

Hierarchical periodic system of chemical compositions of objects of any nature and its relation to the Periodic system of chemical elements

T.G. Petrov

To Vasiliy Ilyich Lebedev's memory

RHA method (*R*-rank formula, *H*-entropy, *A*-anentropy) developed for the unique description - indexing – of chemical compositions, and method of ordering obtained indices is briefly described. Lexicographic ordering of rank formulas is based on the Periodic system of elements as the alphabet. This provides for the formation of an hierarchic periodic system of compositions as a universal classification of chemical composition of natural and artificial objects, as well as the objects to be discovered or synthesized. A particular role of entropy characteristics in the study of compositional changes is emphasized. A comparison of the original and generated systems is made by 17 positions. The main features of utilizing the hierarchic periodic system of compositions are listed. Refs. 16 names.

Key words: Periodic system of elements, hierarchic periodic system of compositions, alphabets, rank formula, information entropy, anentropy, separation, mixing.

Chemical composition is a necessary material basis for formation of a real object with its structure and the entire set of properties. An indefinitely large variety of chemical compositions of geological, biological, technical, and space objects, complexity of compositions, their multicomponent nature, and in most cases the lack of natural boundaries between the compositions of objects that have different names, has moved to find a way of indexing, unique description and ordering of data on the chemical composition of objects of any nature. Such a method named *RHA* is published [1] as the first option for the general classification of geochemical systems: minerals, rocks, ores, water, gases, as well as artificial model objects.

Later on, this method was related to the adjacent areas of knowledge: information theory, thermodynamics, combinatorics, linguistics, semiotics; properties of the resulting system, its capabilities and expanding areas of application were studied. Universality of the proposed language for unique description and ordering of chemical compositions of objects of any nature is shown [2,3]; a large number of opportunities to use it in different branches of geology [4,5] and other fields of knowledge not related to chemistry and geology [2] is revealed; a way to fold large amounts of information in *RHA* form is found, the possibilities to use it in description of

compositional changes are shown [3,6], an education book for students studying geology is created [3].

Purpose of the paper: briefly describe *RHA* information language and show the similarities and differences between the well-known ordering system of *individual chemical elements*, the Periodic system of elements, and the ordering system of *sets of different elements* (mixtures or compounds), namely the hierarchic periodic system of *compositions*, developed on the basis of *RHA* information language.

Let us give a brief description of *RHA* language and the ways of chemical compositions ordering.

The first parameter *R*, rank formula, is a sequence of event symbols of chemical elements by decrease of their atomic contents¹ p_i in the analysis. The use of *atomic* contents (measured in unit fractions) enables a uniform description of matter compositions at any level of its organization (atoms, ions, molecules, molecule mixtures, amorphous matters). Atomic contents correspond to the occurrence frequency (or probability) of the *i*-th sort atom in the analysis.

Equal sign is put between the element characters, if the differences between their contents exist, but do not exceed 15 relative percent, that is at $p_n/p_{n+1} < 1.15$. For a strict equality of contents, characters are written by their order in the Periodic system of elements.

For different lengths of element lists in analyses, for comparability of further calculations, rank formula length *n* is standardized, while *n* is a **measure of detailed study** of the analyzes sampling. The lower the detail, the more neglected information, compositions are less distinguishable. The higher the detail, the greater revealed variety, but both original data and *R*, *H* and *A* parameters are less reliable due to the growing relative errors of minor elements determination.

The second parameter *H* is C. Shannon information entropy, which is calculated for the chosen *n* by $H = -\sum p_i \ln p_i$ formula, where p_i is atomic content. Keeping in mind that p_i is measured in unit fractions, $\sum p_i = 1$. Maximum entropy value is realized at equiprobable p_i distribution. Information entropy is a **complexity measure of chemical composition** [7,2,3], or the measure of *closeness* of all contents of chemical elements in the composition to each other. According to [8, 9] *H* is an analog of thermodynamic entropy of mixing. (There are other interpretations [10]). Normalization to 0÷1 interval is made by $En = H / \ln n$ formula. This eliminates the coincidence of standard symbols (*H*) for information entropy and hydrogen.

