Ordinal Explanation of the Periodic System of Chemical Elements

Eric R. Scerri

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Piotr Wojciechowski
The University of Texas at El Paso, piotr@utep.edu

Ronald R. Yager

Follow this and additional works at: https://scholarworks.utep.edu/cs_techrep

Part of the Computer Engineering Commons

Comments:


Recommended Citation
https://scholarworks.utep.edu/cs_techrep/426

This Article is brought to you for free and open access by the Computer Science at ScholarWorks@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.
ORDINAL EXPLANATION OF THE PERIODIC SYSTEM OF CHEMICAL ELEMENTS

ERIC R. SCERRI
Department of Chemistry, Bradley University
Peoria, IL 61625, USA
E-mail: scerri@bradley.edu

VLADIK KREINOVICH
Department of Computer Science, University of Texas at El Paso
El Paso, TX 79968, USA
E-mail: vladik@cs.utep.edu

PIOTR WOJCIECHOWSKI
Department of Mathematical Sciences, University of Texas at El Paso
El Paso, TX 79968, USA
E-mail: piotr@math.utep.edu

and

RONALD R. YAGER
Iona College, 715 North Avenue
New Rochelle, NY 10801-1893, USA
E-mail: yager@panix.com

Received
Revised

Textbooks often claim that quantum mechanics explained the periodic system: namely, the actual configuration of electronic orbits that is responsible for the element’s chemical properties can be described as the one that minimizes the total energy, and the energy of each configuration can be computed by using quantum mechanics. However, a careful analysis of this explanation reveals that, in addition to the basic equations of quantum mechanics, we need some heuristic rules that do not directly follow from quantum physics. One reason why additional heuristics are necessary is that the corresponding numerical equations are extremely difficult to solve, and as we move to atoms with larger and larger atomic numbers \( Z \), they become even more difficult. Moreover, as \( Z \) grows, we must take relativistic effects into consideration, and this means going from partial differential equations to even more mathematically difficult operator equations.

In this paper, we show that if instead of the (often impossible) numerical optimization, we consider the (available) ordinal information, we can then explain the observed periodic system.

Key words: Periodic table; Chemistry; Ordinal Information; Invariance.
1. Introduction

Periodic table. Chemists have known for a long time that some elements have similar chemical properties. The resulting organization of elements into groups became more and more comprehensive until in 1869, M. I. Mendeleev organized all the known chemical elements into a neat periodic table.

The original table had gaps which, Mendeleev predicted, correspond to yet undiscovered chemical elements. When these elements were actually discovered, the periodic table was universally recognized.

It is desirable to explain the periodic table based on quantum mechanics. The periodic table, as well as many other properties of atoms and molecules, remained a heuristic idea until the early 20th century, when quantum theory and later quantum mechanics appeared. As early as 1913, Bohr used his quantum theory to provide a reasonably successful theoretical explanation for the periodic system. This was followed by more accurate versions by himself and Stoner in 1922 and 1924 respectively. The emergence of quantum mechanics in 1925–26 rather interestingly did not provide any improved qualitative explanation. However quantitative predictions on the energies of many-electron atoms which had not been possible in Bohr’s old quantum theory now became available.

The quantum mechanical explanation is roughly as follows. An atom consists of a small nucleus surrounded by the dynamic cloud of electrons. The number of the electrons is equal to the atomic number \( Z \) of an atom. Schrödinger equation enables us to describe different possible electronic configurations, and to compute the energy of each configuration. In some of these configurations, the energy is higher; in other configurations, the energy is lower. If an atom is in one of the states with the non-smallest energy, then, if left alone, it can (and will) change into the state with smaller energy, emitting photon on the way. On the other hand, if an atom is in the state with the lowest possible energy, it cannot change this state without an external force. Therefore, stable states of the atoms correspond to the lowest possible energy levels of their electron configurations.

