

Philosophy of dynamic development

**Encyclopaedia of relations
and characters, their
evolution and history**

Part I: Natural laws

Marinus Dirk Stafleu

www.mdstafleu.nl

© 2018, 2019, 2022 M.D.Stafleu

Contents

- 1. Contours of a Christian philosophy of dynamic development**
 - 2. Sets**
 - 3. Symmetry**
 - 4. Periodic motion**
 - 5. Physical characters**
 - 6. Organic characters**
 - 7. Inventory of behaviour characters**
-

Chapter 1

Contours of a Christian philosophy of dynamic development

1.1 The idea of law

1.2. Relations

1.3. Characters and character types

1.4. Interlacement of characters

1.5. Natural evolution and cultural history

Encyclopaedia of relations and characters. 1. Contours of a Christian philosophy of dynamic development

1.1. The idea of law

Relations, characters, and development are key concepts in the philosophy described in this treatise. The introductory chapter 1 of the *Encyclopaedia of relations and characters* sketches some contours of a Christian philosophy of dynamic development, intended to be a twenty-first-century update of Herman Dooyeweerd's and Dirk Vollenhoven's philosophy of the cosmonomic idea.^[1] Its religious starting point is the confession that God created and sustains the world according to natural laws and normative principles. Besides the idea of law (1.1), the ideas of relations (1.2) and of characters (1.3) and their interlacements (1.4) will be systematically investigated. Sec. 1.5 explains the dynamic part of this philosophy, distinguishing natural evolution from human history.

This encyclopaedia is not alphabetically ordered, but follows the successive modal aspects as developed by Vollenhoven and Dooyeweerd and amended by the present author.

The idea of law (originally 'wetsidee' in Dutch, translated by Dooyeweerd in 1953 as 'cosmonomic idea') is the realist religious view confessing that God created the world developing according to laws and values which are invariable because He sustains them. Christians know God through Jesus Christ, who submitted himself to the *Torah*, the Law of God. The idea of natural law as used in the physical sciences since the seventeenth century confirms this idea of law. Natural laws are not *a priori* given, but partial knowledge thereof can be achieved by studying the law conformity of the creation, which, in contrast to the eternal God, is in every respect temporal, in a perennial state of dynamic development, ensuring an open future.

The modern idea of law arose together with the rise of modern science.^[2] The idea that invariant laws govern nature is relatively new. The rise of science in the seventeenth century implied the end of Aristotelian philosophy, having dominated the European universities since the thirteenth century. According to Aristotle, four causes, form, matter, potentiality, and actuality determine the essence of a thing and the way it changes naturally. Each thing, plant, or animal has the potential to realise its destiny, if not prohibited by the circumstances. The aim of medieval science was to establish the essence or nature of things, plants and animals, their position in the cosmic order, and their use for humanity.

Although essentialism is still influential, since the seventeenth century it became replaced by the search for laws. The medieval distinction of positive law, given by people, from (mostly moral) natural law, ordained by God, was hardly ever applied in science. In a scientific context, the word law was introduced about 1600.

Johann Kepler was the first to formulate laws as generalizations in the form of a mathematical relation. Apparently, Kepler's first law (planets move in elliptical paths with the sun at their focus) does not differ very much from the view, accepted since Plato, that the orbits of the celestial bodies are circular, albeit with the earth at their centre. After all, both circles and ellipses are geometrical figures. But Plato put circular motion forward as being the *essential form* of celestial motion, not as a generalization from observations and calculations. From Hipparchus of Nicaea and Claudius Ptolemy up to Nicholas Copernicus, astronomers have tried to reconcile the observed motions with a combination of circular orbits. In his elaborate analysis of Tycho Brahe's observations of the planet Mars during two decennia, Kepler found its orbit to be an ellipse, with the sun in a focus rather than at the centre. He assumed this could solve many problems for the other planets, too. Plato's circular uniform motion was a rational hypothesis (*a priori*), imposed on the analysis of the observed facts. Kepler's elliptical motion was a rational generalization of fairly accurate observations, a mathematical formulation (*a posteriori*) of a newly discovered natural law.

Since antiquity, astronomers knew very well that planets as seen from the earth have variable speeds. They applied various tricks to adapt this observed fact to the Platonic idea of uniform circular motion. Kepler accepted *changing* velocities as a fact, and connected these to the planet's varying distance to the sun as expressed in its elliptical path. He established a constant relation, his second law: as seen from the sun, a planet sweeps equal areas in equal times.

The introduction of the area law is the first instance of a method to become very successful in natural science. It related change to a constant, a magnitude that does not change. It means formulating several conservation laws, of energy, linear and angular momentum, of electric charge, etc. These laws impose restraints on any changes to occur.

During the seventeenth and eighteenth century, natural laws were considered instruments of God's government. This could be interpreted either in a rationalistic sense, when natural laws were considered both necessary and irrefutable, based on *a priori* principles (as with René Descartes or Immanuel Kant); or in a voluntaristic way, such that the world is as God willed it, but God could have made the world differently; or in an empiricist way: the laws are not irrefutable but can become known from empirical research (as with Isaac Newton, Robert Boyle, and John Locke).

Most classical physicists were faithful Christians, and many adhered to some variety of natural theology, assuming that God ordained the natural laws at the creation. At the end of the nineteenth century scientists started to take distance from this view, either because they became atheists, or because they asserted it to be theological or metaphysical, beyond the reach of physics. Therefore they avoided the metaphor of law, gradually replacing it by another expression of regularity. They never ceased to study regular patterns in nature. The word *law* remained in use mainly for the results of classical physics, in particular if expressed in a mathematical formula.

Realist scientists usually respond positively to the question of whether natural laws have an existence independent of humankind. Aimed at finding regularities, the empirical method is firmly rooted in the prevalent scientific worldview. Laws discovered

in the laboratory are declared universal, holding for the whole universe at all times. Otherwise, theories of astrophysical or biological evolution cannot be taken seriously. With the purpose of studying the law-conformity of reality, science takes the existence of laws as a point of departure not to be proved. Natural laws are not invented but discovered.

In contrast, rationalist, positivist and post-modern relativist philosophers assert that natural laws are invented by scientists. Rationalists like René Descartes and Immanuel Kant assumed natural laws to be necessary products of human thought. Positivists like Ernst Mach considered natural laws to be logical-economic constructs, intended to create some order in the otherwise chaotic reality consisting entirely of observable phenomena. And post-modern philosophers hold that natural laws are social constructs, agreed upon by interested groups of scientists. They can neither explain the coherence of the natural sciences nor the successful application of natural laws in technology and medical practice. Their opinions effectively maintain that scientists are free and autonomous law-givers even with respect to natural laws. They are at variance with naturalistic determinism, to be discussed presently, assuming that without exception everything is completely submitted to natural laws. Both contradict the realist view of laws being ordained by the Creator.

As long as natural laws were considered instruments of God's government, law conformity was easily identified with causality. The laws were considered to be causes, with God as the first cause. Immanuel Kant and his followers were of the opinion that the principle of causality is nothing but the presupposition of law conformity of all natural phenomena.

Isaac Newton assumed that the natural laws were not sufficient to explain God's interference with the creation. Without His help the solar system could not be stable. When a century later Pierre-Simon Laplace proved that all planetary movements known at the time satisfied Newton's laws, the idea that God would correct the natural laws was pushed to the back stage of theological discussions about miracles. At present, causality is seen as a relation between events, one being the cause of the other, subject to laws. But a law itself is no longer considered a cause.

In the eighteenth and nineteenth century, natural laws were often identified with laws of force, interpreted in a deterministic way. Determinism is sometimes confused with causality (and with law conformity). In physics causality always implies some form of interaction, for instance in experimental situations: if you do this, that will happen.

In the seventeenth century physical causality was accepted without criticism. This changed with the publications of David Hume in the eighteenth century. He stated that any causal connection between two events is unprovable and possibly an illusion. He believed that causality follows from a psychological motive – the need of humans and animals to predict the effects of their behaviour, making decisions possible. Immanuel Kant tried to save the rationality of causality. He stated that causality, like space and time, is a necessary category of thought, necessary because people could not order their sensorial experience otherwise in a rational way.

Determinism assumed that nature itself is ruled by causality entirely.[3] Pierre-Simon Laplace asserted that we ought to regard the present state of the universe as the effect of its anterior state and as the cause of the one that is to follow.

Believing nature to be completely determined by unchangeable natural laws, determinism has always been an article of faith rather than a well-founded theory. In the twentieth century it is refuted by the discovery and analysis of radioactivity and by the development of quantum physics and chaos theory. Scientists agree that things and events are subject to laws leaving a margin of indeterminacy, contingency or chance, individuality and uniqueness. Still, the worldview of many people makes them believe in determinism, contradicting scientific facts and methods.

Twentieth-century science has made clear that lawfulness and randomness coexist, as conditions for an open future. Many laws concern probabilities. Lawfulness does not imply determinism. It appears that laws allow of individual variation. Quantum physics, chaos theory, natural selection, and genetics cannot be understood without the assumption of random processes. Nevertheless, both law and individuality are absolutized in various worldviews. Determinism is upheld contrary to all evidence of random processes, in particular by naturalist science writers believing that everything can and must be reduced to material interactions. As far as applied to human acts, this reductionist determinism contradicts common sense and human responsibility. In contrast, some evolutionists believe that the biological evolution is a pure random process, not subject to any law. It seems difficult to accept that lawfulness and individuality do not exclude each other. In every respect, the dynamic development of reality has both a law side and a subject-and-object side, as will be discussed below.

A realistic view of natural laws not only implies their existence, but also the possibility to achieve knowledge of them.[4] It distinguishes the laws, which govern nature, being independent of mankind, from law statements as formulated by scientists. Newton's law of gravity is a law statement having various alternatives, whereas the law of gravity is a natural law ruling the planetary motions and the fall of material bodies. The first is *formulated* by Newton and dates from the seventeenth century; the latter he *discovered*, but it dates from the creation. Until the beginning of the twentieth century, Newton's law statement was considered to be true, but since Albert Einstein's general theory of relativity, it is considered approximately true. The Newtonian expression is sufficient to solve many problems, and is often preferred because of its relative simplicity. For a similar reason one may prefer Galileo's law of fall, which Newton showed to be an approximation of his own statement of the law of gravity. Realists consider a law statement as true (or approximately true) if it is a reliable expression of the corresponding natural law. Positivists maintain that a law statement is true if it confirms to observable facts. Realists would call this a *criterion* for the truth of a law statement.

1.2. Relations

The view that anything is related to everything else is far less controversial than the idea of law, but as a philosophical theme it is equally important. The diversity of temporal reality cannot be reduced to a single principle of explanation. Like a prism refracts the light of the sun into a spectrum of colours, time refracts the unity and totality of reality into a wide variety of temporal relations: among things and events; among people; between people and their environment and all kinds of objects; between individuals and associations; and between associations among each other. Also the relations of people with their God display the same diversity.

