How Life Emerges from Inanimate Matter

How Life Emerges from Inanimate Matter:

From Atoms to a Living Individual

By

Bernard Korzeniewski

Cambridge
Scholars
Publishing

How Life Emerges from Inanimate Matter:
From Atoms to a Living Individual

By Bernard Korzeniewski

This book first published 2023

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-1980-5
ISBN (13): 978-1-5275-1980-0

I dedicate this book to Stanisław Lem, the man who has had
the greatest impact on my intellectual development.

TABLE OF CONTENTS

Why is Structure Possible at all in the Physical World?

Why Does Structure Originate?

How Could Life Originate Spontaneously?

How Has the Complex Structure and the Function of Living Individuals Evolved?

What are the Regularities and Ways in the Course of Evolution of Living Organisms?

What are Life and Living Individuals?

Biological Individual Vs. Psychical Individual — Essence of Self-Consciousness

Conclusions

Does the Flow of Time Really Occur?

Artificial Life Based on Supra-Molecular Level

INTRODUCTION

Let us consider a complex mechanism (a functional system) built by a person, such as a car. We do not have the slightest problem in answering the question: what purpose does the car serve? Of course, to reach a designated place fast, comfortably, and safely. We also know perfectly well the role of particular parts and the sub-systems of the car. The engine propels, the steering wheel controls the direction, the wheels turn, the car body protects other parts/sub-systems, the crankshaft transfers the drive from the engine to the wheels, and the gas pedal decides the speed. The structure and function of the car are organized hierarchically, which means larger parts/sub-systems are composed of smaller parts/sub-systems. For example, the cylinders, fuel injectors, and spark plugs control the engine. Those, in turn, are composed of still smaller elements. The car and its parts/sub-systems have been purposefully designed to fulfill the tasks and functions subordinate (ancillary) in relation to it. It is easy to understand all of this; alternatively, comprehending particular details and aspects of the construction and operation of the car requires some technical knowledge. Each element of the car's structure and function possesses its clearly specified purpose. In other words, we know perfectly well the car’s purpose.

Let us now consider another complex system, which originated from billions of years of biological evolution—the frog. It is organized hierarchically, executes various tasks, and contains numerous systems and sub-systems. So, we have here organic macromolecules (first of all, proteins and nucleic acids: DNA and RNA), cellular organelles (e.g., mitochondria), whole cells (together with their components and complex structures), tissues, organs, and organ systems. Thanks to progress in biology, we know a lot about how they are built, how they function, and what purpose they serve. Thus, the heart pumps blood that delivers oxygen and nutrients to all body cells while receiving carbon dioxide and the waste products of metabolism. Muscles serve locomotion, vocalization (producing sounds), and catching insects (tongue, jaw muscles). The brain directs the whole frog’s behavior, integrates stimuli from the surroundings by sense organs, undertakes decisions concerning actions and implements them (through stimulation of muscle contraction), and accumulates memory. Sense organs (receptors) collect data from the surroundings, which the peripheral nervous system transfers to the brain. On the other hand, skeletal muscles (effectors), excited through peripheral motor nerves, implement the brain's directives. The digestive system serves the intake, digestion, and assimilation (absorption into the blood circulation system) of food (nourishment). The skin insulates and protects the internal organs, and also, as the frog is an amphibian, it participates significantly in respiration. The kidneys extract and densify the waste products of metabolism, especially nitrogen metabolism (urea in adults, ammonia in tadpoles), which are then expelled. Particular organ systems and organs are composed of parts and sub-systems that perform service and subordinate functions—it is enough to see an anatomical section of the kidney or brain to become familiarized with the mechanisms of their operation. The structure and function of all organs are based on living cells—specific for particular organs and tissues (frequently as different as epithelium cells, neurons, and muscle cells). Most cells contain organelles—a nucleus, mitochondria, endoplasmic reticulum, and macromolecule complexes, such as a cytoskeleton or ribosomes serving protein synthesis.

All parts/sub-systems of the organism of a frog support and mutually serve each other. The heart, together with the circulatory system, enables the functioning of the brain, muscles, and digestive system, the digestive system sources nutrients for all organs, the muscles enable locomotion directed at gaining food and fleeing from a predator, and the brain adequately supervises the behavior of the frog in response to stimuli from the surroundings, enabling it to survive together with all its organs. The same reasoning can be applied to organs and organ systems and every cell (of one of several hundred types). Ultimately, all parts/sub-systems of an individual frog serve all other parts/sub-systems. This property is called *causal closure*.

All of this seems simple and understandable. A problem appears only when we ask what the role, function, or purpose of a living individual in general and our frog in particular is. Certainly, this superior, long-term aim is not the survival of itself, as every complex organism eventually dies, either because of senility, in the jaws of a predator, or due to illness (e.g., infection by parasites), environmental change (e.g., drying out of a water reservoir) or by accident (only simple organisms such as bacteria can potentially be "forever young" and divide *ad infinitum*). Of course, living individuals strive to survive "until a certain time." But what time? Many organisms die "voluntarily" just after reproduction, for which they use all accessible energy resources, e.g., salmon, after spawning (preceded by an exhausting journey upstream against the river flow), or one-day ephemera after copulation and lying fertilized eggs (they spend their whole life in water in the larval stage and as adults do not even have a functional digestive system). It "pays" for the mantis male to be eaten by the female after copulation, as he supplies building material and energy to his progeny in this way.[1] Certainly, the frog is not purposefully directed to be an element of the ecosystem in which it lives, for instance, to serve as food for storks or to eat insects to control their number. Just the opposite, it struggles to avoid being eaten by other organisms, either by hiding or fleeing, and it is not interested in controlling the number but in the abundance of insects. In turn, insects do what they can to not be consumed by a frog. The opinion that living organisms tamely accept the role of an element of an ecosystem is extremely naïve (leaving aside the fact that it implies ignorance and a lack of understanding of biology). The truth is completely different: living individuals struggle for the best possible place in the ecosystem through competition and evolutionary optimization. Every individual is selfishly directed to itself and not for "the good of the ecosystem" or "the good of the species," as nothing like this exists. Also, the frog does not serve to be a supplier of legs for the French. Leaving aside an extreme anthropomorphism of such a point of view, frogs (and their kin) evolved for hundreds of millions of years before the French (and storks) appeared, and most frog species live in tropical forests nowadays, far away from humans. In the face of this, what is the superior purpose of this incredibly complex ensemble of cells/organs/organ systems entering into the composition of a living individual? What does such an individual, for instance, the frog, strive toward? What is the rationale for its existence in its whole complexity of structure and function? In other words: ***what purpose does the frog serve?***