Broadly speaking, chemical properties of an atom, i.e., properties related to its interaction with other atoms, are determined by its outer electrons. Normally, we are interested in the chemical properties of atoms in stable states (excited atoms, e.g., atoms irradiated by a mighty laser beam, sometimes enter into useful non-standard chemical reactions, but this is a different story). So, to find the chemical properties of an atom with a given atomic number (and thus, to describe its place in the periodic table), we must:

- solve the corresponding Schrödinger equation,
- find the state with the lowest energy,
- and then, crudely speaking, describe the outer electrons in the resulting stable (minimum-energy) configuration.
Ideally, we should be able to explain the periodic table by using explicit solutions of Schrödinger equations. However, the explicit solution is only known for the simplest atom of H, with one electron. Therefore, usually, we use the known H solution to make predictions about more complicated atoms.

Simplified quantum explanation only works for 18 first elements. In H, possible states of an electron can be characterized by four numbers called quantum numbers:

- the principal quantum number \( n \) can take any positive integer value \( n = 1, 2, \ldots \);  
- the orbital (azimuthal) quantum number \( l \) can take integer values from 0 to \( n - 1 \);  
- the magnetic quantum number \( m \) (also denoted by \( m_l \)) can take integer values from \(-l\) to \(l\);  
- finally, the spin \( \sigma \) (also denoted by \( m_s \)) can take two possible values: \(-1/2\) and \(+1/2\).

For the hydrogen atom, the energy of a state depends only on \( n \); energy monotonically grows with \( n \). Thus, the lowest possible energy is attained for the states with \( n = 1 \), the next largest for \( n = 2 \), etc.

We can use the same description to approximately describe electrons in other atoms. At first glance, it may seem like the smallest possible energy of a multi-electron atom is attained when all the electrons are in the same lowest-energy state, with \( n = 1 \). However, it is known that no two electrons can be in the same state (this is called Pauli principle). There are only two states that correspond to \( n = 1 \): the states for which \( l = 0 \), \( m = 0 \), and \( \sigma = \pm1/2 \). Thus, at most two electrons can be at this level; the third and following electrons must go to the next energy level that corresponds to \( n = 2 \), etc. The larger \( n \), the further the electron from the nucleus. Thus, the outer electrons are exactly the electrons from the highest possible level.

One can easily check that at each level \( n \), there are exactly \( 2n^2 \) different elements. Thus, in this simplified description, chemical elements have the following properties:

- First, we would have hydrogen, then has 1 electron in level \( n = 1 \) (where two electrons can be placed). So, if a H atom interacts with some other atom, hydrogen can easily accept an additional electron and place it in this level. It cannot as easily accept two or more electrons, because there is no space for them in this level. Alternatively, it can give away an electron.

- Second, we have helium (He), with two electrons on the most energetically stable level. There is no space left on this level, the shell is full, so He is, basically, chemically neutral.
• Third, comes Li that has exactly one electron on its outer orbit \((n = 2)\), and it can easily give away this extra electron.

• Li starts the sequence of 8 elements in which the second layer is gradually filled until we reach another neutral element: neon Ne \((Z = 10)\).

• After Ne, we start filling the third layer \((n = 3)\). In our simplified description, we expect to see \(2n^2 = 18\) elements on this level, but in reality, electrons add smoothly to the third level only until we reach Ar \((Z = 18)\). Then, for K \((Z = 19)\), instead of further filling the third level, electrons start filling level \(n = 4\), and only returns to the third level later.

Bohr’s heuristic rule and how it explains the periodic table. This difference between the electronic configurations that correspond to our simplified model, and the observed configurations means that for elements with \(Z \geq 19\), the hydrogen model, in which the energy of an electron depends only on the principle quantum number \(n\), is no longer applicable, and the energy must depend on other quantum numbers as well.

Due to fundamental symmetry reasons, this energy should not depend on the magnetic quantum number \(m\) or on the spin \(\sigma\) (e.g., states with different values of \(\sigma\) can be obtained from each other by a symmetry \(\vec{r} \rightarrow -\vec{r}\) that should not change the energy). Thus, the energy \(E\) of an electron can only depend on the quantum numbers \(n\) and \(l\): \(E = E(n, l)\).

To explain the periodic system, Bohr proposed the following empirical rule 3:

• the energy of a state increase with \(n + l\);

• for two states with the same value of \(n + l\), the one with larger \(n\) has the larger energy.

This rule of Bohr’s is also called the Aufbau rule, or the Madelung rule, after the researcher who popularized it in his monograph 27.

