Substantia
An International Journal of the History of Chemistry

Firenze University Press
www.fupress.com/substantia

Citation: G. Restrepo (2019) Compounds Bring Back Chemistry to the System of Chemical Elements. *Substantia* 3(2) Suppl. 4: 115-124. doi: 10.13128/Substantia-739

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Compounds Bring Back Chemistry to the System of Chemical Elements

Guillermo Restrepo

Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
E-mail: guillermo.restrepo@mis.mpg.de

Abstract. The periodic system of chemical elements was historically devised by assessing order and similarity relationships among the elements from their compounds, that is, using the accumulated results of chemical practice and knowledge. However, the current approach to the system is based on an ontology of isolated atoms where similarities, especially, are addressed through resemblances of electronic configurations. Here we show how the historical approach can be combined with computational tools for data analysis to build up the system based on the compounds reported by chemists. The approach produces well-known similarities of chemical elements when applied to binary compounds. The results come from the analysis of 4,700 binary compounds of 94 chemical elements, whose resemblances are quantified based on the elements they form compounds with and the proportions of those combinations. It is found that similarities do not always correspond to columns of the conventional periodic table and that besides robust similarities such as those of alkali metals, halogens and lanthanoids, there are other mixed similarities involving transition metals and actinoids, some of which were already known for a long time. These similarities are described. Finally, the advantages and disadvantages of the electronic and the compound approach to the system are discussed. It is concluded that the current data availability and computational facilities make possible to think of a periodic system closer to the chemical milieu of compounds, bringing chemistry back to the system.

Keywords. Compound, substance, periodic system, chemical space, similarity.

INTRODUCTION

While some of the formulators of the periodic system[1] were after numerical relationships among atomic weights,[2] Julius Lothar Meyer (1830 –1895) and Dmitri Ivanovich Mendeleev (1834-1907) were especially interested in systematizing chemical knowledge.[3] They sought to highlight relationships

[1] According to van Spronsen (reference 1), there were at least six formulators: Alexandre-Emile Béguyer de Chancourtois, John Alexander Reina Newlands, Julius Lothar Meyer, William Odling, Gustavus Detlef Hinrichs and Dmitri Ivanovich Mendeleev.
[2] This is especially visible in Newlands' and Odling's approaches. See references 2 and 3, respectively.
[3] The importance of textbook writing in the process of formulating the periodic system for Meyer and Mendeleev has been stressed by Gordin (reference 4) among other historians.

among chemical elements. The two relations supporting their sketches were similarity and order,[5] which built up a system for chemical elements; where "system" is understood in the ample sense of a set of related objects,[6] in this case chemical elements.[4]

The providers of order and similarity were in fact chemical compounds. Atomic weights, which constituted the order criterion; were determined by finding the smallest common weight of large numbers of compounds containing the reference element in question.[7] Similarity was based on resemblance in the composition of substances. As Mendeleev stated it in 1905: "if CO_2 and SO_2 are two gases which closely resemble each other both in their physical and chemical properties, the reason of this must be looked for not in an analogy of sulphur and carbon but in that identity of the type of combination, RX_4, which both oxides assume".[8] He concludes: "The elements, which are most chemically analogous, are characterized by the fact of their giving compounds of similar form RX_n".[8]

Gathering together chemical compounds constitutes a *chemical space*, which spans all energetically stable atomic ensembles.[5] By chemical space we designate all material species chemists experiment with, ranging from substances that can be stored in "bottles" such as liquids, solids or gases, to atomic clusters held together by van der Waals interactions. Throughout history, chemists have explored such a space by synthesis or extraction of new compounds. As chemists report their findings of new substances in the scientific literature, a suitable proxy for knowing how fast the exploration of the space has been carried out is the rate of reports of new chemical substances. We recently demonstrated that the chemical space has been historically explored in an exponential fashion with an annual growth rate of 4.4%,[10] indicating that about every 16 years chemists have doubled the number of substances since 1800, which was the starting point of the study reported in reference 10. This magnitude can be better expressed by the fact that the new substances reported in 2015 amount to the total of those reported between 1800 and 1950, i.e. the production of 2015 is equivalent in magnitude to the production of 150 years of new substances.[6]

This rapid growth poses a challenge to the periodic system and raises different questions: what was the chemical space in the 1860s when the system was formulated? What is the current chemical space and how does it affect the periodic system? We recently explored the space in the 1860s and found that several of the classes of similar elements known at that time could actually be obtained by analyzing the resemblance of the elements through their compounds through our approach, confirming the fact that Mendeleev and Meyer had indeed mapped the chemical space of their time.[13] In the current paper we analyze the question of the relationship between the current space and the periodic system and the implications for teaching the system.

CLASSIFYING THROUGH THE CHEMICAL SPACE

A classification of the chemical elements based upon the known chemical space up to 2011 was reported in 2012[14] through the analysis of 4,700 binary compounds,[7] which accounted for 94 chemical elements (Figure 1).[8] By binary compounds we mean substances made of two elements, e.g. water, ammonia and methane, but not sulfuric acid and fullerene, for instance.

