending by those amounts.

6.3. Inference for differences between quantile slopes

Pairs bootstrap inference assesses whether differences between separately fitted quantile slopes can plausibly be attributed to sampling variation. The τ=0.90 minus τ=0.10 slope contrast is approximately 284.534 francs per 1,000 francs of income, with a 95% bootstrap interval of [187.809, 369.984]. The τ=0.75 minus τ=0.25 contrast is also positive. These results support heterogeneous slopes within the specified linear conditional-quantile model.

Table 6. Pairs-bootstrap intervals for quantile slope contrasts

Slope contrast	Estimate†	95% CI lower	95% CI upper
τ=0.90 − τ=0.10	284.534	187.809	369.984
τ=0.75 − τ=0.25	169.911	109.000	244.390

† Joint pairs-bootstrap contrasts per 1,000 francs of income.

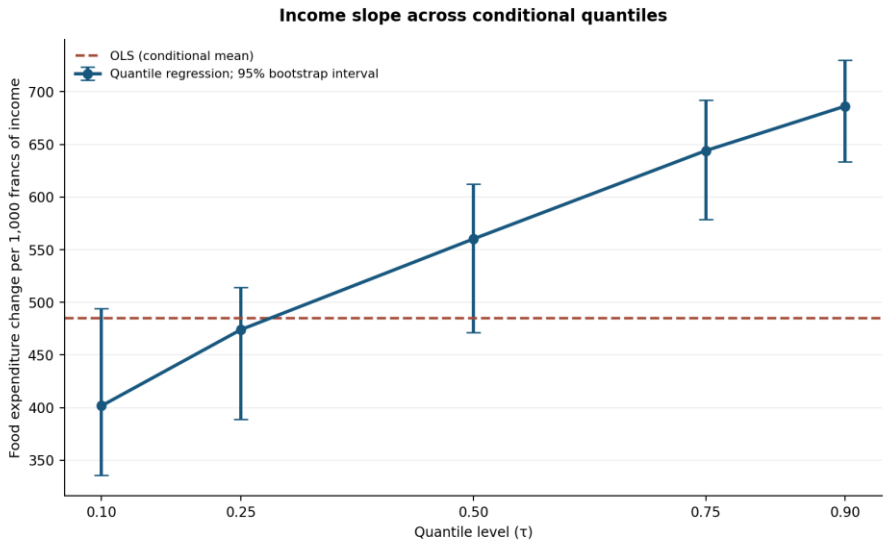


Figure 3. Income slopes and 95% pairs-bootstrap intervals.

6.4. Economic interpretation of conditional dispersion

Quantile estimates provide not only different slopes but also intervals for potential outcomes at fixed income levels. The fitted conditional 10th–90th percentile band is about 99.476 francs wide at an income of 500 francs and 384.010 francs wide at 1,500 francs. The widening of this band provides a direct model-based illustration of increasing conditional dispersion with income.

Table 7. Conditional expenditure quantiles at two income levels

Income	Q0.10	Q0.50	Q0.90	Q0.90 − Q0.10
500	311.02	361.57	410.50	99.48
1,500	712.79	921.75	1,096.80	384.01

Conditional outcome interval, not a coefficient confidence interval.

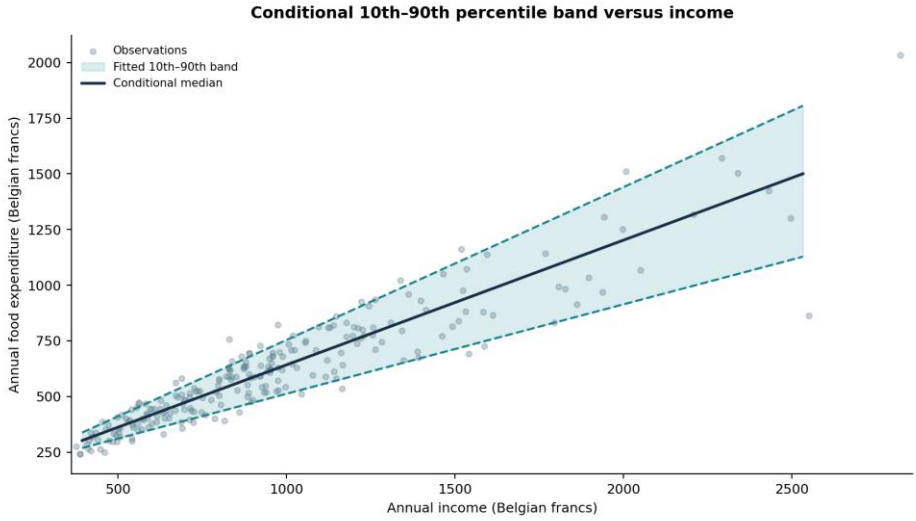


Figure 4. Conditional 10th–90th percentile expenditure band by income.
The band targets the central 80% of the conditional expenditure distribution. It should not be confused with the 95% bootstrap confidence intervals for coefficients. Extrapolating linear quantile lines beyond observed income support is unreliable; the figures are restricted to the data-supported range.

6.5. Out-of-sample loss comparison

Out-of-sample pinball loss is calculated on held-out data separately for each target quantile. Both the linear and log–log quantile models achieve substantially lower losses than a baseline that forecasts only the unconditional training-sample

quantile. This indicates that income contains predictive information about conditional expenditure quantiles, within the limits of the historical dataset.

Table 8. Mean pinball loss across repeated holdout splits

τ	Unconditional	Linear quantile	Log–log quantile	Linear coverage
0.10	32.72	17.59	15.32	0.100
0.25	67.74	30.56	29.63	0.252
0.50	98.88	37.76	36.87	0.501
0.75	92.10	28.45	28.06	0.740
0.90	61.63	15.18	15.06	0.891

100 shared random holdout splits; lower pinball loss is better.

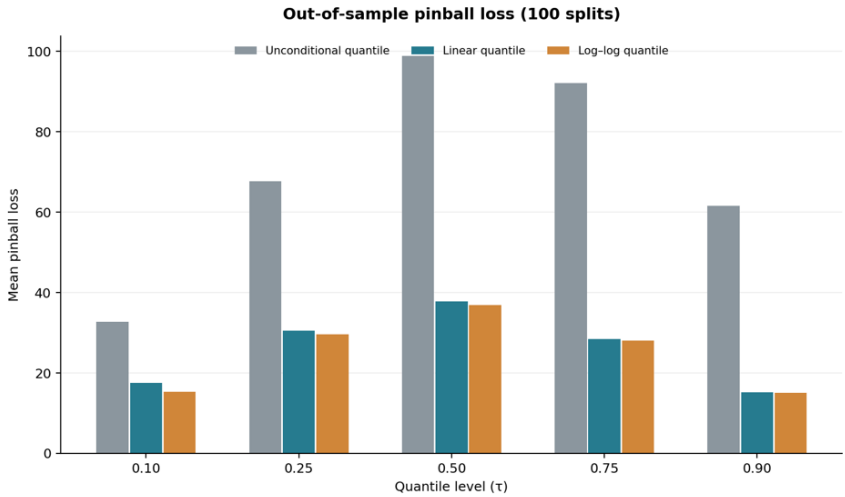


Figure 5. Out-of-sample pinball loss at five quantile levels.

At τ=0.50, mean pinball loss is 98.876 for the unconditional baseline, 37.765 for the linear quantile model, and 36.871 for the log–log model. Although the log–log specification has a slightly smaller mean loss here, dependence across repeated test sets and the limited sample size preclude treating this small gap as definitive superiority.

6.5.1. Paired five-fold out-of-fold comparison

The original 100 repeated splits describe predictive losses but do not constitute 100 independent experiments. In an additional check, the 235 households were partitioned into five fixed random folds. Each household was tested once, and linear and log–log predictions were paired on the same held-out household. The difference below equals linear-model pinball loss minus log–log-model loss; positive values favor log–log regression.

Table 9. Paired model-loss differences under five-fold testing

τ	Linear loss	Log–log loss	Diff. (L−LL)	95% lower	95% upper
0.10	19.28	16.30	2.99	1.25	5.52
0.25	31.79	31.17	0.63	-0.33	2.07
0.50	38.25	37.47	0.78	-0.19	2.02
0.75	27.93	27.62	0.31	-0.34	1.05
0.90	14.76	14.55	0.21	-0.33	0.84

Positive difference favors log–log; intervals are descriptive household bootstraps of fixed out-of-fold predictions.

The 95% ranges use 2,000 household-level bootstrap draws of paired losses at fixed out-of-fold predictions. Because the models are not refitted inside each bootstrap draw, these descriptive intervals do not reflect the full model-training uncertainty. In particular, the interval at $\tau=0.50$ crosses zero; a definitive log–log advantage at the median would be unjustified. The comparison is exploratory and conditional on this historical sample.

6.6. Coverage, prediction bands, and calibration

Using the same 100-split evaluation strategy, 10th and 90th quantiles form a nominal 80% prediction band. Mean test coverage is 79.3% for linear quantile regression and 79.4% for the log–log model. The unconditional baseline attains similar coverage but produces much wider intervals, so both coverage and width need to be considered jointly.

Table 10. Out-of-sample performance of the 80% prediction band

Model	Empirical coverage	Mean width	Inversion fraction
Linear quantile	79.3%	238.82	0.0%
Log–log quantile	79.4%	229.06	0.0%
Unconditional baseline	79.9%	591.27	0.0%

Nominal target 80%; band from Q0.10–Q0.90.

Both models show near-nominal marginal coverage, but this does not imply calibrated forecasts within each low- or high-income subgroup. Stronger validation would require calibration curves within income strata and larger external validation samples (Gneiting et al., 2023).

6.6.1. Conditional calibration within income strata

Income was split into three roughly equally sized groups using full-sample tercile cutoffs of about 715.37 and 1,014.15 francs. These cutoffs are diagnostic group labels only and do not enter model fitting. Each test observation in the 100 repeated splits was assigned an income group; appearance counts are not counts of independent households because the same household can reappear in different holdouts.

Table 11. Calibration of the 80% prediction band by income group

Income stratum	Model	Test appearances	Coverage (%)	Mean band width
Low	Linear	1561	72.5%	116.8
Low	Log–log	1561	76.7%	124.2
Middle	Linear	1602	84.0%	209.9
Middle	Log–log	1602	83.6%	202.8
High	Linear	1537	81.3%	392.9
High	Log–log	1537	77.7%	362.9

Income cutoffs use full-sample tertiles; repeated household test appearances are dependent.

6.8. Extreme-observation and quantile-crossing checks

Excluding three households in the top 1% of the income distribution changes some slopes, particularly in the lower tail. The $\tau=0.10$ slope rises from 401.766 to 454.554, whereas the median slope changes by only about 1%. This indicates greater sensitivity of lower-tail estimates to high-income leverage, without implying that those observations are erroneous.

Table 12. Slope sensitivity after excluding the highest-income 1%

τ	Full slope	Top-1%-trim slope	Change (%)	Log–log elasticity
0.10	401.766	454.554	13.1	0.804
0.25	474.103	489.419	3.2	0.849
0.50	560.181	565.570	1.0	0.877
0.75	644.014	644.014	0.0	0.916
0.90	686.299	661.406	-3.6	0.891

Income threshold at the 99th percentile; three households excluded.

Separately estimated conditional quantile lines can cross and violate monotonic ordering. No crossing was detected among the five fitted linear quantile functions on a 400-point grid spanning observed incomes. This does not guarantee that crossing is impossible for unestimated quantiles or outside the observed income range.

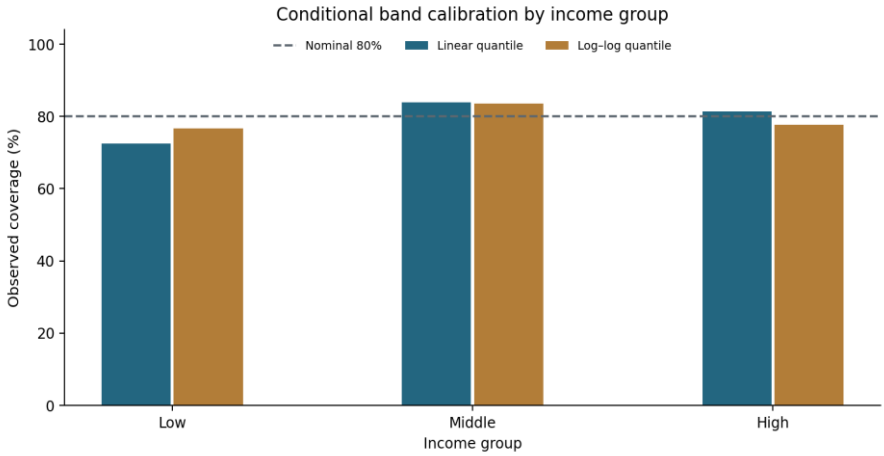


Figure 6. Observed prediction band coverage by income group (nominal 80%).

The nominal 80% band covers 72.5%, 84.0%, and 81.3% of repeated held-out observations in the low-, middle-, and high-income groups under linear quantile regression. The corresponding log–log rates are 76.7%, 83.6%, and 77.7%. Thus, approximately nominal overall coverage masks meaningful discrepancies by income group. Conditional calibration is imperfect, but finite group sizes and repeated use of households prevent treating these rates as exact population parameters.

6.7. The log–log model and Engel's law

In the auxiliary quantile model fitted to log income and log food expenditure, the median slope is approximately 0.877, with a 95% bootstrap interval of [0.782, 0.930] from 600 resamples. The log–log conditional-quantile slope can be interpreted as an elasticity; this estimate is consistent with median food spending varying less than proportionally with income. The pattern aligns with the traditional intuition of Engel's law, but generalization remains limited by functional form and the historical sample.

Because the log transformation is strictly increasing, conditional quantiles on the log scale can be exponentiated back to expenditure units. Nevertheless, a model fitted in log space must be evaluated using pinball loss on the original outcome scale for fair comparison; log-scale and level-scale losses are not directly comparable. All reported out-of-sample losses use original Belgian-franc expenditure units.

7. DISCUSSION

The strongest empirical lesson is that the income–food expenditure relationship is not fully summarized by one conditional-mean slope. Steeper upper-quantile slopes imply that the shape of the conditional expenditure distribution changes with income. In the controlled simulation, this feature arises from a known data-generating process; in the real data, it is quantified through bootstrap slope contrasts.

OLS is not invalid or redundant. It answers a distinct, legitimate question about the conditional mean, whereas quantile regression targets the median and tails. The study therefore does not pretend these targets are identical; it evaluates each quantile using its corresponding loss function.

The modestly lower test loss of the log–log specification at several levels indicates sensitivity to functional form. However, its log-median slope is not on the same scale as a coefficient measured in francs in the level model: the former is an elasticity, the latter an absolute change. Scientific interpretation must retain each coefficient's units.

Near-nominal out-of-sample band coverage is encouraging, but the repeated splits reuse the same 235 historical households. These tests do not establish transferability to other periods, countries, or economic shocks. The report deliberately distinguishes in-sample association, repeated held-out evaluation, and causal conclusions.

The first distinction in relating the findings to the literature is between the target of inference and the scope of the result. Koenker and Bassett (1978) provide a framework for distinct positions of the conditional distribution;

Koenker and Hallock (2001) show on Engel data how conditional dispersion can widen with income. The positive difference between the fitted 0.10 and 0.90 income slopes in this study is consistent with that established pattern, conditional on the same data and a linear conditional-quantile specification. Although the bootstrap interval [187.809, 369.984] excludes zero, it does not capture structural uncertainty from alternative functional forms or unobserved household characteristics.

The second distinction is between goodness of fit and performance for a prediction decision. The pseudo-R^2-type measure of Koenker and Machado (1999) is not a fraction of variance explained that can be compared directly with conventional OLS R^2. By contrast, pinball loss evaluates predictions for a specified quantile. The log–log model achieves a 2.99-franc lower mean loss at $\tau=0.10$ in the five-fold assessment, but the 95% interval for the difference at $\tau=0.50$ contains zero. Thus, a general claim that the logarithmic specification is best at every quantile is not supported. Model choice must depend on quantile level, loss function, and the intended out-of-sample task.

The third distinction is between marginal and conditional calibration. Average coverage of 79.3% for the nominal 80% linear-model band initially appears adequate, whereas coverage falls to 72.5% among lower-income households. From the diagnostic perspective of Gneiting et al. (2023), a pooled coverage statistic can conceal localized systematic error. Because these income strata are used for an explanatory check and the same households recur in repeated test splits, the rates have not been assigned artificial exact-binomial intervals treating appearances as independent observations. The log–log specification raises low-income coverage to 76.7%, yet its high-income coverage falls to 77.7%, revealing a different trade-off.

The fourth distinction concerns data quality and model stability. Identical income–expenditure pairs cannot automatically be classified as erroneous duplicates without household identifiers; collapsing them was used only as a sensitivity exercise. Removing the three highest-income observations changes the fitted 0.10 slope by about 13.1%, indicating the influence of high-income leverage on lower-tail estimation. Similarly, an absence of crossing among five curves does not eliminate the general crossing problem studied by Chernozhukov et al. (2010). The empirical conclusions must therefore remain conditional on the specification and the observed income support.

Economically, the median log–log elasticity of about 0.877 is compatible with food expenditure varying less than proportionately with income in this sample. However, a level-model change per 1,000 francs and a log–log elasticity have different units and should never be compared directly as coefficient magnitudes.

Household composition, relative prices, and expenditure categories are not observed, so an Engel's-law reading should remain descriptive. The strength of the study is not declaring a universally superior estimator from one score; it is integrating a theoretical target, sampling inference, predictive assessment, and sensitivity analysis in a transparent workflow.

CONCLUSION

Quantile regression extends linear analysis beyond the conditional mean to selected positions of the conditional outcome distribution. It is estimated by minimizing asymmetric absolute loss, with median regression as a special case. Rather than merely replacing OLS, it offers a complementary framework for answering different statistical questions.

In the Engel application, income slopes rise markedly from lower to upper quantiles, and the pairs bootstrap indicates a positive contrast despite sampling uncertainty. The widening conditional 10th–90th percentile band within the observed income range provides an economically interpretable measure of changing dispersion.

Out-of-sample pinball loss and near-80% band coverage show why quantile models should be evaluated beyond coefficient tables. Sensitivity checks also reveal the influence of extreme income observations in the lower tail and potential gains from the logarithmic transformation. Additional income-stratum diagnostics show that nominal overall coverage does not imply conditional coverage in every group.

A recommended practical sequence is to define the target conditional quantile; inspect data quality and support; estimate an appropriate quantile model; use joint bootstrap inference for coefficient contrasts; evaluate target-specific loss and calibration on held-out data; and clearly state causal and external-validity limitations. This study provides an executable demonstration of that sequence.

REFERENCES

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. Econometrica, 47(5), 1287–1294. https://doi.org/10.2307/1911963

Buchinsky, M. (1998). Recent advances in quantile regression models: A practical guideline for empirical research. Journal of Human Resources, 33(1), 88–126. https://doi.org/10.2307/146316

Chernozhukov, V., Fernández-Val, I., & Galichon, A. (2010). Quantile and probability curves without crossing. Econometrica, 78(3), 1093–1125. https://doi.org/10.3982/ECTA7880

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69(2), 243–268. https://doi.org/10.1111/j.1467-9868.2007.00587.x

Gneiting, T., Wolffram, D., Resin, J., Kraus, K., Bracher, J., Dimitriadis, T., Hagenmeyer, V., Jordan, A. I., Lerch, S., Phipps, K., & Schienle, M. (2023). Model diagnostics and forecast evaluation for quantiles. Annual Review of Statistics and Its Application, 10, 597–621. https://doi.org/10.1146/annurev-statistics-032921-020240

Koenker, R. (2005). Quantile regression. Cambridge University Press. https://doi.org/10.1017/CBO9780511754098

Koenker, R., & Bassett, G., Jr. (1978). Regression quantiles. Econometrica, 46(1), 33–50. https://doi.org/10.2307/1913643

Koenker, R., & Bassett, G., Jr. (1982). Robust tests for heteroscedasticity based on regression quantiles. Econometrica, 50(1), 43–61. https://doi.org/10.2307/1912528

Koenker, R., & Hallock, K. F. (2001). Quantile regression. Journal of Economic Perspectives, 15(4), 143–156. https://doi.org/10.1257/jep.15.4.143

Koenker, R., & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association,94(448),1296–1310. https://doi.org/10.1080/01621459.1999.10473882

statsmodels developers. (n.d.). Quantile regression and Engel (1857) data [Softwaredocumentation]. https://www.statsmodels.org/stable/examples/notebooks/generated/quantile_regression.html (accessed October 25, 2026).

Chapter 3

In Vitro Behavior of Well-Crystallized Nanocrystalline Boron Nitride Synthesized by Dopant Metal Carbonates in the View of Biocompatibility

Büşra MORAN[1], Muhammed ÖZ[2]

Abstract

Nanocrystalline hexagonal boron nitride (h-BN) has gained significant attention in the biomedical field due to its unique physicochemical properties; however, the influence of crystallinity and specific metal dopants on its biological behavior remains poorly understood. In this study, well-crystallized nanocrystalline hexagonal boron nitride (nh-BN) was successfully synthesized using metal carbonates (Li_2CO_3, $CaCO_3$, and $MgCO_3$) as structural promoters to refine crystal growth and minimize surface defects. The biological safety and functional performance of the synthesized Li-, Ca-, and Mg-nh-BN were comprehensively evaluated following EN ISO 10993 and OECD 487 standards. Cytotoxicity assessments using the MTT assay on L929 fibroblasts revealed exceptional cytocompatibility, with cell viabilities significantly exceeding the 70% threshold, particularly in the Mg-doped group, which demonstrated a stimulatory effect on metabolic activity. Genotoxicity evaluations via the in vitro micronucleus test confirmed that none of the materials induced chromosomal damage or proliferation blockade in CHO cells. Furthermore, the *in vitro* scratch assay demonstrated that Li-nh-BN and Ca-nh-BN significantly accelerated fibroblast migration and wound closure, achieving complete closure within 24 hours. These enhanced cellular responses are attributed to the synergistic effects of high crystallinity and the presence of bioactive Li^+ and Ca^{2+} ions, which likely modulate signaling pathways associated with tissue repair.

Keywords: Nanocrystaline Boron Nitride, Genotoxicity, Cytotoxicity, Wound Healing

[1] Instructor, Scientific Technical App. and Research Center, Hitit University, Çorum, 19030, Türkiye, ORCID:0000-0002-7280-4417

[2]Professor, Chemistry and Chemical Processing Technology, Bolu Abant İzzet Baysal University, Bolu 14900, Türkiye, ORCID: 0000-0003-0049-0161

1. Introduction

Nanostructured materials have emerged over the last decade as highly promising candidates for biomedical and tissue-engineering applications, primarily due to their distinctive physicochemical properties and the capacity to finely tune their biointerfaces. Hexagonal boron nitride (h-BN), a structural analogue of graphene, has gained increasing prominence as a 2D nanomaterial owing to its outstanding material stability, mechanical robustness, and fundamental electrical insulation properties (Pakdel, Zhi, Bando & Golberg, 2012; Matović et al. 2016; Darabdhara et al. 2025). In its nanostructured form, h-BN exhibits a combination of advantageous features including large surface area, biocompatibility, and resistance to oxidative degradation that present it particularly suitable for biomedical and implant-related applications (Atila et al. 2016; Tozar and Karahan 2018). The chemically inert nature of h-BN, coupled with its low cytotoxicity profile and mechanical integrity, has led to its investigation as a scaffold, coating, filler, or nanocarrier in various biomedical systems (Roosta et al. 2016; Kumar et al. 2022). As highlighted by Turhan et al. (2022) and Bahrami et al. 2025, the layered architecture and capacity for surface functionalization of boron nitride nanomaterials open new avenues for biomedical coatings, drug delivery systems, and cell-stimulatory platforms. Beyond biomedical contexts, h-BN has long been employed in industrial applications such as lubricants (Haubner et al. 2002), thermal interface materials (Lin and Connel 2012), and cosmetics (Ansaloni and de Sousa 2013), where its chemical stability and low toxicity allow direct exposure to fluids and biologically relevant environments (Ramirez Leyva et al. 2018; Ma et al. 2020). However, successful translation of h-BN into the biomedical domain requires comprehensive biological evaluation encompassing assessments of cytotoxicity, genotoxicity, inflammatory potential, cellular uptake, hemocompatibility, and the influence on cell migration, proliferation, and differentiation. The biological performance of h-BN is known to depend strongly on its crystal quality, surface chemistry, and the presence of dopants or residual impurities introduced during synthesis (Öz et al. 2016a). Recent advances have focused on incorporating metal-based dopants to tailor the physicochemical surface properties and stability of boron nitride nanomaterials. In particular, metal carbonate doping has been shown to enhance crystallinity, alter surface charge, and improve biointerface compatibility by minimizing reactive surface defects (Öz et al. 2016a). Despite these promising developments, systematic evaluations of the biocompatibility and cytogenetic safety of doped h-BN systems remain limited. In this context, nanocrystalline h-BN materials which are synthesized via the incorporation of magnesium carbonate ($MgCO_3$), calcium carbonate ($CaCO_3$), and lithium

carbonate (Li_2CO_3) is used for estimation of biocompatibility properties. The working hypothesis is that such doping refines crystal growth, reduces surface defects, and introduces benign carbonate moieties that mitigate surface reactivity and enhance biocompatible surface chemistry. These dopant elements were chosen for their biological relevance: magnesium and calcium are naturally abundant in bone and soft-tissue environments, while lithium has been explored for its role in tissue regeneration and neural modulation. Thus, the engineered doped nh-BN nanocrystals are anticipated to demonstrate enhanced biocompatibility and improved cellular interactions relative to undoped or defect-rich counterparts. The synthesized Li-, Ca-, and Mg-doped nanocrystalline h-BN materials (Li-nh-BN, Ca-nh-BN, and Mg-nh-BN) are evaluated in vitro for their biological safety and biomedical potential. All experimental designs adhered to the EN ISO 10993 standards for biological assessment of medical materials. Material extracts were prepared following EN ISO 10993-12, and cytotoxicity, genotoxicity, and fibroblast migration were assessed using MTT, micronucleus, and wound-healing assays, respectively. This comprehensive testing framework aims to elucidate cellular responses and cytocompatibility of metal-carbonate-doped h-BN, providing deeper insight into its suitability for future biomedical and tissue-engineering applications.

2. Experimental

Boron nitride synthesis

Nanocrystalline hexagonal boron nitride (nh-BN) samples were utilized that were sythesized using a modified O'Connor method by Öz et al. 2016a-b. Shortly, a composition that contains 1 g boron oxide, 2 g urea, and 0.9 g metal carbonate [$M_X(CO_3)$, M = Li, Mg, Ca; x = 1 or 2] was pre-heated at 200 °C for 2 h, ground, afterwards nitrided under NH_3 flow (120 mL·min^{-1}) at 1450 °C for 3 h in a tube furnace. The resulting powders were purified in 10% HCl, showered with ethanol, and dried at 100 °C. In the analysis of BN formation, FT-IR spectra Rigaku Multiflex+XRD diffractometer and JEOL 6390-LV SEM were performed in order to determine characteristic properties of nanocrystalline BN.