This two-part rule explains in what order states with different values of \(n\) and \(l\) are filled. Let us describe this order. Electrons with the same values of \(n\) and \(l\) form an orbital; orbitals are denoted by a number-letter combination, e.g., \(1s, 2p\), in which the number is \(n\), and the letter describes \(l\): \(s\) stands for \(0\), \(p\) for \(l = 1\), \(d\) for \(l = 2\), and \(f\) for \(l = 3\).

We start with the smallest possible value \(n + l = 1\), which corresponds to \(n = 1\) and \(l = 0\), i.e., to the orbital \(1s\). Then, we have the only orbital \(2s\) for which \(n+l = 2\) (since \(l < n\), this case \(n = 2\) and \(l = 0\) is the only possibility of obtaining 2 as a sum of \(n\) and \(l\)). Next, come two orbitals \(3s, 2p\) with \(n + l = 3\). According to the second part of Bohr’s rule, the orbital \(2p\) has a smaller energy, and is, therefore, filled first. For \(n + l = 4\), we get the sequence \(3p\), then \(4s\). As a result, the \(4s\) orbital \(4s\) starts filling before \(3d\), as observed experimentally 28.

As a result, we get the following order in which orbitals are filled:

\[
1s < 2s < 2p < 3s < 3p < 4s < 3d < 4p < 5s < 4d < 5p < 6s < 4f < 5d < \ldots
\]
This order, basically, explains the periodic table (with a few exceptions).

**How can we explain Bohr’s rule?** Using Bohr’s rule, we can explain the periodic table. The natural question is: How can we explain Bohr’s rule?

There have been several attempts to explain Bohr’s rule. All these attempts used numbers to explain this rule:

**First attempt to explain Bohr’s rule: by using approximate (heuristic) solutions to the original equations.** In Fermi and Abrahamson, it is shown that the first part of Bohr’s rule can be approximately explained by the Thomas-Fermi continuum (“liquid drop”) model of an atom (see also Klechkovskii, Latter, and Landau and Lifschitz). However, this explanation is very approximate, and for a fixed value of \( n + l \), it does not predict the correct order of orbitals (see, e.g., Latter).

The use of a more precise approximation (due to Hartree and Fock) corrected some wrong predictions, but, on the other hand, led to new problems with the order (see, e.g., Herman and Skillman).

**Second attempt to explain Bohr’s rule: by using new (heuristic) equations.** In Demkov and Ostrovsky, Bohr’s rule is explained by using a heuristic potential field

\[
V(r) = -\frac{2v}{r \cdot R \cdot (r + R)^2}
\]

for some constants \( v \) and \( R \).

This potential function has a non-physical asymptotics \( V(r) \sim r^{-3} \) when \( r \to \infty \) and several other non-physical properties; to avoid this non-non-physicality, a modification of this potential was proposed in Kitagawa and Barut. The resulting explanations show that Bohr’s rule is, in principle, consistent with quantum mechanics, but it leaves us with the necessity to explain where this particular potential came from.

To explain the potential function, with its specific numerical values, is even harder than to explain the original Bohr’s ordering of orbitals.

**Third attempt to explain Bohr’s rule: by using symmetry groups.** Kitagawa and Barut have also shown that both the original and the modified Demkov-Ostrovsky equations are invariant with respect to a transformation group that is larger than the natural group of geometric symmetries.

This fact relates the differential approach with an alternative approach in which, to describe the periodic system, instead of a potential (i.e., a differential equation), researchers fix a symmetry group, and analyze arrangements that correspond to this particular group. This has been done for several different symmetry groups:

- for a sequence \( SU(2) \times SO(4, 2) \supset SU(2) \times SO(4) \supset SU(2) \times SO(3) \supset SU(2) \), by Rumer, Fet, Konopelchenko, et al., and by Kibler et al.;
- for a sequence \( SO(4, 2) \supset SO(3, 2) \supset SO(3) \times SO(2) \), by Barut, Novaro, and Odabasi.
This approach is very interesting and very promising, and it is in line with the general idea of symmetry groups as one of the main tools of modern theoretical physics. However, there are two problems with this approach:

- First, several different symmetry groups are possible, and different groups lead to different filling orders. In Odabasi, it is shown that the particular symmetry groups that lead to the observed periodic system are consistent not only with the *ordinal* information (in which order orbitals are filled), but also with the *numerical* data (e.g., about ionization potentials). However, the need to experimentally choose the symmetry groups defeats the original purpose:
  - we wanted to get a *fundamental* explanation of the periodic table, that is based not on empirical evidence, but on first principles;
  - instead, we get an explanation based on symmetry groups, which, by themselves, have to be chosen *empirically*.

