A Network Study of Chemical Elements: From Binary Compounds to Chemical Trends

Wilmer Leal[a], **Guillermo Restrepo**[a,b], **Andrés Bernal**[c]

[a]*Laboratorio de Química Teórica, Universidad de Pamplona, Pamplona, Colombia*
[b]*Interdisciplinary Research Institute, Universidad de Pamplona, Bogotá, Colombia*
e-mail: `wilmer.leal@unipamplona.edu.co`; `grestrepo@unipamplona.edu.co`

[c]*Grupo de Química Teórica, Universidad Nacional de Colombia, Bogotá, Colombia*
e-mail: `afbernale@unal.edu.co`

(Received September 1, 2011)

Abstract

It has been claimed that relational properties among chemical substances are at the core of chemistry. Here we show that chemical elements and a wealth of their trends can be found by the study of a relational property: the formation of binary compounds. We say that two chemical elements A and B are similar if they form binary compounds AC and BC, C being another chemical element. To allow the richness of chemical combinations, we also included the different stoichiometrical ratios for binary compounds. Hence, the more combinations with different chemical elements, and with similar stoichiometry, the more similar two chemical elements are. We studied 4,700 binary compounds by using network theory and point set topology, we obtained well-known chemical families of elements, such as: alkali metals, alkaline earth metals, halogens, lanthanides, actinides, some transition metal groups and chemical patterns like: singularity principle, knight's move, and secondary periodicity. The methodology applied here can be extended to the study of ternary, quaternary and other compounds, as well as other chemical sets where a relational property can be defined.

1 Mathematical approaches to the chemical elements

Chemical elements and their organization in the periodic table (PT) have been studied using different mathematical tools [1, 2] such as: information theory, order theory, quantum

similarity, group theory, cluster analysis and most recently using a mixture between cluster analysis and point set topology, called chemotopology. An example of the application of information theory to the study of the PT is the one by Bonchev et al. [3], who argued that the electronic and nuclear distribution of an atom is given by a principle of maximum information content. Moreover, Klein et al. [4, 5, 6], using order theory, described the chemical elements as a partially ordered set or "poset", based on the increase or decrease of properties as one moves in a row or column in the conventional PT. Kreinovich et al. [7] also applied order theory, this time to the Madelung rule. Using quantum similarity Carbó-Dorca et al. [8, 9] studied the resemblance of electronic density functions of a set of 20 atoms of different elements. Continuous and discrete groups have been used in chemistry, however, based on the idea that only continuous groups and more specifically Lie groups are relevant to study the periodic system, Novaro [10], Kibler [11] and Ostrovsky [12] developed several studies using those kinds of groups to try to reproduce some periodic trends of the chemical elements. However, they exclusively used physical information such as electronic configurations, ionization energies, etc., disregarding valuable chemical information given by valence, chemical reactions, etc.

Other studies of chemical elements based upon their resemblances using more diverse sets of properties are those of Zhou et al. [13] , Sneath [14] and Restrepo et al. [15, 16, 17, 18]. Zhou et al. studied 50 chemical elements ($Z = 1 - 50$) characterized by seven properties; after classification using cluster analysis, they found families such as halogens, transition metals, alkali metals and some others. Sneath [14] also used cluster analysis and dealt with 69 chemical elements ($Z = 1 - 83$, omitting $Z = 58 - 71$) using a set of 54 properties; among the clusters found there are fragments of chemical families (groups on the periodic table) such as noble gases, chalcogens, halogens, alkali metals, alkaline earth metals, transition metal groups, among others. Most recently, Restrepo et al. studied a set of 72 chemical elements ($Z = 1 - 86$, omitting $Z = 58 - 71$ initially using a set of 31 properties [15], later on enlarged to include 128 properties [19]. In these studies they applied cluster analysis in combination with topology using a method nowadays called chemotopology (details can be found in [20, 21]). Some of the results of those chemotopological studies are the appearance of chemical groups as open sets of a topology. Additionally, evidence of the singularity principle, diagonal relationship and the inert pair effect [19] were found. Perhaps the most important result by Restrepo et al.

is that the topological boundary of nonmetals is made from semimetallic elements. Some of these works as well as others addressing the mathematical study of chemical elements can be found in the book edited by Rouvray and King [22].

We think that a deeper understanding of periodicity in mathematical terms demands greater emphasis on the choice of the property set used to characterize chemical elements. In the present contribution we advance a topological study of the elements using a property set that, we believe, lies closer to the core of the chemical approach to the study of matter.

2   An approach with chemical taste

Chemical properties may be characterized as those that arise from affinity and precursor-product relationships established between substances by means of chemical reactions [23]. Chemical knowledge thus involves a vast network of chemical reactions, whose structure determines chemical similarities, classes, and properties [23]. Yet the current vision, interpretation, and learning of chemical periodicity are addressed mainly through physical ideas and concepts, i.e. concepts that arise without explicitly considering the structure of a chemical reaction network. The fact that a significant part of the periodic trends of the chemical elements discovered by Mendeleev were initially found taking into account similarity in chemical properties is often overlooked.

The works of Restrepo et al. [1, 15, 16, 17, 18, 19, 20, 21, 24] and Bernal [25] have shown that the more chemical properties are included in the characterization of chemical elements, the better agreement of clusters of chemical elements with chemical groups depicted in the periodic table is obtained. This should not be a matter of surprise since the inclusion of chemical properties should yield clusters with more chemical taste, i.e. groups with chemical meaning. These findings motivate us to move further into the realm of strictly chemical properties in the sense proposed by Schummer [23], to find out how far can they advance the enterprise of designing a classification system of the elements that recovers the trends found by Mendeleev.

In the current work we analyzed a small fraction of the network of chemical knowledge, namely a network of chemical elements based upon their presence in binary compounds (of the form AB, with $A \neq B$). The mathematical tool used is network theory, which has had several applications in social sciences [26], computer science [27], biochemistry [28, 29], among many other areas [30, 31]. Estrada et al. [32] have published an attempt to study

chemical elements from a network-theoretical viewpoint, that looked into some topological properties of large-scale organization networks of minerals and chemical elements. The authors found a good correlation between the frequency with which two elements appear together and their abundance in some minerals. However, Estrada et al.'s interest in applying network theory to the set of chemical elements was neither looking for similarities between them nor making a classification of them based upon their trends in the formation of compounds, which is our aim in the current paper.

The present approach combines methods from network analysis, particularly role assignment techniques, with the chemotopological method designed by Restrepo et al. [1, 15, 18, 33]. We first construct a network of chemical elements that accounts for the affinity relations determined by the binary compounds they may form. Then, we define a measure of structural similarity on the vertex set of this network, i.e. on the chemical elements themselves. This measure answers to the similarities in the neighborhoods of each pair of vertices, which in turn relates to similarity in the composition of the binary compounds that different elements form. Last, we apply the chemotopological methodology using this measure of similarity, and analyze the topological properties of some subsets of chemical elements.

3 Methodological foundations

3.1 Network analysis

Some important definitions of network analysis for the ensuing discussion are the following [34]:

Definition 1 *A **graph** G consists of a set V of vertices and a set E of pairs of vertices called edges. Another customary notation for V and E is $V(G)$ and $E(G)$ with cardinalities n and m, respectively. A graph H is a **subgraph** of a graph G if $V(H) \subset V(G)$ and $E(H) \subset E(G)$.*

Definition 2 *Two vertices $u, v \in V$ are called **adjacent** or "neighbors" iff $(u, v) \in E$. The set formed by the neighbors of u is called the **neighborhood** of u and is denoted by N_u. The **degree** of a vertex is the cardinality of its neighborhood. In this work we used (u, v) to represent either an arc or and edge.*

Thus, a network can be defined as follows [35]:

Definition 3 *A **network** consists of a graph $G = (V, E)$ with additional information either on the vertices or on the edges.*

If we look at how vertices are linked to the rest of the network, it is possible to find trends in these connectivities. In trying to formalize this, network theoreticians have come up with the idea of equivalence among vertices. Hence, elements related in a similar fashion, i.e. sharing the same neighbors, are considered as equivalent [36].