We will come to the answer to this question step by step. We will see how the phenomenon of life, in particular its fundamental unit called a living individual, gradually emerges from the functioning of inanimate elements of the physical reality when we pass from single atoms (and earlier—elementary particles), through more and more complex configurations of atoms, to the enormously complicated structure and function of living organisms. This transition should be understood in two ways. First, it is about the biological evolution of life from the stage of spontaneously originated organic compounds and their complexes, through initially very simple, living systems, to all biological organisms appearing presently (and in past geological epochs) on the terrestrial globe and, most probably, on many other planets in the universe. Second, it is about how the properties of single atoms, constituting the building material of biological organisms, lead to, after their appropriate arrangement in a certain superordinate functional structure, the creation of a completely new quality absent in the

---

[1] Of course, it does not pay for the male to be eaten before copulation.

purely physical (and chemical) world, namely the phenomenon of life, based on living individuals, characterized by intrinsic purposefulness and a countless richness of forms. As we are one of the forms of life and (also) living individuals, the solution to this conundrum could help us understand the place of humanity in the (physical) universe.

This book is not addressed to any determined, easy-to-separate circle of readers, e.g., biology students or adepts in the philosophy of nature, and therefore it does not constitute a handbook of the problems dealt within it. Emphasis is put on understanding problems rather than an enumeration of possibly many facts, and before the question of "What?" the priority is given to the question of "How?" More than dazzling with erudition, I prefer a clear (although not always easy to receive) answer to the fundamental questions. This is, first of all, a book written from my point of view, containing what I regard as worth presenting on the considered topic, addressed to all thinking people curious about the world, especially those interested in biology and other natural sciences. Although deeply anchored in contemporary scientific knowledge, it certainly cannot be considered a strictly popular-science item, as I propose in it, especially in Chapter 6, my original interpretation (vision) of the phenomenon of life and a living individual, which has already been partly discussed in my previous works. The cybernetic definition of life (a living individual) is further developed, justified, explained, and confronted with the realities of the physical world. In all earlier chapters, the popular-science motif decidedly predominates, so even if my personal opinion about the essence of life seems too difficult to grasp or is insufficiently convincing, I hope the remaining part of the book will be interesting and deliver some intellectual satisfaction. Finally, in Chapter 7, I shortly summarize my opinions on the essence of mind, subjective psyche, and self-consciousness and their relation to the phenomenon of life. I describe the cybernetic model of the psyche-bearing individual and the formal similarities between them and biological individuals. Generally, I try to justify the theory that the complexity (and, first of all, the complexity at the basis of the mentioned individuals) exists equally realistically and objectively in the external physical world, as does matter, space, and time.

Numerous fundamental questions lie at the basis of the present elaboration. How is the existence of any structure in our universe possible at all? Why, in some cases, does the structure originate? What conditions must be fulfilled in order for this to happen? How can the degree of structural complexity and amount of related information spontaneously increase over long periods of time? What is the specificity of the structure and function of living individuals? What fundamental mechanisms were responsible for their origin, evolution, perfecting, and complications? What is this, the living individual, and what relationships associate it with the phenomenon of life in general? What constitutes the subject of biological evolution? What is the relation between the living individual and the psychical individual? These are just a few of them.

Several major concepts contribute to the leading motif of this book. First, life appears as an emergent property characterized by a determined kind of complexity from the functioning of its elements at a lower physical level. No special factor or agential force is needed for life to come into being. This, which causes the feeling of complete dissimilarity of living systems from inanimate physical systems, and therefore is responsible for the emergence of the biological level, is not so much the degree but rather the *specific kind* of functional complexity of matter constituting the building material of living organisms. For purely physical reasons, living organisms must be dynamic dissipative structures (generally, systems displaced from thermodynamic equilibrium) whose existence is sustained by thermodynamic gradients, unceasing flow of matter, dissipation of energy, and production of entropy. The fundamental unit of life is living individuals that strive to survive in good conditions to the reproduction period, to reproduce and thus transfer their identity to progeny. They (in fact, their identity) can evolve from generation to generation and constitute the unit of natural selection and the subject of Darwinian evolution. The living individual can be best defined in the frame of the cybernetic paradigm as a network of purposeful regulatory mechanisms (first of all, negative feedbacks) supervising the embryonal development (morphogenesis), cognitive functions, behavior, and constancy of internal conditions, directed only and exclusively on the multiplication of itself in a possibly great number of copies, that is reproduction, representing positive feedback. It constitutes an example of the relation of self-applicability (self-reference, recurrence) that is, in my opinion, crucial for understanding not only the phenomenon of life but also of self-consciousness.

The living individual can be conceptualized in a slightly simplified way as a self-controlling, self-copying system; that is, a dynamic *homeostatic replicator*. Life emerges when elements of inanimate matter undergo organization into an appropriately composed SYSTEM, corresponding to the above characteristics.

I would like to thank Andrzej Joachimiak and Ireneusz Ślesak cordially for their critical reading of the manuscript and for detecting several mistakes and obscurities.

CHAPTER 1

WHY IS STRUCTURE POSSIBLE AT ALL IN THE PHYSICAL WORLD?