These relations can be grouped into *relation frames*, in the philosophy of the cosmomic idea called 'law spheres' or 'modal aspects'. In each relation frame, all relations among subjects and objects are governed by one or more laws or principles, characterizing the relation frame concerned. The relation frames are supposed to be mutually irreducible, yet not independent. They show a recognizable serial order. For instance, genetic relations are based on physical interaction. Kinetic relations can be projected on spatial relations, and both can be projected on quantitative relations. Each relation frame presupposes the preceding ones (the spatial frame cannot exist without numbers) and deepens them (spatial continuity expands the denumerable set of rational numbers into the continuous set of real numbers).

Because nothing can exist isolated from everything else, the relation frames constitute conditions for the existence of anything. The relation frames are also aspects of human experience, because experience is always expressed in relations. As a consequence, the relation frames are aspects of being and of experience.

This hypothesis views each relation frame as an aspect of time with its own temporal order. Simultaneity may be considered the spatial order of time, preceded by the quantitative order of earlier and later in a sequence, and succeeded by the kinetic order of uniform succession of temporal moments, the uniform motion from one temporal instant to another. In each relation frame the temporal order functions as a natural law or normative value for relations between subjects and objects, especially among subjects. The relation frames each contain a number of unchangeable natural laws or normative principles, determining the properties of relation networks of subjects and objects.

The temporal order is the law side of a relation frame. The corresponding relations constitute its subject-and-object side. Philosophically speaking, something is a *subject* if it is directly and actively subjected to a given law. An *object* is passively and indirectly (via a subject) subjected to a law. Therefore, whether something is a subject or an object depends on the context. A spatial subject like a triangle has a spatial position with respect to other spatial subjects, subjected to spatial laws. A biotic subject like a plant has a genetic relation to other biotic subjects, according to biotic laws. Something is a physical subject if it interacts with other physical things satisfying laws discovered by physics and chemistry. With respect to a given law, something is an object if it has a function for a subject of that law. Properties of subjects are not subjects themselves (physical properties like mass do not interact), but objects. Hence,

not only the subject-subject and subject-object relations, but even the concepts of a subject and of an object are relational.

The relations receive meaning from the temporal order. Serial order is a condition for quantity, and simultaneity for spatial relations. Periodic motions would be impossible without temporal uniformity. Irreversibility is a condition for causal relations; rejuvenation for life; and without purpose, the behaviour of animals would be meaningless.

Natural relations can be grouped together into six natural relation frames, preceding the normative relation frames to be discussed later.

First, putting things or events in a sequence produces a serial order. This order is expressed by numbering the members of the sequence. The sequential order of numbers gives rise to numerical differences and ratios, being quantitative subject-subject relations. The subjects of the laws belonging to the first relation frame are first of all the numbers themselves: natural and integral numbers, fractions or rational numbers and real numbers, all ordered on the same scale of increasing magnitude. Numbers are subject to laws of addition and multiplication. Everything in reality has a numerical aspect. Expressing some relation in quantitative terms (numbers or magnitudes) one arrives at an exact and objective representation. The numerical relation frame is a condition for the existence of the other frames.

The second relation frame concerns the spatial synchronous ordering of simultaneity. The relative position of two figures is the universal spatial relation between any two subjects, the spatial subject-subject relation. Whereas the serial order is one-dimensional, the spatial order consists of several mutually independent dimensions. In each dimension the positions are serially ordered and numbered, referring to numerical relations. Relative to each of these dimensions, there are many equivalent positions. Independence and equivalence are spatial key concepts, just like the relation of a whole and its parts. The spatial relation frame returns in wave motion as a medium; in physical interactions as a field; in ecology as the environment; in animal psychology as observation space, such as an animal's field of vision; and in human relations as the public domain. Magnitudes like length, distance, area or volume are spatial objects, having a quantitative function for spatial subjects.

The third relation frame records how things are moving and when events occur. Relative motion is a subject-subject relation. Motion presupposes the serial order (the diachronic order of earlier and later) and the order of equivalence (the synchronic order of simultaneity or co-existence), and it adds a new order, the uniform succession of temporal instants. Although a point on a continuous line has no unique successor, it is nevertheless assumed that a moving subject runs over the points of its path successively. Hence, relative motion is an intersubjective relation, irreducible to the preceding two. The law of uniformity concerns all kinds of relatively moving systems, including clocks. Therefore, it is possible to project kinetic time objectively on a linear scale, independent of the number of dimensions of kinetic space.

Contrary to kinetic time, the physical or chemical ordering of events is marked by irreversibility. Different events are physically related if one is the cause of the other,

and this relation is irreversible. All physical and chemical things influence each other by some kind of interaction, by exchanging energy or matter, or by exerting a force on each other. Each physical or chemical process consists of interactions. Therefore, the interaction between two things should be considered the universal physical subject-subject relation. Interaction presupposes the relation frames of quantity, space and motion.

The biotic order may be characterized by rejuvenating and ageing, both in organisms and in populations. An organism germs, ripens and rejuvenates itself by reproduction before it ages. By natural selection, populations rejuvenate themselves before they die out. For the biotic relation frame, the genetic law is universally valid. Each living being descends from another one; all living organisms are genetically related. This applies to the cells, tissues, and organs of a multicellular plant or animal as well. Descent and kinship as biotic subject-subject relations determine the position of a cell, a tissue or an organ in a plant or an animal, and of an organism in one of the biotic kingdoms. Hence, the genetic law constitutes a universal relation frame for all living beings.

The psychic order is being goal-directed. Behaviour, the universal mode of existence of all animals, is directed to future events. Recollection, recognition and expectation connect past experiences and present insight to behaviour directed to the future. Internal and external communication and processing of information are inter- and intra-subjective processes, enabling psychic functioning. Animals are sensitive for each other. By means of their senses, they experience each other as partners; as parents or offspring; as siblings or rivals; as predator or prey. By their mutual sensitivity, animals are able to make connections, between cells and organs of their body, with their environment, and with each other.

After these six natural relation frames, to be investigated in chapters 2-7, ten normative frames will be discussed in chapters 8-19.

These sixteen relation frames are not independent of each other. Except for the final one, all relation frames anticipate the succeeding frames. For instance, the set of real numbers anticipates both spatial continuity and uniform motion. Reversely, each relation frame refers back to preceding frames. The subject-subject relations of one relation frame can be projected onto those of a former one. Numbers represent spatial positions, and motions are measured by comparing distances covered in equal intervals.

These projections are often expressed as subject-object relations. A spatial magnitude like length is an objective property of physical bodies. The possibility to project physical relations on quantitative, spatial and kinetic ones forms the foundation of all physical measurements. Each measurable property requires the availability of a metric: a law for the relations to be measured and their projections. For instance, energy, force and current are generalized projections of physical interaction on quantitative, spatial and kinetic relations respectively.

1.3. Characters and character types

A realist idea of law assumes the existence of invariant natural laws and normative principles. These are not *a priori* stated as in a rationalist philosophy, but discovered as in the empirical sciences. As a consequence, law statements are fallible and revisable. Laws and principles give rise to recognizable clusters of two kinds. *General* laws for relations determine six natural relation frames and ten normative ones. Clusters of *specific* laws form what I shall call characters and character types for individual things and events, artefacts and associations, because that is what they do: characterize. Therefore, relations and characters complement each other. As will be shown, character types can be distinguished with the help of relation frames. Postponing the discussion of normativity, the present section deals with natural characters and their relations.

In the history of science a shift is observable from the search for universal laws, via structural laws, toward characters, determining processes besides structures. Even the investigation of structures is less ancient than might be expected. Largely, it dates from the nineteenth century. In mathematics, it resulted in the theory of groups, later to play an important part in physics and chemistry. Before the twentieth century, scientists were more interested in observable and measurable properties of materials than in their structure. Initially, the concept of a structure was used as an *explanans*, as an explanation of properties. Later on, structure as *explanandum*, as object for research, came to the fore. During the nineteenth century, the atomic theory functioned to explain the properties of chemical compounds and gases. In the twentieth century, atomic research was directed to the structure and functioning of the atoms themselves. Of course, people have always investigated the design of plants and animals. Yet, as an independent discipline, biology established itself not before the first half of the nineteenth century. Ethology, the science of animal behaviour, only emerged in the twentieth century.

Mainstream philosophy does not pay much attention to structures. Philosophy of science is mostly concerned with epistemological problems (for instance, the meaning of models), and with the general foundations of science. A systematic philosophical analysis of characters is wanting. This is strange, for characters form the most important subject matter of twentieth-century research, in mathematics as well as in the physical and biological sciences.

It is quite common to speak of the *structure* of thing-like individuals having a certain stability and lasting identity, like atoms, molecules, plants and animals. However, the concept of a structure is hardly applicable to individual events or processes, which are transitive rather than stable and lack a specific form. A dictionary description of the word 'structure' would be the manner in which a building or organism or other complete whole is constructed, how it is composed from spatially connected parts. In this sense, an electron has no structure, yet it is no less a characteristic whole than an atom. Depending on temperature and pressure, a solid like ice displays several different crystal structures. The typical structure of an animal, its size, appearance, and behaviour depend characteristically on its sex and age, changing considerably during its development. The structure of an individual subject is changeable, whereas its kind remains the same.

A character may be considered a specific structure. A character defined as a cluster of natural laws, values, and norms is not the structure *of*, but the law *for* individuality, indicating how an individual may characteristically differ from other individuals. The character of something includes its structure if it has one. It points out which properties it has and which propensities; how it relates to its environment; under which circumstances it exists; how it comes into being, changes and perishes. In this sense, an electron has no structure, but it has a character. Often, a character implies several structures. For example, the structure of water is crystalline below 0°C, gaseous above 100°C, and liquid in between.

A character often shares its laws (sometimes expressed as objective properties) with other characters. Electrons are characterized by having a certain mass, electric charge, magnetic moment, and lepton number. Positrons have the same mass and magnetic moment, but different charge and lepton number. Electrons and neutrino's have the same lepton number but different mass, charge and magnetic moment. Electrons, positrons and neutrino's are fermions, but so are all particles, which are not bosons. Therefore, it is never a single law, but always a specific cluster of laws that characterizes things or events of the same kind.