Following the Mendeleevian approach to similarity of chemical elements, which states that two elements are similar if they form compounds with common elements in similar proportions, Leal et al.[14] formalized the notion as follows: For a given set of compounds the elements and proportions of combination of each element x are gathered in the neighborhood of the element x, called N_x. For example, if only $BeCl_2$, $MgCl_2$, $BeBr_2$ and $MgBr_2$ are the substances considered, the neighborhoods are: N_{Be}={Cl2/1, Br2/1} = N_{Mg} and N_{Cl} = {Be1/2, Mg1/2} = N_{Br}, which shows the similarity between Be and Mg and between Cl and Br, respectively.[9]

With the neighborhoods for each of the 94 elements, the similarity of every couple of elements was calculated

[4] Interestingly, little emphasis has been made on the periodic system as an actual system. What we stress in reference 5 is that order and similarity are the structure keepers of all possible periodic systems.

[5] As later discussed, by atomic ensembles we mean substances, which may be transient ones. Moreover, in most extreme cases the ensembles do not necessarily require the presence of chemical bonds. More on the chemical space is found in reference 9.

[6] The idea of assessing chemistry growth through the frequency of reports of new substances was initiated by Schummer (reference 11). Quantitative studies of scientific growth began with Solla Price (reference 12), who analyzed the growth of scientific literature in different disciplines. Chemistry was found to be the most rapid growing discipline in terms of published abstracts.

[7] These compounds are a representative sample of the space by 2011, as 4,700 > √12,060,017, where 12,060,017 is the number of known substances by 2011. Details of the annual production of new compounds are reported in reference 10.

[8] The elements analyzed are: H, Li, Be, B, C, N, O, F, Na, Mg, Al, Si, P, S, Cl, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Br, Kr, Rb, Sr, Y, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Cd, In, Sn, Sb, Te, I, Xe, Cs, Ba, La, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu, Hf, Ta, W, Re, Os, Ir, Pt, Au, Hg, Tl, Pb, Bi, Po, At, Ra, Ac, Th, Pa, U, Np, Pu, Am, Cm, Bk, Cf, and Es.

[9] In general, for a compound $x_a y_b$, the neighborhood of x is given by {xa/b, yb/a}.

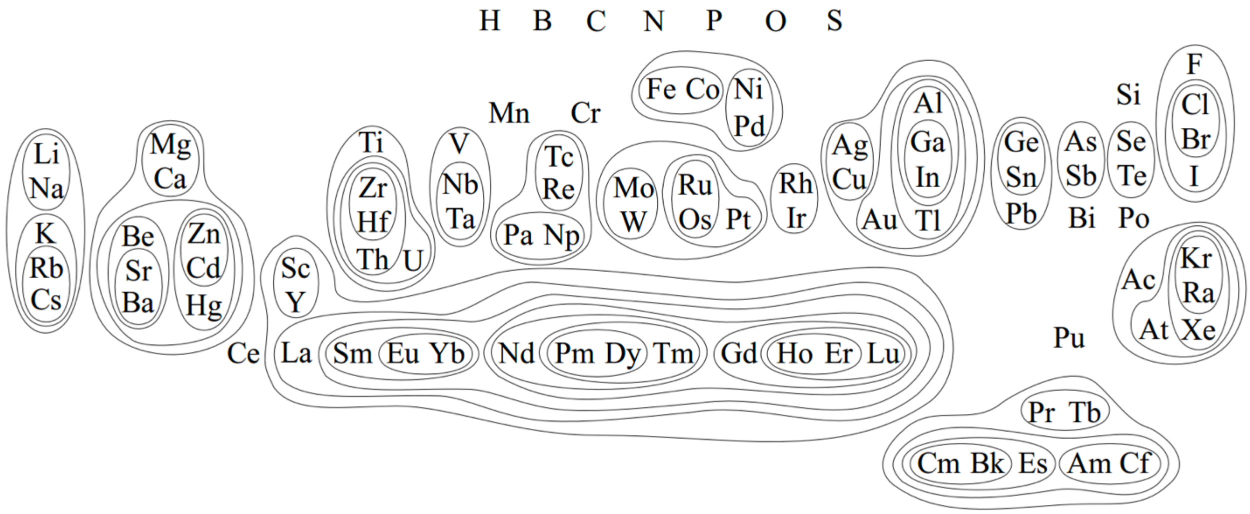


Figure 1. Most relevant similarity classes for 94 chemical elements obtained by analyzing binary compounds. Elements are spread on the plane trying to keep their positions as those depicted in the current middle-form periodic table while at the same time spatially indicating nearness in .behavior, expanding on the traditional grid. Sets and subsets group elements by similarities. Pairs of similar elements are denoted by subsets of two elements. Whenever a subset belongs to a larger subset, this indicates a hierarchical similarity. For example, Rb and Cs are similar elements, which in turn hold a more relaxed similarity regarding K.

as the commonalities of their neighborhoods. In general, the more common neighbors two elements have, the more similar they are (see Appendix 1 for details). This is exemplified with the following compounds:[15] HF, B_2H_6, B_5H_9, $B_{10}H_{14}$, from which the neighborhoods of the elements involved are: N_F= {H1/1}, N_H= {F1/1, B2/6, B5/9, B10/14}, N_B = {H6/2, H9/5, H14/10}. Thus, according to these compounds, hydrogen is more similar to boron than to fluorine, for there are more commonalities with the former than with the latter.