Structure in physics and chemistry

This book should start from the most elementary problems. An inseparable property of living organisms is their complex, hierarchically-organized structure (and function based on it). The fundamental question is why structure, in general, is possible in our physical universe. However, should this question not be considered completely pointless, as it refers to an obvious and understandable matter? After all, we can easily create, in a sense, naturally, various sorts of structures. We can build different constructions from LEGO playing blocks. We can connect plastic balls of various sizes and colors with plastic sticks to obtain models of even very complicated chemical molecules, for instance, proteins or the DNA double helix. We produce computers, cars, skyscrapers, and toothbrushes. Besides, it is enough to take any product of nature, such as a stone, to see that it possesses a determined shape and internal structure (composed of various crystals, minerals, grains, and strands). Then, is there a problem associated with structure? Is it not enough to shape and process components and then appropriately connect them? In the case of a stone, nature created this through certain geological processes.

Unfortunately, it is not so simple. All enumerated, having structure, macroscopic objects are composed of atoms of elements and molecules of chemical compounds. On a microscopic level, where quantum mechanics absolutely rules, there are no “shaped elements” and appropriate “connectors” between them, such as, for instance, plastic balls and sticks in macroscopic models of chemical molecules. Instead, there are only some objects (elementary particles, atoms), and the only thing between them is various interactions (forces) that can be attractive or repulsive. On the other hand, they lack any “rubber bands,” “connectors,” or “sticks.” (Leaving aside the fact that in the microworld, the distinction between objects and the forces generated by them and acting on them is not fully legitimate, which is discussed more broadly at the end of this chapter). How is it possible in such a situation to get any structure?

Well, a fundamental, far from trivial, answer to the question about the possibility of structure coming into existence in the real world is the statement that the key role is played here by the presence of attractive and repulsive forces between physical objects (e.g., elementary particles or atoms), the forces which counterbalance each other when the considered objects are located at a determined distance from each other. Just this distance determines the characteristic size of a given elementary structure.

To better comprehend this, see Fig. 1-1. It demonstrates the dependence of the amount of energy of a certain microscopic object (e.g., an elementary particle and atom), shown in the figure as a little ball, on its distance from the object (again: elementary particle, atomic nucleus, atom) that interacts with (exerts a force on) it, and whose location is indicated by the place of intersection of the vertical *y*-axis with the horizontal *x*-axis (that is zero distance on the *x*-axis). We can see that energy is very high in the nearest neighborhood of the object exerting force; it quickly decreases as a function of the distance from it, reaches a minimum at a certain critical distance, and then increases from negative values tending asymptotically to zero, as the distance increases to infinity. A fundamental physical principle is that such a system tends toward the lowest energy state. In other words, this means that the object (a little ball) on which a force (influence) is exerted will be attracted more strongly when it is situated at a certain distance from the object exerting the force greater than the critical distance (corresponding to minimal energy), the closer to this object it is situated (at very large distances the attraction can be neglected, as it is very close to zero). However, when the distance between the objects is less than critical, a strong repulsion will appear in place of attraction. The system will reach a relatively stable state when the object experiencing the influence finds itself exactly at the critical distance from the object, constituting the source of the force. Additionally, it should be decidedly emphasized that the situation between the objects is fully symmetrical: both objects in any situation exert on each other the same, although oppositely directed force (influence). Therefore, the objects exerting force and experiencing force can be exchanged with each other, and this would change nothing. The apparent lack of symmetry results from the difference of masses between objects: it is simply more difficult to "move from a place" a heavier object using the same force, while it is easier to set in motion lighter objects. Because of this (and only because of this), it seems that, e.g., an atomic nucleus attracts an electron, while the earth attracts a spacecraft, and not inversely.

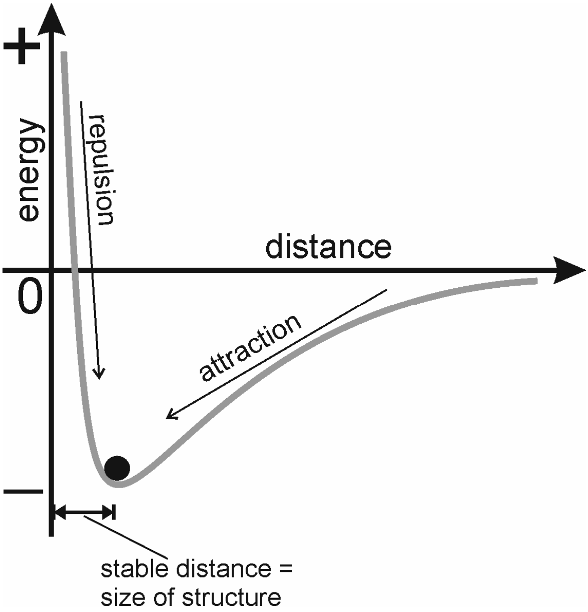

Fig. 1-1. Dependence (gray curve) of the energy of the system on the distance between an object experiencing influence (force) (little black ball) and an object exerting force (situated at distance 0, that is in the place of intersection of the vertical axis y and horizontal axis x). The system spontaneously tends toward minimization of energy, and therefore attraction will predominate at distances greater than the stable (critical) distance, repulsion will prevail at distances less than the stable distance, while the system will remain in equilibrium (or oscillate around it because of disturbances from the surroundings) when the distance between the objects will equal exactly the stable distance.

As a result, the reciprocal interaction between objects constitutes a balance between attractive and repulsive forces. Both forces become weaker with the distance from the objects generating them (in most cases); therefore, the closer the objects are located, the greater force they exert on each other. In order to reach an equilibrium (balance) between them at a certain distance, the strength of repulsion should be stronger at very small distances, but it should fall, with distance, quicker than the force of attraction so that attraction can prevail at greater distances. This property ensures not only the existence of a critical distance between objects, at which attraction and repulsion counter-balance each other, which predicates the existence of the structure itself, but also the dynamic stability and elasticity of this distance. It is expressed by the fact that every disturbance of the stable distance between objects (displacement of the system from the equilibrium state) results (as long as it is not too strong) in an automatic action in the opposite direction, causing the return to the equilibrium state. An increase in the distance will be compensated for by the prevalence of attraction over repulsion, while a decrease in the distance — by an intensified repulsion. Because of this, the distance between objects will oscillate around the stable equilibrium value in an environment full of disturbances (e.g., thermal movements and collisions, where the temperature is proportional to the average kinetic energy of molecules). All of the above statements can seem rather abstract and enigmatic. Therefore, I will try to explain, explicate, and exemplify.