These clusters should not be considered as *definitions* in a logical sense. It is very well possible to define electrons objectively by their properties like mass and charge only. But this definition says very little about the laws concerning other properties, like the electron's spin, magnetic moment or lepton number. The definition does not tell that an electron is a fermion, that it has an antiparticle by which it can be annihilated, or that it belongs to the first of three generations of leptons and quarks. It does not follow from a definition that electrons have the tendency to become interlaced in atoms or metals and in events like oxidation or lightning. It does not depend on a definition that electrons have the disposition to play a part in electric and electronic appliances. Although science needs definitions, theories stating laws are far more important. At the turn of the nineteenth century electrons were identified as charged particles, starting the age of electronics, but the laws for electrons were gradually discovered in a century of painstaking experimental and theoretical research. One can never be sure of understanding the character of a thing or event completely. Human knowledge of most natural kinds is very tentative, even if it were possible to define them fairly accurately by some of their objective properties.

Besides characters, *character types* should be mentioned. An iron atom satisfies a typical character, different from that of an oxygen atom. They have also properties in common, both belonging to the character type of an atom. Because a natural kind is characterized by a cluster of laws partly shared with other kinds, it is possible to find natural classifications, like the periodic system of the chemical elements or the taxonomy of plants and animals. One may discuss the generic character of an atom or the specific character of a hydrogen atom. From a chemical point of view all oxygen atoms have the same character, but nuclear physicists distinguish various isotopes of oxygen, each having its own character. The biological taxonomy of species, genera, and other taxa corresponds to a hierarchy of character types.

A character is not a single law, but a cluster of laws. It determines both a *subjective class* of potential and actual things or events of the same kind, and an *objective*

ensemble of all possible states allowed by the character, describing the possible variation within a class.

The class of all *potential* things or events determined by a character is not restricted to a limited number, a certain place, or a period of time, but their *actual* number, place and temporal existence are usually restricted by circumstances like temperature. As a consequence of the supposition that natural laws are invariant, the class of individuals having the same character must be considered invariant as well. But the individual things and events belonging to this class are far from invariant. Any actual collection of individuals (even if it contains only one specimen) is a temporal and variable subset of the class. In an empirical or statistical sense, it is an example or a sample. A number of similar things may be connected into an *aggregate*, for instance a chemically homogeneous gas of molecules, or a population of interbreeding plants or animals of the same species. An aggregate is a temporal collection, a connected subset of the class defined above. Sometimes it is subject to a cluster of specific aggregate laws (like the gas laws). *Probability* is the relative frequency distribution of possibilities in a well-defined subset of an ensemble, subject to statistical laws. Empirical statistics is only applicable to a specific collection of individuals of the same kind.

As far as the realization of a character depends on external circumstances like the temperature of the environment, it is temporal, too. This is crucial for the understanding of astrophysical and biotic evolution.

A character considered as a cluster of laws determining the nature of a set of individuals allows of a certain amount of variation, giving room to the individuality of the things or events subject to the character. The range of individual variation is relatively small for quantitative, spatial and kinematic characters, larger for physical ones, and even more so for plants, fungi or animals. The set of possibilities governed by a character may be called an *ensemble*. An ensemble's elements are not things or events, but their objective states. An ensemble of objective possible states is as invariant as the corresponding class of potential subjective individuals. It is a set not bounded in number, space or kinetic time. It includes all possible variations of the individuals subject to the same character, whether the possibilities are realized or not. An ensemble reflects the similarity of the individuals concerned, the properties they have in common, and their possible differences, the variations allowed by the character. Variation means that a character allows of various possibilities, either in a specific or in a general sense. For instance, the character of a triangle allows of specific variation with respect to shape and magnitude, as well as its position, which is not specific. The idea of an ensemble is useful whenever an objective representation is available. In biology, the genotype of each organism is objectively projected on the sequence of nucleotides constituting its DNA-molecules, the so-called genetic code.

In three ways typical kinds are connected to the relation frames discussed above. Primarily, each kind is specifically *qualified* by the laws for one of the sixteen relation frames. The universal relation of physical interaction, specified as e.g. electric or gravitational, primarily characterizes physical and chemical things, processes and events. General and specific genetic laws constitute primarily the law clusters valid for living beings and life processes. The psychical relation frame, expressed in its goal-directed behaviour is the primary characteristic of an animal's character. For natural

characters, the qualifying relation frame is the final frame in which the things concerned can be subjects, in the succeeding frames they are objects. (This is not the case for normative characters.)

Each relation frame qualifies numerous characters. A traditional point of view acknowledges only three kingdoms of natural kinds, the physical-chemical or mineral kingdom, the plant kingdom and the animal kingdom. However, the quantitative, spatial and kinematic relation frames characterize clusters of laws as well. A triangle, for instance, has a spatial structure, oscillations and waves have primarily a kinematic character, and mathematical groups are quantitatively qualified.

Except for quantitative characters, a relation frame preceding the qualifying one constitutes the secondary characteristic, called its foundation. In fact, a character is not directly founded in a preceding frame, but in a projection of the primary (qualifying) relation frame on a preceding one. For instance, electrons are secondarily characterized by quantities, not by numbers however, but by physical magnitudes like mass, charge and lepton number. These magnitudes determine to what amount an electron is able to interact with other physical subjects. Atoms, molecules and crystals have a characteristic spatial structure as a secondary characteristic, being as distinctive as the primary (physical) one.

For each primary type one expects as many secondary types as relation frames preceding the qualifying one. For biotically qualified wholes this means four secondary types, corresponding to projections of biotic relations on the quantitative, spatial, kinematic and physical relation frames, respectively. Prokaryotes (bacteria) and some organelles in eukaryotic cells appear to be subject to law clusters founded in a quantitative projection of the biotic relation frame. Being the smallest reproductive units of life, they are genetically related by asexual multiplication, subject to the serial temporal order. In multicellular organisms, eukaryotic cells operate as units of life as well, but eukaryotic cell division starts with the division of the nucleus, having a prokaryotic structure. The character types for eukaryotic cells, multicellular undifferentiated plants and tissues in differentiated plants are founded in symbiosis, being the spatial expression of shared life.

The tertiary characteristic of a character is a *disposition*, the natural tendency or affinity of a character to become interlaced with another one, either because the individuals concerned cannot exist without each other (a eukaryotic cell cannot exist without its nucleus and organelles, and vice versa) or because an individual has a natural tendency to become a constitutive part of another one, in which it performs an objective *function*. Whereas the secondary characteristic refers to *properties*, the tertiary characteristic is usually a *propensity*. A particular molecule may or may not have an actual objective function in a plant, yet the propensity to exert such a function belongs to its specific cluster of laws. Interlacement makes characters dynamic.

Some prokaryotes have the disposition to be part of a eukaryotic cell (cell with a nucleus). In multicellular plants, a eukaryotic cell has the disposition to be a specialized part of a tissue or organ. Plants of a certain species have the propensity to occupy a certain niche, to interbreed and to be a member of a population. A

population has the propensity to change genetically, eventually to evolve into a different species.

Tertiary characteristics imply a *specific subject-object relation* between individuals of different kinds. For instance, with respect to the cluster of laws constituting the structure of an atom, the atom itself is a subject, whereas its nucleus and electrons are objects. The nucleus and the electrons interact with each other, maintaining a physical subject-subject relation, but they do not interact with the atom of which they are constitutive parts. The relation of the atom to its nucleus and electrons is a subject-object relation determined by the laws for the atom. In turn, according to their characters nuclei and electrons have a disposition, a tendency, to become encapsulated within the fabric of an atom.

In physics and chemistry, the characters of atoms and molecules are studied without taking into account their disposition to become interlaced with characters primarily characterized by a later relation frame. But biochemistry is concerned with molecules such as DNA and RNA, having a characteristic function in living cells. Like other molecules these are physically qualified and spatially founded, witness the double-helix structure as a fundamental characteristic property of DNA. But much more interesting is the part these molecules play in the production of enzymes and the reproduction of cells, which is their biotic disposition.

Interlacement is only possible if the two or more subjects involved are somehow correlated to each other. Only because electrons and protons have exactly the same electric charge with opposite sign, atomic nuclei and electrons have the disposition to form electrically neutral, quite stable atoms. Atoms having an affinity to form a molecule adapt their internal charge distribution by exchanging one or more electrons (heteropolar bond); or by sharing a pair of electrons (homopolar bond); or by an asymmetric distribution of the electrons (dipolar bond). The character of a typical event like the emission of light is correlated with the characters of the emitting atom and the emitted photon.

Hence, taking into account its propensities, the specific laws for a physical subject like a molecule not only determine its structure and physical-chemical interactions, but its full *dynamic meaning* in the cosmos as well. The theory of interlacement steers a middle course between reductionism (stressing the secondary, foundational properties of things) and holism (emphasizing the tertiary functions of things in an encompassing whole).

More is said about character interlacement in section 1.4.

Many a thing or process that we experience as an individual unit turns out to be an aggregate of individuals. I shall call an individual thing an *aggregate* if it lacks a characteristic unity. Examples are a pebble, a wood, or a herd of goats. A process is an aggregate as well. It is a chain of connected events. For a physicist or a chemist, a plant is an aggregate of widely differing molecules, but for a biologist, a plant is a characteristic whole. An aggregate consists of at least two individual things, but not every set is an aggregate. The components should show some kind of coherence.

To establish whether something is an individual or an aggregate is not an easy matter. It requires knowledge of the character that determines its individuality. It appears to be important to distinguish between homogeneous and heterogeneous aggregates. A *homogeneous aggregate* is a coherent collection of similar individuals, for instance a wave packet conducting the motion of a photon or an electron; or a gas consisting of similar molecules; or a population of plants or animals of the same species. A *heterogeneous aggregate* consists of a coherent collection of dissimilar individuals, for instance a gaseous mixture like air, or an ecosystem in which plants and animals of various species live together.

1.4. Interlacement of characters

Even apart from the existence of aggregates, an individual thing or event never satisfies the simple character type described in section 1.3. Because of its tertiary characteristic, each character is interlaced with other characters. On the one side, character interlacement is a relation of dependence, as far as the leading character cannot exist without the characters interlaced in or with it. The character of a molecule exists thanks to the characters of its atoms. On the other hand, character interlacement rests on the disposition of a thing or event to become a part of a larger whole. If it actualizes its disposition, it retains its primary and secondary character largely. Sometimes characters are so strongly interlaced that one had better speak of a 'dual character', as in the wave-particle duality (4.3).

(My distinction of two interlaced characters is an elaboration of Herman Dooyeweerd's distinction of radical types, genotypes and variability types.[5])

Several types of character interlacement should be distinguished.

In the first type of interlacement, the whole has a qualifying relation frame different from those of the characters interlaced in the whole. In chapters 4 and 5 we shall meet this phenomenon in the wave-particle duality, where the particle character is physically qualified (particles interact with each other, waves do not) and the wave character is primarily kinetic. As a measure of probability, the wave character anticipates physical interactions.