Once the similarities for all pairs of elements are calculated, clusters of similar elements are built up, for example through hierarchical cluster analysis. This technique looks for the most similar pair of elements and group them together in a first cluster. The new cluster is then included as a new object, where the similarities of the two members of the cluster regarding all the other elements are averaged.[10] In this setting, the most similar couple of elements is found, which may be made either of two elements, or of the cluster of the first merging and a third element. A new cluster is then formed and the process iterates until all elements have been merged.[11]

The outcome of the classification through hierarchical cluster analysis is a nested system of similarity classes that establishes the hierarchy of classes from which the classificatory technique takes its name. In the next section, we discuss the results of applying this methodology to the 4,700 binary compounds.

SIMILARITY LANDSCAPES: FROM CLASSIFICATION TO SYSTEM

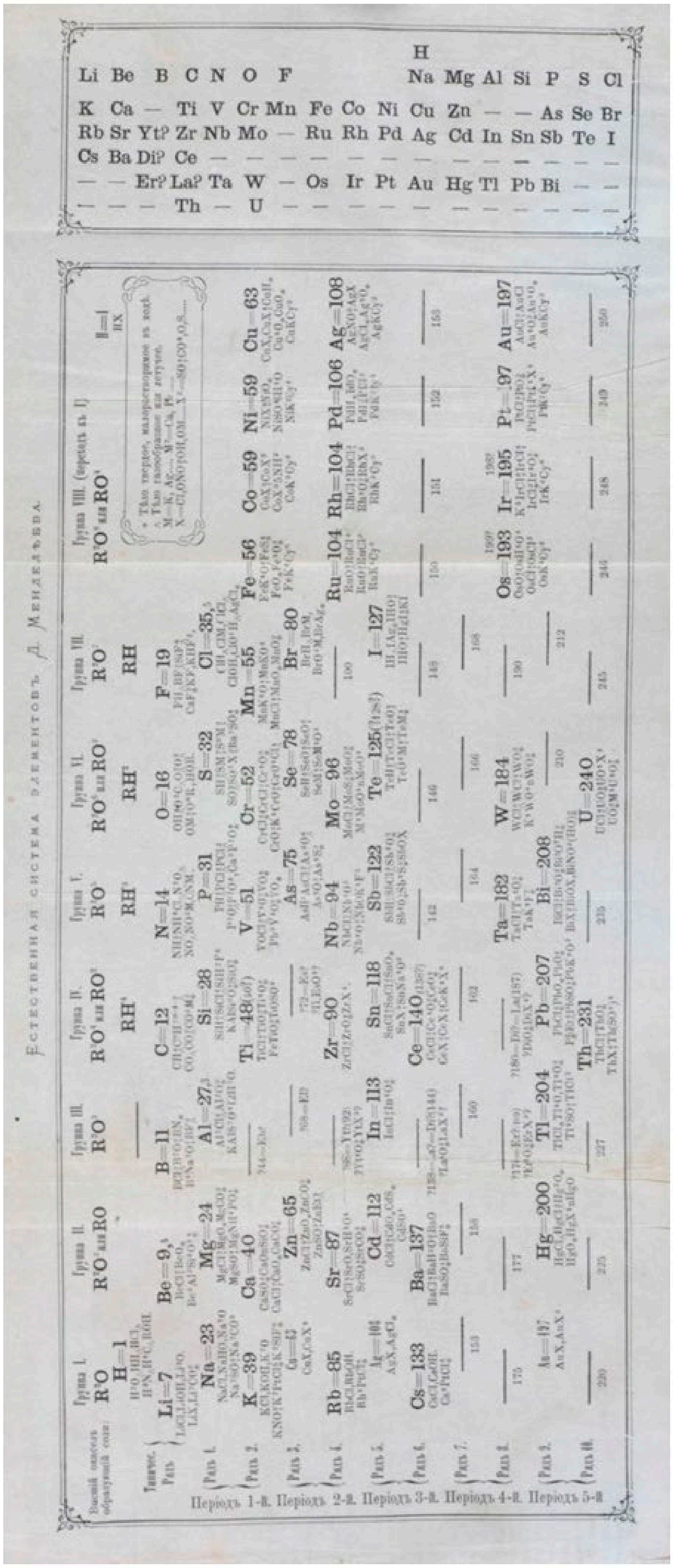
The hierarchy of similarity classes for the 94 chemical elements studied in Leal et al.[14] can be depicted either as a classification tree, as in reference 14, or as a similarity landscape as in reference 15. In the current section, we present a simplified version of the similarity landscape (Figure 1).

Hydrogen is the most dissimilar element, which indicates that other elements combine very differently than hydrogen does. Other dissimilar elements are carbon, oxygen, sulfur, boron, phosphorus, and nitrogen (top of Figure 1).

There are well-known classes of similar elements, e.g. alkali metals and halogens, with opposite chemis-

[10] Merging elements into a cluster and calculating the similarity of the cluster regarding the other elements is equivalent to finding the distance from an object to a set. There are different ways to find such a distance and the selected here of averaging the similarity of the elements of the cluster is called group average methodology. Other approaches are, for instance, the complete linkage, where the similarity of the cluster to the other elements is based on the similarity of the most dissimilar of the elements of the cluster. Further details on these and other grouping methodologies are found in reference 15.

[11] Particular details of the clustering process are found in reference 15.