Biocompatibility experiments

The study is proceeded by first preparing material extracts following the EN ISO 10993-12 standard, ensuring reproducible and biologically relevant extraction conditions. Three doped h-BN materials (Li-nh-BN, Ca-nh-BN, and Mg-nh-BN)) were used. Standard in-vitro assays then assessed cytotoxicity (MTT assay per EN ISO 10993-5), genotoxicity (micronucleus assay per EN ISO 10993-3 / OECD 487) and fibroblast migration/wound healing (scratch assay) to

test vitro anaysis of nh-BN. Results were analysed alongside the material characterization to link biological responses to physicochemical attributes.

Chemicals used for biocompatibility L929 fibroblast cells were the core part of experimental fragment of this study. The complete culture medium consisted of 89% DMEM, 10% fetal bovine serum (FBS; Biological Industries), 1% L-glutamine, and 1% penicillin/streptomycin (Capricorn). For cell detachment, 0.04% trypsin–EDTA solution was used and washing procedures were performed with 1X phosphate-buffered saline (PBS) without calcium or magnesium. Cell counting was carried out using trypan blue (1:1 dilution). Cell viability in cytotoxicity assays was determined using MTT reagent (tetrazolium salt; Serva). Sterile serological pipettes, multiwell plates and culture flasks, (Corning, USA) were preferred for cell culture procedures.

Cell culture preparation for biocompatibility

Frozen L929 cells were rapidly warm up at 37 °C and transferred to a 15 mL Falcon tube under a Class II laminar flow hood. The cells were centrifuged at 250/5 min, refloated in 3 mL of complete culture medium, and seeded into 25 cm^2 flasks. Cultures were reared at 37 °C in a wetted 5% CO_2 incubator. Subculturing continued until a stable cell growth cycle was achieved.

Preparation of samples for biocompatibility

In this study, extracts of Li-nh-BN, Ca-nh-BN, and Mg-nh-BN were organized in accordance with TS EN ISO 10993-12:2021 “Biological evaluation of medical devices-Part 12: Sample preparation and reference materials.” due to inconvenience of direct application of biomaterials or non-feasibility of medical devices. The prepared extracts were sterilized under aseptic conditions prior to use in subsequent biological assays.

Cytotoxicity tests for biocompatibility

The cytotoxic possibility of the test materials was assessed using the MTT assay, following the standard protocol described in EN ISO 10993-5: “Biological evaluation of medical devices- Part 5: Tests for in vitro cytotoxicity.” Briefly, L929 fibroblast cells were seeded in 96-well culture plates at a density of 1×10^4 cells per well in 100 µL of complete culture medium and incubated for 24 hours at 37 °C in 5% CO_2 to allow cell connection and stabilization. After confirming monolayer formation and normal morphology, the medium interchanged with serially diluted material extracts of Li-nh-BN, Ca-nh-BN, and Mg-nh-BN at 5 various compositions ranging from 1:1 (100%) to 1:16. The negative control solely occupied from cells treated with complete culture medium. The culture medium was eliminated, and 50 µL of MTT solution (1 mg/mL) was added to each well after 24 hours of incubation. Plates were incubated approximately 2-

2.5 hours at 37 °C, after which the MTT solution was aspirated, and 100 µL of isopropanol (MTT solvent) was utilized to remove the formazan crystals and absorbance measurement was carried out at 570 nm by a microplate reader.

Cell viability (%) was computed by using the following equation (Karan, Erenler and Moran 2022):

$$\text{Cell Viability (\%)} = \frac{\text{OD}_{570}(\text{sample})}{\text{OD}_{570}(\text{control})} \times 100 \quad (1)$$

Genotoxicity Assessment (Micronucleus Test)

Genotoxic potential was evaluated according to EN ISO 10993-3: "Biological evaluation of medical devices – Part 3: Tests for genotoxicity, carcinogenicity, and reproductive toxicity," in compliance with the OECD Guideline 487 (In Vitro Mammalian Cell Micronucleus Test). Extracts of Li-nh-BN, Ca-nh-BN, and Mg-nh-BN were prepared as described in Preparation of samples for biocompatibility section. All tests were conducted in triplicate. Cells were cultured at 37 °C until reaching subconfluent growth, after which medium was aspirated and changed with 1:1 (100%) extract solution. The cells were exposed to the test materials for 48 hours.

Cytochalasin B (Cyt-B) was added between 44-48 hours of incubation to inhibit cytokinesis and allow binucleated cell formation. Cells were harvested by centrifugation, treated with a hypotonic KCl (0.075M) solution, and incubated for 5 minutes at 37 °C, after 72-hours full incubation time. Until the supernatant clearness, the cells were fixed by using multiple methanol/acetic acid (3:1) fixative. Cell suspensions were put onto clean microscope slides, conditioned in air for dryness, and were exposed to Giemsa solution. Slides were examined under a microscope at 10 fold magnification, and 2,000 binucleated cells per sample were recorded for micronucleus (MN) formation according to the criteria. Only binucleated cells with nuclei of equal size and identical staining intensity were considered. The Cytokinesis-Block Proliferation Index (CBPI) was calculated using the formula provided in OECD 487 (Eq. 2).

$$\text{CBPI} = M_1 + 2M_2 + 3M_{multi}/n, \quad (2)$$

where M_1 to M_{multi} represent counts of cells from one to multiple (more than 2) nuclei and n is the total number of cells. Micronucleus frequency in test groups was compared with that of positive control, and the ratio of micro nucleated binucleated cells was recorded (Jenifer, et al. 2021).

Wound Healing (Cell Migration) Assay

The in vitro wound healing (scratch) assay was employed to evaluate the migration and proliferation potential of L929 fibroblast cells. This assay quantitatively processes the capability of cells to migrate and repopulate a defined cell-free area created on a confluent monolayer. L929 fibroblast cells were planted into 24-well tissue culture plates at a density of 2.5×10^5 cells/mL and incubated at 37 °C in a moistened medium containing 5% CO_2 for 24 hours to allow construction of a uniform confluent monolayer. Following confluence, a linear scratch was set up across the cell layer by a sterile 100 µL pipette tip. Detached cells and debris were eliminated by soft cleaning with PBS.

The control group received only Dulbecco's Modified Eagle Medium (DMEM), whereas the test groups were treated with 100% (1:1) material extracts. All samples were incubated under identical conditions at 37 °C with 5% CO_2. Cell passage into the wounded area was watched and recorded by capturing phase-contrast microscopic images at 0, 8, and 24 hours post-scratch. The extent of wound closure was later quantified by the remaining wound zone at each time point, thereby reflecting the migratory capacity of fibroblasts in response to the test materials.

3. Results and discussion

BN samples synthesized by Öz et al., 2016a, were evaluated for biocompatibility of well-crystallized form of nh-BN samples in vitro. The phase purity and crystallinity of the synthesized BN nanocrystals were confirmed by FT-IR, XRD, SEM and TEM in the mentioned in previous publication. All samples exhibited characteristic diffraction planes at $2\theta \approx 26.7°$, 41.6°, and 55.2°, corresponding to the (002), (100), and (004) peaks of nanocrystalline hexagonal BN (nh-BN), respectively, in agreement with the standard JCPDS card no. 34-0421. These peaks were sharp and well-defined, indicating crystallinity level. The incorporation of metal carbonate dopants ($CaCO_3$, $MgCO_3$, Li_2CO_3) promoted enhanced crystallization, as reflected by increased peak intensities and reduced full width at half maximum (FWHM) values. The average grain size was estimated using the Debye-Scherrer equation, in the range of 16, 24, to 44 nm depending on the dopant species of respectively. Li_2CO_3-assisted synthesis produced the most well-ordered nanocrystals, suggesting a favorable catalytic effect of Li^+ ions during the nitridation process (Öz et al. 2016b).

Fourier-transform infrared (FTIR) spectroscopy established the occupation of B-N stretching and bending vibrations at 1360 cm^{-1} and 780 cm^{-1}, respectively as signal of nanocrystalline hexagonal phase. The absence of broad absorption bands near 3400 cm^{-1} and 1400 cm^{-1} indicated the efficient removal of hydroxyl and carbonate residues during calcination as expressed in the publication. These

results confirm that the dopant metal carbonates were completely decomposed and did not alter the BN lattice integrity.

SEM revealed nanosheet and platelet-like grains with the lengths of 100–200 nm and thicknesses of 5-10 nm. The XRD patterns demonstrated distinct groups indexed to the (002) and (100) planes, further verifying the formation of highly crystalline nh-BN. Doped samples exhibited slightly smoother and more uniform surfaces compared to the undoped h-BN, suggesting dopant-assisted grain growth and defect passivation during thermal treatment (Perdigon-Melon et al 2002; Öz et al. 2016a).

The incorporation of metal carbonates serves as a potential improvement candidate for the physicochemical properties of nh-BN, leveraging the diverse characteristics of various metallic ions. This deliberate doping not only enhances the intrinsic qualities of nh-BN but also imparts unique attributes crucial for addressing challenges in the view of biomedical applications. The quest for innovative materials capable of harmonizing with the intricacies of biological systems has given rise to unconventional solutions in intersection of material science and biomedical research.

Adherence to internationally recognized testing frameworks is critical for any biomaterial destined for medical purpose or implant use. The series of standards encompassed by EN ISO 10993 provides a structured procedure for biological evaluation of medical compounds. For example, EN ISO 10993-12 governs sample preparation and reference materials; EN ISO 10993-5 covers in-vitro cytotoxicity testing; EN ISO 10993-3 addresses genotoxicity, carcinogenicity and reproductive toxicity. These standards provide not only uniformity but regulatory relevance, facilitating eventual translation to clinical-grade materials.

In the current study, material extracts were prepared in compliance with EN ISO 10993-12, ensuring standardized extraction conditions, sterilization, and handling. Thereafter, cytotoxicity was assessed per EN ISO 10993-5 using the MTT assay, genotoxicity per EN ISO 10993-3 following the OECD 487 micronucleus protocol, and cell migration behavior (wound healing assay) to probe fibroblast functional response by nh-BN. By adopting this rigorous approach, we aim to generate data that is comparable across the biomaterials field and support downstream risk-/benefit analyses.

Cytotoxicity Results

The 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay is a colorimetric technique first introduced by Mosmann in 1983 to evaluate cell viability following exposure to chemical, physical, or biological agents (Ferrari, Fornasiero, and Isetta, 1990; Niles et al 2007). It is primarily

employed to assess the proliferation and metabolic activity of adherent cells and has more recently been adapted for evaluating antimicrobial efficacy and microbial growth (Grela, Kozłowska and Grabowiecka, 2018; Oh and Hong, 2022).

The MTT reagent is a yellow, water-soluble tetrazolium salt composed of a positively charged mono-tetrazolium ring surrounded by four nitrogen atoms (Berridge, Herst and Tan, 2005; Ghasemi et al. 2021). The reagent readily permeates cell membranes due to its cationic nature (Sumantran 2011). Mitochondrial dehydrogenase enzymes reduce the tetrazolium salt to insoluble, purple formazan crystals within viable cells. The quantity of formazan formed is directly proportional to the number of metabolically active cells (Otang, Grierson, and Ndip, 2014). Following crystal formation, the formazan product is solubilized using an appropriate solvent such as dimethyl sulfoxide (DMSO) or isopropyl alcohol. In addition, the MTT assay technique given in literature the evaluation of cell viability standardized by EN ISO 10993-5 for determination of biocompatibility of medical devices.

In the present study, cell viability is assessed by comparing the absorbance values of the experimental groups of nh-BN samples with those of the untreated control group. The control group is defined as 100% viable, and a cell viability greater than 70% for the test material is considered indicative of in vitro biocompatibility, as specified in ISO 10993-5. While absorbance measurements provide quantitative analysis results, microscopic measurements by examining cell morphology also provide qualitative analysis results. Both analysis outcomes agreement increases the reliability as well.

According to analysis results, the corresponding cell viability ratios were determined as 148.89 ± 2.11%, 95 ± 4.4%, and 119.22 ± 3.8% at full (1/1) concentrations of the following exposure to Li-nh-BN, Ca-nh-BN, and Mg-nh-BN respectively. In Table 1, cell viability values slightly reduced and were measured as 127.00 ± 6.0%, 98.13 ± 1.23%, and 104.7 ± 0.5%, respectively when the material concentrations were reduced by half (1/2 concentration). Totally, it is understood that tested nanocrystalline hBN materials demonstrated in vitro biocompatibility as their cell viabilities remained well above the 70% threshold value, depending on the above data. Interestingly, Mg-nh-BN exhibited a stimulatory effect on cell proliferation at full concentration, while a moderate decrease in viability was observed at half concentration; however, the values remained higher than those of the control group at all diluted concentrations.

Table 1. Concentration dependent the numerical values of cytotoxic effects of Li-nh-BN, Ca-nh-BN, and Mg-nh-BN

	Li-nh-BN			Ca-nh-BN			Mg-nh-BN		
Concentration	**A**	**% V**	**SD**	**A**	**% V**	**SD**	**A**	**% V**	**SD**
1	0.194	119.224	3.815	0.155	95.445	1.497	0.242	148.891	2.197
1/2	0.170	104.750	0.581	0.160	98.135	1.235	0.207	127.007	6.358
1/4	0.160	98.092	0.891	0.154	94.789	0.716	0.194	119.124	1.111
1/8	0.161	99.085	0.725	0.153	94.000	0.883	0.191	117.472	1.217

A: Absorbance; % V: percent viability; SD: Standard Deviation

A comparative summary of the absorbance results for all concentrations and the control group is presented in Fig. 1, and the corresponding percentage viability values resulting from serial dilution of the full concentrations are provided in Table 1. In addition, microscopic observations further confirmed that cells maintained normal morphology following exposure to all materials. The cells exhibited intact cytoplasmic boundaries and no evidence of morphological damage, supporting the conclusion that the materials are qualitatively biocompatible.

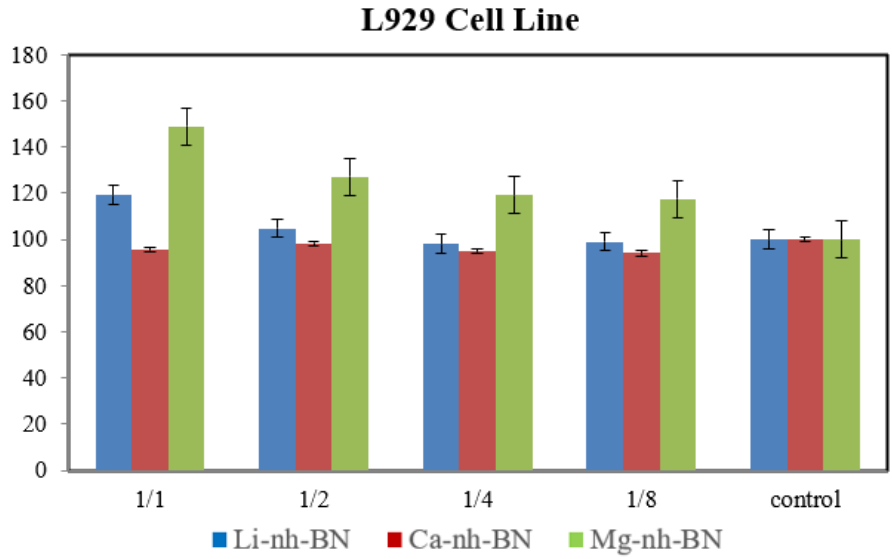


Fig. 1. The cytotoxic effects of Li-nh-BN, Ca-nh-BN, and Mg-nh-BN at different concentrations applied on L929 cells by MTT assay

Genotoxicity Results

The full-concentration extracts of Li-nh-BN, Ca-nh-BN, and Mg-nh-BN were applied to Chinese Hamster Ovary (CHO) cells, and the numbers of binucleated cells as well as binucleated cells containing micronuclei are presented in **Table 2**.

Table 2. In vitro micronucleus test results in CHO cells treatment of BN Li-nh-BN, Ca-nh-BN, and Mg-nh-BN

	Concentration (1/1)			**Control Group**	
	Li-nh-BN	Ca-nh-BN	Mg-nh-BN	N	P
MNBN (%)*	0.6160	0.5086	0.7136	0.6764	9.7708
CBPI**	1.6543	1.6578	1.6619	1.7005	1.5377

*MNBN: Mean micro nucleated binucleated cell frequency; **Cytokinesis-Block Proliferation Index; N: Negative Control; P: Positive Control

In the assay, mitomycin C (Mit-C) was used as a positive control for inducing genotoxic effects in CHO cells. Upon evaluation of the results, no significant difference was observed between the micronucleus frequencies in binucleated cells treated with Li-nh-BN, Ca-nh-BN, and Mg-nh-BN and those of the negative control group. In contrast, the positive control group exhibited a markedly higher frequency of micronucleated binucleated cells (9.77%). The corresponding micronucleus frequencies for Li-nh-BN, Ca-nh-BN, and Mg-nh-BN were determined as 0.71%, 0.50%, and 0.61%, respectively. Representative microscopic images of CHO cells which were analyzed by light miscroscope a Leica inverted fluorescent microscope (400 × magnification) exposed to the full concentrations of each material are shown in **Fig. 2**. These findings indicate that exposure to Li-nh-BN, Ca-nh-BN, and Mg-nh-BN does not induce genotoxic effects under the tested conditions, further supporting the in vitro biocompatibility of these materials.

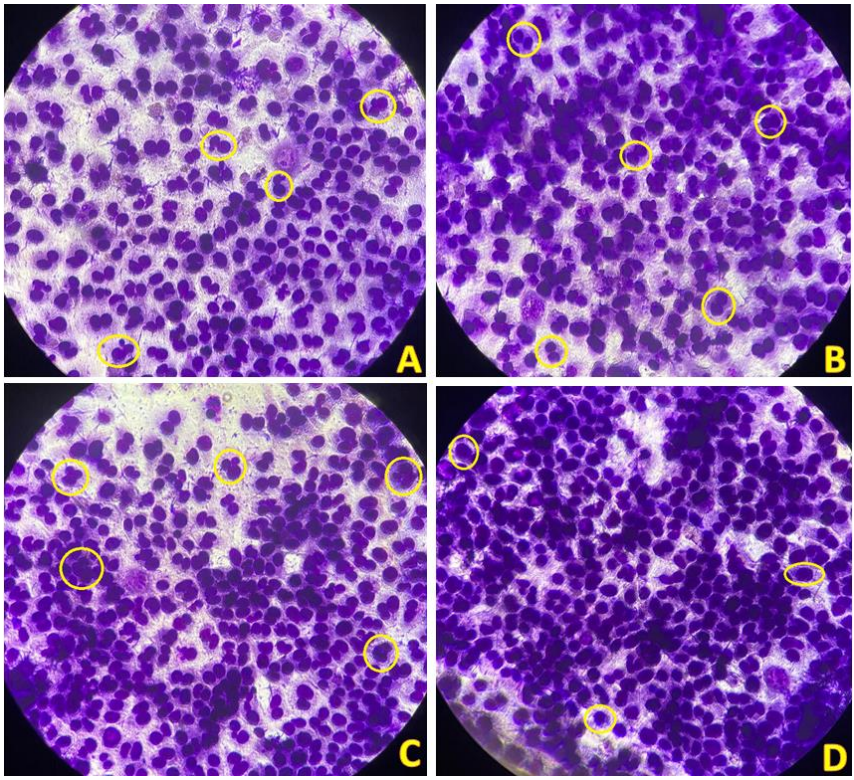


Fig. 2. Microscope image of (A) CHO cells treated with Mitomycin C as positive control (circles indicate multinucleated cells) (B) CHO cells exposed to the 1/1 Li-nh-BN (circles indicate binuclear cells), (C) CHO cells exposed to the 1/1 Mg-nh-BN (circles indicate multinuclear cells) (D) CHO cells exposed to the 1/1 Ca-nh-BN (circles indicate single cells)

Scratch assay (cell migration)

The applied method is based on the creation of an artificial gap on a confluent monolayer of cells using the scratch assay technique. This approach enables direct observation of the ability of cells at the wound edges to migrate and close the artificially generated gap, effectively simulating an *in vitro* wound-healing model. Compared with many other cell-based assays, the scratch assay is straightforward to perform and offers several advantages. It mimics *in vivo* cell migration to a considerable extent and is particularly well suited for investigating the regulation of cell motility governed by extracellular matrix (ECM) components and cell–cell interactions (Chigurupati et al. 2013).

The dynamics of cell migration and wound closure rates were quantitatively evaluated 24 hours post-scratch as summarized in **Fig. 3**. Representative microscopic images illustrating the wound closure process for each material are

presented in **Fig. 4**. The results revealed that the Ca-nh-BN sample exhibited a markedly high wound-healing potential, characterized by extensive cell migration into the denuded area. The Li-nh-BN sample demonstrated an even more accelerated response, showing rapid closure within the first 8 hours and achieving complete closure by the 24-hour mark.

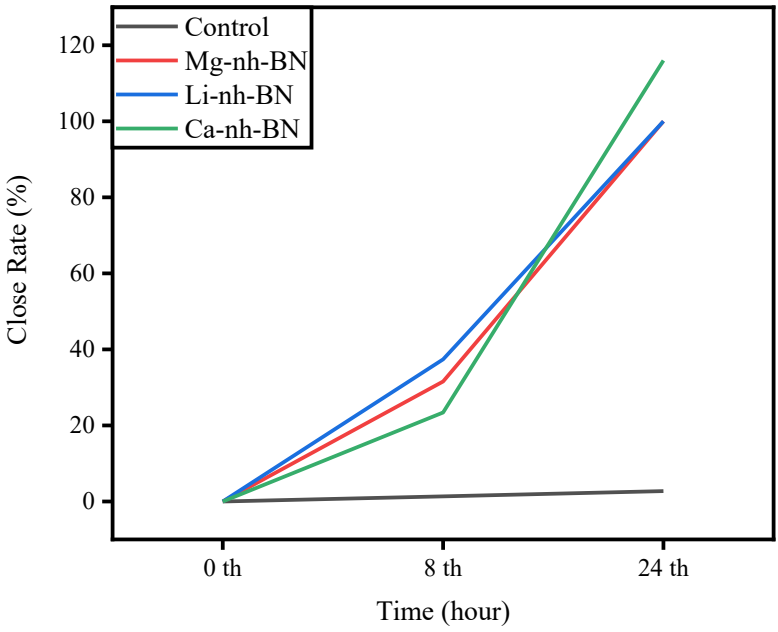


Fig. 3. Quantitative assessment of *in vitro* wound closure percentages for L929 fibroblasts treated with Mg-, Li-, and Ca-doped nh-BN nanomaterials over a 24-hour period.

The enhanced cell migration and wound closure observed for the Ca- and Li-doped h-BN samples can be attributed to dopant-induced modulation of cellular behavior. Calcium ions (Ca^{2+}) play a fundamental role in cellular signaling pathways that regulate cytoskeletal reorganization, focal adhesion formation, and integrin-mediated cell motility (Tsai et al. 2015). Similarly, lithium ions (Li^{+}) are known to influence the inhibition of glycogen synthase kinase-3β (GSK-3β) and the activation of downstream Wnt/β-catenin signaling (Zhu et al. 2014). Both pathways contribute significantly to enhanced cellular proliferation and migration during tissue repair. The observed differences between the Ca- and Li-doped BN samples suggest that dopant composition directly influences the dynamic interplay between proliferation and migration mechanisms. Communally, these findings indicate that dopant-engineered nh-BN

nanomaterials, particularly Li-nh-BN and Ca-nh-BN, promote favorable cellular responses that accelerate *in vitro* wound closure. Such behavior implies high potential for these materials in biomedical coatings, wound-healing scaffolds, and tissue-regenerative interfaces where controlled stimulation of cell motility is essential.

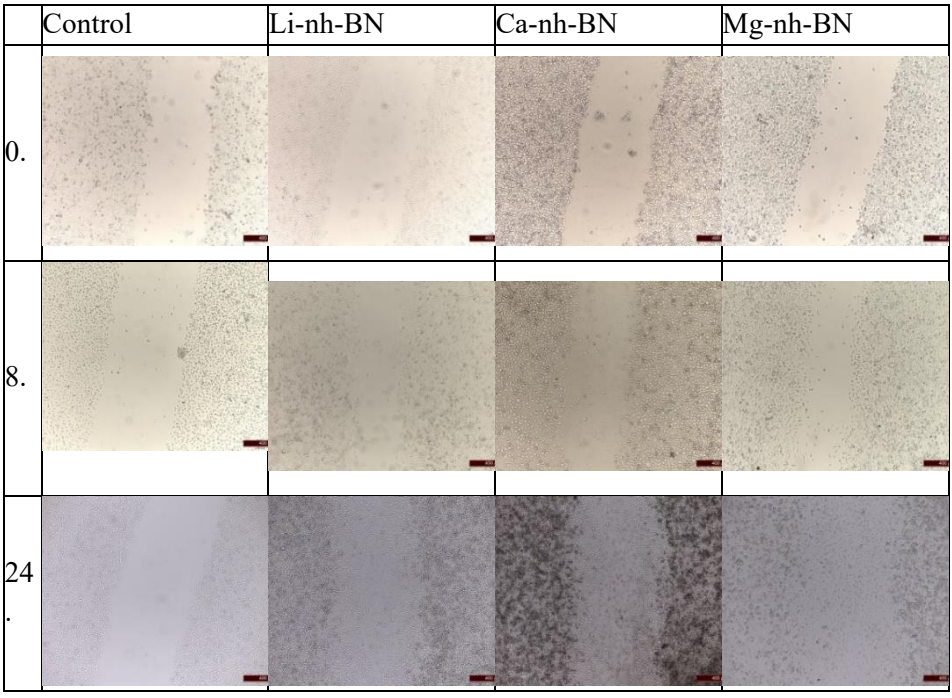


Fig. 4. Time-dependent closure image of the wound areas of the materials as a result of in vitro wound closure and cell migration test after 24 hours of application

The improved *in vitro* response of dopant-assisted nh-BN nanocrystals can be endorsed as the double effects of enhanced crystallinity, optimized surface chemistry, and the presence of biologically relevant metal ions. Besides physicochemical positive alterations, nh-BN contributes to an optimal balance of hydrophilicity and surface charge, promoting protein adsorption and cell adhesion without eliciting inflammatory responses (Darabdhara et al. 2025). Collectively, these results demonstrate that dopant-mediated solid-state synthesis of nh-BN can tune its biointerfacial properties without compromising structural stability.

In summary, the incorporation of metal carbonates as dopant sources during nh-BN synthesis led to the formation of highly crystalline, well-ordered

nanocrystals with excellent in vitro biocompatibility. $MgCO_3$ doped nh-BN exhibited superior cellular viability and protein adsorption behavior, while the other doped samples maintained comparable biocompatibility with slightly enhanced surface uniformity (Pouroutzidou et al. 2021). It is found that the wound closure potential is quite high by $CaCO_3$ doped sample while Li_2CO_3 doped sample shows complete healing after 24 hours. These findings suggest that dopant-assisted h-BN nanocrystals hold strong potential for biomedical facilities such as bioinert coatings, implant surface modifiers, and bone tissue engineering.