A second example is the physically qualified character of a DNA molecule being interlaced with the biotic character of a living cell. The molecule is physically qualified, the cell biotically. Their characters cannot be understood apart from each other. The cell is a biotic subject, the DNA-molecule a biotic object, the carrier of the genome, i.e., the ordered set of genes. A cell without DNA cannot exist, whereas DNA without a cell has no biotic function. The cell and the DNA molecule are mutually interlaced in a characteristic subject-object relation.

We find this type of interlacement in processes as well. For instance, the character of each biotic process is intertwined with that of a biochemical process. The behaviour of animals is interlaced with those of processes in their nervous system.

This may look like supervenience.[6] The idea of supervenience, usually applied to the relation of mind and matter, says that phenomena on a higher level are not always reducible to accompanying phenomena on a lower level. It is supposed that material states and processes invariably lead to the same mental ones, but the reverse is not necessarily the case. A mental process may correspond with various material processes.

As David Papineau observes: 'Supervenience on the physical means that two systems cannot differ chemically, or biologically, or psychologically, or whatever, without differing physically; or, to put it the other way round, if two systems are physically identical, then they must also be chemically identical, biologically identical, psychologically identical, and so on.'[7] For Papineau, this does not imply reductionism: '...I don't in fact think that *psychological* categories are reducible to physical ones.'[8] According to Papineau, in particular natural selection implies that biology and psychology are not reducible to physics, contrary to chemistry and meteorology.[9] But elsewhere Papineau writes: 'Everybody now agrees that the difference between living and non-living systems is simply having a certain kind of physical organization (roughly, we would now say, the kind of physical organization which fosters survival and reproduction)',[10] without realizing that this does not concern a physical but a biotic ordering, and that natural selection, survival and reproduction are not physical but biological concepts.

Character interlacement implies much more than supervenience, which in fact is no more than a reductionist subterfuge.

The second type of interlacement occurs if one or more characters having the same qualifying relation frame but different foundations form a single whole.

For example, the character of an atom is interlaced with the characters of its nucleus and electrons. All these characters are physically qualified. The electron's character is quantitatively founded, whereas the character of the nucleus is spatially founded like that of the atom. However, in the structure of the atom, the nucleus acts like a unit having a specific charge and mass, as if it were quantitatively founded, like the electrons. The (in this sense) quantitatively founded character of the nucleus and that of the electrons anticipate the spatially founded character of the atom. The nucleus and the electrons have a characteristic subject-subject relation, interacting with each other. Nevertheless, they do not interact with the atom of which they are a part, for they have a subject-object relation with the atom, and interaction is a subject-subject relation.

In the third type of interlacement of characters, there is no anticipation of one relation frame to another. For instance, in the interlacement of atomic groups into molecules all characters are physically qualified and spatially founded. For another example, the character of a plant is interlaced with those of its organs like roots and leaves, tissues and cells. Each has its own biotic character, interlaced with that of the plant as a whole. A comparable hierarchy of characters we find in two-, three- or more-dimensional spatial figures. A square is a two-dimensional subject having an objective function as the side of a cube.

Characters of processes are interlaced with the characters of the things involved. Individual things come into existence, change and perish in specific processes. Complex molecules come into existence by chemical processes between simpler molecules. A cell owes its existence to the never ending process called metabolism: respiration, photosynthesis, transport of water, acquisition of food, and secretion of waste, dependent on the character of the cell.

Usually processes occur on the substrate of things, and many thing-like characters depend on processes. Quantum physics assumes that even the most elementary particles are continuously created and annihilated. The question of which is the first, the thing or the process, has no better answer than that of the chicken and its egg. There is only one cosmos in which processes and things occur, generating each other and having strongly interlaced characters.

When a character is interlaced with another one, its properties change without disappearing entirely. If an atom becomes part of a molecule, its character remains largely the same, even if its distribution of charge is marginally adapted.

It is interesting that molecules have properties that the composing atoms do not have. A water molecule has properties which are absent in the molecules or atoms of hydrogen or oxygen. Water vapour is a substance completely different from a mixture of hydrogen and oxygen. This universally occurring phenomenon is called *emergence*.^[11] The theory of emergence states that at a higher level new properties emerge that do not occur at a lower level, the whole is more than the sum of its parts.^[12] In suit of Theodosius Dobzhansky, Ledyard Stebbins^[13] speaks of 'transcendence': 'In living systems, organization is more important than substance. Newly organized arrangements of pre-existing molecules, cells, or tissues can give rise to emergent or transcendent properties that often become the most important attributes of the system'^[14]. Besides the emergence of the first living beings and of humanity, Stebbins mentions the following examples: the first occurrence of eukaryotes; of multicellular animals; of invertebrates and vertebrates; of warm-blooded birds and mammals; of the higher plants; and of flowering plants. According to Stebbins, reductionism and holism are contrary approximations in the study of living beings, with equal and complementary values. It plays a part in discussions between reductionists and holists, not only in biology or in anthropology, but also in physics, where the planned construction of the superconducting supercollider (SSC) about 1990 gave rise to fierce discussions. Supporters like Steven Weinberg assumed that the understanding of elementary particles will lead to the explanation of all material phenomena. Opponents stated that for instance solid state physics owes very little to a deeper insight into sub-atomic processes.^[15]

Emergence is expressed in the symmetry of a system, for instance. A free atom has the symmetry of a sphere, but this is no longer the case with an atom being a part of a molecule. The atom adapts its symmetry to that of the molecule by lowering its spherical symmetry. The symmetry of the molecule is not reducible to that of the composing atoms. Symmetries (not only spatial ones) and symmetry breaks play an important part in physics and chemistry. 'Constraints' like initial and boundary conditions are possible causes of a symmetry break.

The theory of character interlacement leads to an improved insight into the phenomenon of emergence. When a thing gets interlaced with another one its properties change without getting lost completely. When an atom becomes a part of a molecule its character remains recognizable, even if marginally changed. In a molecule an atom may become an ion, for instance, but the nucleus and the inner electrons are hardly influenced by chemical reactions. But the molecule's properties differ considerably from those of the composing atoms. A water molecule has properties irreducible to the properties of hydrogen and oxygen. Water vapour is completely different from a gaseous mixture of hydrogen and oxygen. This widely occurring phenomenon is called the *emergence* of new properties.

This should be distinguished from the emergence of individuals belonging to a different character than those of the composing individuals. A typical example is the formation of a molecule from atoms or molecules, which is only possible if the characters of the composing atoms and molecules have the disposition to become interlaced with each other. Hydrogen and oxygen molecules, both consisting of two atoms, have the disposition to form water molecules only after they have broken their molecular bond. Within the structure of the water molecule, some properties of hydrogen atoms and oxygen atoms are recognizable, but hydrogen and oxygen lack several typical properties of a water molecule.

The phenomenon of the emergence of new characters plays an important part in the natural evolution understood as the realization of characters that did potentially but not actually exist before. Invariant characters come into actual existence if the circumstances permit it. Evolution occurs at the subject side of natural characters, not at their law side. Yet natural evolution is not a completely random process. It is a lawful dynamic development.

Scientific classification is different from the typology of characters based on universal relation frames. Classification means the formation of sets of characters based on *specific* similarities and differences. This is possible because each character is a *set* of laws, which it partly shares with other characters. A set of characters is determined by having some specific laws in common. An example of a specific classification is the biological taxonomy of living beings according to taxa like species or genera. Other examples are the classification of chemical elements in the periodic system, of elementary particles in generations of leptons and quarks, and of solids according to their crystalline structure (sections 5.3, 5.4).

Because specific classifications rest on specific laws, the chemical classification of the elements is hardly comparable to the biological classification of species. The general typology of characters developed in this encyclopaedia is applicable to widely different branches of natural science and may therefore lead to a deepened understanding of characters. Moreover, the typology provides insight into the coherence and the meaning of characters.

Each individual thing is either a subject or an object with respect to any relation frame in a way determined by its primary, secondary, and tertiary characteristics. Individual things and events present themselves in their relations to other things and events, allowing us to establish their identity.

The meaning of a thing or event can only be found in its connection with other things and events, and with the laws valid for them. In addition, the meaning of a character comes to the fore only if we take into account its interlacements with other characters. For instance, it is possible to restrict a discussion of water to its physical and chemical properties. Its meaning, however, will only become clear if we include in the discussion that water is a component of many other materials. Water plays a part in all kinds of biotic processes, and it appeases the thirst of animals and humans. Water has a symbolic function in our language and in many religions. The study of the character of water is not complete if restricted to the physical and chemical properties. It is only complete if we consider the characteristic dispositions of water as well. 'Nowhere else is the intrinsic untenability of the distinction between meaning and reality so conclusively in evidence as in things whose structure is objectively qualified.'^[16]

Likewise, the meaning of individual things and events is only clear in their lawful relations with other individuals. These relations we have subsumed in relation frames, which are of profound significance for the typology of characters. We find the meaning of the cosmos in the coherence of relation frames and of characters, and in particular in the religious concentration of humankind to the origin of the creation, as we have seen before.

The theory of characters as applied in this encyclopaedia rests on the presupposition that a character as a set of laws determines the specific nature of things or processes. Such a set leaves room for individual variation. Hence, the theory is *not deterministic*. Reality has both a law side and a subject side that cannot be separated. Both are always present. In each thing and each process, we find lawfulness besides individuality.

The theory of characters is *not essentialist* either. Essentialism means the hypostatization of being (Latin: *esse*), contrary to the view that the meaning of anything follows from its relations to everything else. According to Herman Dooyeweerd, the 'meaning nucleus' and its 'analogies' with other aspects determine the meaning of each modal aspect. However, this view incurs the risk of an essentialist interpretation, as if the meaning nucleus together with the analogies determines the 'essence' of the modal aspect concerned. In my view, the meaning of anything is determined by its relations to everything else, not merely by the universal relations as grouped into the relation frames, but by the mutual interlacements of the characters as well.

The primary characteristic of each character is not determined by a property of the thing or process itself. Rather, its *relations* with other things or processes, subject to the laws of a relation frame, are primarily characteristic of a character. Besides, the secondary and tertiary characteristics concern relations subject to general and specific laws as well. In particular the tertiary characteristic, the way by which a character is interlaced with other characters, provides meaning to the things and processes concerned. Essentialism seeks the meaning (the essence) of characters in the things and events themselves, attempting to catch them into definitions. In a relational philosophy, definitions do not have a high priority.