Let us start with the structure of an atom. As it is commonly known, the atom is composed of a positively charged nucleus and negatively charged electrons. Why do electrons at all keep themselves in the proximity of the nucleus (they are bound within the atom) and not fly away into the surrounding space? Again, it is well-known that charges of the opposite sign attract each other. This asserts that (apart from conditions, when sufficient energy is supplied from the outside, see below), electrons remain in the immediate neighborhood of the nucleus. Does this suffice to explain the (frequently very complex) structure of the atom? Leaving aside the fact that electrical attraction by itself, which is, after all, identical for every electron, does not illuminate the fact that particular electrons occupy different electron shells and, within them, orbitals of various shapes, the attraction itself should cause electrons to fall on the nucleus and neutralize its charge, because of which the atom would shrink to an electrically-neutral point devoid of any dimensions and structure. However, physical studies indicate unequivocally that nothing like this happens. Moreover, it has been found that a certain repulsive force, which does not allow the excessive approach of electrons to the atomic nucleus, is responsible for this.

This force has more complicated properties than the “classical” repulsive force and is described by quantum mechanics. According to quantum mechanics, an electron bound in an atom does not find itself at any determined point but is as if “blurred” in space: it adopts the shape (form) of an “electron cloud” (called orbital) corresponding to the probability of finding the electron in a given point of space (Atkins and de Paula 2002 326-334; Chandrasekhar 1998, 35-62; Al-Khalili 2012, 52-77; Baggott 2013, 27-53). In the regions where this cloud is dense, the probability that the electron is situated there is large,[1] while in places where the orbital

---

[1] In reality, the electron is not located in any concrete place before the act of measurement (detection) or, strictly speaking, the act of interaction with the surroundings.

undergoes rarefication (dilution), the chance of finding the electron there diminishes. Therefore, even in the case of the simplest orbital (called orbital s), having generally speaking a spherical shape, the greatest density of the electron cloud takes place, not in its center (namely in the atomic nucleus) but on the surface of the sphere of a certain finite radius, determining the distance of this surface from the nucleus. In the direction of the surrounding space, inward and outward of this surface, toward the nucleus or outside, the density of the cloud falls to reach zero both in the atomic nucleus and, asymptotically, at the infinite distance in the emptiness of the universe. All this happens because, according to quantum mechanics, the energy of an electron bound in a given atom can adopt only strictly determined, discrete (discontinuous), quantized values of energy. An electron moves in space as the so-called matter wave. Hence, it can simultaneously pass through two slits in a septum and bend on them: the diffraction phenomenon, and then interfere with itself. A series of electrons passed singly, one by one, through such a pair of slits gives a characteristic pattern of stripes on a screen located behind the slits (Al-Khalili 2012; 78-105; Baggott 2013, 27-53).[2] Orbitals (their shapes, dimensions, and spatial arrangement) result from the fact that only for some energies (and therefore lengths of the matter wave, inversely proportional to energy) of an electron, the “hills” and “valleys” do not suppress or cancel each other (“hills” do not meet “valleys”), but enhance. The “hills” meet the “hills,” and the “valleys” meet the “valleys,” leading to the origination of stable “structures.” Using this metaphor, an electron “goes around” (revolves) the atomic nucleus in such a way that the “hills” of the matter wave associated with this electron overlap with themselves in subsequent “rounds”—if the “hill” would meet the “valley,” they would have canceled each other out. Therefore, the “circumference” of orbitals must, in a sense, correspond to the length of the matter wave of a given electron (the distance between the “hills” and the “valleys”) or its multiplicity (Chandrasekhar 1998, 63-83; Baggott 2015, 105-113). The energy of an electron in an atom is proportional to its distance from the nucleus: the closer to the nucleus, the lower the energy. The lowest possible energy corresponds to the shortest distance of an electron (or, strictly, the smallest distance of the highest density of its electron cloud) from the nucleus. This specific repulsive force, forbidding the electron to come close to the nucleus excessively, and resulting from quantum mechanics (it would be difficult to find its equivalent in the classical world), enables the creation

---

[2] They result from the mutual enhancement of the matter waves being in the same phase (the “hills” of the wave meet “hills,” while the “valleys” meet “valleys”) and with the suppression (quenching) of waves being in the opposite phase (the “hills” overlap with the “valleys”).

of the simplest structure (the nucleus plus electron cloud), determining at the same time the magnitude of this structure (the simplest atom) (Chandrasekhar 1998, 63-83). This constitutes the first exemplification of our general thesis that structure in the physical world is possible thanks to a certain equilibrium between attractive and repulsive forces. At the same time, the attractive force must prevail at a distance greater than some critical distance, while the repulsive force begins to predominate when the distance is lower than the critical distance. This maintains the distance between physical objects on a certain level, determining the size of the structure formed by these objects.