¹ Atomic element contents are the result of normalization of the atomic quantities to unity. The latter are the result of division of elements mass% by their atomic masses.

The third parameter A is anentropy proposed by us [1]; in the simplest representation $A = -\sum \ln p_i/n$, where " $-\ln p_i$ " is contribution to the information made by the i -th symbol of the message. Anentropy is the "**purity**" **measure of composition**, or assessment of the degree of composition closeness to that of the ideally pure single-element substance. It is also an analog of chemical affinity for mixing [9, 10], that is the degree of *composition nonequilibrium* in relation to the equiprobable distribution of elements (not realized in nature.) Normalization to 0÷1 interval is described in [2,3]. Normalized anentropy value acquires An symbol.

An essential requirement to the initial analyses is the *completeness at the given detail*, namely, the analysis must contain all elements, the contents of which exceed that of the n -th element.

A single set of R, En and An values (or, for short, RHA) is a unique alphanumeric designation of composition, its *index*, or code.

Ordering of rank formulas is made as in alphabetical dictionaries. For this, rank formula is taken as "word", in which symbols of chemical elements play the role of "letters". Thus, $OMg=S$ "word" contains three "letters": O, Mg, S. Symbol sequence from the Mendeleev's Periodic system of is taken for an *alphabet* of rank formulas ordering. This method, properly, generates the *hierarchic periodic system of chemical compositions*. Consequently, there is a possibility of unambiguous ordering of ranking formulas, in particular, linear (vertical), as shown in Tab. 1.

First, grouping of rank formulas after their first symbol was made, and then sorting of the groups in accordance with the order of these symbols in the Periodic system of elements. Respectively, in Table 1 in the first rank we have H, O, F, S, Cu, the numbers of which are 1,8,9,16,29. Then, within each group with the same first element (for example, oxygen), record lines are ordered after the second one (Mg, Si, Ca, their numbers are 12,14,20), and so on. Within the groups of *the same* rank formulas, ordering of RHA -indices is made by H decrease, which corresponds to the normal direction of entropy change in separation processes [12]. At the same H , record lines are arranged by A anentropy increase (for the same reasons). All this ensures uniqueness and linearity of the sequence of composition RHA -indices jointly with the emergence of a hierarchical table structure.

Table 1.

Alphabetically ordered (by order of elements in the Periodic system of elements) sampling of *RHA*-indices from the "Chemical compositions of natural objects" database [11]

Rank formula										<i>En</i>	<i>An</i>	Object
H	He	O	C	Ne	N	Mg=	Si=	Fe	S	0.135	0,837	Sun
H	O	C	N	Ca=	P	K=	S	Na	Cl	0.428	0,434	human body
H	O	N	Cl	Si	Li	B=	S	C	Ca	0.278	0,980	water, geyser, Kamchatka
O	C	Ca	Mg	Fe	Si	P	Al	Mn	K	0.561	0,210	carbonatite, Sallanlatva
O	Mg	Si	Fe	Al	Ca	Na	K=	Cr	Ti	0.542	0,301	Mars
O	Mg	Si	Fe	Al	Ca	Na	Cr	K	Ti	0.511	0,305	Earth mantle+crust
O	Si	H	Al	C=	Ca=	Mg=	Fe=	K	Na	0.578	0.166	Quaternary clay
O	Si	H=	Al	Fe	K	Mg	C=	Ca	Ti	0.361	0,401	sandstones, Kazakhstan
O	Si	Na	Mg	Al=	Ca	Fe	Mn	W	Ti	0.286	0,804	quartz, Transbaikalia
O	Si	Mg	Al	Ca=	Fe	Cr	Ti	Mn	Na	0.554	0,274	pyrope Urals
O	Si	Mg	Fe	Al	Ca	Na	Mn	S	K	0.567	0,193	meteorite Zhmerinka
O	Si	Al	Na	K	H	Fe	Ca	Mg	Ti	0.488	0,247	granite, average of 2,485 an.
O	Si	Al	Ca=	Fe	Mg	Ti	Na	K	Mn	0.552	0,236	basalt, Moon
O	Si=	Ca	C	H=	Fe	P	F=	K	Al	0.617	0,138	carbonatite, Malawi
O	Ca=	C	Fe	Mg	P	Si	Al	Sr	Na	0.519	0,278	carbonatite, Kovdor
O	Ca	Fe	P	Mg	Si	Al	Na	Mn	Ti	0.569	0,268	phoscorite, Kovdor
F	Ca	Ba=	Ti	Zr	O	Be=	Al	Bi	Mn	0.281	0,962	fluorite, Transbaikalia
S	Fe	As	Sb	Zn	Pb	Co=	Ni	Bi	Se	0,282	0.967	pyrite, Sibay
Cu	Sn	As	Fe	Sb	Pb	Ni=	Ag	Bi=	Co	0,069	0.526	bronze, knife, Alexeevka