A

Figure 2a. Mendeleev's periodic tables by 1871. A) Folding table in his textbook, D. I. Mendeleev, *Osnovy khimii (The Principles of Chemistry)*, Obshchestvennaia pol'za: St. Petersburg, Part II, 1871.[16]

Reihen	Gruppe I. — R^2O	Gruppe II. — RO	Gruppe III. — R^2O^3	Gruppe IV. RH^4 RO^2	Gruppe V. RH^3 R^2O^5	Gruppe VI. RH^2 RO^3	Gruppe VII. RH R^2O^7	Gruppe VIII. — RO^4
1	H=1							
2	Li=7	Be=9,4	B=11	C=12	N=14	O=16	F=19	
3	Na=23	Mg=24	Al=27,3	Si=28	P=31	S=32	Cl=35,5	
4	K=39	Ca=40	—=44	Ti=48	V=51	Cr=52	Mn=55	Fe=56, Co=59, Ni=59, Cu=63.
5	(Cu=63)	Zn=65	—=68	—=72	As=75	Se=78	Br=80	
6	Rb=85	Sr=87	?Yt=88	Zr=90	Nb=94	Mo=96	—=100	Ru=104, Rh=104, Pd=106, Ag=108.
7	(Ag=108)	Cd=112	In=113	Sn=118	Sb=122	Te=125	J=127	
8	Cs=133	Ba=137	?Di=138	?Ce=140	—	—	—	— — — —
9	(—)	—	—	—	—	—	—	
10	—	—	?Er=178	?La=180	Ta=182	W=184	—	Os=195, Ir=197, Pt=198, Au=199.
11	(Au=199)	Hg=200	Tl=204	Pb=207	Bi=208	—	—	
12	—	—	—	Th=231	—	U=240	—	— — — —

B

Figure 2b. Mendeleev's periodic tables by 1871. B) as in his Table II in his publication, D. Mendeleev, *Die periodische Gesetzmässigkeit der chemischen Elemente*, Ann. Chem. Pharm. 1871, 8 (Supplementband).[17]

tries and ways of combining with other elements. This was noted and detailed by Mendeleev in the table published in his second volume (1871) of *Principles of Chemistry*[16] (Figure 2a), where it is explicitely written that alkali metals combine with oxygen in a 2:1 ratio (R^2O using Mendeleev's notation). In contrast, halogens do it in a 2:7 fashion (R^2O^7). This table, among several other commonalities, shows that alkali metals form hydroxides of the form XOH, being X an alkali metal. One of the commonalities for halogens in this table is that they form compounds RH, where R is a halogen. The table was then simplified to the second table of Mendeleev's 1871 paper on the periodic system[17] (Figure 2b), where only the general formulae for oxides and hydrides remain, but the particular details of the table in Figure 2a are omitted.[12] In our work on the periodic system of 1869 using the known chemical space at that time, we found additional commonalities for alkali metals, e.g. $XAsO_4$, X_2CO_3, X_2SO_4, XNO_3, XCl, and XI.[13] For halogens, we found RC_2H_3O.

Finding alkali metals and halogens as classes of similar elements with the sample of compounds analyzed in reference 14, which include not only oxides and hydrides, indicates that the commonalities of the members of these families extend to most of their combinations with other elements, not only to those with oxygen and hydrogen.

Delving into the details of each of these classes, alkali metals can be divided into two sub-clusters, one of light metals: lithium and sodium and a second of heavier ones: potassium, rubidium, and strontium. Halogens follow a nested similarity structure, chlorine and bromine being the most akin halogens, with some resemblance to iodine. Fluorine is the most dissimilar halogen. The explanation of the strong dissimilarity of fluorine is based upon its small atomic size. This is part of the so-called *singularity principle*, which states that the chemistry of the second period elements is often different from the latter members of their respective groups.[18] Such principle is generally evident in the lack of similarities of carbon, oxygen and the other elements mentioned above and shown at the top of Figure 1.

Alkali-earth metals appear together with group 12 metals. This cluster of eight elements was recognized by Mendeleev as early as 1871 and is characterized by a 1:1 ratio of each element in the cluster with oxygen (RO and R^2O^2 in Mendeleev's 1871 periodic table (Figure 2)). As for alkali metals and halogens, this similarity class indicates that its elements combine in a similar fashion not only with oxygen but also with other elements (more details to be found in reference 14). In the study of the

[12] As suggested by Brigitte Van Tiggelen during the revision of the current paper, this is an early evidence of how the simplification of the table through its dissemination started to leave aside valuable chemical information.

system of 1869 using the known chemical space, other resulting commonalities were RF_2, RCl_2 and RS.[13]

Another cluster of similar elements is the couple of selenium and tellurium, which constitutes the only case of similarity among chalcogens.[13] All other chalcogens constitute single classes. Likewise, arsenic and antimony are the only cluster including pnictogens.[14] This lack of vertical similarity for groups of the periodic table indicates differences among elements members of each group. Although it is true that most chalcogens have RH^2, RO^3, and R^2O^6 combinations as stated by Mendeleev (Figure 2),[16,17] the sample of substances used in Leal et al.[14] shows that there are other combinations disturbing this similarity put forward in 1871.[15] On the other hand, the already discussed singularity principle makes oxygen behave differently in comparison with the other chalcogens, combining with other elements in a rather different way as its homologues do. The same argument applies for pnictogens, with nitrogen behaving differently, but still with RH^3 and R^2O^5 combinations, as noted by Mendeleev (Figure 2).[16,17]