Why, then, do more complicated atoms exist in nature with such diverse structures determining the physicochemical properties of various chemical elements? Why don't all electrons simply gather on the orbital closest to the atomic nucleus, characterized by the lowest energy? Here, a second "repulsive force" enters the stage, namely the so-called Pauli exclusion (prohibition) principle (Atkins and de Paula 2002, 337-344; Chandrasekhar 1998, 63-83; Penrose 2002, 388-389, 428-430, 448-449; Baggott 2013, 34, 65), stating that a given energetic level (more strictly: orbital) can be occupied by only two electrons and only when they have oppositely-directed spins (that is, simplifying a lot, when one of them rotates clockwise, while the other, anti-clockwise; quantum spin is a quantum property that cannot be easily expressed in common language and that is to a certain extent an equivalent of revolving around an axis). Generally speaking, no two electrons can be in an identical quantum state, whose element is orbital, on which a given electron finds itself, and the orientation of its spin. When an atom possesses more than two electrons (this is determined by the positive charge of the nucleus and, therefore, by the number of protons contained in it), they have to look for other higher energy orbitals to occupy and, therefore, a greater distance of the maximum of the density of their electron cloud from the nucleus. Those can be spherical orbitals s, as described above, but with a greater radius, orbitals p having the shape of two tears laid opposite to each other and abutting each other with their sharp ends in the nucleus (the maximum density of the electron cloud is represented here by two points located in the center of the spherical parts of each of the "tears") (Al-Khalili 2012, 154-187; Baggott 2015; 105-107). Three such orbitals are on each energetic level oriented perpendicularly to two other orbitals p, along axes x, y, or z. There are also orbitals d and f with the yet more complicated spatial distribution of the electron cloud. The principles of quantum mechanics determine their shapes. Particular orbitals can be ranked according to the increasing energy associated with them, which determines the succession of the occupation of these orbitals by electrons when the number of electrons in atoms of subsequent elements on the Mendeleyev board increases. Generally, orbitals combine into groups of similar energies called electron shells. This is just the diversity of the shapes of orbitals, the location within various electron shells, and the degree of occupancy of particular orbitals and shells by electrons responsible for such a huge differentiation of (the properties of) the atoms of particular chemical elements. As an additional factor, there can be quoted the third (most classical) repulsive force, namely, the electrical repulsion between identically (negatively) charged electron orbitals, and also the attraction by the atomic nucleus, which causes "the same" orbitals to have different dimensions (and energy values associated with them) in atoms of different elements.

Generally, the combination of attractive and repulsive forces described above is responsible for the complex structure of the atom (as well as its diversity in particular elements) and its size. The greater the number of electrons the atom contains, the more high-energy, and thus distant from the nucleus orbitals they must occupy, and therefore the greater the atom (although this dependence is not so simple, as atomic nuclei containing more protons attract electrons more strongly, which causes the electron shells to shrink).

As has already been said, atoms are comprised of the nucleus and electrons. Electrons are elementary particles. Such particles can be defined as objects *devoid of structure*, that is, having the geometrical form of a point (although, e.g., string theory treats them as one-dimensional strings with a length comparable with the so-called Plank distance, vibrating in various ways) (Greene 2005, 327-375; Greene 2010, 3-22). However, this does not concern atomic nuclei that are not elementary particles and possess structure. The nuclei are composed of a certain number of nucleons: protons (positive charge) and neutrons (electrically neutral). What maintains the structure of the nucleus? Again, the answer is the same as in the case of the atom: the balance between attractive and repulsive forces. This time we will start with repulsion. Protons, like all objects endowed with an electric charge of the same sign (positive in this case), exert reciprocal repulsion on each other. It is counter-balanced by the so-called strong nuclear force, one of the four fundamental physical forces (we distinguish here strong and weak nuclear force, electric or, more generally, the electromagnetic force and gravity), to which nucleons (protons and neutrons) are subject. Strong nuclear influences are mediated by *gluons* (like photons mediate electrical influences) (Al-Khalili 2012, 199-204; Baggott 2013, 54-80; Baggott 2015, 35-60). Interestingly, the strong nuclear force between protons alone does not suffice to offset the electrical repulsion. Therefore, a certain number of neutrons, which deliver additional attraction through strong nuclear force while not contributing to increased repulsive force, is necessary for stable atomic nuclei. Indeed, only the nucleus of hydrogen (its basic isotope, protium),[3] formed by a single proton, can do without neutrons. The nuclei of all heavier elements contain a certain number of neutrons; proportionally, the higher the number, the heavier the element. Unstable (having unstable, radioactively decaying nuclei) isotopes of otherwise stable elements come, among others, from the fact that they lack one or more neutrons in their nuclei.

Does this exhaust the problem of the atom structure? No, because protons and neutrons are not elementary particles, contrary to earlier suppositions. We already know what this means—they possess structure (resembling, to some degree, "mini-atoms"). Namely, protons and neutrons are composed of already true (at least according to the present knowledge) elementary particles—quarks (Al-Khalili 2012, 154-185; Baggott 2013, 54-80). The proton is built of two quarks "up" (u) of a charge +2/3 and one quark "down" (d) of a charge -1/3, which gives the resultant charge of +1. The neutron is composed of one quark u and two quarks d; as it is easy to calculate, this is the reason for the neutron's electrical neutrality. What forces contribute to the existence of the structure of nucleons? The same force is responsible for attraction as in the case of forming the structure of atomic nuclei, namely a strong nuclear force. It acts between nucleons (protons and neutrons) and inside them—between the quarks. The equivalent of the repulsive force is constituted here by the property called asymptotic freedom. It consists in the fact that at very short distances, corresponding exactly to the dimensions of a proton and neutron, the attraction between the quarks decreases with distance, instead of increasing and eventually ceases to act as they approach each other because of which they can freely move within the limits of this distance, forming a chaotic clump (of course, this description is only an imperfect "classical" approximation of the quantum situation).

We see, therefore, that the reason for the existence of the quite complicated structure of the atom is to set itself at a certain distance equilibrium (balance) between the attractive and repulsive forces, and this phenomenon recurs at three levels of hierarchy: between the atomic nucleus and electrons, between nucleons (protons and neutrons in the nucleus), and between the quarks in the nucleons. Therefore, as we already generally understand the nature of the atom structure and the reason for its

---

[3] The protium (^{1}H) nucleus contains only a proton, the deuterium (^{2}H) nucleus: one proton and one neutron, while the nucleus of tritium (^{3}H): one proton and two neutrons.

dimensions, it is time to pass on to the structure at higher levels of the hierarchy of matter organization.