The theory of characters is *not reductionist*. This statement may be somewhat too strong, for there is little objection to raise against 'constitutive reductionism'. This

concept states that all matter consists of the same atoms or sub-atomic particles, and that physical and chemical laws act on all integration levels. According to Ernst Mayr, 'constitutive reductionism ... asserts that the material composition of organisms is exactly the same as found in the inorganic world. Furthermore, it posits that none of the events and processes encountered in the world of living organisms is in any conflict with the physical or chemical phenomena at the level of atoms and organisms. These claims are accepted by modern biologists. The difference between inorganic matter and living organisms does not consist in the substance of which they are composed but in the organization of biological systems.' [17] Mayr rejects every other kind of reductionism. 'Reduction is at best a vacuous, but more often a thoroughly misleading and futile, approach.' [18]

The theory of characters supposes that the laws for physical and chemical relations cannot be reduced to laws for quantitative, spatial, and kinetic relations. However, we have observed already that physical and chemical relations can be *projected* onto quantitative, spatial and kinetic relations. This explains the success of 'methodical reductionism'. It asserts the existence of laws for biotic and psychic relations transcending the physical and chemical laws. It is at variance with a stronger form of reductionism, presupposing that living organisms only differ from molecules by a larger degree of complexity, [19] whether or not supplied by the phenomena of supervenience and emergence. Richard Dawkins calls his view 'hierarchical reductionism', that '... explains a complex entity at any particular level in the hierarchy of organization, in terms of entities only one level down the hierarchy; entities which, themselves, are likely to be complex enough to need further reducing to their own component parts; and so on. It goes without saying - ... - that the kinds of explanations which are suitable at high levels in the hierarchy are quite different from the kinds of explanations which are suitable at lower levels.' Dawkins rejects the kind of reductionism '... that tries to explain complicated things *directly* in terms of the *smallest* parts, even, in some extreme versions of the myth, as the *sum* of the parts...'. I believe that the phenomenon of character interlacement gives a better representation of reality.

The theory of characters cannot be argued on *a priori* grounds. As an empirical theory, it should be justified *a posteriori*, by investigating whether it agrees with scientific results. This we shall attempt to do in the chapters to come.

1.5. Natural evolution and cultural history

The philosophy of dynamic development assumes that for understanding the temporal creation insight in relations and characters is required. Contrary to Dooyeweerd's and Vollenhoven's philosophy of the cosmonomic idea, it states that the modal aspects are not primarily modes of existence or modes of human experience (which they certainly are), but foremost determine relations. With respect to characters, I emphasize much more their tertiary propensities or dispositions than their primary and secondary properties.

This means that relations and characters are not static, but dynamic. The creation develops continuously according to constant laws, generic laws grouped into relation frames and specific laws constituting character types. The *natural* evolution (to be discussed in part I of the *Encyclopaedia*) is a rather slow process, starting some fourteen billion years ago and processing almost invisibly. In the *historical* development (part II), humankind takes an active part. Since its slow start, perhaps two hundred thousand years ago, history is more and more accelerating.

The anticipatory structure of the relation frames accounts for the start of any dynamic development (section 1.2), but the key for its understanding is the tertiary characteristic of characters, their disposition to become interlaced with each other (sections 1.3, 1.4). Because of these dispositions the numerical, spatial and kinetic characters play a decisive part in the evolution of physical and biotic characters.

Herman Dooyeweerd observes that meaning is the mode of being of all that is created.^[20] It becomes manifest whenever something new comes into being. The tertiary characteristics of natural things and events point to the possibility of the emergence of new structures with emerging new properties and propensities. It provides the original characters with meaning, their proper position in the creation. The phenomenon of disposition shows that material things like molecules have meaning for living organisms. It shows that organisms have meaning for animal life. The assumption that God's people are called from the animal world gives meaning to the existence of animals. Both evolution and history display the meaningful development of the creation, the coming into being of ever more structures. Artefacts, in particular written texts, are the most important witnesses of history.^[21] They provide history with an *objective* basis, complementing the *normative* meaning of history, provided by the directive time order in the relation frames,^[22] and the *subjective* attribution of meaning by individuals, associations and unorganised communities in their history shaping transfer of experience.^[23]

As a starting point this implies a realist religious view, confessing that God created the world according to laws which are invariant because he sustains them. We know God through Jesus Christ, who submitted himself to God's laws. Partial knowledge of his laws can be achieved by studying the law-conformity of the creation. This also implies a dynamic view of the creation, as developing continuously, in the natural evolution and in particular in human history.

The theory of evolution may be able to provide necessary conditions for the emergence of human affairs, but by no means sufficient conditions. As a religious statement, we may take the biblical message to reveal that humanity was led out of the animal world, called to be free and challenged to take responsibility for the development of God's creation. This may truly be called the *cultural mandate* of humanity. The relation frames introduced above include the relations of anyone with their true or imagined God. For Christians, these relations exist through Jesus Christ, who came into the world fulfilling God's laws for the creation, leading his people out of the animal world.

-
- [1] Stafleu 2015, 2017.
[2] Stafleu 2019, chapter 6.
[3] Stafleu 2020, chapter 9.
[4] Stafleu 2019, chapter 7.
[5] Dooyeweerd 1953-1958, III, 83, 89-96, 109-128.
[6] Charles, Lennon (eds.) 1992, 14-18.
[7] Papineau 1993, 10:
[8] Papineau 1993, 44.
[9] Papineau 1993, 47; Plotkin 1994, 52, 55; Sober 1993, 73-77
[10] Papineau 1993, 122,
[11] Stafleu 2018.
[12] Popper 1972, 242-244, 289-295; 1974, 142; Popper, Eccles 1977, 14-31; Mayr 1982, 63-64.
[13] Stebbins 1982, 161-167
[14] Stebbins 1982, 167.
[15] Anderson 1995; Weinberg 1995; Kevles 1997; Cat 1998.
[16] Dooyeweerd 1953-1958, III, 107.
[17] Mayr 1982, 60.
[18] Mayr 1982, 63.
[19] Dawkins 1986, 13.
[20] Dooyeweerd *NC*, I, 4.
[21] Stafleu 2011, chapter 3.
[22] Stafleu 2011, chapter 1.
[23] Stafleu 2011, chapter 2.
-

Chapter 2

Sets

2.1. Sets and natural numbers

2.2. Extension of the quantitative relation frame

2.3. Groups as characters

2.4. Ensemble and probability

2.1. Sets and natural numbers

Plato and Aristotle introduced the traditional view that mathematics is concerned with numbers and with space. Since the end of the nineteenth century, many people thought that the theory of sets would provide mathematics with its foundations. According to Ernst Zermelo, for instance, 'set theory is that branch of mathematics whose task is to investigate mathematically the fundamental notions of 'number', 'order', and 'function' taking them in their pristine, simple form, and to develop thereby the logical foundations of all of arithmetic and analysis.'^[1] Since the middle of the twentieth century, the emphasis is more on structures and relations: 'Mathematics is the *deductive* study of structures'.^[2]

In chapter 1, I defined a natural character as a set of natural laws, determining a class of individuals and an ensemble of possible variations. Because classes, ensembles and aggregates are sets, it is apt to pay attention to the theory of sets.

In sections 2.1-2.2 it will appear that with each set at least two relation frames are concerned, according to the tradition to be called the quantitative and the spatial frames. The elements of a quantitative or discrete set can be *counted*, whereas the parts of a spatial or continuous set can be *measured*. Section 2.3 discusses some quantitatively qualified characters, in particular groups. Section 2.4 relates the concept of an ensemble with that of probability.

Numbers constitute the relation frame for all sets and their relations. A set consists of a *number* of elements, varying from zero to infinity, whether denumerable or not, but there are sets of numbers as well. What was the first, the natural number or the set? Just as in the case of the chicken and the egg, an empiricist may wonder whether this is a meaningful question. We have only one reality available, to be studied from within. In the cosmos, we find chickens as well as eggs, sets as well as numbers. Of course, we have to start our investigations somewhere, but the choice of the starting point is relatively arbitrary. Rejecting the view that mathematics is part of logics, I shall treat sets and numbers in an *empirical* way, as phenomena occurring in the cosmos.

At first sight, the concept of a set is rather trivial, in particular if the number of elements is finite. Then the set is denumerable and countable; we can number and count the elements. It becomes more intricate if the number of elements is not finite yet

denumerable (e.g., the set of integers), or infinite and non-denumerable (e.g., the set of real numbers). Let us start with finite sets.

Sets concern all kinds of elements, hence they are closer to concrete reality than numbers. (As a human act, collecting of fruits etc. is one of the oldest means to provide food.) *Quantity* or *amount* is a universal aspect of sets. It is an abstraction like the other five natural relation frames announced in section 1.2. For instance, by isolating the natural numbers we abstract from the equivalence relation.

Equivalence is reflexive ($A \circ A$), symmetric (if $A \circ B$, then $B \circ A$), and transitive (if $A \circ B$ and $B \circ C$, then $A \circ C$). On the other hand, numbers are subject to the order of increasing magnitude. This *sequential order* is exclusive (either $a > b$, or $b > a$), asymmetric (if $a > b$, then $b < a$), not reflexive (a is not larger or smaller than a), but it is transitive (if $a > b$ and $b > c$, then $a > c$). For numbers, the equivalence relation reduces to equality: $a = a$; if $a = b$ then $b = a$; if $a = b$ and $b = c$ then $a = c$. Usually equivalence is different from equality, however.

Two sets A and B are *numerically equivalent* if their elements can be paired one by one, such that each element of A is uniquely combined with an element of B and conversely. All sets being numerically equivalent to a given finite set A constitute the *equivalence class* $[n]$ of A . One element of this class is the set of natural numbers from 1 to n . All sets numerically equivalent to A have the same number of elements n . I consider the *cardinal number* n to be a discoverable property (e.g., by counting or calculating) of each set that is an element of the equivalence class $[n]$. The numbers $1..n$ function as *ordinal numbers* or indices to put the elements of the set into a sequence, to number and to count them. It is a law of arithmetic that in whatever order the elements of a finite set are counted, their number will always be the same.

Sometimes the elements of an infinite set can also be numbered. Then we say that the set is infinite yet denumerable. The set of even numbers, e.g., is both infinite and denumerable. As a set of indices, the natural numbers constitute a universal relation frame for each denumerable set. However, the set of natural numbers is a character class as well. It is relevant to distinguish relation frames from characters, but they are not separable.

Giuseppe Peano's axioms formulate the laws for the sequence N of the natural numbers. The axioms apply the concepts of sequence, successor and first number, but it does not apply the concept of equivalence. According to Peano, the concept of a successor is characteristic for the natural numbers:

1. N contains a natural number, indicated by 0.
2. Each natural number a is uniquely joined by a natural number a^+ , the successor of a .
3. There is no natural number a such that $a^+ = 0$.
4. From $a^+ = b^+$ follows $a = b$.

5. If a subset M of N contains the element 0, and besides each element a its successor a^+ as well, then $M = N$.

Peano took 1 to be the first natural number. Nowadays one usually starts with 0, to indicate the number of elements in the zero set. In the decimal system $0^+ = 1$, $1^+ = 2$, $2^+ = 3$, etc. In the binary system $0^+ = 1$, $1^+ = 10$, $10^+ = 11$, $11^+ = 100$, etc. From axiom 2 it follows that N has no last number.