Other clusters of similar elements are the trio of vanadium, niobium and tantalum, today labeled as group 5 and recognized by Mendeleev as a set of elements having relations RH^3 and R^2O^5.[16,17] Interestingly, the quartet of ferrous metals: iron, cobalt, nickel, and palladium, which are members of the group VIII for Mendeleev[16,17] and the old IUPAC group numbering, or VIIIB in the CAS numbering, forms a cluster.[15] This cluster indicates that these elements have indeed commonalities in terms of the compounds they form, for example RO^4 and R^2O^8.[16-17] In the current group numbering of the periodic table, group VIII corresponds to groups 8 to 10, which include nine elements. The results of Leal et al.[14] actually show that resemblances among these elements are not only restricted to iron, cobalt, and nickel: the trio ruthenium, osmium, and platinum is another case.[16] By considering larger clusters, it is found that ruthenium, osmium, and platinum also have certain resemblance with molybdenum and tungsten.

Interestingly, the pair of similar elements rhodium and iridium, traditionally considered as part of platinum metals,[17] do not appear closely related to the other platinum homologues as usually stated but loosely connected to some lanthanoids and actinoids.

Titanium, zirconium and hafnium, forming group 4 of the current periodic system, constitute a cluster of similar elements, which holds similarity ties with the actinoids thorium and uranium. These transition metal-actinoid resemblances were noted by Seaborg as early as 1945[20] and are based on similarity of combination with other elements where the +4 oxidation state of the metal is the commonality.[18]

The resemblance of transition metals zirconium and hafnium was explained by Goldschmidt through the *lanthanoid contraction*, which is currently understood as the spatial shrinking of lanthanoid atoms as a consequence of the filling of 4*f* shells that contracts 5*p* and 6*s* shells. This contraction makes that Zr^{4+} and Hf^{4+} have roughly the same ionic radii when six-coordinated.[22-24]

As we remarked in our previous work[15], even if the zirconium and hafnium resemblance is known, in some theoretical communities, it is considered an exception caused by "anomalous cancellation of relativistic effects" for elements of the 5th- and 6th-rows of the system.[24,25] In the study by Leal et al. mentioned earlier,[14] it was found that out of the 17 possible pairs of 5th- and 6th-row elements that belong to a group, there are five other pairs sharing similarities: niobium and tantalum; molybdenum and tungsten; technetium and rhenium; ruthenium and osmium; and, finally, rhodium and indium. The first two couples here listed were discussed by Huheey and Huheey on the basis of the very close radii for 5th- and 6th-row species.[23] This resemblance was also discussed in terms of similar oxidation states.[26]

Our work uncovered a cluster of elements belonging in group 13, but which excludes boron. Here, gallium and indium are the most similar elements, which then have resemblance relations with aluminum and finally with thallium. Interestingly, this quartet turns out to be similar to gold and to a lesser extent to the couple of coinage metals cooper and silver.

So far, we have discussed clusters that are only a few elements in length, but there are also larger clusters corresponding to elements that are very similar in terms of the compositions they form. These are the lanthanoids and actinoids. It was found that lanthanoids are more similar among themselves than actinoids. This is caused by a dominant +3 oxidation state, which has been explained on electronic grounds.[15]

Remarkably, rare earths constitute a large cluster of similar elements that groups together scandium and

[13] Group 16 of the conventional periodic table.

[14] Group 15 of the conventional periodic table.

[15] In reference 13 we found that another commonality for chalcogens is XNH_5, being X a chalcogen.

[16] According to Rayner-Canham, ruthenium and osmium become similar as each forms compounds where the +8 oxidation state is favored. The commonality of these two elements with platinum stems mainly from compounds where the +4 oxidation state of the metal is present. Details in reference 19.

[17] By platinum metals is understood: ruthenium, osmium, rhodium, iridium, palladium, and platinum.

[18] Schwarz recently discussed the similarity of early actinoids with some transition metals of the 6th-row of the periodic system (details in reference 21).

yttrium and is relevant to an ongoing IUPAC discussion about the elements that should be recommended as belonging to group 3 of the periodic system.[27] Part of the question is whether scandium and yttrium should be grouped together with lanthanum or with lutetium. The results here discussed show that lanthanum should be placed in group 3 as the element holds similarities with 11 lanthanoids and scandium and yttrium. In contrast, lutetium is more akin to lanthanoids and not so much to scandium and yttrium.[14,15]

In contrast with the strong similarities among lanthanoids, actinoids are tied by a more diverse repertoire of combinations because of a more ample set of available oxidation states that vary from +2 to +6. This has been explained on quantum chemical grounds and is known as the actinoid contraction, which is more irregular than the lanthanoid contraction.[21,26]

Resemblances between transition metals and *f*-elements are not specific to lanthanoids. Actinoids also keep some of these similarity ties, for example with zirconium, hafnium, technetium and rhenium. In particular, uranium is similar to titanium, zirconium and hafnium (group 4) and also to thorium. Similarities of these sorts have been reported by Rayner-Canham and studied by Schwarz and Rich.[18,28]

An actinoid worth mentioning is plutonium, which holds similarities with other actinoids[19] along with lanthanoids terbium and praseodymium. It has been argued that plutonium particularities stem from its peculiar electronic properties resulting from the changing roles of the 5*f* orbitals, which, for example make it equilibrate four oxidation states in solution, something not reported for any other chemical element.[31]

MENDELEEV RETRIEVED − AND MUCH MORE

We have underlined the central role of compounds as providers of order and similarity relationships for the elements in Mendeleev's approach to the periodic system. Using this argument, we analyzed the results of chemical similarity of chemical elements through a sample of their known binary compounds in the early years of the 21st century. The results show that several of the well-known similarities of chemical elements are recovered through this method based on the composition of compounds.