The best known structure on a microscopic level, and at the same time the most directly associated with the complexity of living organisms, is the structure of chemical molecules (Chandrasekhar 1998, 63-83). These molecules are composed of atoms of the same or different elements. As the reader presumably guesses, the structure (its size, the length of the bonds between different atoms) results from the balance (equilibrium) between the attractive and repulsive force. It has already been mentioned that electrons in atoms occupy still-not-filled orbitals with the lowest possible energy (at the same time, one orbital can be occupied maximally by two electrons with opposite spins). In atoms of many (most) elements, the most outer orbitals (or groups of orbitals of the same energy, e.g., three orbitals *p*), called "valence orbitals," are not filled—there are free places for additional electrons on them (Atkins and de Paula 2002, 320-344). Such atoms possess, depending on their electrical properties (determined by the charge of the nucleus and the amount and kind of electron shells), the tendency to either accept the "lacking" electrons on the unfilled orbitals or to get rid of "free" electrons from these orbitals. When electrons are completely taken over from or given back to the surroundings, there originate negative or positive ions (in a non-ionized atom, negative charges of electrons equalize the positive charge of the nucleus). This takes place in water solutions of salts or, e.g., in many crystals (including the crystals of sodium chloride, that is, common table salt). However, frequently, when two atoms meet, and one is slightly more disposed to give away electrons, while the other is somewhat apt to their acceptance, there does not follow a transfer of electrons, but sharing ("communizing"). They become bound to two atoms simultaneously, while they are more shifted toward (there is a greater probability of finding them at) the atom that is more disposed to accept electrons.[4] Instead of two atomic orbitals, one common molecular orbital is formed, completely filled with electrons (Atkins and de Paula 2002, 363-385; Patrick 2000, 17-18). At the same time, if an atom of some element lacks two electrons to fill its most outer (valence) electron shell, it can bind (by sharing electrons with them) two atoms of some other element, of which each has one "spare" electron. The number of electrons that an atom of a given element is disposed to give away or accept is called the valence of this element (atom) (Atkins and de Paula 2002, 363-367). For instance, in the water molecule H_2O, one oxygen atom partly takes over two electrons

---

[4] In the case of bonds between atoms of the same element, for instance carbon, electrons (their "probability clouds") are symmetrically distributed between atoms.

from two hydrogen atoms (as a result, the oxygen atom in the water molecules becomes slightly negatively charged, while the hydrogen atoms become positively charged; the whole molecule becomes *polar* and is called a *dipole*). At the same time, it should be added that the tendency to accept or give away electrons is a relative property: an atom of a certain element taking over electron(s) from atoms of another (more "electropositive") element can give away electron(s) to atoms of yet another (more "electronegative") element. Nevertheless, electrons can be shared (communized) between two atoms of the same element with the same tendency to accept/give away electrons, as it takes place in the case of the oxygen molecule O_2, which is commonly present in the earth's atmosphere. Such bonds, characterized by sharing electrons between (among) two (or more) atoms of the same or different elements, are called *covalent bonds* (March 1992, 3-93; Atkins and de Paula 2002, 363-368).

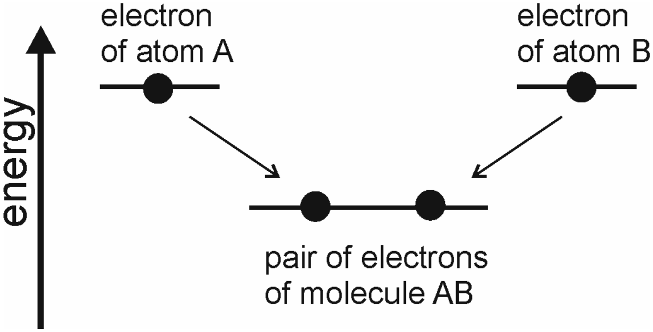

Fig. 1-2. The energy of unpaired electrons on valence orbitals in atoms A and B can be lowered after sharing electrons between atoms, equivalent to formation of a covalent bond and molecular orbital, and thus creation of a molecule of the chemical compound AB.

One could question why the sharing of electrons acts as an attractive force, which leads to bond formation. As mentioned, electrons tend to minimize their energy and occupy the atom orbitals with the lowest possible energy. It appears that forming the molecular orbital within a covalent bond from two atomic orbitals frequently causes the combined energy of electrons on the molecular orbital to be lower than the sum of energies of these electrons on atomic orbitals (the difference in energy is released to the surroundings in the form of heat during the occurrence of chemical reactions, e.g., in the process of combustion) (Atkins and de Paula 2002, 379-384). This happens in spontaneously occurring reactions and is illustrated in Fig. 1-2. One can see that electrons find themselves on relatively high-energy atomic orbitals of atoms A and B after forming a covalent bond between atoms, occupying a relatively low-energy molecular orbital AB. Besides, the above phenomenon concerns not only reactions between single atoms but also between chemical molecules; the only condition of the spontaneous occurrence of such a reaction is that the total energy of electrons on atomic/molecular orbitals of products is lower than their energy on atomic/molecular orbitals of substrates. For example, in the reaction involving the combustion of methane in oxygen, carbon dioxide and water are formed:

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

only molecules participate, and not single atoms. Nevertheless, in many cases, more complex molecules can undergo decomposition into less complex molecules, while these then decompose into atoms if such a reaction is privileged energetically (strictly speaking: thermodynamically—see Chapter 2). A given molecule can also be transformed into another molecule of the same atomic composition (reaction of isomerization). Frequently, for an energetically “unfavorable” reaction that is associated with increased energy of products in relation to the energy of substrates to occur, it must be coupled with another reaction that is energetically (thermodynamically) “privileged,” for example, the reaction of ATP (adenosine triphosphate) hydrolysis to ADP (adenosine diphosphate) and P_i (inorganic phosphate) (Marks et al. 1996, 275-290). Only the sum of these coupled reactions lowers the system's energy. Then, the newly formed chemical molecule is unstable and undergoes a spontaneous decomposition after delivery of sufficiently high energy from the surroundings (e.g., the kinetic energy of molecules proportional to temperature, higher than so-called activation energy).

Therefore, here, our attractive force that enables forming of the structure of molecules of the chemical compounds is a diminution of the system energy, where the excess of energy is released to the surroundings in the form of heat.[5] Similarly, a golf ball placed in a bowl tends to be in the most

---

[5] In principle, in spontaneously occurring reactions, there takes place a negative change of the so-called Gibbs free energy (Nicholls and Ferguson, 2002, 31-50), being a derivative of heat production and increase in the system entropy. For a given chemical reaction, this thermodynamic force can have a negative, positive, or zero value (in the last case we deal with the state of thermodynamic equilibrium), depending on the actual ratio of the concentration of products to the concentration of substrates related to the equilibrium ratio of these concentrations. Our

low-energy state, resting on the bottom of the bowl. On the other hand, what about the repulsive force, causing atoms not to approach each other too much? Here, the matter is simpler—this force is constituted by the aversion of electrons to occupy the same orbitals (to occupy the same places in space and generally the same quantum states). Therefore, when orbitals are fully filled with electrons, orbitals begin to overlap, which causes strong repulsion (the already mentioned Pauli principle). To express this in terms closer to classical mechanics, the negatively charged electron shells of different atoms repulse each other.