The fifth axiom states that the set of natural numbers is unique. The sequence of even numbers satisfies the first four axioms but not the fifth one. The transitive relation 'larger than' is now applicable to the natural numbers. For each a , $a^+ > a$. If $a > b$ and $b > c$, then $a > c$, for each trio a, b, c .

On the axioms rests the method of proof by complete induction: if $P(n)$ is a proposition defined for each natural number $n > a$; and $P(a)$ is true; and $P(n^+)$ is true if $P(n)$ is true; then $P(n)$ is true for any $n > a$.

The natural numbers constitute a character class. Their character, expressed by Peano's axioms, is primarily quantitatively characterized. It has no secondary foundation for lack of a relation frame preceding the quantitative one. Because the first relation frame does not have objects, it makes no sense to introduce an ensemble of possibilities besides any numerical character class.

As a tertiary characteristic, the set of natural numbers has the disposition to expand itself into other sets of numbers (2.2).

The laws of addition, multiplication, and raising powers are derivable from Peano's axioms.[3] The class of natural numbers is complete with respect to these operations. If a and b are natural numbers, then $a+b$, $a \cdot b$ and a^b are natural numbers as well. This does not always apply to subtraction, division or taking roots, and the laws for these inverse operations do not belong to the character of natural numbers.

In 1931, Gödel proved that any system of axioms for the natural numbers allows of unprovable statements.[4] This means that Peano's axiom system is not logically complete.

Using the two ordering relations discussed, 'larger than' and 'numerical equivalence', we can order all denumerable sets. All sets having n elements are put together in the equivalence class $[n]$, whereas the equivalence classes themselves are ordered into a sequence. The sets in the equivalence class $[n]$ have no more in common than the number n of their elements.

The set of natural numbers is the oldest and best-known set of numbers. Yet it is still subject to active mathematical research, resulting in newly discovered regularities, such that '... the differences between mathematics and empirical science have been vastly exaggerated.'[5] 'Even arithmetic contains randomness. Some of its truths can

only be ascertained by experimental investigation. Seen in this light it begins to resemble an experimental science.’[6]

Some theorems relate to prime numbers. Euclid proved that the number of primes is unlimited. An arithmetical law says that each natural number is the product of a unique set of primes. Several other theorems concerning primes are proved or conjectured. For instance, Christian Goldbach’s conjecture, saying that each even number can be written as the sum of two primes in at least one way, dates from 1742, but is at the end of the twentieth century neither proved nor disproved.

In many ways, the set of primes is notoriously irregular. There is no law to generate them. If one wants to find all prime numbers less than an arbitrarily chosen number n , this is only possible with the help of an empirical elimination procedure, known as Eratosthenes’ sieve. From the set of natural numbers 1 to n , starting from 3 the sieve eliminates all even numbers, all triples, all quintets except 5, (the quartets and sextuplets have already been eliminated), all numbers divisible by 7 except 7 itself, etc., until one reaches the first number larger than the square of n . Then all primes smaller than n remain on the sieve. For very large prime numbers, this method consumes so much time that the resolution of a very large number into its factors is used as a key in cryptography.

There are much more sequences of natural numbers subject to a characteristic law or prescription. An example is the sequence of Fibonacci (Leonardo of Pisa, circa 1200). Starting from the numbers 1 and 2, each member is the sum of the two preceding ones: 1, 2, 3, 5, 8, 13, ... This sequence plays a part in the description of several natural processes and structures.[7]

The relation of a set to its elements is a numerical law-subject relation, for a set is a number of elements. By contrast, the relation of a set to its subsets is a whole-part relation that can be projected on a spatial figure having parts. A subset is not an element of the set, not even a subset having only one element. A set may be a member of another set. For instance, the numerical equivalence class $[n]$ is a set of sets. A well-known paradox arises if a set itself satisfies its prescription, being an instance of self-reference. The standard example is the set of all sets that do not contain themselves as an element. Restricting the prescription to the *elements* of the set may preclude such a paradox. This means that a set cannot be a member of itself, not even if the elements are sets themselves.[8] The number of subsets is always larger than the number of elements, a set of n elements having 2^n subsets. A set contains an infinite number of elements if it is numerically equivalent to one of its subsets. For instance, the set of natural numbers is numerically equivalent to the set of even numbers and is therefore infinite. However, the set of all subsets of a given set A should not be confused with the set A itself.

Overlapping sets have one or more elements in common. The intersection A or B of two sets is the set of all elements that A and B have in common. The empty set or zero set is the intersection of two sets having no elements in common. Hence, there is only one zero set. It is a subset of all sets. (This is a consequence of the axiom stating that

two sets are identical if they have the same elements.) If a set is considered a subset of itself, each set has trivially two subsets. (An exception is the zero set, having only itself as a subset).

The union A and B of two sets looks more like a spatial than a numerical operation. Only if two sets have no elements in common, the total number of elements is equal to the sum of the numbers of elements of the two sets apart. Otherwise, the sum is less: If $n(A)$ is the number of elements of A , then $n(A \text{ and } B) = n(A) + n(B) - n(A \text{ or } B)$.

Hence, even for denumerable sets the numerical relation frame is not sufficient. At least a projection on the spatial relation frame is needed. This is even more true for non-denumerable sets (2.2).

Some sets are really spatial, like the set of points in a plane contained within a closed curve. As its magnitude, one does not consider the number of points in the set, but the area enclosed by the curve. The set has an infinite number of elements, but a finite *measure*. A measure is a magnitude referring to but not reducible to the numerical relation frame. It is a number with a unit, a proportion.

This measure does not deliver a numerical relation between a set and its *elements*. It is not a measure of the number of elements in the set. A measure is a quantitative relation between sets, e.g., between a set and its subsets. If two plane spatial figures do not overlap but have a boundary in common, the intersection of the two point sets is not zero, but its measure is zero. The area of the common boundary is zero. In general, only subsets having the same dimension as the set itself have a non-zero measure. We shall see in section 2.2 that all numbers (including the natural ones) determine relations between sets. Only the natural numbers relate countable sets with their elements as well.

Integral calculus is a means to determine the measure of a spatial figure, its length, area or volume. In section 2.4, we discuss probability being a measure of subsets of an ensemble.

For each determination of a measure, each measurement, *real* numbers are needed. That is remarkable, for an actual measurement can only yield a *rational* number (2.2).

The number 2 is natural, but it is an integer, a fraction, a real number and a complex number as well. Precisely formulated: the number 2 is an element of the sets of natural numbers, integers, fractions, real and complex numbers. This leads to the conjecture that we should not conceive of the character of natural numbers to determine a class of *things*, but a class of *relations*. The natural numbers constitute a universal relation frame for all denumerable sets. Peano's formulation characterizes the natural numbers by a sequence, that is a relation as well. We shall see that the integers, the rational, real, and complex numbers are definable as relations. In that case, it is not strange that the number 2 answers different types of relations. A quantitative character determines a *set* of numbers, and a number may belong to several sets, each with its own character. The number 2 is a knot of relations, which is characteristic

for a 'thing'. On the other hand, it responds to various characters, and that is not very 'thing-like'.

However, it is not fruitful to quarrel extensively about the question of whether a number is *essentially* a thing or a relation. Anyway, numbers are individual subjects to quantitative laws.

2.2. Extension of the quantitative relation frame

The natural numbers satisfy laws for addition, multiplication, and taking powers, by which each pair of numbers generates another natural number. The inverse operations, subtraction, division and taking roots, are not always feasible within the set of natural numbers. Therefore, mathematics completes the set of natural numbers into the set of integers and the set of rational numbers. Put otherwise, the set of natural numbers has the disposition of generating the sets of integral numbers and of rational numbers. There remain holes in the set of rational numbers, there are still magnitudes (like the ratio of the diagonal of a square to one of its sides) which cannot be expressed in rational numbers. These holes are to be filled up by the irrational numbers. The various number sets constitute a hierarchy, consisting of the sets of, respectively, natural, integral, rational, real, and complex numbers. Each of these sets has a separate character. A natural number belongs to each of these sets. A negative integer belongs to all sets except that of the natural numbers. A fraction like $\frac{1}{2}$ belongs to each set except the first two sets.

Before discussing the character of integral, rational, real, and complex numbers, I mention some properties.

Starting from its element 0, the set of integral numbers can also be defined by stating that each element a has a unique successor a^+ as well as a unique predecessor a^- , if $(a^+)^- = a$.^[9] Each integer is the difference between two natural numbers. Several pairs may have the same difference. Hence, each integral number corresponds to the equivalence class of all pairs of integers having the same difference. Likewise, each rational number corresponds to the equivalence class of all pairs of natural, integral *and* rational numbers having the same proportion *or* the same difference. If we do not want to relapse into an infinite regress, we had better not identify (in the way of an essentialist definition) an integer or a rational number with an equivalence class. The *meaning* of a number depends on its *relation* to all other numbers and the disposition of numbers to generate other numbers.^[10]

The laws for addition, subtraction, multiplication, and division are now valid for the whole domain of rational numbers, including the natural and integral numbers. It can be proved that the sum, the difference, the product and the quotient of two rational

numbers (excluding division by 0) always gives a rational number. Hence, the set of rational numbers is complete or closed with respect to these operations. After the recognition of the natural numbers as a set of indices, the introduction of negative and rational numbers means a further abstraction with respect to the concept of a set. A set cannot have a negative number of elements, and halving a set is not always possible. The integral and rational numbers are not numbers of sets, but quantitative relations between sets. They are applicable to other domains as well, for instance to the division of an apple. The universal applicability of the quantitative relation frame requires the extension of the set of natural numbers.

Meanwhile, two properties of natural numbers have been lost. Neither the integral nor the rational numbers have a first one, though the number 0 remains exceptional in various ways. Moreover, a rational number has no unique successor. Instead of succession, characteristic for the natural and integral numbers, rational numbers are subject to the order of increasing magnitude. This corresponds to the quantitative subject-subject relations (difference and proportion): if $a > b$ then $a - b > 0$, and if moreover $b > 0$ then $a/b > 1$. For each pair of rational numbers, it is clear which one is the largest, and for each trio, it is clear which one is between the other two.

The classes of natural numbers, integers and rational numbers each correspond to a character of their own. These characters are primarily qualified by quantitative laws and lack a secondary characteristic. We shall see that the character of the rational numbers has the (tertiary) disposition to function as the metric for the set of real numbers.