Definition 4 *Let (V, E) be a graph and $u, v \in V$, we say that u and v are **structurally equivalent**, and note it $u \sim v$, iff $N_u = N_v$.*

It has been proven that structural equivalence is an equivalence relation [37]. As is known, an equivalence relation on a set induces a canonical partition over it. In the case of structural equivalence, classes of this partition are known as **positions**.

Definition 5 *Let (V, E) be a graph and $u \in V$. The set $\bar{u}$ of all vertices structurally equivalent to u is known as a **position** in (V, E).*

Each of the positions, in sociology, represents what is known as a social role. Strictly speaking, a role should be represented by a position in the network, but in practice it is not common to find sets of vertices which are structurally equivalent. The reason is that in practice one hardly finds a subset of vertices connected with exactly the same neighbors. Thus, it is necessary to establish a measure quantifying how similar are two vertices in structural terms. We discuss this later.

3.2   Chemotopology

Broadly speaking, chemotopology consists on the topological analysis of the results of a classification. Such a classification may be performed using different methods and techniques. We have used hierarchical cluster analysis in several studies; in the following we summarize how to go from cluster analysis results to topology. Further information can be found in reference [21].

Hierarchical cluster analysis is a classification technique based on a similarity measure [38] between pairs of elements of a set X of interest and a grouping methodology [39, 40]. The final outcome of a cluster analysis study is a dendrogram, which can be defined as follows:

Definition 6 *A **dendrogram** (also known as a **full binary tree**) on a set X is a graph that represents the resemblances among the elements of the set; it is acyclic, connected, and has the following classes of vertices:*

1. *vertices of degree 1 called **leaves**, which corresponds to the elements of X,*

2. *a single vertex of degree 2 called the **root node**, and*

3. *vertices of degree 3, called **nodes**.*

Restrepo et al. in 2004 showed how to build topologies on a set X from a dendrogram on the same set [15]. The methodology is based on the selection of certain kinds of "branches", particular subgraphs of the dendrogram, to build a topological basis, therefore a topology. Some basic definitions of the method are the following [33]:

Definition 7 *Let D be a dendrogram on X. A subgraph T of D is called a **subtree** if:*

1. *T does not contain the root node, and*

2. *there is a node p of D with a degree different from 1 such that T corresponds to one of the connected graphs obtained by removing p from D.*

Definition 8 *A **n-subtree** is a subtree with n or fewer leaves.*

Definition 9 *A **maximal n-subtree** is a n-subtree for which it is not possible to find another n-subtree containing it.*

A full collection of maximal n-subtrees gives a viable partition in classes of equivalence of the leave set of a dendrogram. In other words, it determines a classification of the elements under scrutiny that is consistent with the input data. The central idea of the chemotopological method is to use such a classification to induce a topology, so that topological invariants related to neighborhood relations in the resulting topological space may be used to describe similarity among the elements being classified.

Definition 10 *Let X be any finite nonempty set and τ a collection of subsets of X, such that:*
1. *$X \in \tau$,*

2. $\varnothing \in \tau$,

3. *For any* $O_i, O_j \in \tau$, $O_i \cap O_j \in \tau$ *and*

4. *For any* $O_i, O_j \in \tau$, $O_i \cup O_j \in \tau$.

Then τ *is called a topology over* X, (X, τ) *is called a* **topological space** *and the elements of* τ *are called* **open sets**.

Definition 11 *Let* $\mathfrak{B}$ *be a collection of subsets of* X, *such that:*

1. $X = \bigcup_{B \in \mathfrak{B}} B$ *and*

2. *If* $B_1, B_2 \in \mathfrak{B}$, *then* $B_1 \cap B_2$ *is the union of elements of* $\mathfrak{B}$.

Then $\mathfrak{B}$ *is a* **topological basis** *for* τ, *where* $\tau = \left\{ \bigcup_{B \in \mathfrak{F}} B \mid \mathfrak{F} \subseteq \mathfrak{B} \right\}$.

Theorem 1 *Let* D *be a dendrogram on* X. *Let* $\mathfrak{B}_n = \{B \subseteq X \mid B$ *is formed by the leaves of a maximal n-subtree of* $D\}$. *Then* $\mathfrak{B}_n$ *is a basis for a topology on* X.

According to theorem 1, (proven in [42]), by selecting different values of n for the maximal n-subtrees, we obtain topological basis, which may be different. Therefore, different topologies may be obtained on the set X. For example, for $n = 1$ the corresponding basis is made from single-element sets.

Recall that our goal is to unveil similarity and dissimilarity relations between chemical elements. The more such relations we uncover, the better the topological model turns out to be. Now, notice e.g. that if we choose $n = 1$ we obtain many clusters minimally populated, indicating that all elements are dissimilar, and giving no similarity relations. On the other hand, for $n = |X|$ we obtain only one cluster highly populated, showing that all elements have similar properties among them, but telling nothing about their dissimilarities. Neither of this options looks very promising. Instead, we want a value of n optimizing both the number of similarities and dissimilarities established, i.e. we need to build a topological basis such that both its cardinality and the cardinality of its elements are as high as possible. In consequence, the best value for n can be found by maximizing the *selection number* S_n [24, 43]:

$$S_n = C_n \prod_i |C_{ni}| \tag{1}$$

where C_n is the number of clusters formed by a cut of the dendrogram in *n-maximal subtrees*, and C_{ni} is the *i*-th cluster in that cut.

As we can see, the value of S_n grows with both, an increase of the number of clusters and an increase of its population, so by maximizing S_n we find the cut that gives the best balance between cluster size and diversity.

Based on the n found, the topological basis $\mathfrak{B}_n$ is built.

Through this methodology, similarity relations among the elements of a set X, that were initially represented by a metric, are translated into the classificatory structure of a dendrogram with leaves X, and then turned into neighborhood relations in the topological space (X, τ) induced by $\mathfrak{B}_n$. In this way topology becomes a tool for speaking about similarity; so that topological properties of subsets of X, that can be easily calculated from $\mathfrak{B}_n$, now give an interesting description of similarity between its elements: some topological properties of subsets of X.

Definition 12 *Let $A \subset X$ and $x \in X$; x is a **closure point** of A iff for all $O \in \tau$, such that $x \in O$, then we have $O \cap A \neq \varnothing$.*

Definition 13 *Let $A \subset X$; the **closure** of A, denoted by $\overline{A}$, is defined as: $\overline{A} = \{x \in X \mid x$ is a closure point of $A\}$.*

The closure of a set A contains those elements of $X \supset A$ which are similar to the elements of A. In a recent publication Restrepo and Mesa [21] take the following case to exemplify the meaning of this property: Suppose a set X collecting substances with activity against either cutaneous or visceral leishmaniasis, now suppose we are interested in those with activity against the cutaneous disease, they constitute the set A. Its closure collects those chemicals which are similar to the ones active against cutaneous leishmaniasis.