It should be stressed once more that covalent bonds are not stiff "sticks" joining the atoms but rather (although this is a broad analogy) springs, which can oscillate around the stable resting state (corresponding to the characteristic bond length). This is equivalent to the ball moving in Fig. 1-1 to the left and right in relation to the lowest location in the depression. These oscillations are caused by disturbances from the surroundings, particularly by the thermal movements of molecules, and are related to their collisions with other molecules endowed with kinetic energy; therefore, their intensity increases with temperature. They can also be excited due to the absorption of a quantum of electromagnetic radiation (photon) of an appropriate wavelength (energy). As long as the energy of such disturbances does not exceed a certain threshold value, the "springs" of the bonds shrink and expand around the resting location, and the amplitude of these oscillations is proportional to the energy of the disturbances. However, when, for example, temperature (that is, the kinetic energy of movements, vibrations, rotations, and collisions) is elevated sufficiently, or the matter is exposed to electromagnetic radiation of sufficiently high energy, the ball from Fig. 1-1 will fall out of the depression, and the spring will be broken. This will lead to the decomposition (or transformation) of the molecule (into) other (frequently simpler) molecules or atoms. In this way, a chemical compound undergoes degradation. The mechanism of melting solid bodies (their transformation into liquids) is quite similar, but here, covalent bonds are not most frequently broken, but rather ionic bonds (discussed earlier) or polar bonds (that are described in the later part of this chapter).

It should be clearly emphasized that Fig. 1-1 constitutes, at least in relation to interactions between atoms and molecules, a significant simplification. If atoms or molecules always attracted each other at a distance greater than the stable critical distance corresponding to minimum energy, all atoms/molecules that could potentially react with each other would instantly do this, forming

---

considerations concern physiological (appearing in living organisms) concentrations of reagents. More details are given in Chapter 2.

molecular bonds and sharing electrons, and no further chemical processes would occur. However, this does not happen; chemical reactions, even those advancing spontaneously that are thermodynamically privileged, occur with a finite, often very low speed. A kind of "hump" in the dependence of the system's energy on the distance between atoms/molecules is responsible for this. This is demonstrated in Fig. 1-3. This hump, the existence of which is responsible for the electrical interaction between negatively charged electron shells and orbitals, so that when atoms/molecules approach each other, a strong repulsion takes place between them until a certain critical limit (distance) is crossed, behind which the strong attraction between the atoms/molecules related to forming molecular orbitals and covalent bonds takes place. However, if so, why is the energetic barrier corresponding to the "hump" ever crossed, or, in other words, why do atoms and molecules react at all? The energetic barrier can be transgressed if energy greater than the so-called *activation energy*, corresponding to the height of the "hump," is supplied to the system (Marks et al. 1996, 104-108; Voet et al. 1999, 288-289; Murray et al. 2012, 57-69). Such energy can be the kinetic energy of atoms/molecules proportional to their mass and velocity and directly associated with the strength of their collisions with other atoms/molecules or of quanta of electromagnetic radiation. This is just the increase of the kinetic energy of atoms/molecules that causes the rise in temperature to speed up chemical reactions—more accelerated atoms/molecules can jump over the discussed "hump" more easily. Does this mean that when the temperature is lower than some critical temperature, the reaction does not proceed at all, and when it is higher, the reaction proceeds at the maximum possible velocity? Not at all, as particular atoms/molecules in a system with a determined temperature have very different speeds and kinetic energies (the distribution of kinetic energy in the population of molecules at a given temperature is described by the bell-shaped Gauss curve). This means that most atoms/molecules have energy close to the average kinetic energy of the system, while atoms/molecules with greater or less energy are much rarer. At lower temperatures, only very few molecules acquire energy greater than the activation energy due to accidental sequences of collisions with other molecules and therefore enter (participate in) chemical reactions. The proportion of such molecules rises very intensively with temperature, which causes the velocity of chemical processes to increase two-three-fold whenever the temperature is elevated by 10 degrees centigrade (or Kelvin degrees)[6] (Voet et al. 1999, 288-289; Murray et al. 2012, 57-69).

---

[6] This is the so-called absolute temperature, whose scale begins from zero 0 Kelvin degrees (K), at which there are no movements of atoms/molecules. In the Celsius scale, this corresponds to the temperature of about -273 degrees centigrade (°C).

One more matter remains yet to be explained. We stated previously that a high temperature can lead to the breaking of chemical bonds, while now we say that elevation of the temperature promotes forming of such bonds. Is this not a contradiction? Fortunately, it is not. Please pay attention to the fact that the "hump" (activation energy) in Fig. 1-3 is much lower than the depth of the depression corresponding to the stable distance related to the covalent bond length. Therefore, a much lower kinetic energy (temperature) is needed for the little black ball (atom/molecule) to jump over the "hump" and fall into the depression than the energy (temperature) necessary to cause (make probable) the ball to jump out from the depression. In other words, a moderate rise in temperature can accelerate forming a given chemical compound, while a very high temperature will bring about its decomposition (in extremely high temperatures, atoms lose their valence electrons, undergo permanent ionization, and the matter adopts the form of plasma; in yet higher temperatures, atoms can disintegrate to protons and neutrons, or even free quarks).