The road from the natural numbers to the real ones proceeds via the rational numbers. A set is denumerable if its elements can be put in a sequence. Georg Cantor demonstrated that all denumerable infinite sets are numerically equivalent, such that they can be projected on the set of natural numbers. Therefore, he accorded them the same cardinal number, called aleph-zero, after the first letter of the Hebrew alphabet. Cantor assumed this 'transfinite' number to be the first in a sequence, aleph 0, aleph 1, aleph 2 ... , where each is defined as the 'power set' of its predecessor, i.e., the set of all its subsets.

The rational numbers are denumerable, at least if put in a somewhat artificial order. The infinite sequence $1/1; 1/2, 2/1; 1/3, 2/3, 3/1, 3/2; 1/4, 2/4, 3/4, 4/1, 4/2, 4/3; 1/5, \dots$ including all positive fractions is denumerable. In this order it has the cardinal number of aleph 0. However, this sequence is not ordered according to increasing magnitude.

In their natural (quantitative) order of increasing magnitude, the fractions lay close to each other, forming a *dense* set. This means that no rational number has a unique successor. Between each pair of rational numbers a and b there are infinitely many others. (If $a < b$ then $a < a + c(b - a) < b$, for each rational value of c with $0 < c < 1$.) In their natural order, rational numbers are not denumerable, although they are denumerable in a different order. Contrary to a finite set, whether an infinite set is countable may depend on the order of its elements.

Though the set of fractions in their natural order is dense, it is still possible to put other numbers between them. These are the irrational numbers, like the square root of 2 and pi. According to the tradition, Pythagoras or one of his disciples discovered that he could not express the ratio of the diagonal and the side of a square by a fraction of natural numbers. Observe the ambiguity of the word 'rational' in this context, meaning 'proportional' as well as 'reasonable'. The Pythagoreans considered something reasonably understandable, if they could express it as a proportion. They were deeply shocked by their discovery that the ratio of a diagonal to the side of a square is not rational. The set of all rational and irrational numbers, called the set of real numbers, turns out to be non-denumerable. I shall argue presently that the set of real numbers is continuous, meaning that no holes are left to be filled.

Only in the nineteenth century, the distinction between a dense and a continuous set became clear.^[11] Before, continuity was often defined as infinite divisibility, not only of space. For ages, people have discussed about the question whether matter would be continuous or atomic. Could one go on dividing matter, or does it consist of indivisible atoms? In this case, *tertium non datur* is invalid. There is a third possibility, generally overlooked, namely that matter is dense.

Even the division of space can be interpreted in two ways. The first was applied by Zeno of Elea when he divided a line segment by halving it, then halving each part, etc. This is a *quantitative* way of division, not leading to continuity but to density. Each part has a rational proportion to the original line segment. Another way of dividing a line is by intersecting it by one or more other lines. Now it is not difficult to imagine situations in which the proportion of two lines segments is irrational. (For instance, think of the diagonal of a square.) This *spatial* division shows the existence of points on the line that quantitative division cannot reach.

In 1892, Georg Cantor proved by his famous diagonal method that the set of real numbers is not denumerable. Cantor indicated the infinite amount of real numbers by the cardinal number C . He posed the problem of whether C equals aleph 1, the transfinite number succeeding aleph 0. At the end of the twentieth century, this problem was still unsolved. Maybe it is not solvable. Maybe it is an independent axiom.

A theorem states that each irrational number is the limit of an infinite sequence or series of rational numbers, e.g., an infinite decimal fraction. A *sequence* is an ordered set of numbers (a, b, c, \dots). Sometimes an infinite sequence has a limit, for instance, the sequence $1/2, 1/4, 1/8, \dots$ converges to 0. A *series* is the sum of a set of numbers ($a+b+c+\dots$). An infinite series too may have a limit. For instance, the series $1/2+1/4+1/8+\dots$ converges to 1.

This seems to prove that the set of real numbers can be reduced to the set of rational numbers, like the rational numbers are reducible to the natural ones, but that is arguable. Any procedure to find these limits cannot be done in a countable way, not consecutively. This would only lead to a denumerable (even if infinite) amount of real numbers. By multiplying a single irrational number like pi with all rational numbers, one finds already an infinite, even dense, yet denumerable subset of the set of real

numbers. Also the introduction of real numbers by means of 'Cauchy-sequences' only results in a denumerable subset of real numbers. To arrive at the set of *all* real numbers requires a non-denumerable procedure. But then we would use a property of the real numbers (not shared by the rational numbers) to make this reduction possible. And this appears to result in circular reasoning, begging the question. (A Cauchy sequence, named after Augustin-Louis Cauchy, is a *sequence* whose *elements* become arbitrarily close to each other as the sequence progresses.)

Suppose we want to number the points on a straight or curved line, would the set of rational numbers be sufficient? Clearly not, because of the existence of spatial proportions like that between the diagonal and the side of a square, or between the circumference and the diameter of a circle. Conversely, is it possible to project the set of rational numbers on a straight line? The answer is positive, but then many holes are left. By plugging the holes, we get the real numbers, in the following empirical way. (This procedure differs from the standard treatment of real numbers.[12])

Consider a continuous line segment AB . We want to mark the position of each point by a number giving the distance to one of the ends. According to the axiom of Cantor-Dedekind, there is a one-to-one relation between the points on a line and the real numbers. These numbers include the set of infinite decimal fractions that Cantor proved to be non-denumerable. Hence, the set of points on AB is not denumerable. If we mark the point A by 0 and B by 1, each point of AB gets a number between 0 and 1. This is possible in many ways, but one of them is highly significant, because we can use the rational numbers to introduce a *metric*. We assign the number 0.5 to the point halfway between A and B , and analogously for each rational number between 0 and 1. (This is possible in a denumerable procedure). Now we define the real numbers between 0 and 1 to be the numbers corresponding one-to-one to the points on AB . These include the rational numbers between 0 and 1, as well as numbers like $\pi/4$ and other limits of infinite sequences or series. The irrational numbers are surrounded by rational numbers (forming a dense set) providing the metric for the set of real numbers between 0 and 1.

A set is called *continuous* if its elements correspond one-to-one to the points on a line segment. It is not difficult to prove that the points on two different line segments correspond one-to-one to each other. On the one hand, the continuity of the set of real numbers anticipates the continuity of the set of points on a line. On the other hand, it allows of the possibility to project spatial relations on the quantitative relation frame.

The set of real numbers is continuous because it does not contain any holes, contrary to the dense set of rational numbers. The above mentioned procedures to divide a segment of a line, or to project the real numbers between 0 and 1 on a line segment, justify the following statement. Divide the ordered set of numbers into two subsets A and B , such that each element of A is smaller than each element of B . Then there is an element x of A or of B , that is larger than all (other) elements of A and smaller than all (other) elements of B . This is called (Richard) Dedekind's cut. The boundary element x can be rational or irrational. This means that the set of real

numbers is *complete* with respect to the order of increasing magnitude, there are no holes left.

The set of real numbers constitutes the quantitative relation frame for spatial relations. Spatial measures like distance, length, area, and angle are projections on sets of numbers. To express spatial relations as magnitudes requires real numbers. Besides spatial relations, kinetic, physical, and chemical magnitudes are expressed in real numbers. This is remarkable, considering the practice of measuring. Each measurement is inaccurate to a certain extent. Therefore, a measurement never yields anything but a rational number. Moreover, computers rely on rational numbers. Hence, the use of real numbers has a *theoretical* background. The assumption that a magnitude is continuously variable is not empirically testable.

2.3. Groups as characters

Mathematics knows several structures that I consider quantitative characters. Among these, the character of mathematical groups expressing symmetries is of special interest to natural science.

A group is a set of elements that can be combined such that each pair generates a third element. In the world of numbers, such combinations are addition or multiplication. Because of the mutual coherence of the elements, a group may be considered an *aggregate*. The phenomenon of isomorphy allows of the projection of physical states of affairs on mathematical ones.

In 1831, Évariste Galois introduced the concept of a group in mathematics as a set of elements satisfying the following four axioms.

1. A combination procedure exists, such that each pair of elements A and B unambiguously generates a new element AB of the group.
2. The combination is *associative*, i.e., $(AB)C = A(BC)$, to be written as ABC .
3. The group contains an element I , the *identity element*, such that for each element A of the group, $AI = IA = A$.
4. Each element A of the group has an *inverse element* A' , such that $A'A = AA' = I$.

A group is called Abelian (after N.H. Abel) or *commutative* if for each A and B , $AB = BA$. This is by no means always the case.

It can be proved that each group has only one identity element, that each element has only one inverse element, and that $I' = I$. Each group has at least one element, I . (Hence, the zero set is not a group.) If a subset of the group is a group itself with the same combination rule, then both groups share the identity element.

It is clear that the elements of a group are mutually strongly connected. They have a relation determined by the group's character, to be defined as AB' , the combination of A with the inverse of B . The relation of an element A to itself is $AA' = I$, A is identical with itself. Moreover, $(AB')' = BA'$, the inverse of a relation of A to B is the relation of B to A .

Each group is complete. If we combine each element with one of them, A , the identity element I is converted into A , and the inverse of A becomes I . The new group as a whole has exactly the same elements as the original group. Hence, the combination of all elements with an element A is a transition of a group into itself. It expresses a symmetry, in which the relations between the elements are invariant. [13]

If two groups can be projected one-to-one onto each other, they are called *isomorphic*. Two groups are isomorphic if their elements can be paired such that $A_1B_1 = C_1$ in the first group implies that $A_2B_2 = C_2$ for the corresponding elements in the second group and conversely. This may be the case even if the combination rules in the two groups are different. The phenomenon of isomorphy means that the character of a group is not fully determined by the axioms alone. Besides the combination rule, at least some of the group's elements must be specified, such that the other elements are found by applying the combination rule.

Isomorphy allows of the projection of one group onto the other one. It leads to the interlacement of various characters, as we shall see in the next few chapters. Hence, isomorphy is a tertiary property of groups, a disposition.

The elements of a group may be numbers, or number vectors, or functions of numbers, or operators transforming one function into another one. In physics, groups were first applied in relativity theory, and since 1925 in quantum physics and solid state physics. Not to everyone's delight, however, see e.g. John Slater about the 'Gruppenpest': '... it was obvious that a great many other physicists were as disgusted as I had been with the group-theoretical approach to the problem.' [14]

Let us first cast a glance at some number groups.

The first examples of groups we find in sets of numbers. Adding or multiplying two numbers yields a third number. With respect to addition, 0 is the identity element, for $a+0=0+a=a$ for any number a . Besides 0, it is sufficient to introduce the number 1 in order to generate the whole group of *integral* numbers: $1+1=2$, $1+2=3$, etc. The inverse of an integer a is $-a$, for $a+(-a)=0$. The relation of a and b is the difference $a-b$. Instead of beginning with 1, we could also start with 2 or with 3, generating the groups of even numbers, threefold numbers, etc. Each of these subgroups is complete and isomorphic with the full group of integers.