Definition 14 *Let $A \subset X$ and $x \in X$; x is an **accumulation point** of A iff for all $O \in \tau$, such that $x \in O$, then we have $(O - \{x\}) \cap A \neq \varnothing$.*

Definition 15 *Let $A \subset X$; the **derived set** of A, denoted by A', is defined as: $A' = \{x \in X \mid x$ is an accumulation point of $A\}$.*

This set contains those elements x of X whose similarity to $A \subset X$ is caused by their similarity to elements of A different to x themselves. In other words, the derived set of A is

made of those elements x of X which are similar to at least one element of A different to x. Therefore, if an element x belongs to a subtree containing just x, and even if that subtree is contained in A, x is not part of the derived set since x is not an accumulation point i.e. x does not accumulate similarity to elements of A around it. Another interpretation of the derived set is that it gathers those elements of X which are similar to elements of $A \subset X$ not because of their own similarity but because of the similarity of their neighbors. The derived set may be regarded as an extreme version of the closure. In fact the derived set is a subset of the closure, gathering all the elements of the closure except those whose similarity is caused only by themselves and not by their similarity to at least one member of the set under study. Taking the example of antileishmaniasis substances, if one calculates the derived set of the active chemicals against cutaneous leishmaniasis, then one finally has those substances of the database which actually accumulate similarity upon the active chemicals against the cutaneous disease. With the derived set one eliminates chemicals which are in the class of active substances against cutaneous leishmaniasis but are not similar to any other chemical of the class i.e. do not accumulate similarity regarding other substances of the class [21].

Definition 16 *Let $A \subset X$; A is a **perfect set** iff $A' = A$.*

Definition 17 *Let $A \subset X$ and $x \in X$; x is a **boundary point** of A iff for all $O \in \tau$, such that $x \in O$, then we have $O \cap A \neq \emptyset$ and $O \cap (X - A) \neq \emptyset$.*

Definition 18 *Let $A \subset X$; the **boundary** of A, denoted by $b(A)$, is defined as: $b(A) = \{x \in X \mid x \text{ is a boundary point of } A\}$.*

We have in this set those elements of X whose neighbourhoods have elements of A and of the complement of A. This set gathers the elements of X whose properties are similar to the ones of the elements in A and also similar to the rest of the elements in X. Taking again the set of substances with activity against either cutaneous or visceral leishmaniasis, the boundary of the chemicals acting against cutaneous leishmaniasis gathers those active substances against both kinds of leishmaniasis [21].

Definition 19 *Let $A \subset X$ and $x \in X$; x is an **interior point** of A iff for all $O \in \tau$, such that $x \in O$, then we have $O \cap (X - A) = \emptyset$.*

Definition 20 *Let $A \subset X$; the **interior** of A, denoted by $\mathring{A}$, is defined as: $\mathring{A} = \{x \in X \mid x$ is interior point of $A\}$.*

This is the set of elements of X whose neighbourhoods are embedded in A. In other terms, it contains the elements of X whose properties are exclusively similar to the elements of A. It can be thought of as the core of A, containing the most representative elements of A. The meaning of this property is different to the one of closure because in the closure similar elements to A are contained, also those which share some similarity with some elements not belonging to A. Considering the example of antileishmaniasis substances, the interior of those with activity against the cutaneous disease are those chemicals which are only active against the cutaneous manisfestation with no activity at all against visceral leishmaniasis [21].

Definition 21 *Let $A \subset X$ and $x \in X$; x is an **exterior point** of A iff for all $O \in \tau$, such that $x \in O$, then we have $O \cap A = \varnothing$.*

Definition 22 *Let $A \subset X$; the **exterior** of A, denoted by $Ext(A)$, is defined as: $Ext(A) = \{x \in X \mid x$ is exterior point of $A\}$.*

It contains the elements of X whose neighbourhoods do not contain elements of A i.e. the elements of X which are completely different to the elements of A. The exterior of the chemicals acting against cutaneous leishmaniasis are those which hold activity only against visceral leishmaniasis. Note that those substances acting simultaneously against both kinds of leishmaniasis are not part of the exterior in question [21].

4 From networks to topologies

The set of chemical elements and compounds can be formally studied by defining a relational property and using network theory [44, 45]. In this sense, we know that in chemistry it is usual to classify substances according to their activity, i.e. how they interact with each other [46]. For instance, in organic chemistry we talk about typical reactions of alkanes, alcohols, carboxylic acids, esters, etc. [47]; i.e. we define families of compounds which are related via chemical reactions with compounds of the same given families, and are regarded as similar in that sense [23]. In the same way, here we consider a network whose vertices are chemical elements, related by their mutual presence in a compound [45]. For convenience, in this article we consider only the subset of binary compounds.

Definition 23 *Let X be a set of chemical elements and B a set of binary compounds. Let $E = \{(x, y)|x, y \in X \wedge \exists \ a \ compound \ x_a y_b \in B \ with \ a, b \in \mathbb{R}^+\}$. We call $G = (X, E)$ the **network of chemical elements** derived from binary compounds.*

Two chemical elements are adjacent in this network if they form a binary compound $x_a y_b$. The neighborhood of a chemical element consists of the set of elements with which it forms a binary compound. Two chemical elements would thus be equivalent if their corresponding vertices are structurally equivalent, that is, if they form binary compounds with the same set of elements. In this way, G formalizes our idea of relating elements based on the compounds they form. However, it misses the large diversity of binary compounds a couple of elements may form. To take one example, let us consider H and B, which form a large number of binary compounds called boranes; according to our formalization they would be represented in our network by a single edge (H, B). A single edge would also be the result of considering H and F i.e. (H, F), two chemical elements forming, as far as reported in the literature, only one binary compound, i.e. HF. Hence, the formalization here discussed misses the variety of binary compounds with different stoichiometries. As a consequence, in the previous example H would be seen as as common element of the neighborhoods of B and F, determining a similarity between these two elements regarding their behavior in chemical combination. And though this is not an entirely wrong take on the matter, it ignores the important chemical idea that valences of the elements present in a chemical combination are just as important as the presence of the combined elements in the compound.

It is thus important to modify the original model of definition 23 in order to include stoichiometrical coefficients of the compounds, aiming at the ultimate goal of re-defining the neighborhoods of chemical elements in such a way that, for each element v, its neighborhood includes information both regarding the elements with which v combines, *and regarding the proportion in which they combine*. We achieve this goal in two steps: first, we define a new network G_s of 'stoichiometrical chemical elements', whose vertices are identified with pairs (x, q), where x is a chemical element and q is the proportion with which it occurs in a binary compound. Edges are introduced joining stoichiometrical elements whenever the corresponding chemical elements occur in a binary compound in the corresponding proportions. Then, on a second step we define the 'chemical neighborhood' of an element v as the union of the graph-theoretical neighborhoods (i.e. as defined in definition 2) in G_s of all stoichiometrical chemical elements of the form (v, q). In this way, each stoichiometrical element (x, q) in the chemical neighborhood of a chemical element

v gives both an element x with which v combines, and the proportion q with which they combine.

Definition 24 *Let X be a set of chemical elements and B a set of binary compounds involving elements in X. We call $X_s = \{(x, \frac{a}{b}), (y, \frac{b}{a}) | x_a y_b \in B\}$ the set of **stoichiometrical chemical elements**. To simplify the notation, we define $x_{\frac{a}{b}} = (x, \frac{a}{b})$. We call $E_s = \{(x_{\frac{a}{b}}, y_{\frac{b}{a}}) | x_a y_b \in B\}$ the **set of edges of stoichiometrical chemical elements** and $G_s = (X_s, E_s)$ the **network of stoichiometrical chemical elements**.*

This enriched chemical network allows us to take stoichiometry into account. However, since we are interested in knowing similarities among chemical elements rather than among stoichiometrical chemical elements, we must now define the 'chemical' neighborhood of an element by reference to the graph-theoretical neighborhoods in G_s of its corresponding stoichimetrical elements.