4. Conclusion

In this study, in vitro assessment of well-crystallized nanocrystalline hexagonal boron nitride (nh-BN) was investigated which were formerly synthesized using dopant metal carbonates as structural promoters. The incorporation of these dopants significantly influenced the crystallization dynamics, leading to enhanced phase purity, increased crystallinity, and refined nanoscale morphology without altering the intrinsic hexagonal BN structure.

Biological evaluations, conducted in strict adherence to EN ISO 10993 and OECD 487 standards, provided compelling evidence of the materials' biocompatibility. The synthesized nh-BN samples exhibited excellent cytocompatibility, with L929 cell viability as above 90 % consistently exceeding the 70 % threshold and showed no signs of genotoxic damage or hemolytic activity. Notably, the Ca and Li-doped variants demonstrated superior functional bioactivity, significantly accelerating *in vitro* wound healing and cell migration. These effects are attributed to the activation of key intracellular signaling pathways, such as the Wnt/β-catenin and integrin-mediated pathways, triggered by the release of bioactive cations.

Overall, this research establishes dopant-engineered nh-BN as a dynamic and safe platform for next-generation biomedical applications. The ability to fine-tune the physicochemical properties and biological interactions of h-BN through deliberate doping opens new horizons for its use in bioactive coatings, tissue engineering scaffolds, and regenerative medicine. Future studies should focus on evaluating long-term cell responses, osteogenic differentiation potential, and in vivo biointegration to validate these nanomaterials for use in implant coatings, bone regeneration scaffolds, and bioinert protective layers.

References

Ansaloni LMS, de Sousa EMB (2013) Boron nitride nanostructured: Synthesis, characterization and potential use in cosmetics. Mater Sci Appl 4(01): 22-28. https://doi.org/10.4236/msa.2013.41004

Atila A, Halici Z, Cadirci E, Karakus E, Palabiyik SS, Ay N, Bakan F, Yilmaz S (2016) Study of the boron levels in serum after implantation of different ratios nano-hexagonal boron nitride–hydroxy apatite in rat femurs. Mater Sci Eng C 58:1082-1089. https://doi.org/10.1016/j.msec.2015.09.041

Bahrami E, Barenji AD, Khabazi ME, Chermahini AN (2025) Boron nitride nanosheets: a comprehensive review of synthesis and multifunctional applications. Result Chem 102655. https://doi.org/10.1016/j.rechem.2025.102655

Berridge MV, Herst PM, Tan AS (2005) Tetrazolium dyes as tools in cell biology: new insights into their cellular reduction. Biotechnol Annu Rev 11:127-152. https://doi.org/10.1016/S1387-2656(05)11004-7

Chigurupati S, Mughal MR, Okun E, Das S, Kumar A, McCaffery M, Seal S, Mattson MP (2013) Effects of cerium oxide nanoparticles on the growth of keratinocytes, fibroblasts and vascular endothelial cells in cutaneous wound healing. Biomaterials 34(9):2194-2201. https://doi.org/10.1016/j.biomaterials.2012.11.061

Darabdhara G, Borthakur P, Boruah PK, Sen S, Pemmaraju DB, Das MR (2025) Advancements in two-dimensional boron nitride nanostructures: properties, preparation methods, and their biomedical applications. J Mater Chem B 13(37):11540-11581. https://doi.org/10.1039/D5TB00469A

Ferrari M, Fornasiero MC, Isetta AM (1990) MTT colorimetric assay for testing macrophage cytotoxic activity in vitro. J Immunol Methods 131(2):165-172. https://doi.org/10.1016/0022-1759(90)90187-Z

Ghasemi M, Turnbull T, Sebastian S, Kempson I (2021) The MTT assay: utility, limitations, pitfalls, and interpretation in bulk and single-cell analysis. Int J Mol Sci 22(23):12827. https://doi.org/10.3390/ijms222312827

Grela E, Kozłowska J, Grabowiecka A (2018) Current methodology of MTT assay in bacteria–A review. Acta Histochem 120(4):303-311. https://doi.org/10.1016/j.acthis.2018.03.007

Haubner R, Wilhelm M, Weissenbacher R, Lux B (2002) Boron nitrides Properties synthesis and applications In Jansen M (ed) High Performance Non-Oxide Ceramics II Structure and Bonding Vol 102 Springer Berlin Heidelberg https://doi.org/10.1007/3-540-45623-6_1

Jenifer A, Senthilarasan K, Arumugam S, Sivaprakash P, Sagadevan S, Sakthivel P (2021) Investigation on antibacterial and hemolytic properties of

magnesium-doped hydroxyapatite nanocomposite. Chem Phys Lett 771:138539. https://doi.org/10.1016/j.cplett.2021.138539

Karan T, Erenler R, Moran Bozer B (2022) Synthesis and characterization of silver nanoparticles using curcumin: cytotoxic, apoptotic, and necrotic effects on various cell lines. Z Naturforsch C 77(7-8):343-350. https://doi.org/10.1515/znc-2021-0298

Kumar AM, Khan A, Hussein MA, Khan MY, Dafalla H, Suresh B, Ramakrishna S (2022) Hybrid nanocomposite coatings from PEDOT and BN-TiO2 nanosheets: Enhanced invitro corrosion resistance, wettability and biocompatibility for biomedical applications. Prog Org Coat 170:106946. https://doi.org/10.1016/j.porgcoat.2022.106946

Lin Y, Connell JW (2012) Advances in 2D boron nitride nanostructures: nanosheets, nanoribbons, nanomeshes, and hybrids with graphene. Nanoscale 4(22):6908-6939. https://doi.org/10.1039/C2NR32201C

Ma Z, Zuo J, Tang C, Wang P, Shi C (2020) Physical properties of a novel phase of boron nitride and its potential applications Mater Chem Phys 252:123245 https://doi.org/10.1016/j.matchemphys.2020.123245

Matović B, Luković J, Nikolić M, Babić B, Stanković N, Jokić B, Jelenković B (2016) Synthesis and characterization of nanocrystaline hexagonal boron nitride powders: XRD and luminescence properties. Ceram Int 42(15):16655-16658. https://doi.org/10.1016/j.ceramint.2016.07.096

Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods 65(1-2):55-63.

Niles AL, Moravec RA, Hesselberth PE, Scurria MA, Daily WJ, Riss TL (2007). A homogeneous assay to measure live and dead cells in the same sample by detecting different protease markers. Anal Biochem 366(2):197-206. https://doi.org/10.1016/j.ab.2007.04.007

Oh YJ, Hong J (2022) Application of the MTT-based colorimetric method for evaluating bacterial growth using different solvent systems. Lwt 153:112565. https://doi.org/10.1016/j.lwt.2021.112565

Otang WM, Grierson DS, Ndip RN (2014) Cytotoxicity of three South African medicinal plants using the Chang liver cell line. Afr J Tradit Complement Altern Med 11(2):324-329. http://dx.doi.org/10.4314/ajtcam.v11i2.16

Öz M, Saritekin NK, Bozkurt Ç, Budak E, Yildirim G (2016b) Effect of Li2CO3 on formation temperature of hBN by modified O'Connor model Cryst Res Technol 51 (12):708–717 https://doi.org/10.1002/crat.201600138

Öz M, Saritekin NK, Bozkurt Ç, Yildirim G (2016a) Synthesis of highly ordered hBN in presence of group I/IIA carbonates by solid state reaction Cryst Res

Technol 51 (6):380–392 https://doi.org/10.1002/crat.201500337

Pakdel A, Zhi C, Bando Y, Golberg D (2012) Low-dimensional boron nitride nanomaterials. Mater Today 15(6):256-265. https://doi.org/10.1016/S1369-7021(12)70116-5

Perdigon-Melon JA, Auroux A, Cornu D, Miele P, Toury B, Bonnetot B (2002) Porous boron nitride supports obtained from molecular precursors Influence of the precursor formulation and of the thermal treatment on the properties of the BN ceramic J Organomet Chem 657 (1–2):98–106 https://doi.org/10.1016/S0022-328X(02)01589-9

Pouroutzidou GK, Liverani L, Theocharidou A, Tsamesidis I, Lazaridou M, Christodoulou E, Beketova A, Pappa C, Triantafyllidis KS, Anastasiou AD, Papadopoulou L, Bikiaris DN, Boccaccini AR, Kontonasaki E (2021) Synthesis and characterization of mesoporous mg-and sr-doped nanoparticles for moxifloxacin drug delivery in promising tissue engineering applications. *Int J Mol Sci 22*(2):577. https://doi.org/10.3390/ijms22020577

Ramirez Leyva JH, Vitale G, Hethnawi A, Hassan A, Perez-Zurita MJ, Ruiz-Esparza GU, Nassar NN (2018) Mechanism of hierarchical porosity development in hexagonal boron nitride nanocrystalline microstructures for biomedical and industrial applications. ACS Appl Nano Mater 1(9):4491-4501. https://doi.org/10.1021/acsanm.8b00765

Roosta S, Nikkhah SJ, Sabzali M, Hashemianzadeh SM (2016) Molecular dynamics simulation study of boron-nitride nanotubes as a drug carrier: from encapsulation to releasing. RSC Advances 6(11):9344-9351. https://doi.org/10.1039/C5RA22945F

Sumantran VN (2011) Cellular chemosensitivity assays: an overview. Cancer Cell Culture: Methods and Protocols 219-236.

Tozar A, Karahan İH (2018) A comparative study on the effect of collagen and h-BN reinforcement of hydroxyapatite/chitosan biocomposite coatings electrophoretically deposited on Ti-6Al-4V biomedical implants. Surf Coat Technol 340:167-176. https://doi.org/10.1016/j.surfcoat.2018.02.034

Tsai FC, Kuo GH, Chang SW, Tsai PJ (2015) Ca2+ signaling in cytoskeletal reorganization, cell migration, and cancer metastasis. *Biomed Res Int 2015*(1):409245. https://doi.org/10.1155/2015/409245

Turhan EA, Pazarçeviren AE, Evis Z, Tezcaner A (2022) Properties and applications of boron nitride nanotubes. Nanotechnol 33(24):242001. https://doi.org/10.1088/1361-6528/ac5839

Weimer AW (2012) Carbide, nitride and boride materials synthesis and processing. Springer Science & Business Media.

Zhu Z, Yin J, Guan J, Hu B, Niu X, Jin D, Wang Y, Zhang, C. (2014). Lithium stimulates human bone marrow derived mesenchymal stem cell proliferation through GSK-3β-dependent β-catenin/Wnt pathway activation. *The FEBS journal*, *281*(23):5371-5389. https://doi.org/10.1111/febs.13081

Chapter 4

The Effect of Distance Geometry on Spectral Clustering: An Analysis through the Minkowski Parameter

Gülay İlona TELSİZ KAYAOĞLU[1], Mustafa EROĞLU[2]

1. Introduction

Clustering is one of the fundamental tasks in unsupervised learning and aims to partition a set of observations into groups according to their similarities. Since no predefined class labels are available in unsupervised learning problems, the relationships among observations play a central role in determining the resulting groups. Distance and similarity measures are therefore important components of clustering methods, as they provide a way of quantifying the degree of closeness or dissimilarity between observations.

Spectral clustering is a graph-based clustering approach that represents observations as vertices of a graph and uses their pairwise similarities to determine the structure of the graph. The method is based on concepts from graph theory and linear algebra and provides an alternative to clustering approaches that operate directly in the original feature space. In particular, spectral clustering can be useful for datasets whose cluster structures are difficult to identify using conventional clustering methods.

The construction of the similarity graph is one of the fundamental stages of the spectral clustering procedure. Given a set of observations, a similarity matrix is constructed to represent the relationships between pairs of observations. This matrix is subsequently used to obtain an appropriate graph Laplacian. The spectral properties of the Laplacian matrix, particularly its eigenvalues and eigenvectors, are then used to transform the original observations into a representation in which the underlying cluster structure can be more clearly identified. The resulting representation can finally be partitioned using a conventional clustering algorithm such as k-means.

The distance measure used when constructing the similarity relationships is therefore an important consideration in spectral clustering. In the standard

[1] Assistant Professor, Mimar Sinan Fine Arts University, Faculty of Arts and Sciences, Department of Mathematics, Istanbul, TURKEY, gulay.telsiz@msgsu.edu.tr, ORCID: 0000-0001-8926-2899
[2] Master's Degree Graduate, Mimar Sinan Fine Arts University, Institute of Science, Department of Mathematics, Istanbul, TURKEY, eroglumustafa208@gmail.com, ORCID: 0009-0006-3040-6169

formulation, the Euclidean distance is frequently employed. However, there is no general reason to assume that Euclidean distance will provide the most suitable representation of similarity for every dataset. Different distance measures can assign different levels of importance to coordinate-wise differences and can consequently lead to different similarity graphs and different clustering results.

The effect of alternative distance measures on clustering performance has previously been investigated, particularly in the context of k-means clustering. For example, different distance measures have been considered as alternatives to the standard Euclidean distance in clustering applications. In our previous study, this issue was examined specifically for spectral clustering. Euclidean, Manhattan (Cityblock), Chebyshev, Canberra, Cosine, and Bray–Curtis distance functions were compared using three different spectral clustering algorithms and three datasets.

That study was motivated by datasets for which the k-means algorithm did not provide satisfactory separation. The purpose was to investigate whether replacing the standard Euclidean distance with alternative distance functions within the spectral clustering framework could improve the resulting clustering performance. The results indicated that the choice of distance function can have a substantial effect on clustering quality. Among the six distance functions considered, the Canberra distance produced the lowest performance on the three datasets, whereas the Manhattan (Cityblock) distance consistently produced the highest clustering performance.

The findings of this previous study raise a further question: rather than comparing several individual distance functions, can the effect of distance geometry on spectral clustering be investigated through a single parameterized family of distances? This question motivates the present chapter.

In this study, the Minkowski distance is considered with different values of its parameter p. The Minkowski distance provides a general framework for measuring the distance between observations and allows the geometry of the distance measure to be modified by changing p. Consequently, investigating different p-values provides an opportunity to examine the sensitivity of spectral clustering to changes in the underlying distance geometry within a unified mathematical framework.

The Minkowski exponent has previously been investigated in different clustering frameworks. De Amorim and Mirkin [6] incorporated the Minkowski metric into weighted k-means, while de Amorim et al. [7] investigated the selection of an appropriate Minkowski exponent based on clustering solutions obtained for different values of p. However, comparatively less attention has been given to systematically examining the effect of p within spectral clustering.

The main objective of this chapter is to investigate how different Minkowski p-values affect the performance of spectral clustering. The same general spectral clustering framework considered in the previous study is employed, while the distance component is extended through the systematic consideration of different Minkowski parameters. To provide a common basis for comparison across different \(p\)-values, the resulting spectral representations are evaluated using the Silhouette score computed with Euclidean distance. Thus, while the Minkowski parameter is varied in the construction of the similarity graph, the distance measure used for evaluating the resulting spectral embeddings is kept fixed.

This extension is particularly relevant because the parameter p provides a systematic way of moving between different notions of distance. Instead of viewing individual distance measures only as separate alternatives, the Minkowski family makes it possible to study changes in distance geometry within a common mathematical framework. By evaluating the resulting spectral embeddings under a fixed reference metric, the analysis focuses on how changes in the Minkowski parameter affect the geometry and separation of the spectral representation. The analysis presented in this chapter consequently complements the previous comparison of individual distance functions and provides a more focused investigation of the relationship between distance geometry and spectral clustering performance.

2. Spectral Clustering

Spectral clustering is a graph-based clustering approach in which observations are represented as vertices and their relationships are encoded by a similarity matrix. Spectral information obtained from the graph Laplacian is then used to transform the original data into a representation in which the cluster structure can be identified more effectively. Unlike methods such as k-means, which rely directly on distances to cluster centers, spectral clustering can therefore be particularly useful for datasets with non-convex or otherwise complex cluster structures [2].

Let $x_1, x_2, \dots, x_n$ denote n observations represented as the vertices of a weighted graph $G = (V, E)$. The corresponding similarity matrix is denoted by $W = \left(w_{ij}\right)_{n \times n}$, where w_{ij} represents the similarity between x_i and x_j. Since the similarity matrix determines the graph structure and consequently the graph Laplacian, the procedure used to calculate pairwise similarities can affect the resulting spectral representation and clustering solution.

2.1. Graph Laplacian

The unnormalized graph Laplacian is defined as $L = D - W$. The Laplacian matrix encodes the connectivity structure of the graph and plays a central role in spectral clustering. Its eigenvalues and corresponding eigenvectors contain information about the structure of the graph and can be used to identify groups of strongly connected observations.

The degree matrix D is defined as a diagonal matrix whose entries are given by $d_i = \sum_{j=1}^{n} w_{ij}$.

Thus, D is defined as the diagonal matrix with $d_1, d_2, \dots, d_n$ on its diagonal and 0 elsewhere.

The spectral clustering procedure exploits the eigenvectors associated with the smallest eigenvalues of the Laplacian. More specifically, if k clusters are to be constructed, the first k eigenvectors corresponding to the smallest k eigenvalues are selected. If these eigenvectors are denoted by $u_1, u_2, \dots, u_k$, they are arranged as the columns of the matrix $U = [u_1, u_2, \dots, u_k]$.

For each observation x_i, the i-th row of U defines a new representation $y_i \in \mathbb{R}^k$.

The observations in this spectral representation are subsequently clustered using the k-means algorithm.

Thus, the overall procedure can be summarized as:
Data → Similarity Graph → Laplacian → Eigenvectors → Spectral Representation → k-means.

This procedure is the common framework underlying the spectral clustering algorithms considered in the present work. The previous study employed three versions of spectral clustering and examined their performance under six different distance functions.

2.2. Normalized and Unnormalized Spectral Clustering

Different spectral clustering algorithms can be distinguished according to the Laplacian matrix used in the spectral decomposition. In this study, three approaches are considered: unnormalized spectral clustering, Shi-Malik normalized spectral clustering, and Ng-Jordan-Weiss normalized spectral clustering. The distinction between these approaches arises from the different forms of the graph Laplacian used in the clustering procedure.

2.2.1. Unnormalized Spectral Clustering

The unnormalized spectral clustering approach uses the standard graph Laplacian $L = D - W$.

The first k eigenvectors corresponding to the smallest eigenvalues of L are calculated. These eigenvectors are assembled into the matrix U, and the rows of U are subsequently clustered using k-means.

The procedure therefore consists of constructing the similarity graph, computing the unnormalized Laplacian, obtaining its first k eigenvectors, and applying k-means to the resulting spectral representation.

2.2.2. Shi–Malik Normalized Spectral Clustering

The normalized approach proposed by Shi and Malik [3] uses the random-walk normalized Laplacian $L_{rw} = D^{-1}L = I - D^{-1}W$.

The use of a normalized Laplacian incorporates the degree information of the graph into the spectral representation. The eigenvectors associated with the smallest eigenvalues of L_{rw} are used to obtain the representation on which the final clustering is performed.

2.2.3. Ng-Jordan-Weiss Normalized Spectral Clustering

The third approach considered is the normalized spectral clustering algorithm associated with Ng, Jordan, and Weiss [4]. This method employs the symmetric normalized Laplacian $L_{sym} = D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$.

The eigenvectors corresponding to the smallest eigenvalues are calculated from this normalized Laplacian. The resulting spectral representation is then used as the input to the k-means algorithm.

The three approaches considered in this chapter therefore share the same general spectral clustering framework but differ in the Laplacian matrix used to represent the graph structure. The algorithmic steps given in the previous study are based on constructing the similarity graph, computing the corresponding Laplacian, obtaining the first k eigenvectors, forming the spectral representation, and clustering these representations using k-means.

This distinction is important for the analysis presented later in the chapter. Since the similarity matrix is constructed from a distance-based measure, changing the distance function changes the graph itself. In the present study, this effect is investigated systematically by replacing a single fixed distance measure with the **Minkowski distance for different p-values**. Thus, both the graph construction and the subsequent spectral representation may vary as p changes.

3. Distance Measures in Spectral Clustering: Previous Results and Motivation for the Minkowski Metric

The choice of distance measure is particularly important in spectral clustering because pairwise distances influence the construction of the similarity graph and, consequently, the graph Laplacian and its spectral representation.

In our previous study [1], six distance functions were compared using three spectral clustering approaches on datasets for which conventional k-means did not provide satisfactory separation. The results showed that the choice of distance function can substantially affect clustering performance, with Manhattan distance providing the highest performance among the considered measures.

These findings motivate the present extension. Rather than comparing individual distance functions independently, the present study considers the Minkowski family, in which the parameter p provides a systematic way of varying the underlying distance geometry. This allows the effect of distance geometry on spectral clustering to be investigated within a common mathematical framework.

3.1. Minkowski Distance

Let $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ be two points in $\mathbb{R}^n$. The Minkowski distance of order p between x and y is defined as

$$d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p},$$

where $p \geq 1$.

The parameter p determines the geometry of the resulting distance. Thus, different choices of p can lead to different pairwise distances even when the same observations are considered. This property makes the Minkowski family particularly useful for studying the effect of distance geometry on clustering algorithms.

Two well-known distance measures arise as special cases of the Minkowski distance. For $p = 1$, the expression becomes the Manhattan distance:

$$d_1(x, y) = \sum_{i=1}^{n} |x_i - y_i|.$$

For $p = 2$, it becomes the Euclidean distance:

$$d_2(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2}.$$

Furthermore, as p approaches infinity, the Minkowski distance approaches the Chebyshev distance:

$$d_\infty(x,y) = max_i|x_i - y_i|.$$

Consequently, the Minkowski family provides a unified framework that contains several commonly used distance measures as particular or limiting cases. This property allows the influence of the distance geometry to be investigated by varying a single parameter rather than changing between unrelated distance definitions.

3.2. Minkowski Distance in the Spectral Clustering Framework

In the spectral clustering framework, the Minkowski distance is used to determine the pairwise relationships between observations. For a selected value of p, the distance $d_p(x_i, x_j)$ is calculated for pairs of observations. These distances are then used in the construction of the similarity graph according to the selected similarity scheme.

Therefore, changing p can alter the entries of the similarity matrix and, consequently, the structure of the weighted graph. Since the graph determines the corresponding Laplacian matrix, changes in p can propagate through the subsequent stages of the spectral clustering procedure:

$$d_p \rightarrow W \rightarrow L \rightarrow \text{eigenvectors} \rightarrow \text{spectral representation} \rightarrow \text{clusters}.$$

This relationship provides the basis for the analysis performed in the following sections. The aim is to determine how different choices of the Minkowski parameter affect the resulting spectral clustering performance and whether certain values of p provide more appropriate representations of the considered datasets.

The following section presents the experimental design and performance evaluation used for the selected Minkowski p-values.

4. Experimental Design and Performance Evaluation

The experimental analysis was designed to investigate the effect of the Minkowski parameter p on spectral clustering performance. The same three

datasets considered in our previous study [1] were used in order to maintain a consistent experimental framework. These datasets were originally selected because their cluster structures could not be satisfactorily separated using the conventional k-means algorithm. Detailed descriptions of the datasets are available in [1].

For each dataset, the Minkowski distance was used in the construction of the similarity graph, and the parameter p was varied from 1 to 10. Thus, ten different distance geometries were considered for each spectral clustering approach. In particular, $p = 1$ corresponds to the Manhattan distance and $p = 2$ corresponds to the Euclidean distance, while larger values of p provide intermediate cases approaching the Chebyshev distance.

The three spectral clustering approaches considered were the unnormalized spectral clustering method based on L, the Shi-Malik normalized method based on L_{rw}, and the Ng-Jordan-Weiss normalized method based on L_{sym}. For each value of p, the corresponding similarity graph and graph Laplacian were constructed, the spectral representation was obtained from the relevant eigenvectors, and the final cluster assignments were determined by k-means.

4.1. Performance Evaluation

Clustering performance was evaluated using the Silhouette score [5]. For each observation i, let a_i denote the average distance between i and the other observations belonging to the same cluster, and let b_i denote the minimum average distance between i and the observations belonging to any other cluster. The Silhouette coefficient of observation i is defined as

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}$$

The overall Silhouette score is obtained by averaging $s(i)$ over all observations. Values close to 1 indicate well-separated clusters, values around 0 indicate overlapping clusters, and negative values indicate possible misclassification.

Since the Minkowski parameter p is the experimental variable in the present study, changing the distance metric used for performance evaluation together with p would make direct comparisons across different values of p more difficult. Therefore, a fixed evaluation metric was used throughout the experiments.

For each value of p, the Silhouette score was computed on the final spectral embedding used as input to k-means, with the resulting k-means assignments

used as cluster labels. Euclidean distance was used in all Silhouette calculations, regardless of the value of p used in constructing the similarity graph. Thus, the spectral representations obtained for different values of p were evaluated using a common distance criterion.

5. Results and Discussion

The results obtained for the three datasets show that the influence of the Minkowski parameter p on the spectral representation depends strongly on the structure of the data. While changes in p produced almost no observable effect for the samples and face datasets, a clear dependence on the Minkowski parameter was observed for the noisy moons dataset. This general pattern was consistent across the three spectral clustering approaches based on L, L_{rw}, L_{sym}.

For both the samples and face datasets, the Silhouette scores were approximately equal to 1.000 for all values of p from 1 to 10 and for all three Laplacian formulations. The very small differences observed among these scores occurred only at numerical precision level and were therefore not regarded as meaningful differences. Since no practically relevant variation was observed, these results are not tabulated separately.

The noisy moons dataset reveals a markedly different response to the Minkowski parameter. As illustrated in Figure 5.1, this dataset has a nonlinear and non-convex cluster structure. The Silhouette scores obtained for different values of p are presented in Table 5.1.

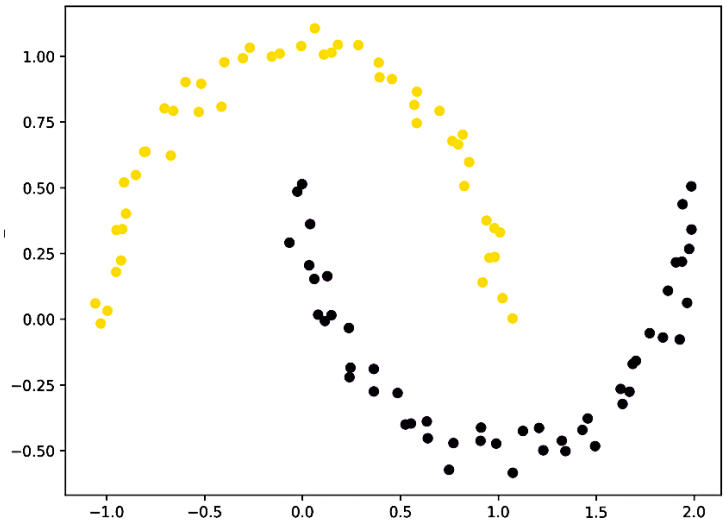


Figure 5.1. Noisy moons dataset with its two underlying cluster structures.

p	L	L_{sym}	L_{rw}
1	0.662	0.660	0.660
2	**1.000**	**1.000**	**1.000**
3	0.809	0.809	0.809
4	0.807	0.806	0.806
5	0.804	0.803	0.803
6	0.804	0.803	0.803
7	0.804	0.804	0.804
8	0.805	0.805	0.805
9	0.807	0.806	0.806
10	0.807	0.806	0.806

Table 5.1. Silhouette scores of the spectral embeddings for the noisy moons dataset using Euclidean distance.