Nevertheless, this depends on the favored (spontaneous) direction of the reaction course. For many reactions, the depression left to the "hump" shown in Fig. 1-3 is much shallower than the broad valley to the right of this "hump." This means that the privileged direction of the reaction is, in this case, the jumping of the ball over the "hump" from the left to the right side and, therefore, the decomposition of a chemical molecule into simpler molecules or atoms. It also happens that the spontaneous direction of a given reaction is reversed when such a reaction (e.g., the synthesis of a complex chemical compound) is coupled with another spontaneously-occurring reaction (e.g., ATP hydrolysis), as mentioned above. This is the strategy ubiquitously applied by living organisms. This coupling of not-spontaneous reactions with spontaneous reactions takes place with the aid of enzymes, which otherwise serve to lower the discussed "hump" in the form of activation energy, that is, to speed up the chemical reactions (Marks et al., 1996, 104-108; Voet et al. 1999, 288-289; Murray et al. 2012, 57-9) (see below, Chapter 3).

Incidentally, the "hump" shown in Fig. 1-3 also appears in the case of the synthesis of the nuclei of heavier elements from nuclei of lighter elements, and the latter—from hydrogen nuclei (that is, protons; neutrons in the nuclei of heavier elements originate from the combination of a proton and an electron,[7] from where their electrical neutrality derives due to the

---

[7] In reality, this is somewhat more complicated. The weak nuclear forces cause the transformation within a proton of one quark u (charge +2/3) into one quark d (charge -1/3), because of which the proton is converted into a neutron. The charge conservation principle requires an emission of boson W^+ (one of particles carrying

mutual canceling out of the positive and negative charge). This synthesis (called thermonuclear synthesis and composed of many stages) occurs in the interior of stars, where gravitation compresses hydrogen to such a degree that enormous pressures and temperatures prevail (Baggott 2015, 121-145; Sparrow 2016, 72-75), which enable the repulsion between positively charged protons that is, in this case, responsible for the existence of the discussed “hump” to be overcome. The huge amount of energy released in the process of thermonuclear synthesis testifies that, generally, despite the presence of a huge “hump,” the process of origination of heavier elements is, *per saldo*, very privileged energetically. This is confirmed by the fact that heavy elements, ejected into cold cosmic space during supernova explosions and forming so-called planetary nebulas (having, in fact, nothing to do with planets) in the terminal stages of the evolution of less massive stars, do not undergo spontaneous decay (decomposition), but participate in forming planets, moons, and (in some cases) living organisms.

The discussed “hump” does not appear in the case of the binding of the electrons in an atom—at the moment when (about 380,000 years after the Big Bang) matter in the expanding universe became sufficiently cool and rarified (diluted), free electrons bound very quickly to hydrogen and helium nuclei (nuclei of only these elements originated in the Big Bang). The remainder after this event is the so-called relic microwave background radiation—the binding of electrons in atoms (the *recombination* phenomenon) caused the universe to become transparent for electromagnetic radiation (earlier it remained in thermodynamic equilibrium with plasma containing free electrons), which since that time propagated freely in space (Baggott 2013, 102-128; Baggott 2015, 61-87).

While just after the recombination, electromagnetic radiation was very hot (high-energy, of a short wavelength), it cooled down greatly due to the expansion of the universe, and today its temperature equals only about 2.7°K.

Let us return to the most interesting in this context atoms and molecules. Already atoms or molecules of simple chemical compounds can organize themselves into relatively large, frequently macroscopic structures such as crystals and metals (Chandrasekhar 1998). Crystals are three-dimensional structures possessing one of several (strictly speaking: 14) distinguished

---

weak nuclear force) (charge +1), which in turn is decomposed to positron (anti-particle of electron) (charge +1) and electron neutrino (charge 0). The positron instantaneously annihilates with an electron (charge -1), due to which high-energy electromagnetic radiation is emitted. So, the aggregate reaction is: proton + electron → neutron + electron neutrino + hν (energy of the quantum of electromagnetic radiation).

kinds of symmetry. This symmetry results from the regular arrangement of atoms or molecules (or ions) in space, due to which the same structural unit is repeated periodically in determined spatial directions. The structure of a crystal is maintained either by ionic bonds (e.g., sodium chloride, that is, table salt) or by covalent bonds (e.g., graphite crystals made of carbon atoms) or else polar bonds (e.g., water ice). If these bonds are broken, the crystal breaks up, determining the fragility of crystals and their little resistance to deformation. Large crystals of table salt or quartz (SiO_2) are commonly known. However, microscopic crystals are ubiquitous in nature. In particular, most rocks forming the earth's crust are produced due to various geological processes—agglomerations of mixed-in and various combinations of crystals of different minerals.

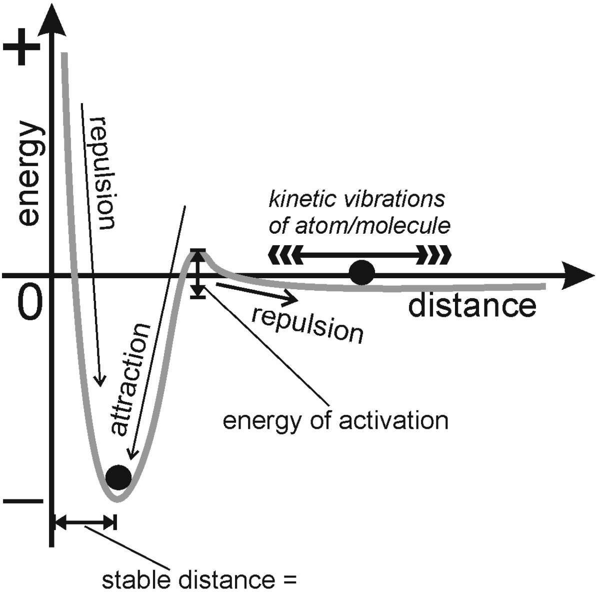

Fig. 1-3. Modified (made more realistic in relation to Fig. 1-1) dependence (gray curve) of the system energy on the distance between atom/molecule being the subject of force (little black ball) and atom/molecule being the source of the force. For the rapprochement (close up) between atoms/molecules and chemical reaction to occur, the "hump" corresponding to the activation energy must be jumped over by the ball (atom/molecule). The energy necessary for this can be delivered by thermal vibrations, movements, and collisions of atoms/molecules in the system. The rest as in Fig. 1-1.