Definition 25 *Let X be a set of chemical elements, B a set of binary compounds involving elements in X, and G_s the corresponding network of stoichiometrical elements. We define the **chemical neighborhood of** $x \in X$ to be the set $C_x = \cup_{a,b} N_{x_{\frac{a}{b}}}$.*

Let us summarize what we have done so far. Following the basic formalism of network analysis, the basic criterion of equivalence between chemical elements would be given by structural equivalence in the network of chemical elements G (definition 23), that is, two chemical elements $u, v \in X$ are equivalent if their neighborhoods N_u, N_v in G are identical. However, the highly relevant stoichiometrical factor is being ignored in this construction. In order to take stoichiometry into account we introduced the network of stoichiometrical elements G_s (definition 24). This network could be seen as a 'refinement' of G, that expands each chemical element in X into a collection of stoichiometrical chemical elements, each corresponding to a different proportion with which the chemical element may combine in a binary compound. Last, by defining the chemical neighborhoods of the chemical elements as the union of the neighborhoods in G_s of its associated stoichiometrical elements, we are ultimately performing the same refinement over the graph-theoretical neighborhoods in G of the chemical elements, thus including stoichiometrical data into the primary descriptors of the chemical behavior of the elements. Identity on this chemical neighborhoods, then, becomes our new criterion of equivalence for chemical elements:

$$u \simeq v \text{ iff } C_u = C_v.$$

However, as we noted above, structural equivalence is a very restrictive criterion. It is highly unlikely that two chemical elements would have exactly the same neighborhoods according to the previous definition. Since our aim is to look for similarities among chemical elements and to gain some helpful flexibility, we replace the above defined criterion of equivalence by a measure of similarity. Clearly, this measure of similarity should also be defined with respecto to the chemical neighborhoods of definition 25, and to be consistent maximum similarity must imply equivalence. This means that we must assess the degree of similarity of two chemical elements by contrasting their neighborhoods, in such a way that the less different neighbors two elements have, the more similar the elements are. In mathematical terms:

Definition 26 *Let X be a set of chemical elements, and $u, v \in X$. The **difference** $\delta(u, v)$ between $u, v \in X$ is equal to the cardinality of the symmetric difference between their chemical neighborhoods C_u, C_v respectively, denoted by $|C_u \ominus C_v|$, where $C_u \ominus C_v = (C_u - C_v) \cup (C_v - C_u)$.*

Hence, $|C_u \ominus C_v|$ counts the number of non-common stoichiometrical chemical elements of the neighborhoods of u and v. It is proven and widely known that the cardinal of the symmetric difference is a metric (see e.g. [48]). Thus, we may use the above defined difference as the metric required to apply the chemotopological method to the network of chemical elements: we run a non-hierarchical clustering algorithm on a set of chemical elements X using the metric function $\delta(u, v)$ to obtain a dendrogram; then we use the selection number to choose a cut of the dendrogram; and last we use that dendrogram cut as a basis for a topology on X. In this way, we build a topological representation of the network of chemical elements that allows us to analyze the similarities among the elements determined by a relational property, i.e. by their presence in binary compounds.

4.1 Data set and results

We gathered 4,700 binary compounds from references [49, 50, 51]. Highly unstable compounds found in the sources such as free radicals (e.g. CH, CH_2, and CH_3) were excluded from the data set. Taking into account that our methodology groups elements according to their structural equivalence in the network, then noble gases would be equivalent to some heavy metals for which binary compounds have not been reported in the literature (actually only He, Ne, Ar do not have reported compounds, Kr and Xe form some halides). That is why in this study we included only chemical elements for which one or more binary compounds have been reported. Hence, when we refer to a chemical family,

we consider it as formed only by the elements belonging to it and additionally forming at least one binary compound. The set X of 94 chemical elements for which at least one binary compound has been reported is shown in definition 27, also the distribution of binary compounds per element is depicted in Figure 1:

Definition 27 *Let*

$$
X = \left\{
\begin{array}{llllllllllll}
\text{H,} & \text{Li,} & \text{Be,} & \text{B,} & \text{C,} & \text{N,} & \text{O,} & \text{F,} & \text{Na,} & \text{Mg,} & \text{Al,} & \text{Si,} \\
\text{P,} & \text{S,} & \text{Cl,} & \text{K,} & \text{Ca,} & \text{Sc,} & \text{Ti,} & \text{V,} & \text{Cr,} & \text{Mn,} & \text{Fe,} & \text{Co,} \\
\text{Ni,} & \text{Cu,} & \text{Zn,} & \text{Ga,} & \text{Ge,} & \text{As,} & \text{Se,} & \text{Br,} & \text{Kr,} & \text{Rb,} & \text{Sr,} & \text{Y,} \\
\text{Zr,} & \text{Nb,} & \text{Mo,} & \text{Tc,} & \text{Ru,} & \text{Rh,} & \text{Pd,} & \text{Pt,} & \text{Cd,} & \text{In,} & \text{Sn,} & \text{Sb,} \\
\text{Te,} & \text{I,} & \text{Xe,} & \text{Cs,} & \text{Ba,} & \text{La,} & \text{Ce,} & \text{Pr,} & \text{Nd,} & \text{Pm,} & \text{Sm,} & \text{Eu,} \\
\text{Gd,} & \text{Tb,} & \text{Dy,} & \text{Ho,} & \text{Er,} & \text{Tm,} & \text{Yb,} & \text{Lu,} & \text{Hf,} & \text{Ta,} & \text{W,} & \text{Re,} \\
\text{Os,} & \text{Ir,} & \text{Pt,} & \text{Au,} & \text{Hg,} & \text{Tl,} & \text{Pb,} & \text{Bi,} & \text{Po,} & \text{At,} & \text{Ra,} & \text{Ac,} \\
& \text{Th,} & \text{Pa,} & \text{U,} & \text{Np,} & \text{Pu,} & \text{Am,} & \text{Cm,} & \text{Bk,} & \text{Cf,} & \text{Es} &
\end{array}
\right\}
$$

be the set of chemical elements for which at least one binary compound has been reported in the literature.

Figure 1 shows the plot of the number of binary compounds formed by each element (which is proportional to its frequency of occurrence in the data set), with elements ordered by increasing atomic number. We can see the following trends in this plot:

- Due to their inert nature, noble gases (crosses) appear as some of the lowest points of the histogram, if we take them as the end of a period, like in the periodic table, we can see trends for some chemical families as described below.

- Alkali metals (circles) appear in the first peaks starting each period, except the first one.

- Following alkali metals (increasing atomic number), in a lower position are the alkaline earth metals (squares).

- As expected, halogens (diamonds) are located on the highest peaks in the histogram due to its high presence in binary compounds.

- Before halogens, in a lower position, are the chalcogens (hexagons).

- If we look at the elements of the second period, we see a different behavior from the other periods, this can be regarded as an evidence of the singularity principle.