Horizontal lines - dividers – are drawn between the elements differing vertically. They show a hierarchic structure of the table. At the same time, the possibility of considering the rank formula itself as a *hierarchically organized list of all taxa* - enumerated classes including this composition representation, becomes obvious. For example, pyrope composition enters the widest *class of the first order* of "oxygenic" substances *R1*: "O taxon". The number of such classes is equal to the number of elements found in nature (let this number be 83). Further, the same pyrope enters a narrower class - *class of the second order R2*: "OSi taxon". There will be 83*82 such classes; then *R3*class: OSiMg (83*82*81 classes) and so on with increasing rank formula detail (*n*) and a corresponding increase in information about its composition registered in the formula. Such table structure causes a group arrangement of slightly differing analyses, as in the case of analyses similarity, to the automatic arrangement of their *RHA*-indices adjacently, one below the other.

Table periodicity is manifested in arrangement of similar composition vertically. There are two types of periodicity. One of them is directly associated with the System of *elements* as the alphabet used to form the System of *compositions*. In it, rank formulas starting with affined elements Li-Na-K-Rb-Cs are dissociated by compositions, the rank formulas of which start with

other elements. For example, between the rank formulas beginning with lithium and sodium, there are rank formulas that begin with Be, B, C, N, O, F and Ne elements. The second type of periodicity is common only to the System of compositions and is due to the fact that the permutation of two adjacent elements in the rank formulas usually preserves the chemical proximity of substances. In this context, let us draw attention to important differences between the alphabets of natural languages and the Periodic system of elements taken as an alphabet. In natural languages, *individual letters*, in general, have no *conceptual meaning*, therefore the words that have the same beginning, as a rule, are not related (*crab, crack, cradle, crane, crayon, crater*), therefore, "crab" and "crane" are not conceptual varieties of some, common to them, "CRA". In contrast to the previous case, in *RHA* information language, rank formulas with *similar beginnings* relate to *similar objects*. Thus, rank formulas starting with OSiAlH, OSiAlNa, OSiAlK relate to H-, Na-, K-varieties of granitoids, and OCCaMg, OCCaP, OCCaSi - to magnesia, phosphorous, silica varieties of carbonate rocks. The same situation arises in case of permutations of adjacent letters. Thus, there is no conceptual association between the words in pairs: *crap-carp* or *carve-crave*, but the pair of beginnings of rank formulas of chemical compositions, for example, OCCa-OCaC carbonatites, there is such association. Compositions of these related objects are spaced apart in the "chemical alphabet" sequence by formulas starting with ON, OF, ONa, OMg, OCl, OK.

As for some conceptual properties of the table, it reflects the closeness of the Earth's composition (according to preliminary data) to that of Mars. Composition of the Zhmerinka meteorite shows a fairly high degree of similarity to the compositions of both Mars and Earth - they up to rank seven(!) differ only by a permutation in the second and third ranks: MgSi ... SiMg. This is also an example of periodicity of the System of compositions (in the System of elements, aluminum is between magnesium and silicon). Rank formula of basalt from the Moon reflects the known increased titanium content.