- Second, a group explanation may not be sufficient to explain the periodic table: even when the symmetry group is fixed, and we consider a reasonable potential that is invariant with respect to this group, we may get an order of filling the orbitals that is somewhat different from what we actually see; see, e.g., Novaro.

**Fourth attempt to explain Bohr's rule:** by solving the corresponding equations of quantum mechanics. There have also been many attempts to numerically solve the equations of quantum mechanics and thus, to explain Bohr’s rule. Such computations have actually been performed; see, e.g., Froese-Fischer. Several methods have been used that go beyond Hartree-Fock approximation, such as *configurations integration* (Shavitt), *cluster methods* (Sinanoglu), *many body perturbation theory* (Wilson), and many others (see, e.g., Wilson). The resulting approximate computations seem to confirm Bohr’s rule in most (but not all) cases. There are two main problems with these computations:

- First, these methods are based on *approximate* computations. To decide which of the two orbitals $a$ and $b$ will be filled first, i.e., whether for the energies of these orbitals, we have $E_a < E_b$ or $E_a > E_b$, we apply these approximate methods to find *estimates* $\hat{E}_a$ and $\hat{E}_b$ for these orbitals and then compare these estimates. Most approximate methods do not provide us with any error estimate, as a result, when $\hat{E}_a > \hat{E}_b$, we do not know:
  - whether actually $E_a > E_b$,
  - in reality $E_a > E_b$, and the reverse inequality is caused by computation errors (as we have already mentioned, this has actually happened with the original Fermi-Thomas approximate computations).
Thus, to make reliable predictions based on data, we need to use approximate methods that return the guaranteed lower and upper bounds for the corresponding energy, i.e., methods that return the interval that is guaranteed to contain the actual energy. The need for such methods is emphasized, e.g., in Scerri. Such interval methods have actually been developed; see, e.g., Fefferman, Seo and references therein. These methods, however, are extremely complicated, are based on very sophisticated mathematics and therefore, have not yet been applied to the problem of explaining the periodic system.

- The above results are based on Schrödinger equations that describe non-relativistic effects. However, as the atomic number \(Z\) increases, relativistic effects become significant. For a single electron, there are known Dirac equations that correctly describe all relativistic effects. However, for multiple electrons, no such simple equation is known; instead, we have to use quantum electrodynamics, a complicated theory for which only perturbation numerical methods are known (e.g., based on Feynman diagrams), and for such methods, no guaranteed bounds are known.

In spite of all these attempts, the problem is still largely open. In short: in spite of all known attempts, the problem of explaining the periodic system is still far from being solved; see, e.g., Novaro, Scerri and references therein.

Additional problem: periodic table for ions. An ion is a atom from which some electrons are taken away, or to which some electrons are added. At first glance, the electronic configuration of an ion should correspond to the one of the corresponding atom, with the external electrons added or deleted. However, in practice, ionization often changes the electronic configuration; e.g., after deleting a few outer electrons, the resulting electron configuration stops being stable, and reorganizes itself into a different orbital structure. Thus, for ions, the corresponding “periodic table” is slightly different than for the ions. This new order was analyzed, e.g., by Goudsmit and Richards, who showed that for highly ionized atoms, we get the original Bohr’s hydrogenic order, in which:

- the energy of a state increase with \(n\);
- for two states with the same value of \(n\), the one with larger \(l\) has the larger energy

(see also Odabasi).