In summary, the previous observations show that oscillations between maximum and minimum frequency in the histogram of Figure 1 follow the trends of chemical periodicity.
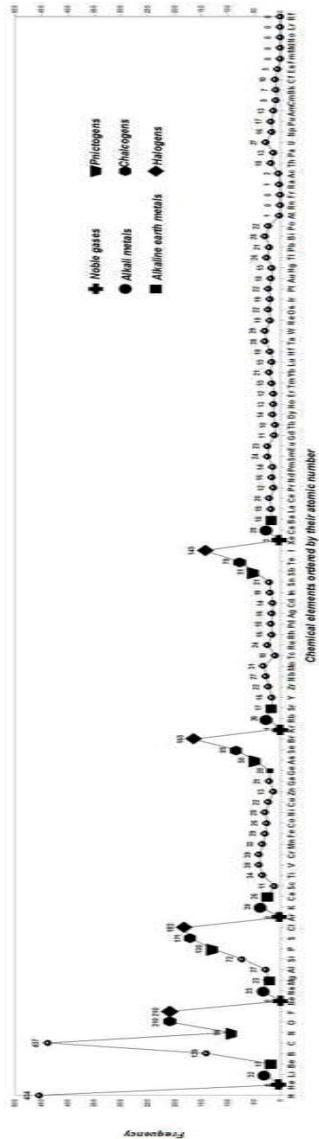
Figure 1: Number of binary compounds (frequency) in which chemical elements (ordered by atomic number) are present.
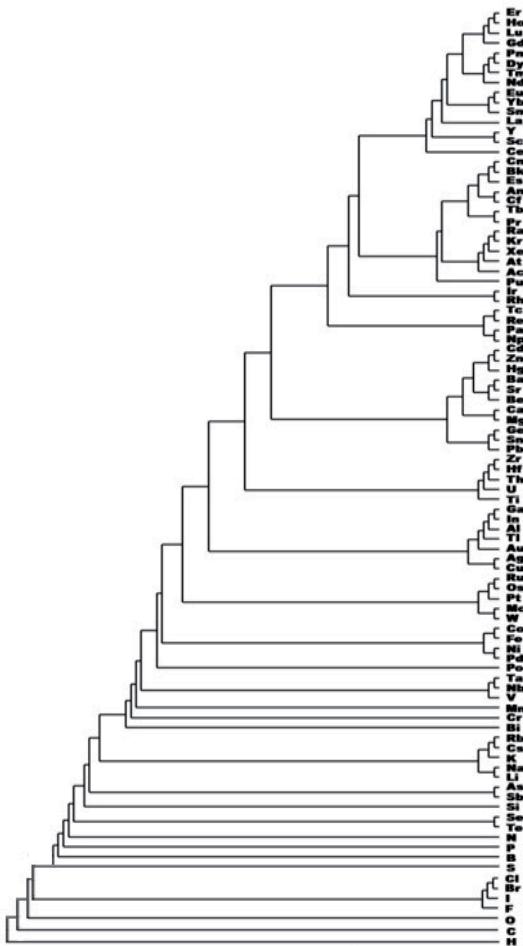
Figure 2: Dendrogram of chemical elements using the cardinality of the symmetric difference as metric and the average union as grouping methodology.

Going ahead with the methodology proposed, we used the difference of definition 26 as a metric and average union as a grouping methodology to built the dendrogram shown in Figure 2. To select an optimal cut in n-subtrees, we used the *selection number* and then generated the corresponding topological basis $\mathfrak{B}_n$. The value of n maximizing S was $n = 4$; therefore we built $\mathfrak{B}_4$. Thus, for the dendrogram shown in Figure 2 we have that

$\mathfrak{B}_4$ is:

$$\mathfrak{B}_4 = \left\{ \begin{array}{c} \{\text{Ge, Sn, Pb}\}, \{\text{Zr, Hf, Th, U}\}, \{\text{Ga, In, Al, Tl}\}, \{\text{Am, Cf}\}, \{\text{As, Sb}\}, \{\text{Ac}\}, \{\text{La}\}, \\ \{\text{Au}\}, \{\text{Ag, Cu}\}, \{\text{Ru, Os, Pt}\}, \{\text{Mo, W}\}, \{\text{Co, Fe, Ni, Pd}\}, \{\text{Ce}\}, \{\text{Cd, Zn, Hg}\} \\ \{\text{Ta, Nb, V}\}, \{\text{Tc, Re, Pa, Np}\}, \{\text{Rb, Cs, K}\}, \{\text{Na, Li}\}, \{\text{Ca, Mg}\}, \{\text{Ba, Sr, Be}\} \\ \{\text{Mn}\}, \{\text{Er, Ho, Lu, Gd}\}, \{\text{Pm, Dy, Tm, Nd}\}, \{\text{Eu, Yb, Sm}\}, \{\text{Bi}\}, \{\text{Y, Sc}\}, \\ \{\text{Cm, Bk, Es}\}, \{\text{Tb, Pr}\}, \{\text{Si}\}, \{\text{Ra, Kr, Xe, At}\}, \{\text{Ir, Rh}\}, \{\text{Se, Te}\}, \{\text{Pu}\}, \\ \{\text{Po}\}, \{\text{Ti}\}, \{\text{Cr}\}, \{\text{Cl, Br, I, F}\}, \{\text{N}\}, \{\text{P}\}, \{\text{C}\}, \{\text{B}\}, \{\text{H}\}, \{\text{S}\}, \{\text{O}\} \end{array} \right\} \quad (2)$$

We can see clusters whose elements belong to well known families in chemistry such as alkali metals, alkaline earth metals and halogens. We see also fragments of groups such as $\{\text{Al, Ga, In, Tl}\}$, $\{\text{Ge, Sn, Pb}\}$, $\{\text{Sc, Y}\}$, $\{\text{Mo, W}\}$, $\{\text{Ir, Rh}\}$ and $\{\text{Cu, Ag}\}$; some families of transition metals like $\{\text{V, Nb, Ta}\}$, and $\{\text{Zn, Cd, Hg}\}$. As columns on the periodic table (groups) are customarily regarded as depicting similar elements, we calculated the number of clusters found by our procedure, matching with columns (or parts of them) of the periodic table. We considered the open sets of the topological basis and found 18 single clusters (with only one element), 10 clusters with two elements, eight with three elements and eight with four elements. Out of these, we took those with more than one element for carrying out the analysis. Thus, 26 clusters might match with columns of the periodic table. We found that 15 of them (58%) are columns or parts of them, while 11 (42%) are not (Figure 3). However, other authors [52, 53, 54, 55, 56] have pointed out some other resemblances among chemical elements not matching with the columns of the periodic table. Some instances of those resemblances are the singularity principle (the chemistry of the second period elements often differs from that of the later members of their respective groups) [52], the diagonal relationships (there are similarities in chemical properties between an element and that at the lower right of it in the periodic table) [52], the inert pair effect (in some groups, the elements following the fifth and sixth periods exhibit oxidation states two values below the maximum of their respective groups) [52], the knight's move (there are similarities between an element of group n and period m with the element in group $n+2$ and period $m+1$ in the same oxidation state) [53, 54], and the secondary periodicity (there are similarities between the properties of the corresponding elements belonging to period m and those belonging to period $m + 2$) [55].

Of these similarity patterns we found instances of the singularity principle, as can be seen on the respective single clusters: B, C, N, and O. This result indicates that no other element forms binary compounds with those B, C, N, and O do and with the stoichiometry they do. Laing [53] has discussed the knight's move relationship, which has been notorious for Zn-Sn, Cd-Pb, and Ag-Tl. Although we did not find those similarities, we did found

others which fulfill knight's move definition; they are: Ru-Pt and Fe-Pd. Instances of the secondary periodicity we found are the resemblances between Xe-Rn, Kr-At, Tc-Pa, Tc-Np, Zr-Th, and Zr-U.