Table shows the entropy characteristics of compositions. One can see that the Sun and a bronze knife from an early site have minimal entropy in Table 1. This is due to the fact that the first element - hydrogen and copper, respectively predominates in their composition. Close *En* values of water, quartz, fluorite (CaF₂), and pyrite (FeS₂) are associated with the same set of stoichiometric indices in their chemical formulas. Relatively low entropy values are determined by simplicity of their chemical formulas. Generally, the closer the values of entropy characteristics, the greater the similarity of ranked distributions character.

Entropy characteristics are of special interest when discussing the genetic problems.

It has been shown that if the system is **separated** into two non-identical in composition parts, the information entropy of at least one of the resulting part is often lower than the entropy of

the original one; at **mixing**, entropy of the resulting system higher than that of at least one of the original ones [12]. These assertions have been proved as theorems and represent the theoretical explanations to the known facts. Namely: at separation, a **decrease of compositions complexity** usually takes place, **at mixing – an increase**. General instability of complex systems [13], for example, solutions that are separated into relatively simple parts under "favourable" conditions is widely known. Thus, a solution is separated into two parts: solution and crystal, which is almost always more simple. Similarly: a crystal of complex composition can break up into two phases, of which at least one will have a lower complexity than the original crystal.

Let us call evolution of compositions occurring with a monotone (increase and decrease for different elements) change in element contents *one-way* processes. Such processes in *HA (EnAn)* diagram are shown by the trajectories that may have entropy maximums and anentropy minimums, but not vice versa. The presence of entropy minimum (and anentropy maximum) on the trajectory is an evidence of process reversal [14]. In small trajectory segments, anentropy changes usually (but not always) inversely correlate with entropy changes. This is understandable. The more first components (which implies entropy decrease), the less the other (the latter contributes to anentropy increase). For more details see [3].

In general, for the given entropy, anentropy values of individual crystal compositions as products of separation at the atomic and molecular level, are almost always higher than anentropy of mixtures of different crystals (rocks), as the latter are the result of less effective selection processes taking place during crystallization-gravitational, hydrodynamic differentiation or any other.

Both entropy characteristics are effectively used to describe the evolution of natural objects composition [3,6], and there are no obstacles to monitor the engineering process products.

But let us go back to rank formulas as a structuring basis of chemical compositions ordering.

From the standpoint of combinatorial analysis, rank formulas can be considered as **arrangements**, that is permutations without repetition of n elements (mentioned above *detail character*) of their total number N (the number of different chemical elements). Successive application of the described rules of alphabetical ordering, in principle, enables to construct a *full table*, that is, a *complete full variety of chemical compositions of all objects existing in Space, in technology, as well as in theoretical constructs*, and, in addition, to reserve *free cells* for all undiscovered substances².

The total number of cells of this table equals to the number of arrangements of N symbols taken n at a time, namely: $N! / (N - n)!$. If we assume $n = 10$ and $N = 83$ (lower estimate), we get

² And those which will never be found and created; and it will be interesting to find out why such substances cannot exist.

$\approx 10^{19}$ rank formulas. In practice, it is impossible to draw such a table, however it is unnecessary. It is enough to suppose its existence as the basis of order during collection, storage and retrieval of information on the chemical compositions in the material world. In natural languages having 30-40 letters, for the existing variations in word length and repetitions of letters in words, the full range of possible words is also combinatorially large, but it does not prevent using the set of words that already exists and is intensely replenished by combinatorial reserve.

Universal coverage described by table-system of all existing and possible compositions makes appropriate a comparison of the above Hierarchical periodic system of compositions and the Periodic system of elements that underlies its structure. Significant similarities between these two systems, as well as some of the differences are reflected in Tab. 2.

Table 2.