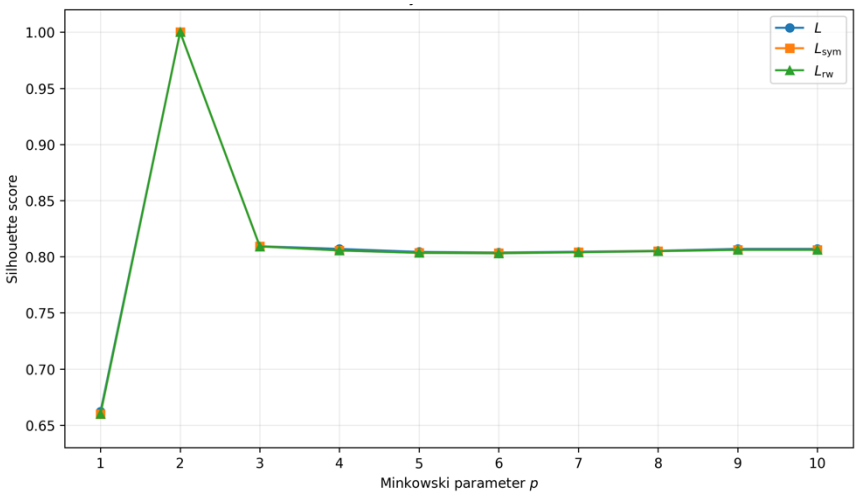


Figure 5.2. Silhouette scores of the spectral embeddings for the noisy moons dataset for different Minkowski parameter values, evaluated using Euclidean distance.

As shown in Table 5.1 and Figure 5.2, the noisy moons dataset exhibits a clear dependence on the Minkowski parameter p. For $p = 1$, the Silhouette scores are approximately 0.66 for all three spectral clustering approaches. A substantial increase is observed at $p = 2$, where the Silhouette scores reach approximately 1.00. For $p = 3$, the scores decrease to approximately 0.81, while for $p \geq 4$ they

remain relatively stable around 0.80. These results indicate that the choice of p affects the geometry of the spectral embedding, with $p = 2$ producing the most clearly separated representation for the noisy moons dataset under the common Euclidean evaluation criterion.

The differences among the three Laplacian formulations are comparatively small for each value of p. In particular, L, L_{rw}, L_{sym} exhibit very similar Silhouette scores across the entire range of p. This suggests that, for the noisy moons dataset, the choice of the Minkowski parameter has a more pronounced effect on the spectral representation than the choice among the three Laplacian formulations considered in this study.

These findings indicate that the effect of the Minkowski parameter is data-dependent rather than uniform. For the samples and face datasets, the spectral representations were essentially insensitive to the choice of p within the investigated range. In contrast, the noisy moons dataset exhibited a clear dependence on p, with $p = 2$ producing a substantially better-separated spectral representation under the common Euclidean evaluation criterion.

An important aspect of this comparison is that the same Euclidean distance metric was used to evaluate all spectral embeddings. Therefore, differences in the Silhouette scores reflect changes in the geometry and separation of the spectral representations induced by different values of the Minkowski parameter, rather than changes in the distance metric used for evaluation.

Overall, the results suggest that varying the Minkowski parameter does not have a uniform effect across datasets. While the samples and face datasets remained largely insensitive to p, the geometry of the spectral representation for the noisy moons dataset changed substantially, with $p = 2$ yielding the clearest separation. These empirical findings indicate that the geometry induced by the Minkowski parameter can influence the spectral representation obtained by spectral clustering. A theoretical investigation of the relationship between the Minkowski parameter, the similarity graph, and the resulting spectral representation could provide a more fundamental explanation of the observed behavior.

References

[1] Telsiz Kayaoğlu, G. İ., & Eroğlu, M. (2024). Farklı Uzaklık Fonksiyonlarının Spektral Kümeleme Algoritmasının Performansına Etkisi. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 26(77), 237–241. https://doi.org/10.21205/deufmd.2024267706

[2] von Luxburg, U. (2007). A tutorial on spectral clustering. Statistics and Computing, 17(4), 395–416. https://doi.org/10.1007/s11222-007-9033-z

[3] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8), 888–905. https://doi.org/10.1109/34.868688

[4] Ng, A. Y., Jordan, M. I., & Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. Advances in Neural Information Processing Systems, 14, 849–856.

[5] Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[6] de Amorim, R. C., & Mirkin, B. (2012). Minkowski metric, feature weighting and anomalous cluster initializing in K-Means clustering. Pattern Recognition, 45(3), 1061–1075. https://doi.org/10.1016/j.patcog.2011.08.012

[7] de Amorim, R. C., Shestakov, A., Mirkin, B., & Makarenkov, V. (2017). The Minkowski central partition as a pointer to a suitable distance exponent and consensus partitioning. Pattern Recognition, 67, 62–72.
https://doi.org/10.1016/j.patcog.2017.02.001

Chapter 5

Current Systematic Botanical Nomenclature of Turkish Plant Names in Tervîhü'l-Ervâh

Harun ŞAHİN[1], Mevlüt ALATAŞ[2]

1. INTRODUCTION

Plants play a critical role in the prevention and treatment of both human and animal diseases. Various compounds and extracts derived from plants are effective in curing many illnesses and constitute the fundamental resources that form the basis of modern pharmaceuticals (Dost and Allahverdiyeva, 2025). The science of herbal treatment, born from the vital necessity of sustaining life, has gradually developed through the compilation of works in different geographies and languages. The translation of these written texts into the respective languages of each nation has allowed this field to expand over a highly vast geographical area (Küçüker and Yıldız, 2018).

Despite the widespread dissemination of these medical works, the plant names contained within them were often distinct or highly localized. This regional nomenclature restricted the utilization of these plants to their native areas and limited their worldwide applicability. To overcome this restriction and ensure that therapeutic plants can be utilized on a global scale, it is necessary to identify their universally valid Latin equivalents within the system of binomial nomenclature (Peru et al., 2025). Determining the contemporary binomial Latin equivalents of plant names found in historical medical texts within the framework of systematic botany will enhance plant awareness and optimize their usage as remedies for diseases. Conducted for this purpose, this study aims to identify the Turkish plant names mentioned in Tervîhü'l-Ervâh and to provide their current Latin equivalents in systematic botany.

Composed between the late 14th and early 15th centuries, Tervîhü'l-Ervâh is the first original medical treatise written in verse mathnawi form in Turkish literature. As Ahmedî's largest mathnawi, the work consists of 10,082 couplets. Owing to its nature as a medical text, it contains a large repertoire of Arabic,

[1]Prof. Dr., Gazi University, Gazi Faculty of Education, Department of Turkish and Social Sciences Education, harunsahin75@gmail.com, Orcid No: 0000-0001-7178-7267
[1]Prof. Dr., Munzur University, Tunceli Vocational School, Department of Plant and Animal Production, mevlutalatas@gmail.com, Orcid No: 0000-0003-0862-0258

Persian, and Turkish plant names alongside the names of various diseases, drugs, and treatments (Özer, 2023).

In order to establish a standardized nomenclature and common usage for these plants, this study identifies the Turkish plant names recorded in Tervîhü'l-Ervâh and provides their contemporary Latin equivalents in systematic botany.

2. MATERIALS AND METHODS

The primary material of this research consists of the medical text titled Tervîhü'l-Ervâh. In the study, the Turkish plant names mentioned in Tervîhü'l-Ervâh were first determined. These names are presented in bold letters at the beginning of the first line in each cell of the table. The plant names written in bold are followed by informative and descriptive explanations regarding the respective plant. Various dictionaries and historical texts were consulted during the writing of these descriptions (Baytop, 1985; Develioglu, 2000; Unat et al., 2004; Önler, 2004; Bilgin, 2006; Şahin, 2007; Paçacıoğlu, 2014; Baytop, 2015; Baytar, 2017; Küçüker and Yıldız, 2018; Gümüşatam, 2022; Çakıcı, 2023).

Following these explanations, the identified plants were classified according to contemporary binomial nomenclature. Taxa with sufficient descriptive information to confirm their specific identity were nomenclatured at the species level, whereas taxa lacking sufficient details to determine the exact species were classified at the genus level. Plant terms expressed in a purely general sense were not evaluated under any specific systematic category.

To evaluate the valid Latin names of the plants and their synonymy statuses, the APG III system (Haston et al., 2009), the List of Plants of Turkey (Vascular Plants) (Güner et al., 2012), the Illustrated Flora of Turkey Volume 1 (Güner et al., 2014), and the Illustrated Flora of Turkey Volume 2 (Güner et al., 2018) were utilized. Additionally, the online databases Bizimbitkiler (2013), IPNI (2024), and POWO (2024) were accessed.

3. RESULTS, AND DISCUSSION

As a result of the research and examinations conducted on Tervîhü'l-Ervâh—the first original verse medical book written in Turkish literature—it was determined that 80 Turkish plant names are mentioned in the text. These local plant names, which also reflect the linguistic characteristics and the rich vocabulary of the period, were classified under binomial nomenclature in systematic botany, and their current Latin equivalents were provided.

While providing the Latin equivalents of these plant names, 68 taxa containing sufficient identification details were nomenclatured at the species and infraspecific levels. On the other hand, 9 taxa with limited descriptive attributes that prevented

species-level determination were classified at the genus level. Finally, 3 taxa expressed in a general sense that lacked adequate information to be categorized at either the species or genus level were left unassigned to any systematic category (Table 1).

Table 1. Plant names mentioned in Tervîhü'l-Ervâh and their current Latin equivalents in binomial nomenclature

acı bāẕām: Bitter almond. A species of almond tree from the Rosaceae family, either thorny or thornless, characterized by thick-shelled fruits and small kernels. ***Prunus amygdalus*** Batsch.
acı ḳışşā: Bitter Armenian cucumber or wild cucumber. A large fruit from the Cucurbitaceae family, featuring a striped and hairy rind, colored green or yellowish with green spots. ***Cucumis melo*** L.
acı turb: Wild radish / bitter radish. A plant from the Brassicaceae family with hairy leaves and white, yellow, or purple flowers. ***Raphanus raphanistrum*** subsp**.** ***sativus*** (L.) Schmalh.
acı mışmış: Bitter apricot. A tree from the Rosaceae family that grows in warm or temperate climates, bearing pinkish-white flowers. The light orange, juicy, fragrant, single-seeded stone fruit of this tree. ***Prunus armeniaca*** L.
aḫ enc̦ir: White fig, Fig. ***Ficus carica*** L.
aḳ fulful: White pepper, Pepper. ***Capsicum annuum*** L.
aḫ ḫaşḫaş: White poppy, Opium poppy. Gelincikgillerden, capsules from which opium is obtained, an annual and herbaceous culture plant whose seeds are oil extracted. ***Papaver somniferum*** L.
aḫ ḫaṭmi̱ : Marshmallow flower. A perennial herbaceous ornamental plant from the Malvaceae family, whose roots and flowers of certain species are used in medicine; hollyhock. ***Althaea officinalis*** L.

aḳ şandal: White sandalwood. A tree from the Santalaceae family with hard and fragrant timber. ***Santalum album*** L.
aḳ ṭabāşi̱r: Tabasheer, bamboo manna, bamboo sugar. ***Saccharum officinarum*** L.
aḳ türbüd: White turpeth, Turpeth root. ***Operculina turpethum*** (L.) Silva Manso
aḫ üzüm: White grape. ***Vitis vinifera* L.**
alma: Apple. ***Malus sylvestris*** (L.) Mill.
Arpa: Barley. A plant from the Poaceae family. Its grains are used in making bread and beer, and are also given to animals as fodder. ***Hordeum vulgare*** L.
ayva: Quince. ***Cydonia oblonga*** Mill.
Bögrülce: Cowpea. A plant resembling a common bean. Its green pods are consumed as vegetables and its seeds feature a dark-spotted hilum. ***Vigna unguiculata*** (L.) Walp.
buġday: Wheat. ***Triticum aestivum*** L.
burçaḫ: Bitter vetch. A forage plant from the Fabaceae family whose seeds are utilized as animal feed. ***Vicia ervilia*** (L.) Willd.
buyan: Licorice root. ***Glycyrrhiza glabra*** L.
Çıġındur: Beet. ***Beta vulgaris*** L.

çuğundur: Chukundur, beet. ***Beta vulgaris*** L.
demür diken: Puncturevine / caltrop. A prostrate plant species found flat on the ground, bearing horn-shaped thorns and small, light-yellow flowers. ***Tribulus terrestris*** L.
dikenlü ot: Thorny herb / prickly weed.
ekşi ayva: Ayva. Sour quince. A variety of quince that is firm, juicy, and low in sugar content. ***Cydonia oblonga*** Mill.
ekşi erük: Sour plum. A variety of plum that is firm, juicy, and low in sugar content. ***Prunus* sp.**
ekşi nār: Sour pomegranate. A slightly thorny shrub from the Punicaceae family with smooth, glossy leaves and large red flowers. ***Punica granatum*** L.
erük: Plum. A tree from the Rosaceae family with white flowers. The juicy, stone-bearing fruit of this tree, which comes in various colors and tastes tart or sweet. ***Prunus domestica*** L.
eşek marulı: A type of wild lettuce, bitter lettuce. ***Lactuca viminea*** (L.) J. Presl & C. Presl.
gökce: Green, fresh. Mistletoe. ***Viscum album*** L.
günlük: Sweetgum tree and the resin (frankincense) obtained from it. ***Liquidambar orientalis*** Mill.
ẖurmā çiçegi: The flower of the date palm tree. ***Phoenix dactylifera*** L.
ısırğan: Stinging nettle. An herb from the Urticaceae family completely covered in stiff hairs that secrete a highly irritating substance known as formic acid when broken. ***Urtica* sp.** L.
ıssı ot: Hot pepper / Pepper. ***Capsicum annuum*** L.

ızḫır çiçegi: Camel grass / Mecca grass. ***Cymbopogon schoenanthus*** (L.) Spreng. (Ayrık otu, Mekke Ayrığı)
ince biḫ: Quince (fine/slender variety). ***Cydonia oblonga*** Mill.
ince cünd-i bi̱dester: Castoreum plant and its root, galbanum/jawashir, and the gum resin derived from it. ***Ferula elaeochytris*** Korovin
ince kettān: Flax plant and the fiber/gum derived from it. ***Linum usitatissimum* L.**
ince rāziyānec: Fennel (fine variety). ***Foeniculum vulgare*** Mill.
it üzümi: Black nightshade. An annual herbaceous plant from the Solanaceae family used in the manufacturing of certain medicines; dog's grape, fox's grape. ***Solanum nigrum*** L.
ḳabaḳ: Squash / Pumpkin. A climbing or trailing plant from the Cucurbitaceae family with yellow flowers and numerous cultivated varieties. ***Cucurbita pepo*** L.
ḳamış: Reed. A jointed plant from the Poaceae family with a hard stem that grows in wetlands and damp areas. ***Phragmites australis*** (Cav.) Trin. ex Steud.
ḳara erük: Black plum. ***Prunus domestica*** L.
ḳara kūhi̱ erük: Black mountain plum, wild mountain plum. ***Prunus cocomilia*** Ten.
ḳavun: Melon. A trailing plant from the Cucurbitaceae family bearing large fruits. ***Cucumis melo*** L.
keber: Caper / cat's claw. ***Capparis spinosa*** L.

kenger saḳızı: Gundelia / Tumble thistle. A wild plant from the Asteraceae family with thorny leaves. A type of gum resin made from the latex of the gundelia plant.

Gundelia tournefortii L.

ḳızıl gül: Red rose.

Rosa gallica L.

ḳızıl mercimek: Red lentil. The small, round, and flattened seed of this plant, which possesses high nutritional value and can be red, yellow, or green.

Lens culinaris Medik.

kūḫi̱ erük: Mountain plum.

Prunus cocomilia Ten.

ot: Herb, plant; raw medicinal material.

pazu: Chard. A plant from the Amaranthaceae family whose leaves are consumed as a vegetable; wild beet, wild spinach.

Beta vulgaris L. var. ***cicla*** L.

saḳız: Gundelia gum / Kenger gum.

Gundelia tournefortii L.

sarmaşuḳ: Ivy. A plant from the Araliaceae family with dark green, variously shaped leaves that climbs upright or attaches to flat surfaces using small adventitious roots extending from its stems and branches.

Hedera helix L.

sarmısaḳ: Garlic. A cultivated plant from the Amaryllidaceae family containing aromatic oils in its leaves, stem, and underground bulb. The cloves of this plant used as a spice.

Allium sativum L.

şarmusaḳ: Garlic.

Allium sativum L.

şaru ehli̱le: Yellow myrobalan / Chebulic myrobalan. A plant whose seeds are used in medicine as a laxative. It has varieties such as black myrobalan and yellow myrobalan.

***Terminalia* sp.**

ṣaru kişnı̲z: Yellow coriander. Coriander herb. ***Coriandrum sativum*** L.
semiz otı: Purslane. An herbaceous plant from the Portulacaceae family with fleshy and slightly tart leaves eaten as a vegetable. ***Portulaca oleracea*** L.
sögüd: Willow. ***Salix* sp.**
sögüt: Willow. A large tree from the Salicaceae family growing in wet areas, with alternate leaves whose lower surfaces are often covered in fine down. ***Salix* sp.**
südlügen: Spurge. A plant from the Euphorbiaceae family whose leaves, stems, and roots contain a milky, acrid, and burning sap. ***Euphorbia* sp.**
ṭaġ erügi: Blackthorn / Sloe. A highly tart, firm type of plum with a large stone. ***Prunus spinosa*** L.
ṭaru: Millet. A drought-resistant plant from the Poaceae family; proso millet. ***Panicum miliaceum*** L.
ṭatlu bāz̲ām: Sweet almond. The almond tree and its edible nut. ***Prunus amygdalus*** Batsch.
ṭatlu nār: Sweet pomegranate. The pomegranate tree and its sweet fruit. ***Punica granatum*** L.
ṭatlu rummān: Sweet pomegranate. ***Punica granatum*** L.
ṭatlu üzüm: Sweet grape. ***Vitis vinifera*** L.
teke sakalı: Goat's beard / Salsify. An herbaceous plant from the Asteraceae family whose roots are consumed as a vegetable. ***Tragopogon porrifolius*** L.
ṭopalaḫ: Nut grass / Nut sedge. A plant from the Cyperaceae family from whose leaves a green dye is extracted.

Cyperus rotundus L.
Tütin: Tobacco. ***Nicotiana tabacum*** L.
tütün: Tobacco. An herbaceous plant from the Solanaceae family containing nicotine in its chemical composition. ***Nicotiana tabacum*** L.
üzüm: Grape. The fruit of the grapevine, grouped in clusters and consumed either fresh or dried. ***Vitis vinifera*** L.
yabān soğanı: Wild onion / Sea squill. A perennial plant from the Asparagaceae family whose bulb is processed to extract substances used in medicine. ***Drimia maritima*** (L.) Stearn
yarpuz: Pennyroyal. A species of wild mint from the Lamiaceae family. A highly aromatic plant with short petioles and varying degrees of hairiness. ***Mentha pulegium*** L.
yaşıl māzū: Green arborvitae / thuja. ***Thuja* sp.**
yavşan: Wormwood. A fragrant, bitter plant from the Asteraceae family bearing blue or white flowers, whose boiled water is consumed traditionally as medicine. ***Artemisia* sp.**
yelencük: Agarwood / Aloe wood. ***Aquilaria malaccensis*** Lam.
yılğun: Tamarisk. A genus of trees or shrubs from the Tamaricaceae family native to the Mediterranean region. ***Tamarix* sp**.
yosun: Moss / Algae. The general name given to primitive thalloid plants growing mostly in water or on trees and stones.
yüzerlik: Syrian rue / Harmal. A plant from the Nitrariaceae (formerly Zygophyllaceae) family with alternate leaves and white flowers, whose bitter seeds resemble sesame and are used traditionally for therapeutic purposes and as incense. ***Peganum harmala*** L.

Seeking alternative treatment modalities alongside synthetic drugs, contemporary modern medicine has adopted a path that references historical experiences in preventive medicine (Gümüşatam, 2022). To this end, traditional and modern medicine have begun to be utilized collaboratively in the treatment of diseases. China and Japan are among the countries where traditional and complementary medicine applications are widely integrated (Karahancı et al., 2015). It is known that modern and traditional medical practices are used concurrently in 95% of hospitals in China, where there are more than 2,500 hospitals dedicated to traditional methods and more than 35,000 healthcare personnel operating in this sector (Watanabe et al., 2011). These data underscore the profound importance of leveraging historical works written in archaic languages within the fields of the history of science and medicine.

Representing this executed on behalf of the history of science and medicine, this study identifies the Turkish plant names recorded in Tervîhü'l-Ervâh one of the most important historical herbal references and provides their current Latin equivalents in systematic botany, thereby offering a significant contribution to the fields of medicine and traditional folk healing.

By providing the current Latin nomenclature of these plants, this study facilitates a more comprehensive global understanding of the flora that prevents and treats diseases, thereby enabling their utilization as a source of healing across broader geographical regions.

Furthermore, this study will serve as a foundational reference text for future corresponding research in the history of medicine, pharmacy, botany, folk medicine, and clinical therapeutics.

References

Baytar, İ. (2017). *Tercüme-i Müfredât-ı İbn Baytar*. İstanbul: Health Sciences University Publications.

Baytop, A. (1985). Tıbbi Bitkiler Atlası. İstanbul: İstanbul University Press.

Baytop, T. (2015). *Dictionary of Turkish Plant Names.* Ankara: Turkish Language Association Publications.

Bilgin, A. (2006). *Medicinal Plants Used in Medicine Production in the Ottoman Period. Health in the Ottomans I.* Istanbul: Biofarma Publications.

Bizimbitkiler. (2013). *Ourplants, Nezahat Gökyiğit Botanical Garden and Flora Research Association Publication.* (06. 10. 2026, http://www.bizimbitkiler.org.tr).

Çakıcı, B. (2023). *Analysis of Plant Names in Four Medical Texts Written in Old Anatolian Turkish.* Master's Thesis. Bursa: Bursa Uludağ University, Institute of Social Sciences.

Devellioğlu, F. (2000). *Ottoman-Turkish Encyclopedic Dictionary.* Ankara: Aydın Bookstore.

Dost, A., Allahverdiyeva, Ş. (2025). A study on the plant names in Nüzhetü'l-Ebdân fî Tercemeti Gâyeti'l-İtkân (251a-497b). *Mecmua,* 10(19), 96-122.

Gümüşatam, G. (2022) Herbal Treatment Methods and Terminology in Sixteenth-Century Ottoman Medicine According to the Risale-i Mu'âlece. Studies in *Ottoman Science,* 23(2), 261-290.

Güner, A., Aslan, S., Ekim, T., Vural, M., Babaç, M.T. (2012). *List of Turkish Plants - Vascular Plants.* Istanbul: Nezahat Gökyiğit Botanical Garden and Flora Research Association Publications.

Güner, A., Ekim, T. (2014). *Illustrated Flora of Türkiye Volume 1.* Istanbul: ANG Foundation Nezahat Gökyiğit Botanical Garden, Türkiye İş Bankası Cultural Publications.

Güner, A., Ekim, T. (2018). *Illustrated Flora of Türkiye Volume 2.* Istanbul: ANG Foundation Nezahat Gökyiğit Botanical Garden, Türkiye İş Bankası Cultural Publications.

Haston, E., Richardson, J.E., Stevens, P.F., Chase, M.W., Harris, D.J. (2009). The Linear Angiosperm Phylogeny Group (LAPG) III: A Linear Sequence of The Familiesin APG III. *Botanical Journal of the Linnean Society,* 161, 128-131.

IPNI (2024). *International Plant Names Index, Harvard University Herbaria & Libraries and Australian National Herbarium. The Royal Botanic Gardens, Kew.* (10. 10. 2026, http://www.ipni.org).

Karahancı, O.N., Öztoprak, Ü.Y., Ersoy, M., Ünsal, Ç. Z., Hayırlıdağ, M., Büken, N.Ö. (2015). Comparison of the Regulation on Traditional and

Complementary Medicine Practices and the Draft Regulation. *Turkish Journal of Bioethics*, 2(2), 117-126.

Küçüker, P., Yıldız, Y. (2018). A Study on the Plant Names in the Translation of the Curriculum Vitae of Ibn Baytar (150b-295a) (II). *International Journal of Turkish Dialect Studies,* 2(1), 109-137.

Önler, Z. (2004). Turkish Plant Names in 14th-15th Century Medical Texts. *Kebikeç,* 18, 273-301.

Özer, O. (2023). A Study on Turkish Plant Names in Tervihü'l-Ervah. *International Journal of Philology Bengü,* 3(1), 19-28.

Paçacıoğlu, B. (2014). *Dictionary of Drug and Plant Names.* Sivas: Cumhuriyet University Publications.

Peru, C., Şahin, H., Alataş, Y.K., Ünlü, M., Alataş, M. (2025). Plant Names in the Kitâb-ı Maʿcûn and Their Current Latin Equivalents in Binomial Nomenclature. (New Approaches In Natural Sciences And Mathematics: Theory, Method, And Practice, Chapter 4). İzmir. Duvar Publications.

POWO (2024). *Plants of the World Online, Facilitated by the Royal Botanic Gardens, Kew.* (10. 10. 2026, https://powo.science.kew.org/).

Şahin, H. (2007). Plant Names in the 16th Century in the Example of the Câmi'ü'l-Fürs. *Turkish Studies*, 2(2), 570-602.

Unat, E.K., İhsanoğlu, E., Vural S. (2004). *Dictionary of Ottoman Medical Terms.* Ankara: Turkish Historical Society Publications.

Watanabe, K., Matsuura, K., Gao, P., Hottenbacher, L., Tokunaga, H., Nishimura, K., Imazu, Y., Reissenweber, H., M Witt C. M. (2011). Traditional Japanese Kampo Medicine: Clinical Research Between Modernity and Traditional Medicine-The State of Research And Methodological Suggestions For The Future. *Evidence-Based Complementary and Alternative Medicine*, 513842.

Chapter 6

Degradation Mechanisms of Mordant Red 19, Carminic Acid, and Curcumin Molecules Using the DFT Method.