Perhaps the most interesting result of a non-vertical, in fact horizontal, similarity on the periodic table is that of the lanthanides. All of them are similar to at least another lanthanide, except by La and Ce, which appear in unitary clusters. For the case of actinides, we found two self similarities: the clusters $\{Cm, Bk, Es\}$ and $\{Am, Cf\}$. It is worth noting that all the other actinides, except Ac and Pu are similar to some transition metals, as Rayner-Canham has pointed out based upon experimental facts [54]. These similarities are part of the already discussed secondary periodicity.

It is also interesting to note the similarities $\{Fe, Co, N, Pd\}$, $\{Ru, Os, Pt\}$, and $\{Rh, Ir\}$. Besides the first two sets being instances of the knight's move, they constitute the group VIII, in the old IUPAC group numbering or VIIIB in the CAS numbering. But above all, the VIII group of Mendeleev's periodic table [56]. The similarities among these nine elements was what led Mendeleev to group them together.
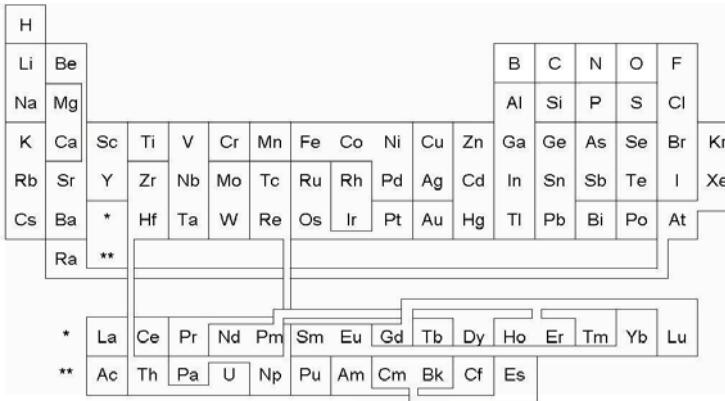
Figure 3: Open sets of the topological basis (similarity classes).

We can contrast the above results with those of Zhou et al., some clusters founds were: $\{Cl, Br\}$, $\{Zn, Cd\}$, $\{Li, Na, K, Rb\}$, $\{Mg, Ca\}$, $\{Ar, Kr, Ne, He\}$ and $\{Cu, Ag\}$, which highlights some fragments of families such as alkali metals and noble gases. On the other hand, some results obtained by Sneath were: $\{He, Ne, Ar, Kr, Xe\}$, $\{Cr, Mn, Fe, Co, Ni\}$, $\{Cl, Br\}$, $\{Cu, Ag, Au\}$, $\{Li, Na, K, Rb, Cs\}$, and $\{Mg, Ca, Sr, Ba\}$ which show fragments belonging to the alkali metals, noble gases and unlike the work by Zhou et al., alkaline earth metals and clusters of transition metals $\{Cu, Ag, Au\}$. In the study by Restrepo et.

al., unlike the studies of Zhou et al. and Sneath, they went further and developed the methodology to provide the set of study with a topology. In that study they obtained fragments of chemical families such as alkali metals, alkaline earth metals, noble gases and unlike the work by Zhou et al. and Sneath, fragments of families such as: pnictogens, chalcogens, halogens and relationships as the singularity principle, the inert pair effect and the diagonal effect. When contrasting Zhoe et al., Sneath and Restrepo et al. works with the current one, it is found that, in the current work there are more resemblances matching those depicted in the groups of the periodic table than in the previous works mentioned. The current results are specially interesting for transition metals, where the number of similarities found for these elements matched to a big extent those found in the periodic table (groups or columns). Due to the enlarged set of chemicals considered in this study, bigger than the sets studied by Zhou et al., Sneath and Restrepo et al., we could, for the first time using cluster analysis and in general chemotopology, analyze lanthanides and actinides, which turned out to be similar among themselves and not depicting the typical vertical similarity (group) on the periodic table.

In the following we describe the topological properties found for different families of chemical elements.

Topological properties of halogens (H) and other perfect sets

Let $H = \{F, Cl, Br, I\}$.

- **Closure:**
  $\overline{H} = \{F, Cl, Br, I\}$

- **Interior:**
  $\mathring{H} = \{F, Cl, Br, I\}$

- **Exterior:**
  $Ext(H) = X - H = H^c$

- **Boundary:**
  $b(H) = \varnothing$

- **Derived set:**
  $H' = \{F, Cl, Br, I\}$

In this case $H' = H$, which indicates that halogens form a perfect set according to definition 16. The same topological property is found for the following chemical sets: alkaline metals, alkaline earth metals, lanthanides, V group, Cr group, VIII group (old IUPAC numbering), Cu group, Zn group, B group, C group, pnictogens, and chalcogens, as well as $\{B, C, N, O\}$, among others. In topological terms, this means that, for any $x \in H$ (or in any other perfect set), each point adherent to $\{x\}$ is included in H. In chemical terms, this means that the set of halogens (or any other perfect set) includes
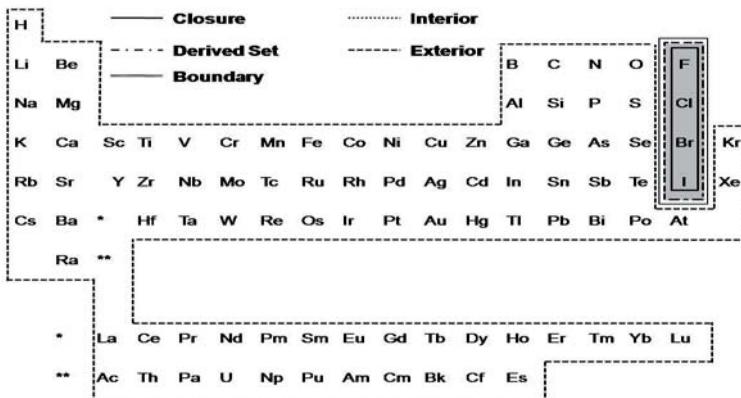
Figure 4: Topological properties of the halogens

every element that is 'comparably similar' to some element of the set. In this way, perfect sets are 'closed' (in the strict topological sense) and 'separated' from the rest of the space of chemical elements, so they may be regarded as **robust chemical families**.

This is quite clear for the case of halogens, which are all similar among themselves regarding their behavior when conforming binary compounds: each halogen combines with the same set of other elements in a similar ratio. This kind of combination is not shared with any other element out of the halogens, so that they clearly conform a robust family. But this fact only applies for perfect sets B that are also basic open sets, as that implies that no proper subset of B is perfect. Consider, on the other hand, the set of halogens and alkaline metals, which is also perfect. In this case, we have a perfect set conformed by two radically different families of elements. Yet it is also reasonable to regard it as a robust chemical family, as there is no 'comparably similar' element in its complement. For instance, if we were to remove any element (e.g. fluorine) from this set, it would no longer be a perfect set.

The set $S = \{B, C, N, O\}$, besides being a robust family, is interesting because any subset of S is also a perfect set. In other words, regardless of whether we take the whole set, or $\{B, C\}$, or $\{C, S, O\}$, or whatever subset of S we may please, no elements 'comparably similar' to those of the set being considered are left out. This means that elements in S are exceptionally separated from the rest of the elements, not by properties they share, but by characteristic properties that distinguish each individual element in the set. In other words, this set comprises 'singularities' of the space of chemical elements. Thus, here we have a topological image of the singularity principle.

But, besides robust families there are some other chemical sets of interest. In the following we show their topological properties:

Topological properties of metalloids (M)

Let $M = \{$B, Si, Ge, As, Sb, Te, Po$\}$.