Comparing properties of the Periodic system of elements and the hierarchic periodic system of compositions

	Subject of comparison	System of Elements (E)	System of Compositions (C)
1	Organization objects in the System	Chemical elements, simple matters	Chemical elements, compounds, compositions, mixtures.
2	Physical content of the System (table) cell	Assembly of atoms having the given nucleus charge. Simple pure substance.	Assembly of chemical compositions having the given rank formula. Any substances.
3	Indistinguishable in a cell	Structural phase states of simple substances	Structural phase states of any substances
4	Distinguishable in a cell	Isotopes, the limits between them are natural, discrete.	Entropy characteristics of contents distribution in composition: natural limits between the distributions are absent.
5	Principle of distinguishable ordering within a cell	Increase in the atomic mass of isotopes.	Decrease in information entropy of element distribution. At equal H , increase of the elements distribution anentropy.
6	Setting (assigning) the detail of cell content consideration	Specified by requirements in the data placed into a cell.	Specified by requirements in the degree of detail (n) during compositions study.
7	Cells number in the system. Cardinality of the set of elements Me . Cardinality of the set of compositions Mc .	Me – determined by physics development level ($Me = 109?$, in nature 83?).	Mc – determined by Me value and chosen detail (n) of analytical materials consideration $Mc = Me!/(Me-n)!$.
8	Principle of cells ordering in the System	Linear – atomic charge value.	Linear vocabulary (lexicographic). Alphabet for rank formulas ordering - the System of elements.

9	Periodicity	Present	Present
10	Periodicity source	Similarity in outer atomic shell structure.	Similarity in outer atomic shell structure and element symbol permutations.
11	Form of periodicity manifestation	Table geometry – a combination of rows and columns with similar elements arrangement in a column.	Arrangement of similar compositions in groups, with other compositions between.
12	System hierarchic character	?	Present. Consequence: hierarchic structure of the whole System of compositions.
13	Form of hierarchic character manifestation	?	Each class of the n -th order enters a single class of the $(n-1)$ -th order.
14	Mean atomic mass of the System objects in their ordered sequence	Increases statistically	Increases statistically
15	Occurrence of the System objects in nature	Decreases periodically in a sequence of chemical elements.	Decreases periodically in a sequence of rank formulas.
16	Reflection of the System objects genesis in a cell content	In atom structure complexity.	In complexity and purity of chemical composition.
17	Visibility of the System – table	Exists.	Partial. For the complete System is absent due to its sizes. Exists for fragments and samplings, see item 7.

Key features of the hierarchic periodic System of Composition and its advantages when dealing with material composition of geological objects are as follows:

1. *Ordering information* on the chemical composition of objects of any nature, whether they are gases, liquids or solid bodies, natural and artificial.
2. Establishing data banks with the ordered storage and *factographic search* for composition analogs.
3. *Removing the barriers* between professionals dealing with chemical compositions of object of different types.
4. *Ease of perception* of large amounts of chemical analytical data, due to their clear algorithmic structuring (especially manifested during the first acquaintance with analytical materials).
5. *Objects identification after their composition* – in the presence of a complete enough data bank
6. *Assistance in improving the terminology* and classification of objects, in particular, through their discovery as standing in *RHA* table "out of order", among "strangers".
7. *Assessment of completeness and variety of data samplings*, as well as of originality and triviality of *individual compositions* against a particular sampling or data bank.
8. *Compressing rank formula arrays* and representing information on the elements distribution in ranks in the sampling as a generalized rank formula.
9. Quantitative and graphic *representation* of change in multicomponent chemical compositions as the *separation and mixing* processes taking place in natural and laboratory conditions. Such a representation of processes enabling to *trace in a single diagram with a fixed coordinate system the trajectory of the processes of change of chemical (and other) compositions of any objects* is inaccessible for other methods known to the author.
10. A sequence of rank formulas of theoretical mineral compositions, as an *alphabet of mineral compositions*, can be the basis for *RHA-classification of mineral compositions of crystalline rocks* [1,3,15], which, subject to a small number of rules, in principle will emerge as a commonly understood *self-organizing system* developed as required by different researchers.