Bahar EREN[1]

ABSTRACT

In this study, the degradation mechanisms of the molecules Mordant Red 19, carminic acid, and curcumin were investigated. Of these three molecules, Mordant Red 19 (MR19) is an organic compound belonging to the azo dye class and is the only synthetic dye among them. Carminic acid, has an anthraquinone structure; although used as a natural red coloring agent, it is derived from an insect rather than a plant source. Curcumin, on the other hand, is a food coloring produced from the plant known as *zerdeçal* (turmeric) in Türkiye. These three colorants, one of plant origin, one of animal origin, and one obtained through synthetic production have been selected from different groups.

All the dyes to be examined theoretically are organic. It is also known that organic pollutants are present in surface waters at low concentrations. Hydroxyl radicals are degradative agents that attack biomolecules. The aim of the study is to monitor new hydroxylated products formed through reaction with the OH radical which can be far more dangerous than the substance released into the environment and to elucidate the degradation reaction.

For this purpose, geometric optimizations were performed using the DFT/B3LYP/6-31G(d) basis set an implementation of Density Functional Theory (DFT), in order to identify all possible reaction pathways for the molecules under theoretical investigation. The Gauss View 6 molecular visualization software and the Gaussian 16 program were used for orbital calculations involving all fragments.

Keywords: DFT, Mordant Red 19, Carminic acid, Curcumin, coloring agent, dyeing agent

[1] Dr.; Tekirdag Namik Kemal University, Vocational School of Social Sciences, Turkey.
beren@nku.edu.tr ORCID No: 0000-0002-1523-274X

INTRODUCTION

Textile dyeing entails significant environmental pollution and poses health risks to personnel exposed to harmful chemicals. Many colorants derived from synthetic sources can be harmful (Manzoor and Sharma, 2020). Consequently, interest in natural dyes has grown substantially in recent years. Of the three molecules examined in this study, the one shown in Figure 1(a) is an organic compound belonging to the azo dye class and is the only synthetic dye among them. The molecule in Figure 1(b) features an anthraquinone structure; although used as a natural red dye, it is derived from an insect rather than a plant source. The molecule in Figure 1(c) is a food colouring produced from the plant known in our country as turmeric.

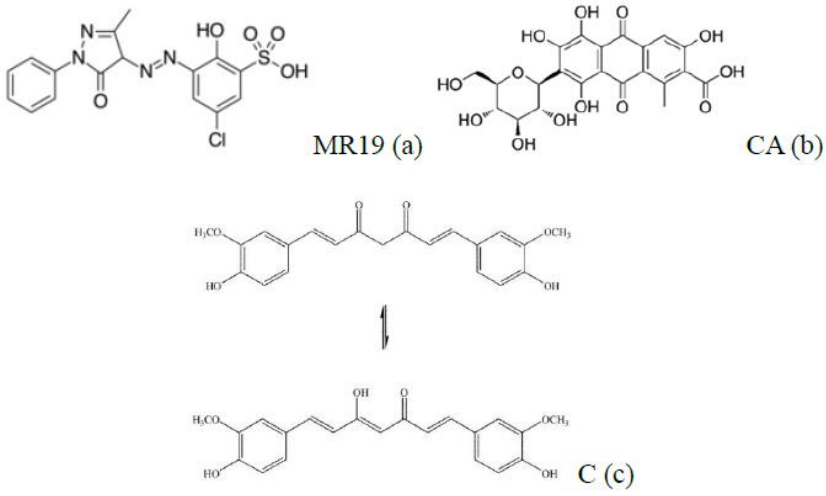


Figure 1. Two-dimensional structures of (a): Mordant Red 19 (MR19), (b): Carminic acid (CA), and (c): Curcumin (C)

Mordant Red 19 is an organic compound classified as an azo dye (Hunger, 2005). It is a mordant used in textile dyeing, typically in combination with chromium. It generally exists in the form of a sodium salt (Baumann and Federkiel, 1957). A mordant, or dye fixative, is a substance used to fix or bind dyes to fabrics. It achieves this by forming a coordination complex with the dye, which then binds to the fabric or tissue. They can be used to dye fabrics or to intensify stains in cell or tissue preparations. Azo dyes are organic compounds containing the functional group R−N=N−R′, where R and R′ are typically aryl or substituted aryl groups. They constitute a commercially important family of azo compounds that is, compounds containing the C-N=N-C linkage. Azo dyes are

synthetic dyes and do not occur naturally. Most azo dyes contain only one azo group, although some referred to as "diazo dyes" and "triazo dyes," respectively contain two or three azo groups. Azo dyes account for 60–70% of all dyes used in the food and textile industries. They are widely used in the processing of textiles, leather goods, and certain foods. Chemically related derivatives of azo dyes include azo pigments, which are insoluble in water and other solvents (McNaughty and Wilkinson, 1997; Hunger et al., 2009).

Carminic acid, the coloring agent derived from the cochineal insect, is obtained by brushing the insects off cacti into bags and subsequently killing and drying them using agitation, hot water, steam, or sunlight. Used to dye wool and silk red, this acid is prone to fading. Carminic acid is a colorant belonging to the anthraquinone class (Li et al., 2009). It exists as a dark red liquid. Carminic acid is water-soluble, and its color hue depends on the pH of the solution; it appears orange in acidic solutions and violet in alkaline solutions, shifting toward red as the pH rises from 5 to 7. It is a natural, water-soluble colorant that is resistant to degradation over time (Altuğ, 2009).

Curcumin (C) is a food coloring derived from the plant known in Turkey as turmeric (Zerdeçal). It is a bright yellow chemical substance produced by plants of the Curcuma longa species. It is used as a dietary supplement, a cosmetic ingredient, a flavoring agent, and a food coloring (in products such as curry powders, mustards, butter, and cheeses). Its use as a food additive to provide an orange-yellow color in processed foods is approved in the European Union under the code E 100, and it has been approved by the FDA for use as a food coloring in the USA. Curcumin contains a seven-carbon linker and three main functional groups: an α, β-unsaturated β-diketone moiety and an aromatic *o*-methoxy-phenolic group. The aromatic ring systems which are phenolic in nature are connected by two α, β-unsaturated carbonyl groups. It is a diketone tautomer that exists in the enol form in organic solvents and in the keto form in water. Due to its hydrophobic nature, curcumin is sparingly soluble in water but readily soluble in organic solvents (Farooqui et al., 2019; Curcumin E100).

Little regard is typically paid to the environmental fate of colorants whether during production, during the dyeing process, or when discarded as waste. This study aims to shed light on this uncertainty. Whether utilized or discarded after their shelf life expires, these organic compounds enter the natural cycle and the environment; this occurs because due to a lack of sufficient public awareness in our country waste (whether chemical or general) is often sent to the same disposal streams. Compounds that are partially or completely water-soluble can constitute a significant group of air and water pollutants. To remove these residues from air

and water, it is first necessary to elucidate their degradation mechanisms (Eren and Yalçın Gürkan, 2024).

During the degradation reactions of organic pollutants, products that are more harmful than the original pollutants may be formed. Therefore, there is a significant need to identify intermediate products in air and water treatment. The reactions responsible for the atmospheric removal of organic pollutants involve OH and NO_3 radicals, O_3, and Cl atoms. Among these, reactions with the OH radical are the most effective (Eren and Yalçın Gürkan, 2022; 2023a,b).

In this research, molecules containing electronegative atoms such as O, N were preferred. The necessity of water-soluble particles was considered, as the interactions of the molecules with the OH radical would be investigated, and studies would be conducted in both the gas and water phases. During the degradation of the molecules, the structures around the electronegative atom within the molecule, the presence of stable closed rings, double bonds, hydrophobic or hydrophilic structures at the molecular ends, and easily separable groups were examined individually (Eren and Yalçın Gürkan, 2022; 2023a,b).

All dyes to be studied theoretically are organic. It is also known that low concentrations of organic pollutants are present in surface waters. Hydroxyl radicals are degraders and specific detectors that attack biomolecules at different rates. A hydroxyl radical attack on an aromatic compound leads to the formation of a new hydroxylated product. (Eren and Yalçın Gürkan, 2017). The initial aim was to track new products that might be even more dangerous than the product released into nature and to elucidate the degradation reaction. Biomolecules, which are hydroxyl scavengers at various rates, are specific detectors for hydroxyl radicals due to their hydroxylation capabilities. Any hydroxyl radical attacking an aromatic compound leads to the formation of a hydroxylated product, and these newly formed products can be far more harmful than the original products at the beginning of the process, therefore monitoring these products is crucial. The OH radical acts as an electrophile in its reaction with any organic molecule and therefore readily binds to unsaturated bonds, whereas the O radical is a nucleophile and therefore cannot interact with these bonds. If an aliphatic side chain is attached to an aromatic molecule, the OH radical prefers the aromatic ring for binding. (Anbar and Neta, 1967; Buxton et al., 1988; Halliwell et al., 1998; Aschmann et al., 2005, 2010; Ashmann and Atkinson, 2006; Atkinson et al., 1988; Atkinson and Arey, 2003; Eren and Yalçın Gürkan, 2017).

MATERIAL and METHODS

To determine all possible reaction pathways of the theoretically studied molecules, geometric optimizations were performed using the DFT/B3LYP/6-31G(d) basis set of Quantum Mechanical Density Functional Theory (DFT), an electronic structure method. For all fragments, the Gauss View 6 molecular representation program and the Gaussian 16 program were used for orbital calculations (Gaussian 16, 2009). The fragmentation reaction energy is affected by the water molecules in the aqueous medium. In addition, the geometric stretching in solutions is induced by H_2O. That is, the presence of a dielectric medium such as H_2O causes relaxation in the geometry related to the solute and has a stabilizing effect that reduces the energy for this mechanism (Eren and Gürkan, 2017). Therefore, in this study, the CPCM solvent model from the COSMO (conductor-like screening solvation model) software package in Gaussian 16 was used to explain the solvent effect of H_2O on the Molecule + ·OH reaction energy (Gaussian 16, 2009).

The Gaussian 16 program was used to calculate the energies of atoms and molecules and to perform geometric optimizations. Gauss View 6 visualizes molecules, allowing us to rotate, move, and modify them as desired. It enables us to graphically examine the results calculated by the Gaussian program. These results can be categorized as optimized molecular structures, molecular orbitals, electrostatic potential surfaces, atomic charges, IR, Raman, and NMR spectra, and normal mode animations depending on vibration frequencies (Gaussian 16, 2009).

If it is predicted that the longest bond and the widest bond opening will break first in the fragmentation pathway for any given molecule, this is determined by observing the energy values and environments of electronegative atoms. It should be remembered that double-bonded or closed-loop structures are more stable than others, and that in the event of fragmentation, breaking from these stable structures will occur in the final stage (Eren and Gürkan, 2023a,b). The energy values of the possible reaction pathways for each studied molecule and its fragments have been calculated in both the gas phase and in water phases. Each fragment is included in the study as a result of the interaction of the parent molecules with the OH radical. The reason why the energy values calculated for each molecule are presented in atomic mass units (Au) is that when converted to SI units, the part after the decimal point makes it difficult to observe the energy difference between the fragments.

Research Findings

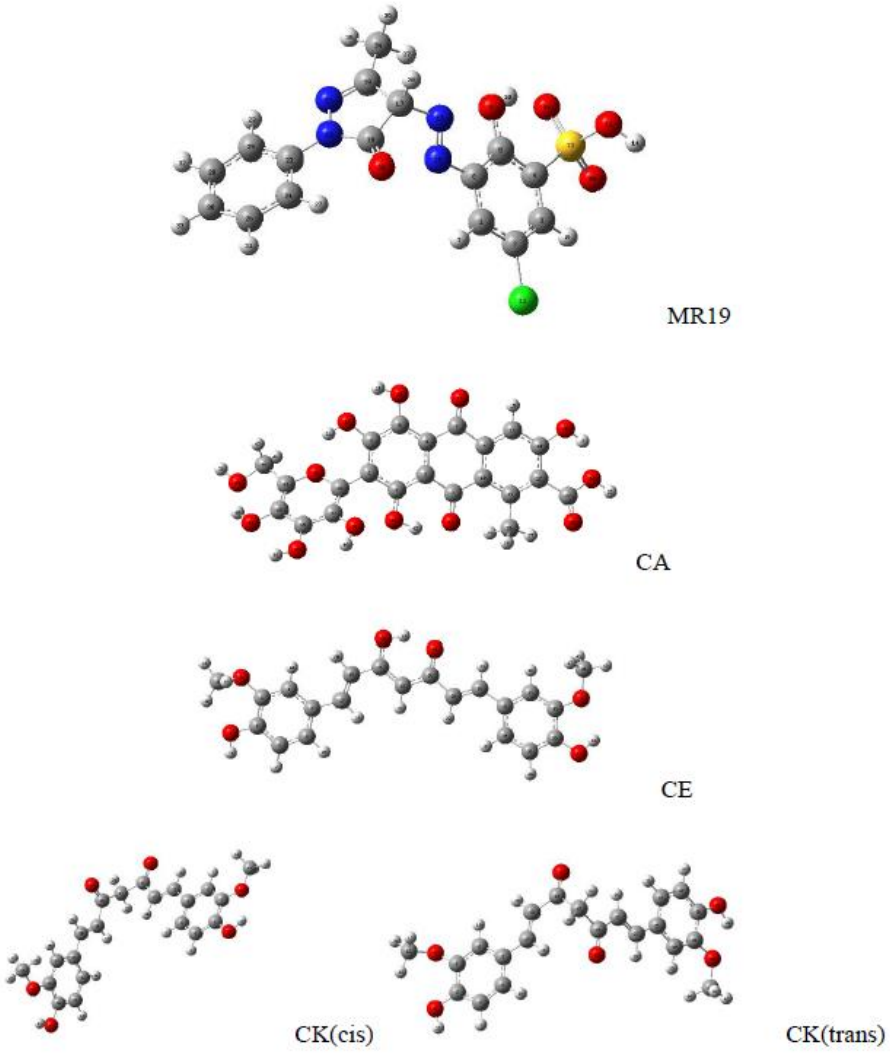


Figure 2. Three-dimensional structures of the Mordant Red 19 (MR19), Carminic acid (CA), molecules, and the enol (CE), keto cis CK (cis), and keto trans CK (trans) forms of the Curcumin molecule (C) optimized using the DFT method (C is represented by gray, O by red, S by yellow, N by dark blue, Cl by green, and H by white)

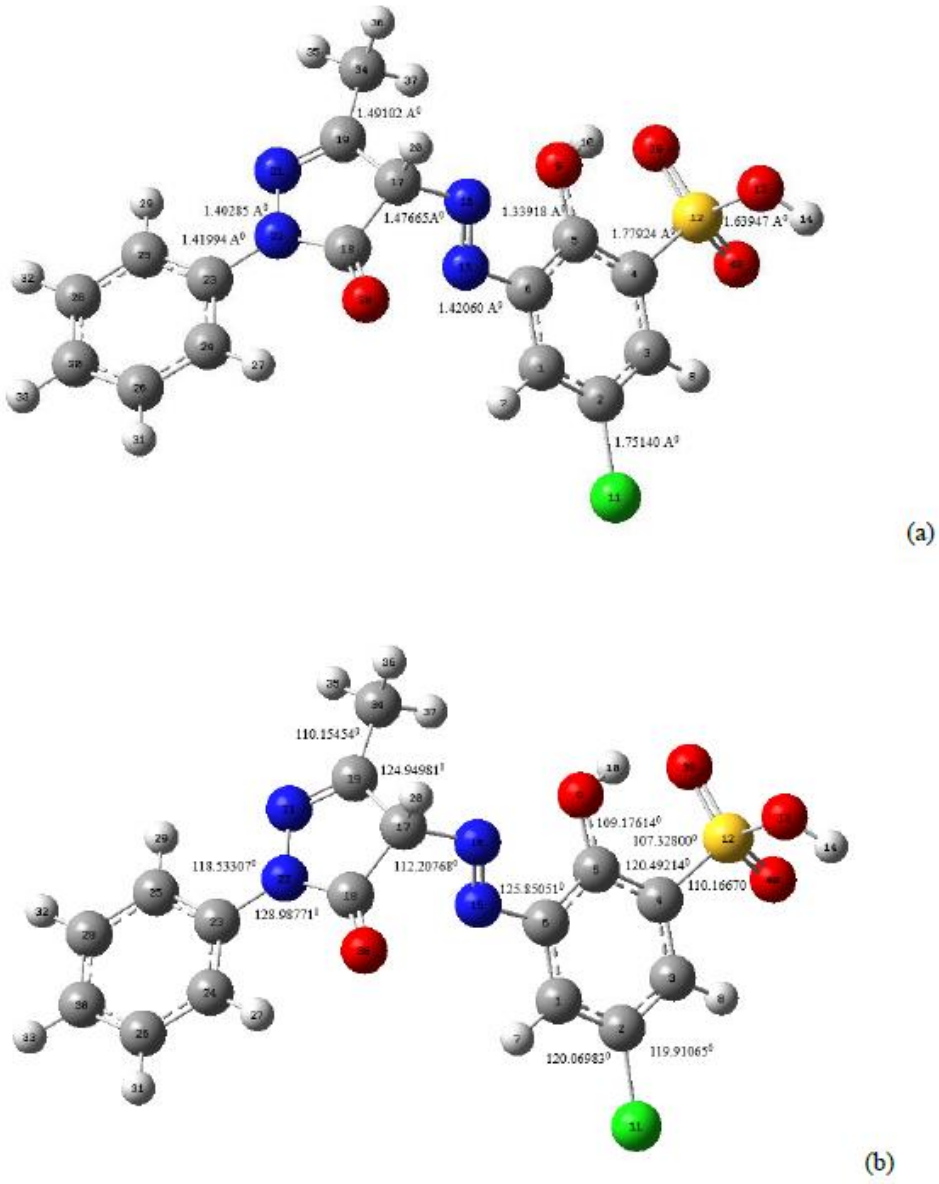


Figure 3. Bond lengths (a) and bond angles (b) between atoms in the optimized structure of the Mordant Red 19 (MR19) molecule (C is represented by gray, O by red, S by yellow, N by dark blue, Cl by green, and H by white)

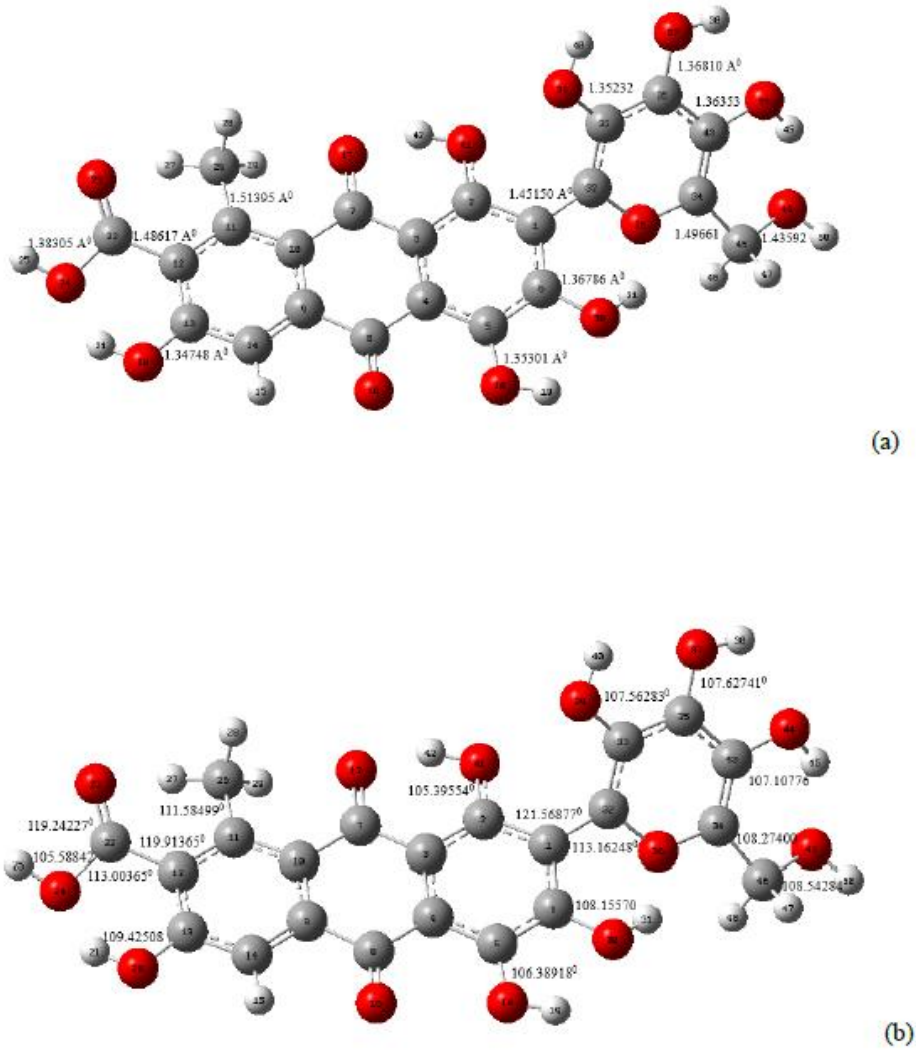


Figure 4. Bond lengths (a) and bond angles (b) between atoms in the optimized structure of the Carminic acid (CA) molecule (C is represented by gray, O by red, and H by white)

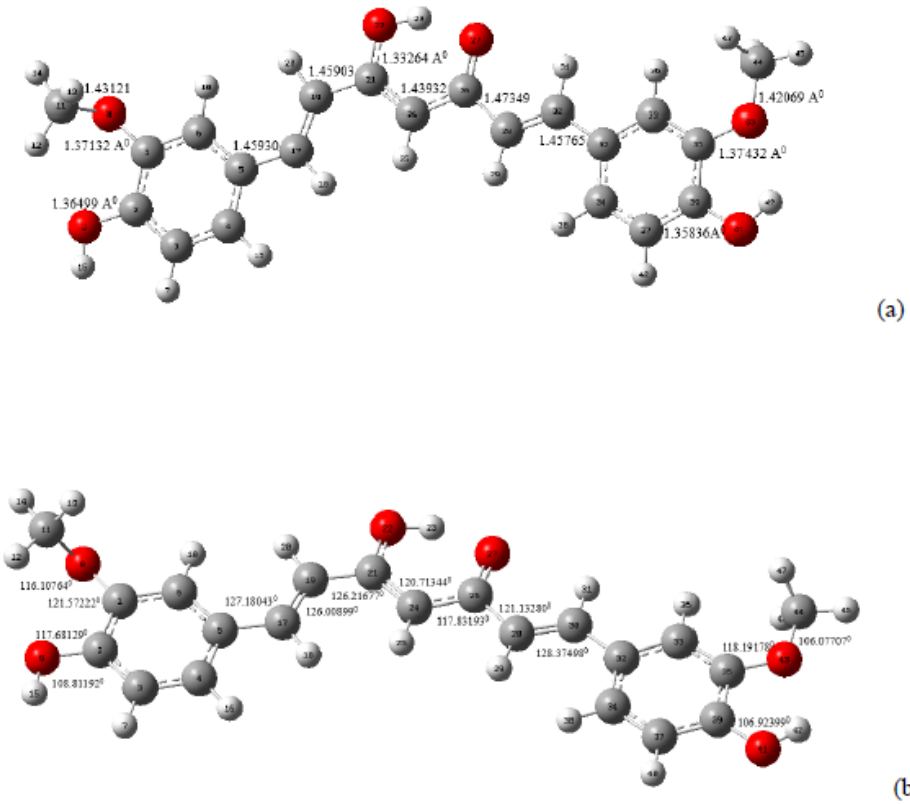


Figure 5. Bond lengths (a) and bond angles (b) between atoms in the optimized structure of the enol (CE) form of the Curcumin molecule (C) (C is represented by gray, O by red, and H by white)

Table 1. Calculated gas-phase Energy (ΔE), Enthalpy (ΔH), and Gibbs free (ΔG) energy (Au) values for the MR19, CA, and CE molecules; bond lengths (A^0) and bond angles (^{0}) between atoms within the molecule; and Mulliken charges of electronegative atoms

	Gas Phase (Au)	Bond Lengths	(A^0)	Bond Angles	(^{0})	Mulliken Yükleri
						O_9 -0.627985
MR19		$C_{23}N_{22}$	1.41994	$C_{23}N_{22}N_{21}$	118.53307	Cl_{11} 0.010144
		$N_{22}N_{21}$	1.40285	$C_{19}C_{34}H_{35}$	110.15454	S_{12} 1.228546
ΔE	-2070.587552	$C_{19}C_{34}$	1.49102	**$C_{17}C_{19}C_{34}$**	**124.94981**	**O_{13} -0.615005**
ΔH	-2070.586607	$C_{17}N_{16}$	1.47665	$C_{17}N_{16}N_{15}$	112.20768	N_{15} -0.278534
ΔG	-2070.669693	$N_{15}C_6$	1.42060	**$N_{15}C_6C_5$**	**125.85051**	N_{16} -0.158487
		C_5O_9	1.33918	$C_5O_9H_{10}$	109.17614	C_{17} -0.205046
		C_4S_{12}	**1.77924**	**$C_5C_4S_{12}$**	**120.49214**	N_{21} -0.277649
		$S_{12}O_{13}$	**1.63947**	$C_4S_{12}O_{39}$	107.32800	**N_{22} -0.462438**
		C_2Cl_{11}	**1.75140**	$C_4S_{12}O_{40}$	110.16670	C_{24} -0.172669
				$Cl_{11}C_2C_3$	**119.91065**	C_{25} -0.162474
				C_1C_2 Cl_{11}	**120.06983**	C_{26} -0.146172
				$C_{23}N_{22}C_{18}$	**128.98771**	C_{28} -0.140654
						C_{30} -0.121611
						C_{34} -0.516983
						O_{38} -0.493660
						O_{39} -0.536321
						O_{40} -0.501910
						C_3 -0.495903
		C_5O_{18}	1.35301	$C_{34}C_{46}O_{49}$	108.27400	C_4 -0.743455
CA	-1824.827404	C_6O_{30}	1.36786	$C_{43}O_{44}H_{45}$	107.10776	C_{10} -0.473440
	-1824.826460	$C_{13}O_{20}$	1.34748	$C_{35}O_{37}H_{38}$	107.62741	C_{12} -0.543142
	-1824.920894	$C_{22}O_{24}$	1.38305	$C_{33}O_{39}H_{40}$	107.56283	C_{14} -0.479929
		$C_{33}O_{39}$	1.35232	**$C_2C_1C_{32}$**	**121.56877**	**O_{16} -0.884025**
		$C_{35}O_{37}$	1.36810	**$C_1C_2C_{35}$**	**113.16248**	**O_{17} -0.695273**
		$C_{43}O_{44}$	1.36353	$C_{13}O_{20}H_{21}$	109.42508	**O_{18} -0.680921**
		$C_{34}C_{46}$	**1.49661**	**$C_{11}C_{12}C_{22}$**	**119.91365**	**O_{20} -1.010058**
		$C_{46}O_{49}$	**1.43592**	$C_5O_{18}H_{19}$	106.38918	**O_{23} -0.817597**
		C_1C_{32}	**1.45150**	$C_6O_{30}H_{31}$	108.15570	**O_{24} -0.859875**
		$C_{12}C_{22}$	**1.48617**	$C_6O_{30}H_{31}$	105.39554	**O_{30} -0.591533**
		$C_{11}C_{26}$	**1.51395**	**$C_{12}C_{22}O_{24}$**	**113.00365**	O_{36} -0.235332
				$C_{12}C_{11}C_{26}$	**119.72390**	**O_{37} -0.932675**
						O_{39} -0.535204
						O_{41} -0.581698
						O_{44} -0.919924
						O_{49} -0.663100
						C_3 -0.199423
	CE	C_1O_8	1.37132	$C_1O_8C_{11}$	116.10764	C_4 -0.199931
	-1263.162551	C_2O_9	1.36499	**$C_2C_1O_8$**	**121.57222**	C_6 -0.238185
	-1263.161607	**O_8C_{11}**	**1.43121**	$C_1C_2O_9$	117.68129	**O_8 -0.518692**
	-1263.248885	**C_5C_{17}**	**1.45930**	$C_2O_9H_{15}$	108.81192	**O_9 -0.632779**
		$C_{35}O_{43}$	1.37432	**$C_5C_{17}O_{19}$**	**127.18043**	C_{11} -0.216433
	CK (cis)	$C_{39}O_{41}$	1.35836	**$C_{17}C_{19}O_{21}$**	**126.00899**	C_{17} -0.170006
	-1263.164637	**$O_{43}C_{44}$**	**1.42069**	**$C_{19}C_{21}O_{24}$**	**126.21677**	C_{19} -0.181175
CE	-1263.163693	**$C_{19}C_{21}$**	**1.45903**	**$C_{21}C_{24}O_{25}$**	**120.71344**	**O_{22} -0.632888**
	-1263.252900	**$C_{26}C_{28}$**	**1.47349**	$C_{24}C_{25}O_{28}$	117.83193	C_{24} -0.295181
		$C_{30}C_{32}$	**1.45765**	**$C_{25}C_{28}O_{30}$**	**121.13280**	**O_{27} -0.607751**
	CK(trans)	**$C_{24}C_{26}$**	**1.43932**	**$C_{28}C_{30}O_{32}$**	**128.37498**	C_{28} -0.195454
	-1263.159722	$C_{21}C_{22}$	1.33264	$C_{39}O_{41}H_{42}$	106.92399	C_{30} -0.168036
	-1263.158778			$C_{33}O_{43}C_{44}$	118.19178	C_{33} -0.271594
	-1263.246568			$O_{43}C_{44}H_{45}$	106.07707	C_{34} -0.185168
						C_{37} -0.170774
						O_{41} -0.645208
						O_{43} -0.553374