- **Closure:**
  $$\overline{M} = \left\{ \begin{array}{l} \text{B, Si, Ge, As, Sb,} \\ \text{Te, Po, Sn, Pb, Se} \end{array} \right\}$$

- **Interior:**
  $$\mathring{M} = \{\text{B, Si, As, Sb, Po}\}$$

- **Exterior:**
  $$Ext(M) = X - \overline{M}$$

- **Boundary:**
  $$b(M) = \{\text{Ge, Sn, Pb, Se, Te}\}$$

- **Derived set:**
 $$M' = \{\text{Sn, Pb, As, Sb, Se}\}$$
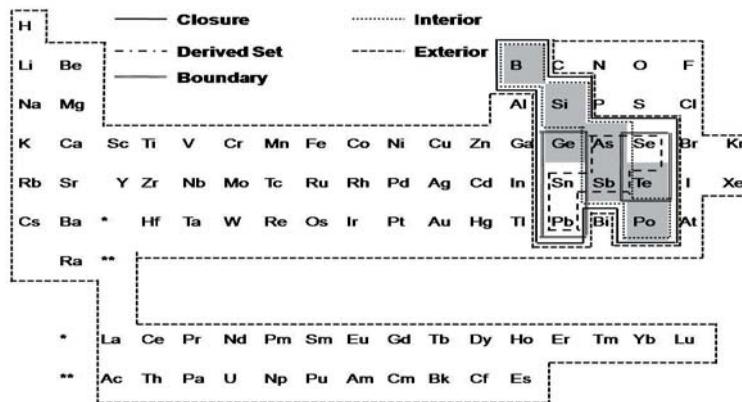
Figure 5: Topological properties of metalloids.

It turns out that the closure of M is the set formed by the elements of $M \cup \{$Sn, Pb, Se$\}$, which indicates that metalloids relate each other in a similar fashion to metals Sn, Pb and non-metal Se. Although metalloids, as a set, are similar to some metals and non-metals, it is possible to find representatives of metalloids, which are those belonging to the interior, i.e. B, Si, As, Sb, and Po. Those elements having the particularity of combining in the way metalloids do and also combining similarly as non-metalloids are Ge, Sn, Pb, Se, and Te, which are metalloids' boundary.

Topological properties of actinides (AC)

Let $AC = \{\text{Pu}, \text{Am}, \text{Cm}, \text{Bk}, \text{Cf}, \text{Es}, \text{Ac}, \text{Th}, \text{Pa}, \text{U}, \text{Np}\}$

- **Closure:**
  $\overline{AC} = \left\{ \begin{array}{l} \text{Pu}, \text{Am}, \text{Cm}, \text{Bk}, \text{Cf}, \text{Es}, \text{Ac}, \\ \text{Th}, \text{Pa}, \text{U}, \text{Np}, \text{Zr}, \text{Hf}, \text{Tc}, \text{Re} \end{array} \right\}$

- **Interior:**
  $\overset{\circ}{AC} = \{\ \text{Ac}, \text{Pu}, \text{Am}, \text{Cm}, \text{Bk}, \text{Cf}, \text{Es}\ \}$

- **Exterior:**
  $Ext(AC) = X - \overline{AC} = (\overline{AC})^c$

- **Boundary:**
  $b(AC) = \{\ \text{Th}, \text{Pa}, \text{Np}, \text{Zr}, \text{Hf}, \text{Tc}, \text{Re}, \text{U}\ \}$

- **Derived set:**
  $AC' = \left\{ \begin{array}{l} \text{Pu}, \text{Am}, \text{Cm}, \text{Bk}, \text{Cf}, \text{Es}, \text{Th}, \\ \text{Pa}, \text{Np}, \text{Zr}, \text{Hf}, \text{Tc}, \text{Re} \end{array} \right\}$
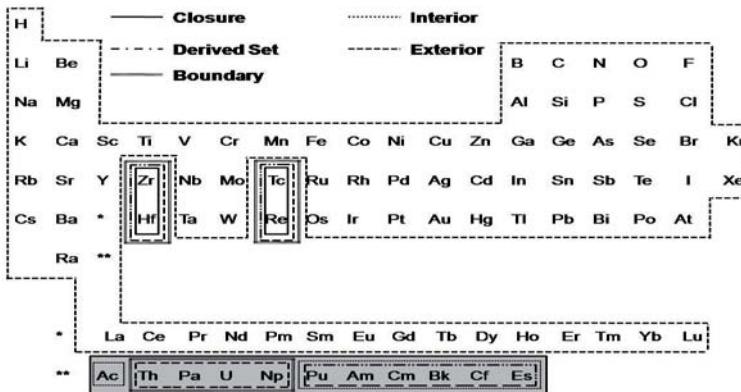

Figure 6: Topological properties of actinides.

The topological properties of actinides show the existence of chemical elements not belonging to the actinides but combining in a similar way with other elements as actinides do, they are Zr, Hf, Tc, and Re.

5   Conclusions

Previous similarity studies of chemical elements have shown that groups of chemical elements (on the periodic table) are clusters of similar elements. These similarities have been found through the characterization of chemical elements by their chemical, physical and physicochemical properties. Some studies combine all these properties, some others focus on some of them. All in all, the conclusion regarding a group as a set of similar elements remains, independently of the kind of properties discussed in use. In the current manuscript we did not use neither physical nor physicochemical properties, we ran

a study based only upon a relational property: two chemical elements are related if they are found on a binary compound. This property, according to Schummer, is a chemical property. The results show that chemical groups can be found with this unique property and the idea of a group as a set of similar elements remains. What is worth nothing is the chemical taste we gave to the similarity; we went from physical and physicochemical similarity (of previous works by our group) to chemical similarity. With this study it is possible to claim that two chemical elements are similar if they combine with other (same) elements forming binary compounds with similar stoichiometries. This kind of similarity on chemical combination grounds was in fact one of the key properties Mendeleev considered when finding chemical periodicity.

Besides finding groups and fragments of groups of the periodic table, we also found instances of the singularity principle, the knight's move and of the secondary periodicity.

In topological terms, we found that a wealth of groups of the periodic table constitute robust families, here defined as perfect sets, which means that the elements belonging to a chemical family have only similar properties to themselves, and different from those presented by elements belonging to another family. We found that the topological boundary of semimetallic elements contains, on the one hand, some metals and, on the other, some nonmetals.

From the mathematical viewpoint we linked network theory with chemotopology by proving that the set of positions of a network is a topological basis. This result opens the field of network theory to further chemotopological analyzes.

6 Acknowledgements

References

[1] G. Restrepo, L. Pachón, Mathematical aspects of the periodic law, *Found. Chem.* **9** (2007) 198–214.

[2] W. Schwarz, S. G. Wang, Icon of chemistry: The periodic system of chemical elements in the new century, *Angew. Chem.* **48** (2009) 3404–3415.

[3] D. Bonchev, Periodicity of the chemical elements and nuclides: An information–theoretic analysis, in: R. B. King, D. Rouvray (Eds.), *The Mathematics of the Periodic Table*, Nova, New York, 2006, pp. 161–188.

[4] D. Klein, Similarity and dissimilarity in posets, *J. Math. Chem.* **18** (1995) 321–348.

[5] D. Klein, D. Babić, Partial orderings in chemistry, *J. Chem. Inf. Comput. Sci.* **37** (1997) 656–671.

[6] D. Klein, Prolegomenon on partial orderings in chemistry, *MATCH Commun. Math. Comput. Chem.* **42** (2000) 7–21.