Regarding the similarities obtained, and contrary to the general message of current textbooks,[30] *resemblances are not always vertical on the periodic table*. Besides the well-known vertical similarities of the alkali metals, halogens, aluminum-group and copper-group, horizontal resemblances were detected such as those of 4th-row platinum metals, lanthanoids, actinoids. To which mixed similarities can be added, e.g. lanthanoids and scandium and yttrium (rare earths); and actinoids with some transition metals. Interestingly, Mendeleev had noted as early as 1869 that "in certain parts of the system the similarity between members of the horizontal rows will have to be considered, but in other parts, the similarity between members of the vertical columns."[31] Hence, chemically speaking, similar elements are close to each other on the table but vertical proximity is not the only and most relevant similarity scheme.

The results here discussed agree with the classification of elements presented in specialized chemical books such as the classic *Chemistry of the Elements*,[32] where the classification is the basis for the distribution of the material presented in the book. It is worth noting that the same pedagogical aim rooted on a chemical system was sought for by Meyer and Mendeleev when writing their respective chemistry textbooks.[4] This presentation of chemical knowledge is therefore expected from books rooted in chemical information, which contrasts with the current simplistic approaches of introductory chemistry textbooks, based on electronic resemblance of free atoms. We have also shown how quantum chemistry concepts can be used to make sense of the similarity results obtained through compounds.

It has been claimed that the motivation for developing a periodic system was to make sense of the large amount of information about compounds and their reactions that had been gathered by mid 19th century.[4,13] However, the exponential growth of chemical substances made it difficult for 19th century chemists to assess similarities through all known compounds, even if efforts of gathering chemical information in a systematic fashion had begun during that time as evidenced in the different editions of the famous *Gmelins Handbuch der anorganischen Chemie* and *Beilsteins Handbuch der organischen Chemie*, which by 1869 included more than 11 thousand substances.[13] These handbooks plus the *Patent Chemistry Database* are now available in digital form in *Reaxys*™, a large electronic database that is updated on regular basis from material published in more than 15,000 scientific journals and patents. Another database gathering chemical information is *SciFinder*™. Therefore, the method here presented can be computationally applied to those databases in order to shed light on the similarity structure of the chemical space at a particular historical period of the available chemical space.

19th century approaches to similarity, the growth of

[19] Curium, berkelium, einstenium, americium, californium, and actinium.

the chemical space, combined with the emerging atomistic ontology at the end of the century and the advent of quantum mechanics at the beginning of the 20th century, led to analyze and assess similarities among chemical elements through resemblances on the energetic distributions of valence shell electrons.[33] This is the root of the current over-emphasized textbook introduction to the periodic system through electronic configurations of free atoms.[34] However, as some authors have remarked,[34,35] these configurations are rather dissimilar to those of the bounded atoms present in compounds, which are the actual relevant species for chemistry.

ELECTRONIC AND COMPOUNDS: TWO APPROACHES TO A SYSTEM OF THE CHEMICAL ELEMENTS

The approach discussed in the current paper therefore constitutes an alternative way to introducing the periodic system to students, with more chemical "flavor" than what has become the traditional electronic approach.[15] "Compound" is the fundamental concept of chemistry that is part of the bulk level by Nelson for describing chemistry.[38] By bulk level, we mean the approach to chemical education based on compounds and chemical reactions, often performed at chemistry laboratories with bulk matter or material that consists of large numbers of atoms, molecules, or ions. We have indeed currently two options to approach the study of the periodic system and its teaching.

The first approach, the *electronic approach*, now largely in use in chemical education and practice, requires possible molecular ensembles as input to calculate properties. However, current quantum chemical approaches are not able to systematically treat chemical species with the same levels of theoretical accuracy to end up with properties that can be compared leading to classifications of elements. This poses an interesting and worthwhile challenge to quantum chemistry which is computationally difficult, for the number of compounds populating the chemical space is extremely large:[20] even the simplest quantum chemical methods would require too much time to finally end up with values for various material properties.

To make matters worse, in teaching, the electronic approach cannot be introduced as here described because the periodic system is normally presented in the first year of chemistry studies, where quantum chemistry concepts are still to be developed and taught. One could, however, approximate the approach using quantum chemical results of isolated atoms in their ground state energy, which brings back the problem of a "fantasy chemistry"[21] far from the chemistry of bonded atoms forming compounds with electronic configurations different from those of isolated atoms.