Table 2. Gas-phase Energy (ΔE), Enthalpy (ΔH), and Gibbs free energy (ΔG) values (Au) for the MR19 molecule and its fragments

MR19	**MK192**	**MR192**	**MR196**	**MR197**
ΔE -2070.587552	-2031.295055	-1995.357751	-1839.617225	-1800.321296
ΔH -2070.586607	-2031.294111	-1995.356807	-1839.616281	-1800.320352
ΔG -2070.669693	-2031.375756	-1995.436727	-1839.689767	-1800.390114
MR191	**MR193**	**MR198**	**MR1932**	**MR1983**
-1610.987552	-1446.831315	-1390.734744	-1371.609386	-1315.517572
-1610.986608	-1446.830371	-1390.733800	-1371.608442	-1315.516627
-1611.066016	-1446.902084	-1390.783864	-1371.678732	-1315.565086
MR1931	**MR1982**	**MR194**	**MR1981**	**MK194**
-987.229355	-931.133617	-912.006504	-766.958416	-681.030903
-987.228411	-931.132673	-912.005559	-766.957472	-681.029959
-987.296330	-931.179385	-912.072359	-766.996347	-681.084007
MR1953	**MR1951**	**MR1954**	**MR1955**	**MR199**
-571.607983	-450.060415	-340.637490	-301.342801	-232.143516
-571.607039	-450.059471	-340.636546	-301.341857	-232.142571
-571.656358	-450.101938	-340.674493	-301.376032	-232.175364

Table 3. Gas-phase Energy (ΔE), Enthalpy (ΔH), and Gibbs free energy (ΔG) values (Au.) for the CA molecule and its fragments

CA	**CA26**	**CA39**	**CA49**	**CA18**
ΔE -1824.827404	-1785.554326	-1749.627785	-1749.624852	-1749.624420
ΔH -1824.826460	-1785.553382	-1749.626841	-1749.623908	-1749.623475
ΔG -1824.920894	-1785.644605	-1749.718824	-1749.716578	-1749.715323
CA44	**CA30**	**CA20**	**CA41**	**CA24**
-1749.612871	-1749.611906	-1749.611472	-1749.610153	-1749.578091
-1749.611927	-1749.610962	-1749.610528	-1749.609209	-1749.577147
-1749.705245	-1749.703500	-1749.703808	-1749.702875	-1749.670953
CA46	**CA32**	**CA1**	**CA2**	**CA3**
-1710.331493	-1217.272861	-1177.999240	-1142.023115	-1028.725619
-1710.330549	-1217.271916	-1177.998296	-1142.022171	-1028.724675
-1710.419473	-1217.340443	-1178.063581	-1142.090280	-1028.786717
CA4	**CA5**	**CA6**	**CA7**	**CA10**
-1142.056588	-914.228353	-991.632178	-1142.057661	-688.587197
-1142.055644	-914.227409	-991.631234	-1142.056717	-688.586253
-1142.122676	-914.283538	-991.693206	-1142.122868	-688.635688

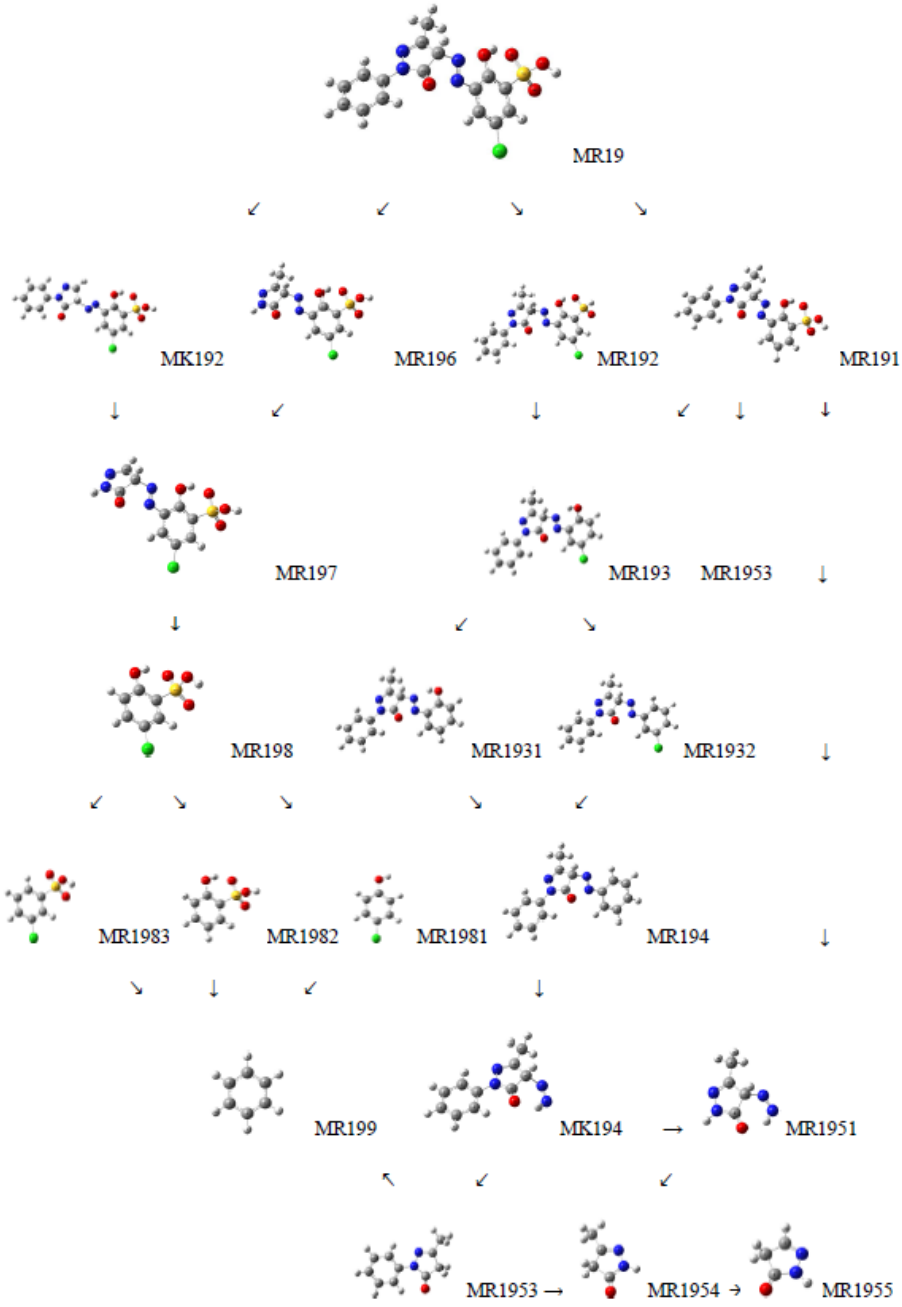


Figure 6. Degradation pathway of MR19 molecule and its fragments

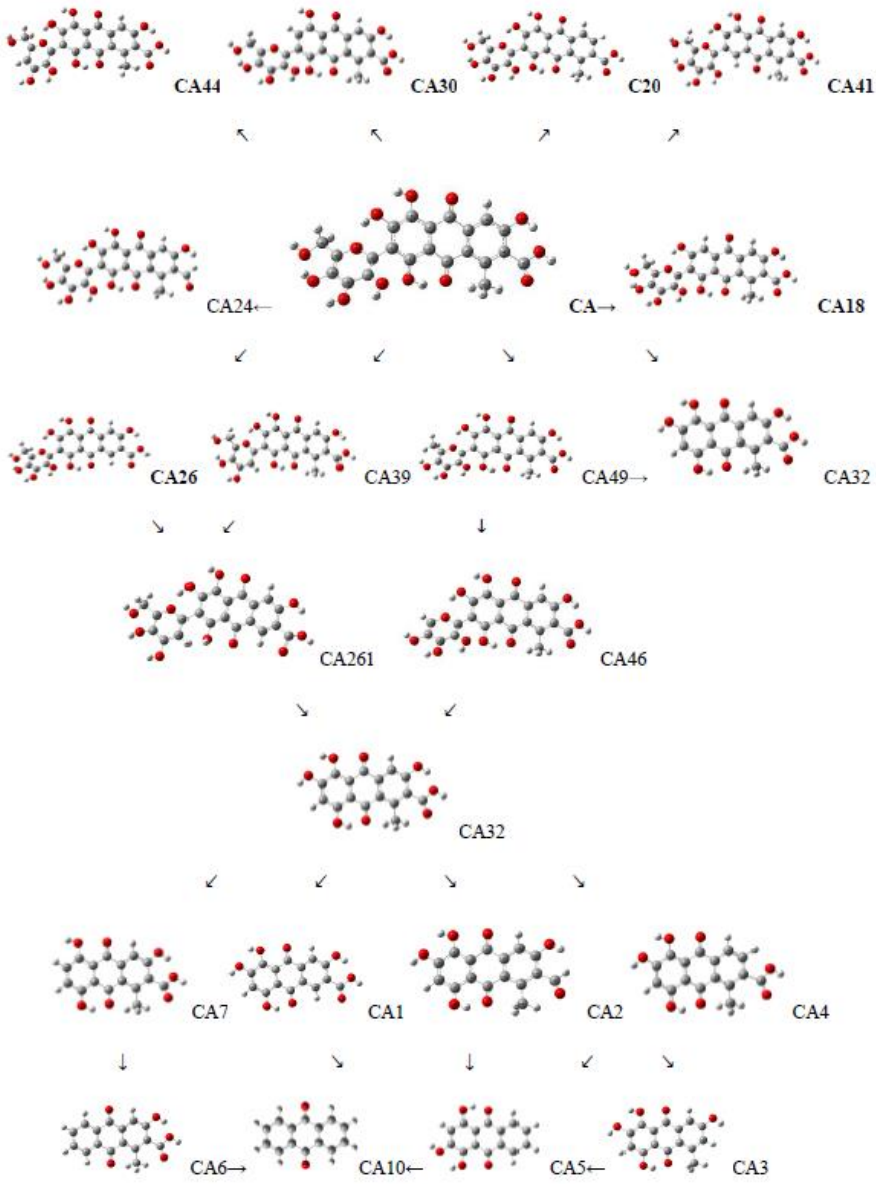


Figure 7. Degradation pathway of the CA molecule and its fragments

Table 4. Gas-phase Energy (ΔE), Enthalpy (ΔH), and Gibbs free energy (ΔG) values (Au.) for the CE molecule and its fragments

CE	**CE1**	**CE2**	**CE12**	**CE11**
ΔE -1263.162551	-1223.885446	-1187.958436	-1184.616790	-1148.673736
ΔH -1263.161607	-1223.884501	-1187.957492	-1184.615846	-1148.672792
ΔG -1263.248885	-1223.968507	-1188.041715	-1184.695854	-1148.754129
CE21	**CE22**	**CE13**	**CE24**	**CE23**
-1112.743848	-1073.467527	-1073.461496	-998.256488	-883.769382
-1112.742904	-1073.466583	-1073.460552	-998.255543	-883.768438
-1112.824782	-1073.544873	-1073.539833	-998.331345	-883.838416
CE14	**CE16**	**CE15**	**CE17**	**CE18**
-842.484097	-803.215501	-765.114263	-767.279007	-728.004237
-842.483153	-803.214556	-765.113318	-767.278063	-728.003293
-842.549734	-803.278224	-765.173771	-767.342084	-728.063093
CE19	**CE25**	**CE26**	**CE27**	**CE28**
-725.845630	-652.791941	-575.421986	-462.085689	-386.872353
-725.844686	-652.790997	-575.421042	-462.084745	-386.871409
-725.901229	-652.848748	-575.473114	-462.132162	-386.916327
CE29	**CE30**			
-309.507758	-232.143513			
-309.506814	-232.142569			
-309.546012	-232.175362			

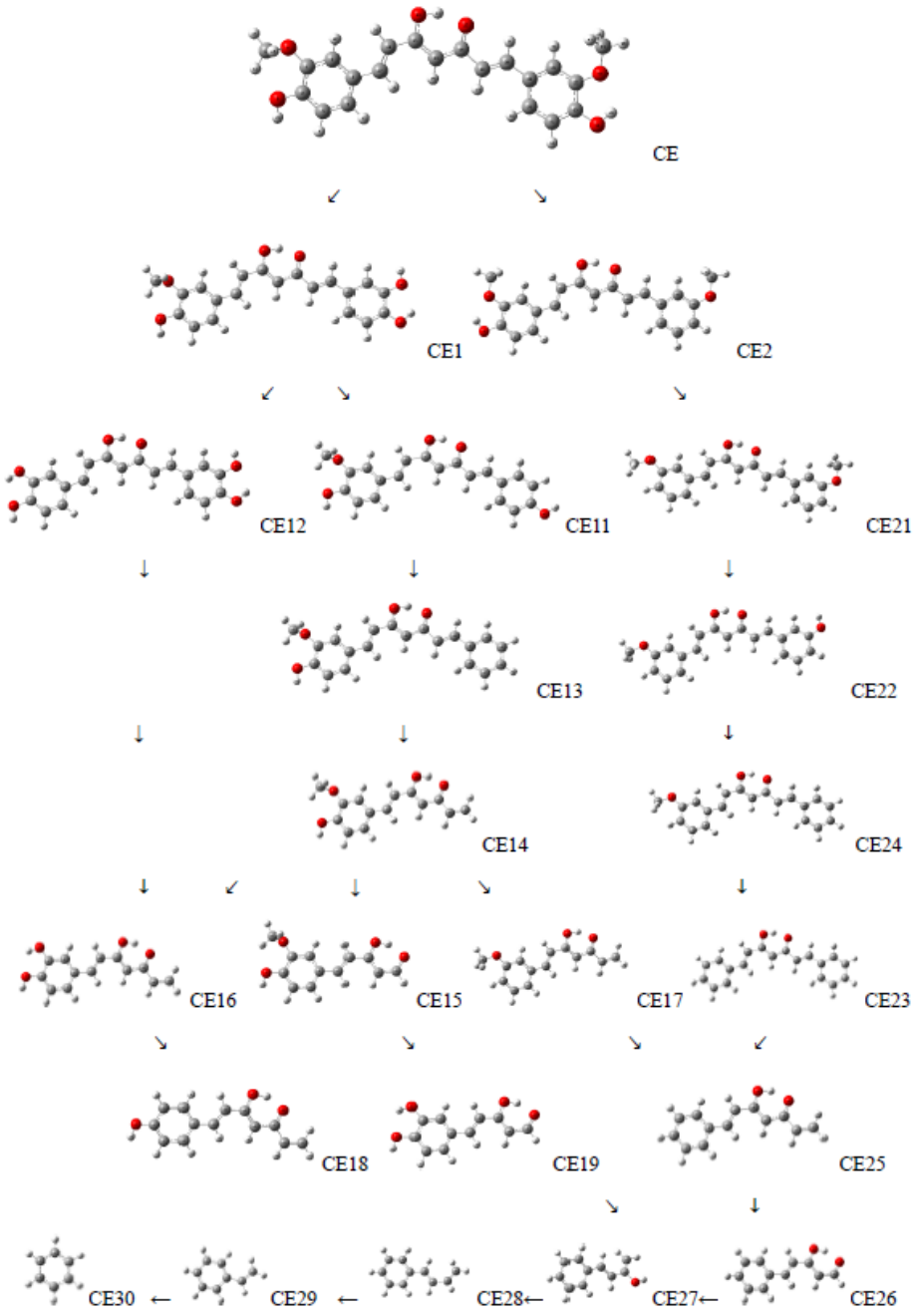


Figure 8. Degradation pathway of the CE molecule and its fragments

CONCLUSION and DISCUSSION

An examination of the bond lengths between the atoms of the MR19 molecule shown in Figure 3(a) and Table 1 suggests that the bonds specifically C_4S_{12}, C_2Cl_{11}, $S_{12}O_{13}$, and so forth are expected to break in that order, starting with the longest bond. These long bonds are highlighted in bold in Table 1. The fragments MR193, MR191, and MR192, shown in Figure 6, are formed precisely through the breaking of these bonds.

An examination of the bond angles between the atoms of the MR19 molecule in Figure 3(b) and Table 1 reveals the widest bond angles to be in descending order $C_{23}N_{22}C_{18}$, N15C6C5, $C_{17}C_{19}C_{34}$, $C_5C_4S_{12}$, $C_1C_2Cl_{11}$, and so forth; these wide bond angles are highlighted in bold in Table 1. The first wide bond angle signals that the benzene ring attached to atom N_{23} will break off and dissociate from the molecule. The second wide bond angle indicates that the benzene ring substituted with four R groups attached to atom N_{15} will break off and dissociate from the molecule. The third widest bond angle, $C_{17}C_{19}C_{34}$, points to the detachment of the methyl group attached to atom C_{10}. The $C_5C_4S_{12}$ bond angle indicates that the C_4S_{12} bond, which links the SO_3 group, will break, causing it to dissociate from the molecule. Finally, the wide $C_1C_2Cl_{11}$ bond angle indicates that the C_2Cl_{11} bond, connecting the Cl_{11} atom, will break.

An examination of the Mulliken charges of the electronegative atoms in the MR19 molecule (Table 1) reveals the following order of electronegativity: O_9, O_{13}, O_{39}, O_{38}, N_{22}, etc. Given that the O_{39} and O_{40} atoms are bonded to the S atom, and the N_{15} and N_{16} atoms are linked by double bonds, it is unlikely that the cleavage of these stable bonds would occur during the initial stages of the molecule's fragmentation. O_9 and O_{13} are observed to be the most electronegative atoms: consequently, the C_5O_9 and $S_{12}O_{13}$ bonds are expected to break in the first step. The N_{22} atom, ranking fifth in electronegativity, indicates that the benzene ring attached to it will undergo cleavage.

When the energy values in Table 2 are examined and ordered starting from the lowest that is, the most stable fragment, the sequence observed is MK192, MR192, MR196, ... The MK192 fragment is predicted to be the first fragment formed, with an energy of -2031.295055 Au. This fragment arises from the detachment of a methyl substituent from the molecule following the cleavage of the C_{10}–C_{34} bond. Although the C_{19}–C_{34} bond length in Table 1 is not among the long bond lengths highlighted in dark color, the $C_{17}C_{19}C_{34}$ bond angle is one of the wide bond angles. This first pathway intersects with MR197 a degradation product of MR196 that occupies the fourth position in this energy series at a later stage of the degradation chain.

Based on its MR192 value of 1995.357751 Au, it has been predicted to be a secondary fragment formed without the need for external energy. The MR192 fragment is formed through the cleavage of the S_{12}–O_{13} bond within the SO_3H substituent, resulting in the loss of an OH group. In Table 1, the S_{12}–O_{13} bond is identified as one of the long bonds, and cleavage at this site is indeed expected. O_{13} is additionally marked in bold among the electronegative atoms in Table 1, with a value of -0.615005.

Although the MR191 fragment does not possess the lowest energy value, it is nevertheless depicted as a dissociation product because the detachment of the Cl atom from the molecule were such a cleavage to occur would constitute the final step. In the second stage of degradation, this fragment binds to MR193 which is a fragment of MR192 from the second arm and becomes incorporated into one of the main arms. This degradation pathway, proceeding via multiple routes with a similar mechanism, culminates in the MR199 and MR1955 fragments. Mordant Red 19 is a synthetic azo dye and all degradation pathways of this molecule are shown in Figure 6. When azo dyes break down whether in the body or in the environment, they can release aromatic amines that may be mutagenic, genotoxic, or carcinogenic; MR1955 is one such aromatic amine.

Figure 7 illustrates the fragmentation steps for all branches of the carminic acid molecule; specifically, there are eleven distinct branches. Among these, the first fragment expected to form and the one with the lowest energy is the fragment designated as CA26. Based on the low-energy ranking, the CA26 fragment is expected to be followed by fragments numbered 39, 49, 18, 44, 30, 20, 41, 24, 46, and 32, in that order.

The numbering is based on the atom indices in the initial optimized state of the CA molecule. Specifically, the CA26 fragment is formed by the cleavage of the C_{11}–C_{26} bond that is, the detachment of the methyl group bonded to the C_{26} atom from the molecule. In Table 1, the longest bond angle the C_{11}–C_{26} bond at 1.51395 Au already signals that this bond will break. The wide bond angle of 119.72390° for $C_{12}C_{11}C_{26}$ in Table 1 also confirms this. Instead of determining and displaying the degradation pathway for every low-energy fragment on a single page, the three fragments with the lowest energies respectively numbers 26, 39, and 49 were selected.

The reasoning steps followed during the fragmentation of the MR19 molecule along this fragmentation pathway were as follows: the bond lengths, bond angles, and Mulliken charges in Table 1 were examined; the energy values in Table 3 were ordered starting from the lowest value; and the fragmentation pathway presented in Table 7 was determined by evaluating all the data. Figure 7 shows

the degradation mechanism of the CA molecule, which forms the closed-ring CA10 structure as the final part of the reaction.

Two-dimensional representations of the keto and enol forms of the curcumin (C) molecule are shown in Figure 1, while three-dimensional DFT-optimized structures are shown in Figure 2. Table 1 presents the energy values for the enol form of curcumin (CE) and the two versions of its keto form CK (cis) and CK (trans). The enol form was selected for the degradation pathway, and the CE molecule was analysed. Figure 5 and Table 1 show the bond lengths and bond angles of the atoms within the CE molecule, while Table 1 also presents the Mulliken charges of the electronegative atoms among them. When these values are examined together with the energy values in Table 4, the degradation pathway shown in Figure 8 emerges. Since the degradation steps for MR19 and CA molecules have already been clearly described, the same repetition has been avoided for CE. When the structures in Figure 1 and Figure 2 are examined, the symmetric structure of the Curcumin (C) molecule distinguishes it from other two molecules. Except for the keto or enol structure at the centre of the molecule, the two end sections are identical to each other.

MR19 (an azo dye used in textile dyeing), CA (a colorant derived from the cochineal insect), and CE (a food coloring derived from the plant known as turmeric in Türkiye) are the three different coloring agents selected for this study. Degradation mechanisms have been determined using DFT for three molecules with an anthraquinone structure Alizarin Red S, Alizarin, and Purpurin which are organic colorants that produce a natural red color (Eren, 2026). In addition to these substances with an anthraquinone structure, the MR19, CA, and CE molecules under investigation represent examples of both synthetically produced colorants and those derived from plant and animal sources. Little attention is generally paid to the environmental fate of dyes during their production and dyeing processes, or when they are disposed of as waste. This study aims to shed light on this uncertainty through the degradation pathways shown in Figures 6, 7, and 8.

Acknowledgment

This study was supported by the Tekirdag Namik Kemal University Scientific Research Projects Coordination Unit (NKÜBAP) under project number ***NKUBAP.00.GA.25.669****.*

The HP Z04 Workstation used in this study was obtained from research project number ***NKUBAP.00.GA.23.506****.*

References

Altuğ, T. (2009). Gıda Katkı Maddeleri. İçinde: (268). İzmir: Sidas Yayını.

Anbar M, Neta P (1967). Int J Appl Radiat Isot, 18: 493.

Aschmann, S M, Tuazon, E C ve Atkinson R (2005). J. Phys. Chem. A, 109: 11828,2282.

Aschmann S M, Atkinson R (2006). J. Phys. Chem. A, 110: 13029.

Aschmann S M, Tuazon E C, Long W D, Atkinson R (2010). J. Phys. Chem. A, 114: 3523.

Atkinson R, Aschmann S M, Goodman M A, Winer A M (1988). Int. J. Chem. Kinet., 20: 273.

Atkinson R, Arey J (2003). Chem. Rev., 103: 4605.

Baumann, H., Federkiel, W. (1957) "Verfahren zum Färben von natürlichen Fasern und Fasern aus Polyamiden oder Polyurethanen [Process for the dyeing of natural fibres and fibres of polyamides or polyurethanes]". BASF AG.

Buxton G. V., Greenstock C. L., Phillips Helman W., Ross A. B. (1988). *J. Phys. Chem. Ref. Data*, *17*(2), 513.

Curcumin, E 100, page 9". Specifications for food additives listed in Annexes II and III to Regulation (EC) No 1333/2008 of the European Parliament.

Eren B. (2026). Theory, Research and Discussions in Natural Sciences and Mathematics, İzmir: Duvar Publishing ,16-29

Eren B., Yalcin Gurkan Y. (2024). Bulgarian Chemical Communications. 56, 242-246.

Eren, B., Yalcin Gurkan, Y. (2023a) Innovative Research in Engineering, Duvar Publishing, İzmir Turkey,4,49.