[7] E. Scerri, V. Kreinovich, P. Wojciechowski, R. Yager, Ordinal explanation of the periodic system of chemical elements, *Int. J. Uncertainty Fuzziness Knowledge Based Sys.* **6** (1998) 387–400.

[8] D. Robert, R. Carbó–Dorca, On the extension of quantum similarity to atomic nuclei: Nuclear quantum similarity, *J. Math. Chem.* **23** (1998) 327–351.

[9] D. Robert, R. Carbó–Dorca, General trends in atomic and nuclear quantum similarity measures, *Int. J. Quantum Chem.* **77** (2000) 685–692.

[10] O. Novaro, Group theory of the periodic table, in: R. B. King, D. Rouvray (Eds.), *The Mathematics of the Periodic Table*, Nova, New York, 2006, pp. 217–236.

[11] M. Kibler, A group–theoretical approach to the periodic table: Old and new developments, in: R. B. King, D. Rouvray (Eds.), *The Mathematics of the Periodic Table*, Nova, New York, 2006, pp. 237–264.

[12] V. Ostrovsky, Group theory applied to the periodic table of the elements, in: R. B. King, D. Rouvray (Eds.), *The Mathematics of the Periodic Table*, Nova, New York, 2006, pp. 265–312.

[13] X. Z. Zhou, K. H. Wei, G. Q. Chen, Z. X. Fan, J. J. Zhan, Fuzzy cluster analysis of chemical elements, *Chinese J. Appl. Chem..* **17** (2000) 167–168.

[14] P. H. A. Sneath, Numerical classification of the chemical elements and its relation to the periodic system, *Found. Chem.* **2** (2000) 237–263.

[15] G. Restrepo, H. Mesa, E. Llanos, J. L. Villaveces, Topological study of the periodic system, *J. Chem. Inf. Comput. Sci.* **44** (2004) 68–75.

[16] G. Restrepo, J. L. Villaveces, Topological study of the chemical elements, in: *Proceedings of the Third Joint Sheffield Conference on Chemoinformatics*, The University of Sheffield, Sheffield, 2004.

[17] G. Restrepo, J. L. Villaveces, From trees (dendrograms and consensus trees) to topology, *Croat. Chem. Acta.* **78** (2005) 275–281.

[18] G. Restrepo, H. Mesa, E. Llanos, J. L. Villaveces, Topological study of the periodic system, in: R. B. King, D. Rouvray (Eds.), *The Mathematics of the Periodic Table*, Nova, New York, 2006, pp. 75–100.

[19] G. Restrepo, E. Llanos, H. Mesa, Topological space of the chemical elements and its properties, *J. Math. Chem.* **39** (2006) 401–416.

[20] G. Restrepo, H. Mesa, J. L. Villaveces, On the topological sense of chemical sets, *J. Math. Chem.* **39** (2006) 363–376.

[21] G. Restrepo, H. Mesa, Chemotopology: Beyond neighbourhoods, *Curr. Comput. Aid Drug.* **7** (2011) 90–97.

[22] D. Rouvray, R. B. King (Eds.), *The Mathematics of the Periodic Table,* Nova, New York, 2006.

[23] J. Schummer, The chemical core of chemistry I: A conceptual approach, *HYLE – Int. J. Phil. Chem.* **4** (1998) 129–162.

[24] E. Uribe, M. Daza, G. Restrepo, Chemotopological study of the fourth period mono–hydrides, *WSEAS Trans. Inf. Sci. Appl.* **2** (2005) 1085–1090.

[25] A. Bernal, E. Daza, Chemical properties, elements and periodicity, *J. Sci. Educ.* **10** (2009) 126–127.

[26] S. Wasserman, K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge Univ. Press, New York, 1994.

[27] M. Newman, The structure and function of complex networks, *SIAM Rev.* **45** (2003) 167–256.

[28] H. Jeong, B. Tombor, R. Albertl, Z. N. Oltvai, A. Barabási, The large–scale organization of metabolic networks, *Nature* **407** (2000) 651–654.

[29] A. Barabási, Z. N. Oltvai, Network biology: Understanding the cell's functional organization, *Nat. Rev. Genet.* **5** (2004) 101–113.

[30] Y. Ahn, J. Bagrow, S. Lehmann, Link communities reveal multiscale complexity in networks, *Nature* **466** (2010) 761–764.

[31] J. Montoya, S. Pimm, R. Solé, Ecological networks and their fragility, *Nature* **442** (2006) 259–264.

[32] E. Estrada, The complex networks of earth minerals and chemical elements, *MATCH Commun. Math. Comput. Chem.* **59** (2008) 605–624.

[33] H. Mesa, G. Restrepo, On dendrograms and topologies, *MATCH Commun. Math. Comput. Chem.* **60** (2008) 371–384.

[34] N. Biggs, E. Lloyd, R. Wilson, *Graph Theory, 1736–1936*, Oxford Univ. Press, New York, 1999.

[35] W. Nooy, A. Mrvar, V. Batagelj, *Exploratory Social Network Analysis with Pajek*, Cambridge Univ. Press, New York, 2005.

[36] U. Brandes, T. Erlebach, *Network Analysis: Methodological Foundations*, Springer, New York, 2005.

[37] J. Lerner, Role assignments, in: U. Brandes, T. Erlebach (Eds.), *Network Analysis: Methodological Foundations*, Springer, New York, 2005, pp. 216–252.

[38] E. Deza, M. Deza, *Dictionary of Distances*, Elsevier, Oxford, 2006.

[39] R. R. Sokal, P. Sneath, *Principles of Numerical Taxonomy*, Freeman, San Francisco, 1993.

[40] I. H. Witten, E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, Elsevier, San Francisco, 2005.

[41] B. Mendelson, *Introduction to Topology*, Dover, New York, 1990.

[42] G. Restrepo, *Búsqueda de la Estructura Matemática de la Ley Periódica*, MSc thesis, Universidad Industrial de Santander, Bucaramanga-Colombia, 2004.

[43] G. Restrepo, R. Brüggemann, Ranking regions through cluster analysis and posets, *WSEAS Trans. Inf. Sci. Appl.* **2** (2005) 976–981.

[44] E. J. Llanos, Private Communication.

[45] W. Leal, G. Restrepo, Chemical elements, their sociology, *J. Sci. Educ.* **10** (2009) 126-127.

[46] E. Scerri, *The Periodic Table: Its Story and Its Significance*, Oxford Univ. Press, New York, 2007.

[47] M. Smith, J. March, *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, Wiley, New York, 2007.

[48] W. Leal, G. Restrepo, A. Bernal, L. Esquivel, On the topology of formal concept analysis, in preparation.

[49] N. Greenwood, A. Earnshaw, *Chemistry of the Elements*, Elsevier, London, 1997.

[50] NIST Chemistry WebBook. http://webbook.nist.gov/chemistry/, 2009

[51] The Chemical Thesaurus. http://www.chemthes.com/, 2009

[52] G. Rayner–Canham, T. Overton, *Descriptive Inorganic Chemistry*, Freeman, New York, 2002.

[53] M. Laing, Periodic patterns, *J. Chem. Educ.* **78** (2001) 877–877.

[54] G. Rayner–Canham, Periodic patterns, *J. Chem. Educ.* **77** (2000) 1053–1056.

[55] V. N. Ostrovsky, What and how physics contributes to understanding the periodic law, *Found. Chem.* **3** (2001) 145–182.

[56] D. I. Mendeleev, The periodic law of the chemical elements, *J. Chem. Soc.* **55** (1889) 634–656.