For more details on the proven capabilities of *RHA* method see [6]. Materials related to the issue are posted in the Internet at <http://www.geology.pu.ru/> in "Scientific research" *RHA* method section.

PETROS-2 software providing the work according to *RHA* method and solving a number of other tasks was compiled by S.V. Moshkin [16].

Acknowledgment.

The author thanks S.V.Moshkin for long-term work on compilation and support of PETROS software, as well as V.V. Gordienko, V.N. Dech, N.I. Krasnova, and S.V. Chebanov for a number of advices that helped to improve the text.

Summary

Petrov T.G. Hierarchical periodic system chemical compounds of objects of any nature and its communication with Periodic System of Chemical Elements.

The information language RHA intended for indexing and ordering of data on chemical compounds of objects of any nature is briefly described. It includes one discrete characteristic of elements distribution in structure - rank formula (R) and two continuous - information entropy (H) and anentropy (A). Ordering rank formulas is made by a lexicographic principle with use as the alphabet of Periodic System of Chemical Elements. The Hierarchical Periodic System of Chemical Compounds is as a result received. Their comparison is lead.

References

- Petrov T.G.* Substantiation of a variant of general classification of geochemical systems, Vestnik of Leningrad State University. 1971. No.18. Pp. 30-38.
- Petrov T.G.* Information language for description of compositions of polycomponent objects. // Scientific - Technical Information. Ser. 2. 2001. No. 3. Pp. 8-18.
- Petrov T.G., Farafonova O.I.* Information-component analysis. RHA method. (Educational guidance) SPb. 2005. 168 p.
- Petrov T.G.* Rank-entropy approach to description of compositions of geological objects and their changes (exemplified by geological cenology). // General and applied cenology. 2007. No.5. Pp. 27-33.
- Petrov T.G., Farafonova O.I., Sokolov P.B.*, Informational - entropic characteristics of composition of minerals and rocks as a reflection of crystallization process.// Zapiski VMO 2003. No.2, Pp. 33-40.
- Petrov T.G.* RHA method as a solution of systematization problem of analytical data on the material composition of geological objects. Otechestvennaya geologiya (National geology). 2008. No. 4. Pp. 98-105.
- Petrov T.G.* On the complexity measure of geochemical systems from the information theory position.// AS USSR Proceedings. 1970. Vol.19. No.4. Pp. 1094-1096.
- Moelwyn-Hughes E.A.* Physical chemistry. M., IIL. 1962. 1148 p.
- Prigogin I., Defay R.* Chemical thermodynamics. Novosibirsk. Nauka. 1966. 510 p.
- Ebeling V.* Structure Formation in Irreversible Processes. Introduction to the Theory of Dissipative Structures. M.-Izhevsk. 2004. 256 p .
- Petrov T.G., Knizel A.A., Usanova S.S.* Information retrieval system "Chemistry of natural objects"// In. Magmatic and metamorphic formations of Middle Asia and petrological criteria of ore content. Tashkent: AS branch. 1983. Pp.153-155.
- Shurubor Yu.V.* On one property of complexity measure of geochemical systems.// AS USSR Proceedings. 1972. Vol.205. No.2. Pp. 453-456.
- Prigogin I.* From existing to emerging. M. KomKniga. 2006. 296 p.
- Gordienko V.V., Petrov T.G.* Studying of rare-metal pegmatites using RHA language. // Zapiski VMO (RMS Proceedings). Part 110. Iss.5. 1981. Pp. 546-558
- Krasnova N., Petrov T., Balaganskaya E., et al.* Introduction to phoscorites: occurrence, composition, nomenclature and petrogenesis. Ed. F.Wall, A.A. Saitzev. Foskorites and carbonatites from the mantle to main: the key example of Kola alkaline province. Publ.Min.Soc. of GB et Irl. L. 2004. p. 45-74.
- Moshkin S.V., Shelemotov A.S., Petrov T.G., Krasnova N.I.* "PETROS-2" - a software complex for petrochemical data processing // In: Geochemistry of magmatic rocks. Kola Branch of RAS. Apatity. 2003. Pp.115-116.