Eren, B., Yalcin Gurkan, Y. (2023b*). International Journal of Advanced Natural Sciences and Engineering Research,* 7, 31.

Eren B., Yalcin Gürkan Y. (2022). *Bulgarian Chemical Communications*, 54, 224.

Eren B, Yalcin Gurkan Y (2017). J. Serb. Chem. Soc., 82(3), 277–287

Fay I.W. (1919) The chemistry of the coal-tar dyes. D. Van Nostrand Company: New York.

Gaussian16 & Gauss View 6 & TCP Linda Programs, (2009).

Halliwell B., Grootveld M., Gutteridge J. M. C. (1998). *Methods Biochem. Anal.,* 33: 59.

Farooqui, T., Akhlaq A. Farooqui, A.A. (2019). Chapter 2 - Curcumin: Historical Background, Chemistry, Pharmacological Action, and Potential Therapeutic Value. Curcumin for Neurological and Psychiatric Disorders. Neurochemical and Pharmacological Properties, 23-44.

Hunger, K., Mischke, P., Rieper, W., Raue R., Kunde, K., Engel, A. (2005) "Azo Dyes" in Ullmann's Encyclopaedia of Industrial Chemistry. Wiley-VCH, Weinheim.

Hunger, K., Mischke, P., Rieper, W., Zhang, S. (2009). Azo Dyes, 5. Developing Dyes. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. 1–19.

Li, N., Tan, S.N, Cui, J., Guo, N., Wang, W., Zu, Y. G., Jin, S., Xu, X. X., Liu, Q., Fu, Y. J. 2014. "PA-1, a novel synthesized pyrrolizidine alkaloid, inhibits the growth of Escherichia coli and Staphylococcus aureus by damaging the cell membrane", The Journal of Antibiotics, 67, 689-696.

Manzoor, J., Sharma M. (2020), Impact of Textile Dyes on Human Health and Environment 8,163.

McNaughty, A.D., Wilkinson, A. (1997), the Gold Book: Compendium of Chemical Terminology. (2nd ed.) IUPAC.

Chapter 7

Osculating Helices in the Multiplicative Euclidean Space. $\mathbb{E}_*^3$ and Their Duality with Normal Helices

Derya SAĞLAM[1], Umut SELVİ[2]

Abstract

An osculating helix is a space curve that carries a vector field lying in its osculating plane, with constant coordinates with respect to the Frenet frame, which keeps a constant angle with a fixed direction; together with normal and rectifying helices, these curves form a family of generalized helices introduced by Lucas and Ortega-Yagües. In this chapter we study osculating helices in the multiplicative Euclidean space $\mathbb{E}_*^3$, in which multiplicative (non-Newtonian) calculus plays the role of ordinary calculus, when the osculating field is orthogonal to the fixed direction. We show that the axis of such a curve is forced to be the direction of an explicit vector field built from the curvature, the torsion and the Frenet frame, and we derive from this fact the natural equation of multiplicative osculating helices. We then show that integrating the binormal of a nonplanar curve, in the multiplicative sense, produces a curve with the same multiplicative Darboux vector. This yields a precise duality between osculating and normal helices, from which we recover the natural equation of multiplicative normal helices and a description of osculating helices by means of multiplicative cylinders. Finally, we compute in closed form the osculating helix dual to a normal helix on a multiplicative circular cylinder; its ratio of curvature to torsion is a multiplicatively linear function of the arc length. All natural equations are also written in ordinary arithmetic.

Keywords: Osculating helix; normal helix; integral of the binormal; duality; multiplicative Darboux vector; multiplicative Euclidean space; non-Newtonian calculus

2020 Mathematics Subject Classification: Primary 53A04; Secondary 11U10, 08A05.

[1] Prof. Dr., Depatment of Mathematics, Ankara Hacı Bayram Veli University, ORCID: 0000-0003-4323-1901, derya.saglam@hbv.edu.tr

[2] Res. Ast., Depatment of Mathematics, Ankara Hacı Bayram Veli University, ORCID: 0000-0003-1015-0056, umut.selvi@hbv.edu.tr

1 Introduction

The osculating plane of a space curve at a point is the plane that fits the curve best near that point; it is spanned by the unit tangent T and the principal normal N. Two classical families of curves are defined by means of vectors of this plane: the cylindrical helices, or curves of constant slope, whose tangent is inclined at a fixed angle to some direction [18, p. 33], and the slant helices, for which the principal normal behaves in the same way [8, 10]. A unit vector of the osculating plane with constant Frenet coordinates, $W = \cos\theta\, T + \sin\theta\, N$, interpolates between these two situations. Following Lucas and Ortega-Yagües [11], a curve is called an osculating helix when such a field is inclined at a constant angle to a fixed direction V. Normal and rectifying helices are defined in the same way with fields of the normal and of the rectifying plane, and the whole family is described in [11] in terms of Frenet-constant vector fields, that is, fields that are carried rigidly by the Frenet frame. The name should not be confused with the osculating circular helix of a curve at a point, which belongs to the theory of contact [18, p. 42].

When W is orthogonal to V, the natural equations of osculating and normal helices were obtained in [11], where it was also observed that the two classes are dual to each other: integrating the binormal of a normal helix produces an osculating helix, and conversely. The approach of [11] has since been extended beyond the Euclidean 3-space: timelike osculating, normal and rectifying helices in the Minkowski space $\mathbb{E}_1^3$ were characterized in [13], and helices with Frenet-constant vector fields in the Euclidean space $\mathbb{E}^4$ were studied in [15]. Slant helices have also been investigated in other ambient spaces, such as the dual space $\mathbb{D}^3$ [12] and the dual Lorentzian space $\mathbb{D}_1^3$ [14].

Non-Newtonian calculi, introduced by Grossman and Katz [6], replace the arithmetic of the real numbers by the arithmetic induced by a bijection. The multiplicative calculus, generated by the exponential function, turns sums into products; its basic properties and several applications can be found in [2, 17] (which, however, use the geometric version of the multiplicative derivative; see Section 2.1), and Georgiev developed the corresponding differential calculus [5] and differential geometry [4]. In the multiplicative Euclidean space $\mathbb{E}_*^3$, curves of exponential type are described by elementary invariants, and several classical problems have recently been studied there, such as rectifying curves [1], general (cylindrical) and slant helices [7], and normal helices [16].

In this chapter we investigate the multiplicative osculating helices whose osculating Frenet-constant field is multiplicatively orthogonal to the axis. Section 3 establishes that the axis of such a helix is necessarily parallel to an explicit field

Y_θ built from the curvature, the torsion and the Frenet frame (Lemma 3.1), and we deduce the natural equation (Theorem 3.2) from the requirement that Y_θ keep a constant direction; this argument differs from the one used in [11]: it yields the unit axis (11) in closed form, without integrating differential equations for the coefficients of the axis. In Section 4 we prove that the multiplicative integral of the binormal of a nonplanar curve has the same multiplicative Darboux vector as the curve itself (Proposition 4.1). The duality between osculating and normal helices (Theorem 4.2) is an immediate consequence, and it yields the natural equation of multiplicative normal helices (Corollary 4.3) and a description of osculating helices by means of multiplicative cylinders (Corollary 4.4). In Section 5 we compute in closed form the osculating helix dual to a normal helix on a multiplicative circular cylinder (Figure 1).

The coordinate-wise logarithm identifies $\mathbb{E}_*^3$ with $\mathbb{R}^3$ together with all the structures used here (Remark 2.3), so most results of this chapter correspond to Euclidean statements of [11] about the logarithmic picture of the curve and are best described as their multiplicative counterparts. Two ingredients do not appear in [11], even in the Euclidean setting: the invariance of the Darboux vector under the binormal construction, together with the sign that it requires (Proposition 4.1), and the closed-form osculating helix of Section 5.2. The proofs, however, are carried out intrinsically, with the single exception of Corollary 4.4, which combines [11, Theorem 2.2] with the transfer principle, and each characterization is also translated into ordinary arithmetic.

2 Notation and preliminaries

2.1 The logarithmic picture of multiplicative calculus

Let $\mathbb{R}_* = (0, \infty)$. The basic device of this chapter is the logarithmic picture: a positive number a is represented by $\ln a$, a point $x = (x_1, x_2, x_3)$ with positive coordinates by $\Lambda x = (\ln x_1, \ln x_2, \ln x_3)$, and a positive function f defined on an interval $I \subset \mathbb{R}_*$ by $\hat{f}(u) = \ln f(e^u)$, $u \in \ln I$. Multiplicative arithmetic, calculus and geometry [1, 4, 5, 6] can be summarized by saying that the logarithmic picture of each multiplicative operation is the corresponding ordinary operation. For numbers this gives

$$a +_* b = ab, \qquad a -_* b = \frac{a}{b}, \qquad a \cdot_* b = e^{\ln a \ln b},$$
$$a /_* b = e^{\ln a / \ln b} \quad (b \neq 1),$$

and $(\mathbb{R}_*, +_*, \cdot_*)$ is an ordered field with zero $0_* = 1$ and unit $1_* = e$, isomorphic to $(\mathbb{R}, +, \cdot)$ through ln. We write $-_* a = 1/a$, $n_* = e^n$ for $n \in \mathbb{R}$, $a^{n_*} = e^{(\ln a)^n}$,

$|a|_* = e^{|\ln a|}$, and $F_*(x) = e^{F(\ln x)}$ for the multiplicative version of a real function F; thus $\sin_*\theta = e^{\sin(\ln\theta)}$, $\cos_*^{2_*}\theta = (\cos_*\theta)^{2_*}$, and so on. A multiplicative angle $\theta \in \mathbb{R}_*$ has the ordinary angle $\ln\theta$ as its logarithmic picture; for example, the multiplicative right angle is $(\pi/2)_* = e^{\pi/2}$. All algebraic and trigonometric identities remain valid in multiplicative form.

For functions, the multiplicative derivative and the multiplicative integral [1, 5] are characterized by

$$\widehat{f^*} = \hat{f}', \qquad *\int_a^b f(s) \cdot_* d_*s = \exp\left(\int_{\ln a}^{\ln b} \hat{f}(u)\, du\right), \tag{1}$$

that is, $f^*(s) = e^{s(\ln f)'(s)}$ and $*\int_a^b f(s) \cdot_* d_*s = \exp\left(\int_a^b \ln f(s)\, ds/s\right)$. This is the bigeometric derivative introduced by Grossman and Katz [6], which is also used in [1, 7]; it should not be confused with the geometric multiplicative derivative $e^{(\ln f)'(s)}$ of [2, 17], of which it is the s-th power. Consequently every rule of ordinary calculus has a multiplicative version: the product, quotient and chain rules hold with the multiplicative operations, $f^* \equiv 0_*$ on an interval exactly when f is constant, and the fundamental theorem of calculus takes the form

$$(*\int_{s_0}^{s} f(u) \cdot_* d_*u)^* = f(s), \quad *\int_a^b g^*(s) \cdot_* d_*s = g(b) -_* g(a). \tag{2}$$

Vector-valued functions are differentiated and integrated componentwise, and (2) holds for them as well. To express multiplicative statements in ordinary arithmetic we shall use the rule

$$\ln f^*(s) = s\,(\ln f)'(s). \tag{3}$$

2.2 The space $\mathbb{E}_*^3$ and the Frenet apparatus

The multiplicative Euclidean space $\mathbb{E}_*^3$ is $\mathbb{R}_*^3$ with the operations $x +_* y = (x_1y_1, x_2y_2, x_3y_3)$ and $a \cdot_* x = \left(x_1^{\ln a}, x_2^{\ln a}, x_3^{\ln a}\right)$, the zero vector $0_* = (1,1,1)$ and the canonical basis $e_1 = (e,1,1)$, $e_2 = (1,e,1)$, $e_3 = (1,1,e)$. The multiplicative scalar product, length and vector product are defined by prescribing their logarithmic pictures [1, 4]:

$$\ln\langle x, y\rangle_* = \langle \Lambda x, \Lambda y\rangle, \qquad \ln \| x \|_* = \| \Lambda x \|, \qquad \Lambda(x \times_* y) = \Lambda x \times \Lambda y.$$

Hence $\Lambda: \mathbb{E}_*^3 \to \mathbb{R}^3$ is an isomorphism of all these structures. Orthogonality, angles, lines, planes and orientation are defined as in $\mathbb{R}^3$, and each formula of vector algebra in $\mathbb{R}^3$ remains true in $\mathbb{E}_*^3$ with the multiplicative operations.

Let $\alpha: I \subset \mathbb{R}_* \to \mathbb{E}_*^3$ be a curve with sufficiently many multiplicative derivatives, parametrized by multiplicative arc length ($\| \alpha^* \|_* = 1_*$) and such that α^{**} never equals 0_*. Its multiplicative Frenet apparatus [1] is

$$T = \alpha^*, \qquad \kappa = \| \alpha^{**} \|_*, \qquad N = (1_* /_* \kappa) \cdot_* \alpha^{**},$$
$$B = T \times_* N, \qquad \tau = \langle N^*, B \rangle_*,$$

and it satisfies the multiplicative Frenet–Serret equations

$$T^* = \kappa \cdot_* N, \qquad N^* = -_* \kappa \cdot_* T +_* \tau \cdot_* B, \qquad B^* = -_* \tau \cdot_* N. \tag{4}$$

The logarithmic picture $\hat{\alpha}(u) = \Lambda\big(\alpha(e^u)\big)$ is a unit speed curve in $\mathbb{R}^3$; its Frenet vectors are $\hat{T}, \hat{N}, \hat{B}$, and its curvature and torsion are $\ln\kappa$ and $\ln\tau$. More generally, the Frenet apparatus of a curve that is not parametrized by multiplicative arc length is obtained by computing, with the usual formulas, the Frenet apparatus of its logarithmic picture. The curve is nonplanar when τ never takes the value 0_*, and $\rho = \tau /_* \kappa$ denotes its multiplicative Lancret curvature.

The multiplicative Darboux vector of α is $D = \tau \cdot_* T +_* \kappa \cdot_* B$. A direct computation with (4) and the rules $T \times_* N = B$, $N \times_* B = T$, $B \times_* T = N$ shows that

$$T^* = D \times_* T, \qquad N^* = D \times_* N, \qquad B^* = D \times_* B, \tag{5}$$

so that D is the multiplicative angular velocity of the Frenet frame (see [9, pp. 43–45] for the Euclidean notion).

2.3 Frenet-constant fields and helices

Lemma 2.1. *A vector field W along α has constant coordinates with respect to $\{T, N, B\}$ if and only if $W^* = D \times_* W$.*

Proof. Write $W = w_1 \cdot_* T +_* w_2 \cdot_* N +_* w_3 \cdot_* B$. By the product rule and (5),

$$W^* = w_1^* \cdot_* T +_* w_2^* \cdot_* N +_* w_3^* \cdot_* B +_* D \times_* W.$$

Hence $W^* = D \times_* W$ exactly when $w_1^* = w_2^* = w_3^* = 0_*$, that is, when the coordinates are constant. □

Vector fields with constant Frenet coordinates are called Frenet-constant, or $\mathcal{F}$-constant. By Lemma 2.1, they depend on the curve only through its Darboux vector, a fact that will be essential in Section 4. Since $\langle W, W \rangle_*^* = 2_* \cdot_* \langle D \times_* W, W \rangle_* = 0_*$, $\mathcal{F}$-constant fields have constant norm.

Definition 2.2. Following [11], we call α a *multiplicative helix* with *axis* V, where $V \neq 0_*$ is a constant vector, if the multiplicative angle between V and some unit $\mathcal{F}$-constant field W along α does not depend on s. The helix is *osculating*, *normal* or *rectifying* according as W is perpendicular to B, to T or to N.

Every unit osculating $\mathcal{F}$-constant field can be written as

$$W_\theta = \cos_*\theta \cdot_* T +_* \sin_*\theta \cdot_* N \tag{6}$$

for a constant multiplicative angle θ. For $\theta = 0_*$ it is T and for $\theta = e^{\pi/2}$ it is N, so the osculating fields interpolate between the fields defining the multiplicative cylindrical and slant helices of [7]. Replacing W_θ by $-_* W_\theta$ we may assume that $e^{-\pi/2} \leq \theta \leq e^{\pi/2}$. We shall always consider the case in which W is orthogonal to the axis, $\langle W, V\rangle_* = 0_*$.

Remark 2.3. The logarithmic picture of D is the Darboux vector of $\hat{\alpha}$. Together with Lemma 2.1 this shows that Λ maps the $\mathcal{F}$-constant fields of α onto those of $\hat{\alpha}$; hence the multiplicative osculating, normal and rectifying helices with axis V are exactly the curves whose logarithmic pictures are Euclidean helices of the same kind, in the sense of [11], with axis ΛV. This transfer principle allows one to carry results of [11] to $\mathbb{E}^3_*$ and back. Nevertheless, except in Corollary 4.4, we give multiplicative proofs, and we use (3) to write each result in ordinary arithmetic.

3 The natural equation of osculating helices

In this section α denotes a nonplanar curve of unit multiplicative speed, and θ a constant. Together with the field W_θ of (6) we use the unit osculating field

$$U_\theta = -_* \sin_*\theta \cdot_* T +_* \cos_*\theta \cdot_* N. \tag{7}$$

The fields W_θ, U_θ and B form a positively oriented orthonormal frame, with $W_\theta \times_* U_\theta = B$, $U_\theta \times_* B = W_\theta$ and $B \times_* W_\theta = U_\theta$. Since $N = \sin_*\theta \cdot_* W_\theta +_* \cos_*\theta \cdot_* U_\theta$, the equations (4) give

$$\begin{aligned} W_\theta^* &= \kappa \cdot_* U_\theta +_* \sin_*\theta \cdot_* \tau \cdot_* B, \\ U_\theta^* &= -_* \kappa \cdot_* W_\theta +_* \cos_*\theta \cdot_* \tau \cdot_* B, \\ B^* &= -_* \tau \cdot_* (\sin_*\theta \cdot_* W_\theta +_* \cos_*\theta \cdot_* U_\theta). \end{aligned} \tag{8}$$

The key observation is that the axis of an osculating helix can be read off from the curve.

Lemma 3.1. *Put*

$$Y_\theta = W_\theta \times_* W_\theta^* = \kappa \cdot_* B -_* \sin_*\theta \cdot_* \tau \cdot_* U_\theta. \tag{9}$$

If V is a constant vector with $\langle W_\theta, V\rangle_* = 0_*$, then $V = f \cdot_* Y_\theta$ for some function f. In particular, the axis of a multiplicative osculating helix whose field W_θ is orthogonal to it is parallel to Y_θ.

Proof. Differentiating $\langle W_\theta, V\rangle_* = 0_*$ we get $\langle W_\theta^*, V\rangle_* = 0_*$. By (8) the coordinate of W_θ^* along U_θ is $\kappa \neq 0_*$, so W_θ and W_θ^* are linearly independent, and V, being orthogonal to both, is parallel to their cross product. Finally, $W_\theta \times_* U_\theta = B$ and $W_\theta \times_* B = -_* U_\theta$ give the expression of Y_θ in (9). □

The field Y_θ never vanishes, because its coordinate along B is $\kappa > 0_*$, and $\| Y_\theta \|_*^{2_*} = \kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*}$.

Theorem 3.2 (cf. [11, Theorem 3.1])**.** *Assume that α is nonplanar and has unit multiplicative speed, and fix a constant $\theta \neq 0_*$ with $e^{-\pi/2} \leq \theta \leq e^{\pi/2}$. Then the following three conditions are equivalent.*

(a) Some nonzero constant vector is multiplicatively orthogonal to W_θ; in other words, α is a multiplicative osculating helix with W_θ perpendicular to its axis.

(b) The field Y_θ of (9) has a constant direction.

(c) The functions κ and τ are related by

$$\left(\tau /_* \left(\kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*}\right)\right) \cdot_* (\kappa /_* \tau)^* = \cot_*\theta. \tag{10}$$

In that case the unit vector

$$V_\theta = (\kappa \cdot_* B -_* \sin_*\theta \cdot_* \tau \cdot_* U_\theta) /_* (\kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*})^{(1/2)_*} \tag{11}$$

is constant, and every axis orthogonal to W_θ is a multiple of it. Written with ordinary operations, (10) is

$$\frac{\ln\tau}{\ln^2\kappa + \sin^2(\ln\theta)\ln^2\tau}\; s\,\frac{d}{ds}\left(\frac{\ln\kappa}{\ln\tau}\right) = \cot(\ln\theta). \tag{12}$$

Proof. (a)⇒(b). Let $V \neq 0_*$ be a constant vector orthogonal to W_θ. By Lemma 3.1, $V = f \cdot_* Y_\theta$; as V and Y_θ never vanish, neither does f, and $Y_\theta = (1_* /_* f) \cdot_* V$ has the constant direction of V.

(b)⇒(a). If Y_θ has a constant direction, then $V_\theta = Y_\theta /_* \| Y_\theta \|_*$ is a constant unit vector, perpendicular to W_θ because $Y_\theta = W_\theta \times_* W_\theta^*$ is.

(b)⇔(c). A nowhere vanishing field Y has a constant direction if and only if Y^* is parallel to Y at every point. Indeed, if $Y^* = h \cdot_* Y$, then the derivative of the unit field $\tilde{Y} = Y /_* \| Y \|_*$ is also parallel to $\tilde{Y}$; being orthogonal to $\tilde{Y}$, it is zero. Conversely, if $\tilde{Y}$ is constant, then $Y^* = \| Y \|_*^* \cdot_* \tilde{Y}$. Differentiating (9) with (8), the terms along W_θ cancel and we obtain

$$Y_\theta^* = (\kappa^* -_* \sin_*\theta \cdot_* \cos_*\theta \cdot_* \tau^{2_*}) \cdot_* B \\ -_* (\cos_*\theta \cdot_* \kappa \cdot_* \tau +_* \sin_*\theta \cdot_* \tau^*) \cdot_* U_\theta.$$

Both Y_θ and Y_θ^* lie in the plane spanned by U_θ and B, so they are parallel exactly when the determinant formed by their coordinates, namely

$$\Delta_\theta = (-_*\sin_*\theta \cdot_* \tau) \cdot_* (\kappa^* -_* \sin_*\theta \cdot_* \cos_*\theta \cdot_* \tau^{2_*}) \\ +_* \kappa \cdot_* (\cos_*\theta \cdot_* \kappa \cdot_* \tau +_* \sin_*\theta \cdot_* \tau^*),$$

vanishes. Expanding and collecting terms we find

$$\Delta_\theta = \sin_*\theta \cdot_* (\kappa \cdot_* \tau^* -_* \tau \cdot_* \kappa^*) +_* \cos_*\theta \cdot_* \tau \cdot_* \left(\kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*}\right).$$

By the quotient rule, $\kappa \cdot_* \tau^* -_* \tau \cdot_* \kappa^* = -_*\tau^{2_*} \cdot_* (\kappa /_* \tau)^*$. Hence $\Delta_\theta = 0_*$, that is, Y_θ^* is parallel to Y_θ, if and only if

$$\sin_*\theta \cdot_* \tau^{2_*} \cdot_* (\kappa /_* \tau)^* = \cos_*\theta \cdot_* \tau \cdot_* \left(\kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*}\right),$$

and, as $\sin_*\theta \neq 0_*$, $\tau \neq 0_*$ and $\kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*} > 0_*$, this is equivalent to (10). The statement about the axis follows from Lemma 3.1, and (12) is obtained from (10) by taking logarithms, using (3) and $\ln(\kappa /_* \tau) = \ln\kappa/\ln\tau$. □

Remark 3.3. (1) *The excluded value* $\theta = 0_*$. For $\theta = 0_*$ we have $W_\theta = T$ and $Y_\theta = \kappa \cdot_* B$, which has a constant direction only if B is constant, that is, only if α is planar. Accordingly, the determinant Δ_θ in the proof reduces to $\tau \cdot_* \kappa^{2_*}$, which never vanishes for a nonplanar curve. Thus a nonplanar curve is never an osculating helix with $W = T$ orthogonal to the axis.

(2) *The values* $\theta = e^{\pm\pi/2}$. Then $W_\theta = \pm N$, $\cot_*\theta = 0_*$, and (10) becomes $(\kappa /_* \tau)^* = 0_*$. Moreover $U_\theta = \mp T$ and $Y_\theta = \tau \cdot_* T +_* \kappa \cdot_* B = D$. Hence the osculating helices with $W = \pm N$ orthogonal to the axis are the multiplicative cylindrical helices, whose axis is the direction of the Darboux vector, as predicted by the multiplicative Lancret theorem [7, Theorem 4.12].

(3) *Equivalent form.* Since $(\kappa /_* \tau)^* = -_*\rho^* /_* \rho^{2_*}$, equation (10) can also be written as $\left(\kappa^{2_*} /_* \left(\kappa^{2_*} +_* \sin_*^{2_*}\theta \cdot_* \tau^{2_*}\right)\right) \cdot_* \rho^* = -_*\cot_*\theta \cdot_* \tau$.

(4) *Monotonicity.* If $\theta \neq e^{\pm\pi/2}$, the right-hand side of (10) is different from 0_*, so $(\kappa /_* \tau)^*$ never vanishes: along a multiplicative osculating helix whose osculating field is orthogonal to the axis and which is not a cylindrical helix, the ratio $\kappa /_* \tau$ is strictly monotone.

4 Duality between osculating and normal helices

4.1 The multiplicative integral of the binormal

In this section α is again nonplanar and of unit multiplicative speed. Since τ is continuous and never equals 0_*, its sign is constant; we put $\varepsilon = 1_*$ if $\tau > 0_*$ and $\varepsilon = -_* 1_*$ if $\tau < 0_*$, so that $|\tau|_* = \varepsilon \cdot_* \tau$ and $\varepsilon^{2_*} = 1_*$. For $s_0 \in I$, the *binormal dual* of α is the curve

$$\bar{\alpha}(s) = * \int_{s_0}^{s} B\,(u) \cdot_* d_* u, \qquad s \in I, \tag{13}$$

where the multiplicative integral is computed componentwise. Any multiplicative primitive of B differs from $\bar{\alpha}$ by a multiplicative translation and has the same properties.

Proposition 4.1. *With the notation above, the following hold.*

(i) $\bar{\alpha}$ is parametrized by multiplicative arc length, and its Frenet apparatus is

$$\bar{T} = B, \qquad \bar{N} = -_* \varepsilon \cdot_* N, \qquad \bar{B} = \varepsilon \cdot_* T, \qquad \bar{\kappa} = |\tau|_*, \qquad \bar{\tau} = \kappa.$$

In particular, $\bar{\alpha}$ is nonplanar and has positive multiplicative curvature.

(ii) $\bar{\alpha}$ and α have the same multiplicative Darboux vector, $\bar{D} = D$. Consequently, a vector field along I is $\mathcal{F}$-constant for $\bar{\alpha}$ if and only if it is $\mathcal{F}$-constant for α.

(iii) At each $s \in I$, the normal plane of α is parallel to the osculating plane of $\bar{\alpha}$, the osculating plane of α to the normal plane of $\bar{\alpha}$, and the two rectifying planes are parallel.