The second approach to the study of the system and its teaching is the *compound approach*, discussed in this paper, which requires managing the rapidly growing chemical space, currently recorded in electronic databases. Here, obtaining similarity classes of chemical elements requires formulas of the compounds reported and the application of classification algorithms, whose complexity, in general, does not depend on the size of the compounds nor on their number of elements. It is, in this sense, independent from both size and electronic theories, which is its advantage compared to the electronic approach.

In teaching, the compound approach would require knowledge on how to operate on the chemical space, which, as noted by Schummer,[37] requires data analysis techniques to make sense of the information stored in databases. One can hope that sooner or later, chemical databases will include the possibility of running data analysis studies on the cloud in such a way that clicking on "give me the system of elements" button, one can retrieve the shape of the system constructed with the available chemical knowledge.[22]

FUTURE PERSPECTIVES: MAPPING SIMILARITIES AND CREATING CHEMICAL SYSTEMS

For now, a more realistic approach to the systems from the compounds is through random samples of the space, easy to handle in personal computers.[23] Another option is to run projects with enough computational facilities, able to store the complete chemical space at a given time and to process its information. This approach is currently followed in our research group, whose ini-

[20] Up to March 2019, for example Reaxys reported 31,134,633 chemical species.

[21] Expression taken from Peter Schuster at the Mathematics in Chemistry Meeting (Leipzig 2016), when objecting classification results of chemical elements not meeting well-known similarities.

[22] Actually, the technicalities of the "button" should read "give me the system of elements according to the available chemical space for period p (a range of years) using the merging methods A, B, ..." A very recent instance of how data analysis techniques applied to chemical information are making their path in current chemistry is the publication of the first chemistry book written entirely by a machine (reference 38). It contains a survey on lithium-ion batteries based on 150 papers published between 2016 and 2018.

[23] A similar approach was followed by Schummer when analyzing the growth of chemical compounds at the end of the 1990s. Details in reference 10.

tial results have analyzed the temporal evolution of the growth of the chemical space since 1800 up to 2015. A third option to apply the compound approach is through classification of the compounds of the space in such a manner that one can select representative compounds of the classes to run similarity studies. This approach requires further research on the chemical space and on its mathematics. Further work to develop appropriate tools in this direction is currently carried out in our research group.

Even if we are advocating for a more data-driven approach to the system of elements through their compounds, it is not free of subtleties. It brings to the surface another fundamental question of chemistry. What is a chemical compound? Strikingly, as noted by several authors,[39,40] even its fundamental role in the edifice of chemistry, there is no consensus on what this concept is.

At first glance, it looks like the compound approach to chemical similarity here discussed cannot stand the test of time, for it relies on compounds, which are especially scarce for the heavy elements. Moreover, for these elements the few compounds that are obtained are synthesized in a one-atom-at-a-time fashion, which is different from the bulk process of the traditional chemistry.[15,41] This sparks not only a clash of chemical traditions, but also the mixture of two different ontological levels for types of compounds. By contrast, the computational methods that operate on chemical databases overcome these problems, for it is actually based, beyond compounds, on their mathematical generality, i.e. their composition and stoichiometry, not on their mode of existence or acquisition. Both composition and stoichiometry can be extracted from either bulk or atomic aggregate compounds; it does not matter whether the substances have been synthetized through wet-lab techniques, or in a one-atom-at-a-time fashion, or even estimated through quantum chemical approaches.[42]

We have shown that a sample of the current chemical space is the natural source of information about similarity among chemical elements. These similarities, when combined with the traditional order of elements by atomic number, provide what we see as the current structure of the periodic system. This methodology is nothing else than Mendeleev's methodology applied to the current chemical space, now assisted by computational tools of data analysis. Applying the same pedagogical motivation that was the hallmark of Mendeleev has produced results that ought to be introduced in contemporary chemistry classrooms together with the electronic understanding of elements in order to bring chemistry back into the periodic system.

APPENDICES

Appendix 1: Similarity calculation

If N_x and N_y are the neighborhoods of elements x and y, respectively; the similarity $s(x, y)$ between x and y is calculated as $s(x, y) = |N_x \cap N_y|/|N_x \cup N_y|$, where $|X|$ represents the number of elements in the set X. Thus, $0 \leq s(x, y) \leq 1$ and values close to one indicate similar elements, whereas those close to zero, very dissimilar elements.

ACKNOWLEDGMENTS

The author is indebted to Andrés Bernal, Michael Gordin, Wilmer Leal, Eugenio J. Llanos, Geoff Rayner-Canham, Eric Scerri, Joachim Schummer, W. H. Eugen Schwarz, and José Luis Villaveces for their valuable discussions and ideas. Thomas Endler is thanked for designing the graphical abstract.