Thus, for highly ionized atoms, we have a different order of filling orbitals:

\[1s < 2s < 2p < 3s < 3p < 3d < 4s < 4p < 4d < 4f < 5s < 5p < 5d < 5f < \ldots\]

(In Barut, this order was explained by a different sequence of transformation groups \(SO(4, 2) \supset SO(4, 1) \supset SO(4) \supset SO(3)\).)
So, instead of the problem of explaining a single periodic system, we actually face a more complicated problem: to explain different "periodic systems" that correspond to different levels of ionization, systems that range from Bohr's order to the hydrogenic order.

2. Our main idea

Informal explanation. Our goal is to explain the ordinal information, the information describing in which order different orbitals are filled. Previously, numerical methods were used to explain this order, but the precise numerical values are hard to obtain:

- in non-relativistic approximation, we know the differential equations; they are easy to write down, but we are unable to solve them;

- if we take relativistic effects into consideration, then even the equations become very complicated.

Since we do not know the corresponding numerical values, we suggest avoiding numbers altogether and using only ordinal information.

Formulation of the problem in mathematical terms. In mathematical terms, each orbital is represented by a pair of natural numbers \((n, l)\) with \(n > l\). For example, \(1s\) is \((1, 0)\), \(2s\) is \((2, 0)\), \(2p\) is \((2, 1)\), etc.

The periodic table, as formulated in terms of orbitals, represents a linear order on the set of all such pairs \((n, l)\) (for which \(n > l\)). In physical terms, \((n, l) < (n', l')\) means that, in general, the orbital \((n, l)\) starts filling before the orbital \((n', l')\).

Comment. What we call a linear order is sometimes called a total order, meaning that for every two elements \(x\) and \(x'\), if \(x\) is different from \(x'\), then either \(x < x'\), or \(x' < x\).

The natural requirement on the order in which the orbitals (shells) are filled is that this order should be local, i.e., that the order between two shells should depend on how these shells are related to each other and should not depend on how many shells are hidden inside. For example, if if \(2p\) is filled before \(3s\), then

- \(3p\) should be filled before \(4s\),
- \(4p\) should be filled before \(5s\), etc.

In mathematical terms, this means that the relation between the two shells \((n, l)\) and \((n', l')\) should depend only on the differences \(n - n'\) and \(l - l'\), but not on the absolute values of \(n\) and \(l\). In other words,
• if we add (or delete) a inner layers, i.e., if we shift both values \(n\) and \(n'\) (i.e., replace them with \(n + a\) and \(n' + a\) for some \(a\)), the order should not change; and

• similarly, if we shift both values \(l\) and \(l'\) (i.e., replace them with \(l + b\) and \(l' + b\) for some \(b\)), the order should not change.

In other words, we can define a local order as an order that satisfies the following property: for every \(n, n', l, l', a,\) and \(b\), \((n, l) < (n', l')\) if and only if \((n + a, l + b) < (n' + a, l' + b)\).

Both Bohr's order and the hydrogen order (characterizing ions) are local. A natural question is: how to describe all linear (total) orders which are local (i.e., which satisfy with the above invariance property)?

3. Definitions and the main result

Previously we dealt with the set of non-negative integers \((n, l)\) with \(n > l\) and \(n > 0\), and were interested in a certain type of total orders on the set. The object of this section is to completely classify those orders. Let us start with a formalization of previously introduced concepts.

**Definition.** Let \(S\) denote the set of all pairs \((n, l)\) of non-negative integers with \(n > 0\) and \(n > l\). We will say that a total order < on the set \(S\) is shift-invariant if the following property is satisfied: for every two elements \((n, l)\) and \((n', l')\) of \(S\) and for every integers \(a\) and \(b\) such that \((n + a, l + b)\) and \((n' + a, l' + b)\) are also in \(S\),

\[(n, l) < (n', l') \text{ if and only if } (n + a, l + b) < (n' + a, l' + b).\]

**Theorem.** Every shift-invariant order can be described as follows: On a 2D plane \(\mathbb{R}^2 = \{(x, y) | x \in \mathbb{R}, y \in \mathbb{R}\}\), let \(r\) be a ray (half-line) coming from the origin, and let \(r'\) be the opposite ray. Then \((n, l) > (n', l')\) if and only if the difference \((n - l, n' - l')\) is positive, and the set \(P\) of positive elements is defined as follows:

• if the ray \(r\) is of an irrational slope, then the set \(P\) coincides with all the points lying in (precisely) one of the half-planes determined by the line \(rr'\);

• If the ray \(r\) is of rational (or infinite) slope, then there are four possible sets \(P\): the set \(P\) contains one of the half-planes as before and, additionally, one of the rays \(r\) or \(r'\).
Comment. A ray can be described by a single parameter — its slope \( k \) — as \( r = \{(x, k \cdot x) | x > 0\} \). Bohr’s order corresponds to \( k = -1 \), and the hydrogen order corresponds to \( k = -\infty \). In general, we, in essence, get a 1-parametric family of orders that describes the transition from Bohr’s order (for neutral atoms) to the hydrogen order (that describes highly ionized atoms). It is reasonable to assume that the intermediate values of the parameter \( k \) describe intermediate degrees of ionization.

4. Proof

This proof uses terms and results about ordered algebraic structures; see, e.g., Fuchs 13.

Lemma. The set \( S \) with a shift-invariant order forms a commutative, cancellative totally-ordered semigroup under a coordinate-wise addition.

Proof of the Lemma. That \( S \) forms a commutative cancellative semigroup is easy to show. We must verify that \( S \) is an ordered semigroup. To this end let us take four elements of \( S \), \((n_1, l_1), (n_1', l_1')\), \((n_2, l_2)\) and \((n_2', l_2')\) such that

\[
(n_1, l_1) < (n_1', l_1')
\]

and

\[
(n_2, l_2) < (n_2', l_2').
\]

By using the shift-invariant property with \( a = n_2 \) and \( b = l_2 \) first and the second time with \( a = n_2' - n_2 \) and \( b = l_2' - l_2 \), we have

\[
(n_1, l_1) + (n_2, l_2) = (n_1 + n_2, l_1 + l_2) < (n_1', n_2', l_1' + l_2') < (n_1' + n_2', l_1' + l_2') = (n_1', l_1') + (n_2', l_2').
\]

Lemma is proven.

Corollary. The ordered semigroup \( S \) is order-embedded into a totally-ordered group \( \mathbb{Z} \oplus \mathbb{Z} \).

Proof of the Corollary. By Corollary 6 Chapter 10.4 in 13, the ordered semigroup \( S \) is uniquely embedded into the totally-ordered group \( G \) of quotients of \( S \). Obviously in our case \( G = \mathbb{Z} \oplus \mathbb{Z} \).

Let us now prove our main theorem. Due to the corollary, all our attention concentrates on the totally-ordered group \( \mathbb{Z} \oplus \mathbb{Z} \). These groups, in turn, even in a greater generality (the order is a lattice and the group is \( \mathbb{Q} \oplus \mathbb{Q} \), where \( \mathbb{Q} \) denotes the set of all rational numbers) have been completely classified in Section 4 in Holland and Szekely 16. The result can be rephrased in the following form:
Theorem. (Holland and Szekely) Let $G = \mathbb{Z} \oplus \mathbb{Z}$. Then $G$ can be totally ordered precisely in one of the following ways: In the Cartesian plane $\mathbb{R} \times \mathbb{R}$, choose a ray $r$ emanating from the origin. Denote by $r'$ the opposite ray to $r$.

- If the ray $r$ is of an irrational slope, then there are exactly two total orders of $G$ associated with $r$, namely the positive elements of $G$ are all the points lying in (precisely) one of the half-planes determined by the line $rr'$.

- If the ray $r$ is of rational (or infinite) slope, then there are exactly four different total orders of $G$ associated with $r$. This time the positive elements lie in one of the half-planes as before, and, additionally, in (precisely) one of the rays $r$ or $r'$.

Since every order on $G$ uniquely determines the shift-preserving ordering of $S$, we obtain the desired classification of shift-invariant orders on $S$. The theorem is proven.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCCW-0089, by NSF grants No. DUE-9750858 and EEC-9322370, and by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518.

The authors are greatly thankful to numerous researchers, especially to Nelson H. P. Beebe, Yuri N. Demkov, M. Kibler, Vladimir G. Konopelchenko, Ramon E. Moore, Yuri B. Rumer, and Luis A. Seco, for valuable discussions and references.

References