(iv) The binormal dual of $\bar{\alpha}$ is $\varepsilon \cdot_* \big(\alpha -_* \alpha(s_0)\big)$; thus, up to a multiplicative isometry, the construction is involutive.

Proof. (i) By (2), $\bar{\alpha}^* = B$, which is a unit vector. By (4), $\bar{T}^* = B^* = -_* \tau \cdot_* N = |\tau|_* \cdot_* (-_* \varepsilon \cdot_* N)$, and $-_* \varepsilon \cdot_* N$ is a unit vector; hence $\bar{\kappa} = |\tau|_*$ and $\bar{N} = -_* \varepsilon \cdot_* N$. Then $\bar{B} = \bar{T} \times_* \bar{N} = -_* \varepsilon \cdot_* (B \times_* N) = \varepsilon \cdot_* T$, and

$$\bar{\tau} = \langle \bar{N}^*, \bar{B} \rangle_* = -_* \varepsilon^{2_*} \cdot_* \langle N^*, T \rangle_* = -_* \langle N^*, T \rangle_* = \kappa,$$

because $\langle N^*, T \rangle_* = -_* \kappa$ by (4).

(ii) By (i), $\bar{D} = \bar{\tau} \cdot_* \bar{T} +_* \bar{\kappa} \cdot_* \bar{B} = \kappa \cdot_* B +_* (|\tau|_* \cdot_* \varepsilon) \cdot_* T = D$, since $|\tau|_* \cdot_* \varepsilon = \varepsilon^{2_*} \cdot_* \tau = \tau$. The second assertion follows from Lemma 2.1, which characterizes the $\mathcal{F}$-constant fields as the solutions of $W^* = D \times_* W$.

(iii) By (i), the plane spanned by N and B is spanned by $\bar{N}$ and $\bar{T}$, the plane spanned by T and N is spanned by $\bar{B}$ and $\bar{N}$, and the plane spanned by T and B is spanned by $\bar{B}$ and $\bar{T}$.

(iv) By (i), $\bar{B} = \varepsilon \cdot_* \alpha^*$, and (2) gives $* \int_{s_0}^{s} \bar{B}\,(u) \cdot_* d_* u = \varepsilon \cdot_* \big(\alpha(s) -_* \alpha(s_0)\big)$. □

In the logarithmic picture, $\bar{\alpha}$ is obtained by integrating the binormal of $\hat{\alpha}$ with respect to arc length. Proposition 4.1(ii) explains in a conceptual way why the Frenet-constant fields of the two curves coincide: both frames rotate with the same angular velocity. The sign ε in (i) cannot be dispensed with. In [11, Section 3.2] the Frenet apparatus of the Euclidean binormal dual is written, without loss of generality, as $\bar{T} = B, \quad \bar{N} = -N, \quad \bar{B} = T, \quad \bar{\kappa} = \tau$; these formulas are correct when $\tau > 0$, but for $\tau < 0$ the last one would give a negative curvature, and the logarithmic picture of (i) with $\varepsilon = -_* 1_*$ has to be used instead. Both cases occur in Section 5, where the torsion of γ changes sign at $v = 1$.

4.2 The duality theorem and its consequences

Theorem 4.2. *Let $\bar{\alpha}$ be the binormal dual of a nonplanar curve α of unit multiplicative speed, and let $V \neq 0_*$ be constant. Then α is a multiplicative normal (respectively osculating, rectifying) helix with axis V if and only if $\bar{\alpha}$ is a multiplicative osculating (respectively normal, rectifying) helix with the same axis V and the same $\mathcal{F}$-constant field W. Moreover, if $W = \cos_*\theta \cdot_* N +_* \sin_*\theta \cdot_* B$, then $W = \cos_*\bar{\theta} \cdot_* \bar{T} +_* \sin_*\bar{\theta} \cdot_* \bar{N}$ with*

$$\bar{\theta} = \theta -_* (\pi/2)_* \ \text{ if } \varepsilon = 1_*, \qquad \bar{\theta} = (\pi/2)_* -_* \theta \ \text{ if } \varepsilon = -_* 1_*. \tag{14}$$

Proof. By Proposition 4.1(ii), a field W is $\mathcal{F}$-constant along α if and only if it is $\mathcal{F}$-constant along $\bar{\alpha}$. The helix condition only involves W, V and the angle between them, while the type of the helix is determined by the plane in which W lies; by Proposition 4.1(iii), passing from α to $\bar{\alpha}$ exchanges normal and osculating planes and preserves rectifying planes. For the last assertion, Proposition 4.1(i) gives $B = \bar{T}$ and $N = -_*\varepsilon \cdot_* \bar{N}$, so that $W = \sin_*\theta \cdot_* \bar{T} -_* \varepsilon \cdot_* \cos_*\theta \cdot_* \bar{N}$. For $\varepsilon = 1_*$ this equals $\cos_*(\theta -_* (\pi/2)_*) \cdot_* \bar{T} +_* \sin_*(\theta -_* (\pi/2)_*) \cdot_* \bar{N}$, and for $\varepsilon = -_* 1_*$ it equals $\cos_*((\pi/2)_* -_* \theta) \cdot_* \bar{T} +_* \sin_*((\pi/2)_* -_* \theta) \cdot_* \bar{N}$. □

Theorem 4.2 makes precise the duality between multiplicative normal and osculating helices. As a first application, we deduce the natural equation of multiplicative normal helices, which are studied in [16], from Theorem 3.2.

Corollary 4.3 (cf. [11, Theorem 2.1])**.** *For a nonplanar curve α of unit multiplicative speed and a constant $\theta \in \big(e^{-\pi/2}, e^{\pi/2}\big)$, the normal field $\cos_*\theta \cdot_* N +_* \sin_*\theta \cdot_* B$ is multiplicatively perpendicular to a nonzero constant vector precisely when*

$$\Big(\kappa /_* \big(\cos_*^{2_*}\theta \cdot_* \kappa^{2_*} +_* \tau^{2_*}\big)\Big) \cdot_* (\tau /_* \kappa)^* = -_* \tan_*\theta, \tag{15}$$

that is, $\frac{\ln\kappa}{\cos^2(\ln\theta)\ln^2\kappa+\ln^2\tau}\, s\, \frac{d}{ds}\left(\frac{\ln\tau}{\ln\kappa}\right) = -\tan(\ln\theta)$ *in ordinary arithmetic.*

Proof. By Theorem 4.2, α has this property if and only if $\bar{\alpha}$ is a multiplicative osculating helix with field $W = \cos_*\bar{\theta} \cdot_* \bar{T} +_* \sin_*\bar{\theta} \cdot_* \bar{N}$ orthogonal to the axis, $\bar{\theta}$ being given by (14). Replacing W by $-_*W$ if necessary, which replaces $\bar{\theta}$ by $\bar{\theta} +_* \pi_*$ (angles being taken modulo a full turn) and does not change $\sin_*^{2_*}\bar{\theta}$ and $\cot_*\bar{\theta}$, we may assume that $e^{-\pi/2} \le \bar{\theta} \le e^{\pi/2}$; moreover $\bar{\theta} \neq 0_*$ because $\theta \neq e^{\pm\pi/2}$. Since $\bar{\alpha}$ is nonplanar, Theorem 3.2 applies, and the condition becomes

$$\left(\bar{\tau} /_* \left(\bar{\kappa}^{2_*} +_* \sin_*^{2_*}\bar{\theta} \cdot_* \bar{\tau}^{2_*}\right)\right) \cdot_* (\bar{\kappa} /_* \bar{\tau})^* = \cot_*\bar{\theta}.$$

By Proposition 4.1(i), $\bar{\tau} = \kappa$, $\bar{\kappa}^{2_*} = \tau^{2_*}$ and $(\bar{\kappa} /_* \bar{\tau})^* = \varepsilon \cdot_* (\tau /_* \kappa)^*$, while (14) gives $\sin_*^{2_*}\bar{\theta} = \cos_*^{2_*}\theta$ and $\cot_*\bar{\theta} = -_*\varepsilon \cdot_* \tan_*\theta$ in both cases. Multiplying by ε and using $\varepsilon^{2_*} = 1_*$ we obtain (15). □

The second consequence is a geometric description of osculating helices. By a multiplicative cylinder in the direction of a unit vector V we mean a surface $X(t,z) = \beta(t) +_* z \cdot_* V$, where β is a curve in a multiplicative plane orthogonal to V; its logarithmic picture is a Euclidean cylinder.

Corollary 4.4. *Let α be nonplanar with unit multiplicative speed, and let V be a constant unit vector. The curve α is a multiplicative osculating helix with axis V and osculating field perpendicular to V exactly when its binormal dual $\bar{\alpha}$ is contained in a multiplicative cylinder generated by lines in the direction of V and, along $\bar{\alpha}$, the angle between $\bar{N} = -_*\varepsilon \cdot_* N$ (hence also N) and the surface normal is constant.*

Proof. By Theorem 4.2, α is such a helix if and only if $\bar{\alpha}$ is a normal helix, in the multiplicative sense, whose $\mathcal{F}$-constant field is perpendicular to its axis V. In the Euclidean setting, [11, Theorem 2.2] identifies these normal helices with the curves of cylinders generated by lines parallel to the axis along which the principal normal meets the surface normal at a fixed angle. By the transfer principle of Remark 2.3, the same characterization holds for multiplicative normal helices and multiplicative cylinders. Applying it to $\bar{\alpha}$ and using $\bar{N} = -_*\varepsilon \cdot_* N$ we obtain the statement. □

Put differently, up to a multiplicative isometry, every osculating helix of the kind considered here arises from the binormal construction (13) applied to a curve on a multiplicative cylinder whose principal normal and surface normal enclose a fixed angle, and conversely.

5 An explicit osculating helix

We now use the duality to write down an explicit multiplicative osculating helix. Fix $\vartheta \in (0, \pi/2)$ and put $\theta = e^{\vartheta}$ and $k = \tan\vartheta = \ln(\tan_*\theta)$.

5.1 A normal helix on a multiplicative circular cylinder

Consider the curve

$$\gamma(v) = \Big(\cos_* v,\ \sin_* v,\ -_*\big(\cot_*\theta \cdot_* \cosh_*(\tan_*\theta \cdot_* v)\big)\Big), \quad v \in \mathbb{R}_*, \quad (16)$$

that is, $\gamma(v) = \big(e^{\cos(\ln v)},\ e^{\sin(\ln v)},\ e^{-\cot\vartheta\cosh(k\ln v)}\big)$. It is contained in the multiplicative circular cylinder $\{x \in \mathbb{E}^3_*: \ln^2 x_1 + \ln^2 x_2 = 1\}$, generated by lines in the direction of e_3. Its logarithmic picture is $\hat{\gamma}(u) = (\cos u, \sin u, -\cosh(ku)/k)$, $u = \ln v$, whose speed is $\cosh(ku)$. Consequently, the multiplicative arc length of γ measured from $v = 1$ is

$$\sigma(v) = \cot_*\theta \cdot_* \sinh_*(\tan_*\theta \cdot_* v), \quad \text{that is,} \quad \ln\sigma = \frac{\sinh(k\ln v)}{k}. \quad (17)$$

A computation with the Frenet formulas of $\hat{\gamma}$ gives, with $u = \ln v$,

$$\begin{aligned}
\hat{T}_\gamma &= \frac{1}{\cosh(ku)}\big(-\sin u,\ \cos u,\ -\sinh(ku)\big),\\
\hat{N}_\gamma &= \Big(\sin\vartheta\tanh(ku)\sin u - \cos\vartheta\cos u,\\
&\qquad -\sin\vartheta\tanh(ku)\cos u - \cos\vartheta\sin u,\ -\frac{\sin\vartheta}{\cosh(ku)}\Big),\\
\hat{B}_\gamma &= \frac{\cos\vartheta}{\cosh(ku)}\big(-k\cosh(ku)\cos u - \sinh(ku)\sin u,\\
&\qquad -k\cosh(ku)\sin u + \sinh(ku)\cos u,\ 1\big),
\end{aligned}$$

and

$$\begin{aligned}
\kappa_\gamma &= \sec_*\theta\ /_*\ \cosh_*^{2_*}(\tan_*\theta \cdot_* v),\\
\tau_\gamma &= -_*\sinh_*(\tan_*\theta \cdot_* v)\ /_*\ \cosh_*^{2_*}(\tan_*\theta \cdot_* v),
\end{aligned} \quad (18)$$

that is, $\ln\kappa_\gamma = \sec\vartheta/\cosh^2(ku)$ and $\ln\tau_\gamma = -\sinh(ku)/\cosh^2(ku)$. In particular, γ is nonplanar on each of the intervals $0 < v < 1$ and $v > 1$, with $\tau_\gamma > 0_*$ on the first one and $\tau_\gamma < 0_*$ on the second. Moreover,

$$\cos\vartheta\ \hat{N}_\gamma + \sin\vartheta\ \hat{B}_\gamma = (-\cos u,\ -\sin u,\ 0),$$

so the $\mathcal{F}$-constant normal field $W = \cos_*\theta \cdot_* N_\gamma +_* \sin_*\theta \cdot_* B_\gamma = \big(e^{-\cos(\ln v)}, e^{-\sin(\ln v)}, 1\big)$ is multiplicatively orthogonal to e_3, and γ is a multiplicative normal helix with axis e_3. This curve is the multiplicative counterpart of the normal helices on a circular cylinder in [11, Section 2.3]; note that the third coordinate of (16) carries a minus sign. With the plus sign used in

[11], the torsion of the logarithmic picture changes sign and the normal field orthogonal to the axis is obtained for the opposite angle, so the minus sign is the one compatible with (18) and with the normal field above. Expressed through σ, formulas (18) become

$$\ln\kappa_\gamma = \frac{\cos\vartheta}{\cos^2\vartheta + \sin^2\vartheta\,\ln^2\sigma}, \qquad \ln\tau_\gamma = \frac{-\sin\vartheta\cos\vartheta\,\ln\sigma}{\cos^2\vartheta + \sin^2\vartheta\,\ln^2\sigma},$$

so that $\tau_\gamma /_* \kappa_\gamma = \text{-}_*\sin_*\theta \cdot_* \sigma$. By [1, Proposition 4.2], γ is therefore congruent to a multiplicative rectifying curve, the multiplicative analogue of the curves studied by Chen [3]. In fact, a direct computation shows that the multiplicative principal normal component of its position vector vanishes identically, so the curve itself is a multiplicative rectifying curve.

5.2 The dual osculating helix

By (2), a multiplicative primitive of B_γ with respect to σ is the curve ψ whose logarithmic picture is

$$\begin{gathered}\hat{\psi}(u) = \cos\vartheta\,(p(u)\cos u - q(u)\sin u,\ p(u)\sin u + q(u)\cos u,\ u),\\ p(u) = \cos 2\vartheta\sinh(ku), \qquad q(u) = \sin 2\vartheta\cosh(ku),\end{gathered} \tag{19}$$

as can be checked by verifying that $d\hat{\psi}/du = \cosh(ku)\,\hat{B}_\gamma$. Geometrically, the projection of $\hat{\psi}(u)$ onto the (x_1, x_2)-plane is the vector $\cos\vartheta\,\big(p(u), q(u)\big)$ rotated through the angle u, while its height is $u\cos\vartheta$. In the actual coordinates of $\mathbb{E}^3_*$, $\psi(v) = \big(e^{\hat{\psi}_1(\ln v)},\ e^{\hat{\psi}_2(\ln v)},\ v^{\cos\vartheta}\big)$, and its third coordinate is $\cos_*\theta \cdot_* v$.

By Theorem 4.2, ψ is a multiplicative osculating helix with axis e_3 on each of the intervals $0 < v < 1$ and $v > 1$, its osculating $\mathcal{F}$-constant field being W. Proposition 4.1(i) and (18) give its Frenet apparatus. For $v > 1$, where $\varepsilon = \text{-}_* 1_*$,

$$\begin{gathered}T_\psi = B_\gamma, \qquad N_\psi = N_\gamma, \qquad B_\psi = \text{-}_* T_\gamma,\\ \kappa_\psi = \sinh_*(\tan_*\theta \cdot_* v) /_* \cosh_*^{2_*}(\tan_*\theta \cdot_* v),\\ \tau_\psi = \sec_*\theta /_* \cosh_*^{2_*}(\tan_*\theta \cdot_* v),\end{gathered} \tag{20}$$

and $W = \cos_*\bar{\theta} \cdot_* T_\psi +_* \sin_*\bar{\theta} \cdot_* N_\psi$ with $\bar{\theta} = (\pi/2)_* \text{-}_* \theta$. For $0 < v < 1$ one has instead $N_\psi = \text{-}_* N_\gamma$, $B_\psi = T_\gamma$, $\kappa_\psi = \big|\tau_\gamma\big|_*$ and $\bar{\theta} = \theta \text{-}_* (\pi/2)_*$. At $v = 1$ the curvature of ψ equals 0_*; this point separates the two osculating helices.

As a check of Theorem 3.2, consider $v > 1$ and use the logarithmic picture. There $\ln\kappa_\psi/\ln\tau_\psi = \cos\vartheta\sinh(ku) = \sin\vartheta\,\ln\sigma$ by (17), so that

$$\kappa_\psi /_* \tau_\psi = \sin_*\theta \cdot_* \sigma, \qquad \left(\kappa_\psi /_* \tau_\psi\right)^* = \sin_*\theta. \tag{21}$$

In (21) the multiplicative derivative is taken with respect to the multiplicative arc length. Moreover, $\ln^2\kappa_\psi + \cos^2\vartheta \ln^2\tau_\psi = (\sinh^2(ku) + 1)/\cosh^4(ku) = 1/\cosh^2(ku)$, and therefore

$$\frac{\ln\tau_\psi}{\ln^2\kappa_\psi + \cos^2\vartheta \ln^2\tau_\psi} \sin\vartheta = \sec\vartheta \sin\vartheta = \tan\vartheta = \cot\left(\frac{\pi}{2} - \vartheta\right),$$

which is (12) with $\ln\bar{\theta} = \pi/2 - \vartheta$, because $\sin^2(\pi/2 - \vartheta) = \cos^2\vartheta$. The interval $0 < v < 1$ is handled in the same way. In terms of its own multiplicative arc length σ, the curvature and torsion of ψ on $v > 1$ are

$$\ln\kappa_\psi = \frac{\sin\vartheta\cos\vartheta \ln\sigma}{\cos^2\vartheta + \sin^2\vartheta \ln^2\sigma}, \qquad \ln\tau_\psi = \frac{\cos\vartheta}{\cos^2\vartheta + \sin^2\vartheta \ln^2\sigma}. \tag{22}$$

Remark 5.1. By (21), the ratio $\kappa_\psi /_* \tau_\psi$ is a multiplicatively linear function of the multiplicative arc length. This is the dual counterpart of the relation $\tau_\gamma /_* \kappa_\gamma = -_*\sin_*\theta \cdot_* \sigma$ satisfied by γ, curvature and torsion being exchanged by Proposition 4.1(i). By Remark 3.3(4), the ratio $\kappa /_* \tau$ is strictly monotone along every osculating helix whose osculating field is orthogonal to the axis and that is not a cylindrical helix; for ψ it is even linear.

A multiplicative homothety $x \mapsto c \cdot_* x$, $c > 0_*$, multiplies the logarithmic picture of a curve by $\ln c$. It divides $\ln\kappa$ and $\ln\tau$ by $\ln c$ and multiplies the logarithm of the arc length by $\ln c$, so it leaves (12) unchanged and transforms osculating helices into osculating helices with the same angle and axis. Figure 1 shows the curve $e^{0.3} \cdot_* \psi$ for $\vartheta = \pi/12$. The logarithmic picture of ψ lies on the surface of revolution about the third coordinate axis given in cylindrical coordinates by $r^2 = \cos^2\vartheta\,(\cosh^2(ku) - \cos^2 2\vartheta)$, $x_3 = u\cos\vartheta$, which is also drawn in the figure after the same homothety.

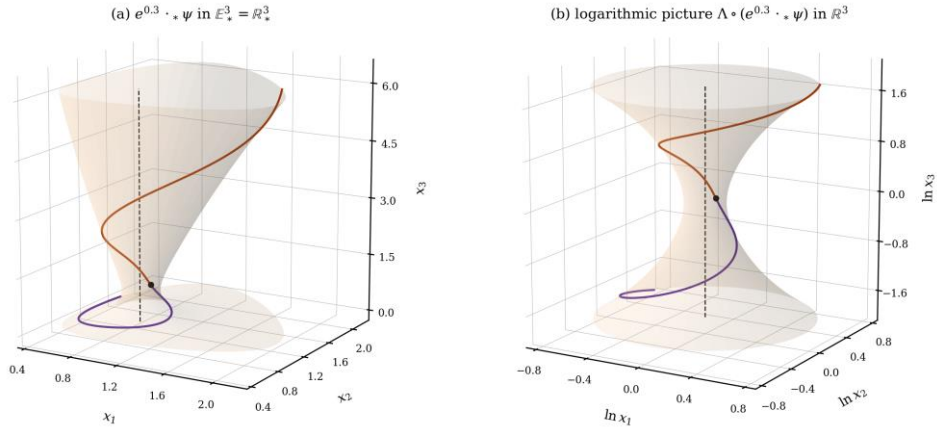


Figure 1. The multiplicative osculating helix $e^{0.3} \cdot_* \psi$, where ψ is given by (19) with $\vartheta = \pi/12$, for $e^{-2\pi} \le v \le e^{2\pi}$. The branches $0 < v < 1$ and $v > 1$ are drawn in different colors, and the dot marks the point $v = 1$, where the curvature equals 0_*. The dashed line is the axis, i.e. the multiplicative line $\{(1,1,z): z \in \mathbb{R}_*\}$ spanned by e_3; the shaded surface is the multiplicative surface of revolution about it on which the curve lies. (a) Actual coordinates in $\mathbb{E}^3_* = \mathbb{R}^3_*$. (b) The logarithmic picture.

6 Concluding remarks

We have studied multiplicative osculating helices in $\mathbb{E}^3_*$ with osculating Frenet-constant field perpendicular to the axis. The axis of such a curve is forced to be the direction of the field $Y_\theta = \kappa \cdot_* B \,-_* \sin_* \theta \cdot_* \tau \cdot_* U_\theta$ (Lemma 3.1), and requiring this direction to be constant leads to the natural equation (10) (Theorem 3.2). The binormal construction preserves the multiplicative Darboux vector (Proposition 4.1), which explains the duality between osculating and normal helices (Theorem 4.2); through it we recovered the natural equation of normal helices (Corollary 4.3) and described osculating helices by means of multiplicative cylinders (Corollary 4.4). The explicit example of Section 5 is typical: by [11, Proposition 2.3] and the transfer principle, along every multiplicative normal helix on a multiplicative circular cylinder generated by lines parallel to its axis, the ratio of torsion to curvature is a multiplicatively linear function of the arc length, so its binormal dual has a multiplicatively linear ratio of curvature to torsion on each interval where the torsion does not vanish.

As explained in Remark 2.3, the logarithm transports most of these statements to Euclidean results of [11], and in this sense the chapter provides their multiplicative counterparts; the invariance of the Darboux vector and the closed-form osculating helix of Section 5.2, however, do not appear in [11]. The

multiplicative point of view is well suited to curves of exponential type, such as those of Figure 1(a), whose invariants become elementary. When the osculating field meets the axis at a constant angle other than $e^{\pi/2}$, Theorem 4.2 still applies, because Definition 2.2 does not require the angle to be right; thus these helices are, up to multiplicative isometries, the binormal duals of the normal helices with the same field and axis, and their natural equation can be obtained by requiring the spherical curve described by the field in the logarithmic picture to have constant geodesic curvature. It remains to be seen whether duality constructions of this type exist for other moving frames or in other non-Newtonian geometries.

References

[1] Aydın ME, Has A, Yılmaz B. A non-Newtonian approach in differential geometry of curves: multiplicative rectifying curves. Bulletin of the Korean Mathematical Society 2024; 61 (3): 849–866. doi: 10.4134/BKMS.b230491

[2] Bashirov AE, Kurpınar EM, Özyapıcı A. Multiplicative calculus and its applications. Journal of Mathematical Analysis and Applications 2008; 337 (1): 36–48. doi: 10.1016/j.jmaa.2007.03.081

[3] Chen BY. When does the position vector of a space curve always lie in its rectifying plane? American Mathematical Monthly 2003; 110 (2): 147–152. doi: 10.2307/3647775

[4] Georgiev SG. Multiplicative Differential Geometry. New York, USA: Chapman and Hall/CRC, 2022.

[5] Georgiev SG, Zennir K. Multiplicative Differential Calculus. New York, USA: Chapman and Hall/CRC, 2022.

[6] Grossman M, Katz R. Non-Newtonian Calculus. Pigeon Cove, Massachusetts, USA: Lee Press, 1972.

[7] Has A, Yılmaz B. On non-Newtonian helices in multiplicative Euclidean space. Fundamentals of Contemporary Mathematical Sciences 2025; 6 (2): 196–217. doi: 10.54974/fcmathsci.1644427

[8] Izumiya S, Takeuchi N. New special curves and developable surfaces. Turkish Journal of Mathematics 2004; 28 (2): 153–163.

[9] Kreyszig E. Differential Geometry. Toronto, Canada: University of Toronto Press, 1959.

[10] Lucas P, Ortega-Yagües JA. Slant helices in the Euclidean 3-space revisited. Bulletin of the Belgian Mathematical Society – Simon Stevin 2016; 23: 133–150. doi: 10.36045/bbms/1457560859

[11] Lucas P, Ortega-Yagües JA. A generalization of the notion of helix. Turkish Journal of Mathematics 2023; 47 (4): 1158–1168. doi: 10.55730/1300-0098.3418

[12] Sağlam D. On dual slant helices in $\mathbb{D}^3$. Advances in Mathematics: Scientific Journal 2022; 11 (7): 577–589. doi: 10.37418/amsj.11.7.2

[13] Sağlam D, Bada D. Some characterizations of timelike helices with the F-constant vector field in Minkowski space $\mathbb{E}_1^3$. Advances in Mathematics: Scientific Journal 2024; 13 (3): 263–279. doi: 10.37418/amsj.13.3.1

[14] Sağlam D, Özkan S, Özdamar D. Slant helices in dual Lorentzian space $\mathbb{D}_1^3$. Natural Science and Discovery 2016; 2 (1): 3–10. doi: 10.20863/nsd.04519

[15] Sağlam D, Selvi U, Babadağ F, Atasoy A. Helices with F-constant vector fields in the Euclidean space $\mathbb{E}^4$. AIMS Mathematics 2026; 11 (2): 4299–4334. doi: 10.3934/math.2026173

[16] Selvi U, Sağlam D. Normal helices in the multiplicative Euclidean space $\mathbb{E}_*^3$. In: Theory, Research and Discussions in Natural Sciences and Mathematics. Duvar Yayınları, 2026, pp. 55–79.

[17] Stanley D. A multiplicative calculus. PRIMUS 1999; 9 (4): 310–326.

[18] Struik DJ. Lectures on Classical Differential Geometry. Courier Corporation, 2012.