REFERENCES

1. J. W. van Spronsen, *The Periodic System of Chemical Elements: A History of the First Hundred Years*, Elsevier, **1969**.
2. J. A. R. Newlands, *Chem. News* **1864**, *10*, 59.
3. W. Odling, *Quart. J. Sci.* **1864**, *1*, 642.
4. M. D. Gordin in *Nature Engaged: Science in Practice from the Renaissance to the Present*, Ed. M. Biagioli, Palgrave Macmillan US; New York, **2012**, Chapter 3, pp. 59-82.
5. W. Leal, G. Restrepo, *Proc. R. Soc. A* **2019**, *475*: 20180581.
6. L. Bertalanffy, *General System Theory*, George Braziller, **1968**.
7. E. R. Scerri, *The Periodic Table, Its Story and Its Significance*, Oxford University Press, **2007**.
8. D. Mendeleev in *Mendeleev on the Periodic Law: Selected Writings, 1869-1905*, Ed. W. B. Jensen, Dover; New York, **2002**; Paper 13, pp. 253-314.
9. G. M. Keserű, T. Soós, C. O. Kappe, *Chem. Soc. Rev.* **2014**, *43*, 5387.
10. E. J. Llanos, W. Leal, D. H. Luu, J. Jost, P. F. Stadler, G. Restrepo, *P. Natl. Acad. Sci. USA* **2019**, *116*, 12660.
11. J. Schummer, *Scientometrics* **1997**, *39*, 107.
12. D. J. Solla Price, *Little Science, Big Science*, Columbia University Press, **1963**.
13. W. Leal, E. J. Llanos, P. F. Stadler, J. Jost, G. Restrepo, *ChemRxiv* **2019**, Preprint.
14. W. Leal, G. Restrepo, A. Bernal, *MATCH Commun. Math. Comput. Chem.* **2012**, *68*, 417.

15. G. Restrepo in *Elements Old and New: Discoveries, Developments, Challenges, and Environmental Implications*, Eds.: M. A. Benvenuto, T. Williamson, ACS Symposium Series; American Chemical Society; Washington, DC, **2017**; Chapter 5, pp. 95-110.
16. D. I. Mendeleev, *Osnovy khimii (The Principles of Chemistry*), Obshchestvennaia pol'za: St. Petersburg, Part II, **1871** (folding table)
17. D. Mendeleev, *Ann. Chem. Pharm*. **1871**, *8*, 133-229 (Supplementband)
18. D. Mendeleev, *J. Chem. Soc*. **1889**, *55*, 634.
19. G. Rayner-Canham in *Mendeleev to Oganesson: A Multidisciplinary Perspective on the Periodic Table*, Eds: E.Scerri and G. Restrepo, Oxford University Press; New York, **2018**; Chapter 10, pp. 195-205.
20. G. T. Seaborg, *Chem. Eng. News* **1945**, *23*, 2190.
21. J-B. Liu, G. P. Chen, W. Huang, D. L. Clark, W. H. E. Schwarz, J. Li, *Dalton Trans*. **2017**, *46*, 2542.
22. K. H. Wedepohl, *Naturwiss*. **1996**, *83*, 165.
23. J. E. Huheey, C. L. Huheey, *J. Chem. Educ*. **1972**, *49*, 227.
24. P. Pyykkö, J-P. Desclaux, *Acc. Chem. Res*. **1979**, *12*, 276.
25. P. Pyykkö, *Chem. Rev*. **2012**, *112*, 371.
26. B. Fricke, W. Greiner, J. T. Waber, *Theoret. Chim. Acta* **1971**, *21*, 235.
27. URL: www.iupac.org/project/2015-039-2-200 (Accessed 23 April 2019)
28. W. H. E. Schwarz, R. L. Rich, *J. Chem. Educ*. **2010**, *87*, 435.
29. S. K. Cary, M. Vasiliu, R. E. Baumbach, J. T. Stritzinger, T. D. Green, K. Diefenbach, J. N. Cross, K. L. Knappenberger, G. Liu, M. A. Silver, A. E. DePrince, M. J. Polinski, S. M. Van Cleve, J. H. House, N. Kikugawa, A. Gallagher, A. A. Arico, D. A. Dixon, T. E. Albrecht-Schmitt, *Nat. Commun*. **2015**, *6*, 6827.
30. R. Chang, *General Chemistry*, McGraw-Hill, **2008**.
31. D. Mendeleev in *Mendeleev on the Periodic Law: Selected Writings, 1869-1905*, Ed. W. B. Jensen, Dover; New York, **2002**; Paper 2, pp. 18-37.
32. N. N. Greenwood, A. Earnshaw, *Chemistry of the Elements*; 2nd ed., Elsevier, **2005**.
33. G. Restrepo in *What is a chemical element? A collection of essays by chemists, philosophers, historians and educators*, eds: E. Scerri and E. Ghibaudi, Oxford University Press: New York, **2019** (forthcoming).
34. C. K. Jørgensen, *Angew. Chem. Int. Ed*. **1973**, *12*, 12.
35. S-G. Wang, W. H. E. Schwarz, *Angew. Chem. Int. Ed*. **2009**, *48*, 3404.
36. P. G. Nelson, *Chem. Educ. Res. Pract.* **2002**, *3*, 215.
37. J. Schummer, *Educación Química*, **1999**, *10*, 92.
38. B. Writer, *Lithium-Ion Batteries*, Springer, **2019**.
39. J. Schummer, *HYLE Int. J. Phil. Chem*. **1998**, *4*, 129.
40. J. van Brakel in *Handbook of the Philosophy of Science*, Volume 6: Philosophy of Chemistry, Eds: R. F. Hendry, P. Needham, A. I. Woody, Elsevier; Amsterdam, **2012**; pp. 190-229.
41. M. Schädel, *Angew. Chem. Int. Ed*. **2006**, *45*, 368.
42. P. Pyykkö, *Phys. Chem. Chem. Phys*. **2011**, *13*, 161.