The rational, real, and complex numbers, too, each form a complete addition group, but the natural numbers do not constitute a group. The natural numbers form a class with a quantitatively qualified character, expressed by Giuseppe Peano's axioms (2.1) or an alternative formulation. However, this character does not include the laws for subtraction and division, because the set of natural numbers is not complete with respect to these operations.

The mentioned groups are infinite, but there are finite groups of numbers as well. Two numbers are called 'congruent modulo x ' if their difference is an integral multiple of x . The four integral numbers 0, 1, 2, and 3 form a group with the combination rule of 'adding modulo 4'. If the sum of two elements would exceed 3, we subtract 4 (hence $3+2=1$, and $4=0$). If the difference would be less than 0 we add 4 (hence $2-3=3$). This group is isomorphic to the rotation group representing the symmetry of a square. Likewise, the infinite but bounded set of real numbers between 0 and 2π constituting the addition group modulo 2π is isomorphic to the rotation group of a circle.

In the multiplication of numbers, 1 is the identity element. For each number a , $1 \cdot a = a \cdot 1 = a$. The inverse of multiplication is division, $1/a$ being the inverse of a . The relation between a and b is their proportion a/b . Introducing the positive integers as elements, we generate the group of positive rational numbers. The full set of rational numbers is not a group with multiplication as a combination rule, because division by 0 is excluded, hence 0 would be an element without an inverse. Likewise, the set of positive real numbers is a multiplication group, but the set of all real numbers is not.

Addition and multiplication are connected by the distributive law: $(a+b)c=ac+bc$. Some addition and multiplication groups are combined into a structure called a *field*, having two combination rules. Three number fields with an infinite number of elements are known, respectively having the rational, real, and complex numbers as elements. Because division by zero is excepted, I do not consider a field a character, but an interlacement of two characters.

For a given positive real number a all numbers a^n form a multiplication group, if the variable exponent n is an element of the set of integral, rational, real, or complex numbers. The character of this group depends on the fact that the integral, rational, real, or complex numbers each form an addition group. The combination of two elements of the power group, the product of two powers, arises from the addition of the exponents: $a^n \cdot a^m = a^{(n+m)}$. The identity element of this multiplication group is $a^0=1$ and the inverse of a^n is a^{-n} . The group is isomorphic with the addition group of integral, rational, real, or complex numbers.

Each addition group, multiplication group, and power group is a character class. Their characters are primarily numerically qualified. They have no secondary foundation, and their tertiary disposition is to be found in many interlacements with spatial, kinetic, physical, and chemical characters (chapters 3-5).

Sometimes, a variable spatial, kinetic, or physical property or relation turns out to have the character of a group, isomorphic to a group of numbers. If that magnitude may be

positive as well as negative (e.g., electric charge) this is an addition group. If only positive values are allowed (e.g., length or mass), it is a multiplication group. In other cases, the property or relation is projected on a vector group (e.g., velocity or force). Isomorphy is not trivial. Sometimes one has to be content with a weaker projection, called homomorphy. An example is Friedrich Mohs' scale, indicating the relative hardness of minerals by numbers between 0 and 10: if A is harder than B , A gets a higher numeral. It makes no sense to add or to multiply these ordinal numbers.

If a property or relation is isomorphic to a group of numbers, it is called *measurable*.^[15] Since antiquity, its importance is expressed in the name *geometry* for the science of space. The law expressing the measurability of a property or relation is called its *metric*. Measurable magnitudes isomorphic to a number group allow us to perform calculations, which is the basis of the mathematization of science.

Measurability is not trivial. A physical magnitude is only measurable if a *physical* combination procedure is available, which can be projected on a *quantitative* one. To establish whether this is the case requires experimental and theoretical research.

Relativity theory demonstrates that a kinematic or physical combination rule in a group cannot always be projected on addition or multiplication. In the case of one-dimensional motion, the combination rule for two velocities v and w is not $v+w$ (as in classical kinematics), but $(v+w)/(1+vw/c^2)$, where c equals the speed of light in a vacuum. For small velocities, the numerator is about 1, and the classical formula is approximately retained. The meaning of this formula becomes clear by taking v or w equal to c : if $w=c$, the combination of v and w equals c . A combination of velocities smaller than that of light never yields a velocity exceeding the speed of light. The formula also expresses the fact that the speed of light has the same value with respect to each moving system. In the Lorentz-group the speed of light is the unit of speed ($c=1$), having the same value in all inertial frames (3.3). This, of course, was the starting point for the formula's derivation. The elements of the group are all velocities which magnitude is at most the speed of light.

Vectors play an important part in mathematics and in physics. With all kinds of vectors, like position, displacement, velocity, force, and electric field strength, the numerical vector character is interlaced. Spatial, kinetic and physical vectors are isomorphic with number vectors. Besides, mathematics acknowledges tensors, matrices and other structures.

A number vector $\mathbf{r}=(x,y,z,\dots)$ is an ordered set of n real numbers, called the components of the vector. Number vectors are subject to laws for addition and subtraction, by applying these operations to the components apart. For example, the difference between two vectors is $\mathbf{r}_2-\mathbf{r}_1=(x_2-x_1,y_2-y_1,z_2-z_1, \dots)$. The set of all number vectors with the same number of components is an addition group, the zero vector $\mathbf{0}=(0,0,0, \dots)$ being its identity element. Each vector multiplied by a real number yields a new vector within the group. If c is an ordinary number, $\mathbf{b}=\mathbf{c}\mathbf{a}=c(a_1,a_2,a_3, \dots)=(ca_1,ca_2,ca_3, \dots)$.

However, division by zero being excluded, this does not define a combination procedure for a group.

Besides the zero vector as the identity element, the set contains unit vectors. In a unit vector, one component is equal to 1, the others are equal to 0. Any vector can be written as a linear combination of the unit vectors: for each number vector, $\mathbf{a}=(a_1, a_2, a_3, \dots)=a_1(1,0,0, \dots)+a_2(0,1,0, \dots)+a_3(0,0,1, \dots)+ \dots$. The set of unit vectors constitutes the *base* of the set of vectors. For number vectors, the base is unique, but in other cases, a group of vectors may have various bases. For spatial vectors, e.g., each co-ordinate system represents another base. With the help of functions, other orthonormal bases for number vectors can be constructed, as is practised in quantum physics.

The *scalar product* of two number vectors can be used to determine relations between vectors. The scalar product of the vectors \mathbf{a} and \mathbf{b} is: $\mathbf{a}\cdot\mathbf{b}=a_1b_1+a_2b_2+a_3b_3+\dots$. The square root of the scalar product of a vector with itself ($\mathbf{a}\cdot\mathbf{a}=a_1^2+a_2^2+a_3^2+\dots$) determines the *magnitude* of \mathbf{a} . Each component of the vector \mathbf{a} is equal to its scalar product with the corresponding unit vector, e.g., $a_1=\mathbf{a}\cdot(1,0,0\dots)$. Analogous to the spatial case, this is called the *projection* of \mathbf{a} on a unit vector. If the scalar product is zero we call the vectors orthogonal, anticipating the spatial property of mutually perpendicular vectors. For instance, the unit vectors are mutually orthogonal. This multiplication of vectors is not a combination rule for groups, because the product is not a vector. (The 'vector product' is an anti-symmetric tensor, having n^2 components, of which $\frac{1}{2}(n-1)n$ components are independent. In a three-dimensional space, this yields exactly three independent components. Hence a vector product looks like a vector (only in three dimensions). However, it is a 'pseudovector'. At perpendicular reflection, a real vector reverses its direction, whereas the direction of a pseudovector is not changed.)

Apart from being real, the components of number vectors may be rational or complex, or even functions of numbers. These anticipate spatial vectors representing relative positions. An important difference is that spatial vectors are in need of a co-ordinate system, with an arbitrary choice of origin and unit vectors (3.1). Hence, number vectors are not identical with spatial vectors determining positions or displacements. *A fortiori*, this applies to kinetic or physical vectors, representing velocities or forces. Rather, the character of number vectors has the disposition to become interlaced with the characters of spatial, kinetic, or physical vectors.

A special case is the set of *complex numbers*, two-component vectors with a specific arithmetic. Also known as $\mathbf{c}=a+bi$, a complex number $\mathbf{c}=(a, b)$ is a two-dimensional number vector having real components a and b . The complex numbers for which $b=0$ have the same properties as real numbers, hence for convenience one writes $a=(a,0)$. This makes the set of real numbers a subset of the set of complex numbers. The unit vectors are $\mathbf{1}=(1,0)$ and $\mathbf{i}=(0,1)$, the imaginary unit. The complex numbers form an addition group.

The vector $\mathbf{c}^*=(a, -b)$ is called the complex conjugate of $\mathbf{c}=(a, b)$. The magnitude of \mathbf{c} is the square root of $\mathbf{c}\mathbf{c}^*=(a, b)(a, -b)=a^2+b^2$ and is a real number. The complex numbers

form an addition group with the combination rule: $(a, b) + (c, d) = (a+c, b+d)$. The identity element is $\mathbf{0} = (0, 0)$, and $-(a, b) = (-a, -b)$ is the inverse of (a, b) .

Complex numbers have the unique property that their multiplication yields a complex number. This is not the case for other number vectors. The product of the complex numbers (a, b) and (c, d) is $(a, b) \cdot (c, d) = (ac - bd, bc + ad)$, which is a complex number. Clearly, $i^2 = (0, 1)(0, 1) = (-1, 0) = -1$. The inverse operation also gives a complex number, but division by zero being excluded, this does not result in a group. As observed, the set of complex numbers is a field, an interlacement of two characters, subject to two combination rules.

Unlike the real numbers, the complex numbers cannot be projected on a line in an unambiguous order of increasing magnitude, because different complex numbers may have the same magnitude. However, they can be projected on a two-dimensional 'complex plane'. The addition group of complex numbers is isomorphic to the addition group of two-dimensional spatial vectors.

If we call w the angle with the positive real axis (for which $a > 0, b = 0$), then a complex number having magnitude c can be written as an imaginary exponential function $c \cdot \exp.iw = c(\cos w + i \sin w)$. The product of two complex numbers is now $cd \cdot \exp.i(v+w)$ and their quotient is $(c/d) \cdot \exp.i(v-w)$. In the complex plane, the unit circle around the origin represents the set of numbers $\exp.iw$. Multiplication of a complex number with $\exp.iw$ corresponds with a rotation about the angle w .

Interesting is that some theorems about real numbers can only be proved by considering them a subset of the set of complex numbers. The characters of real and complex numbers are strongly interlaced.

Mathematical *functions* may also have a character, a specific set of laws. A function is a prescription, connecting a set of numbers $[x]$ onto another set $[y]$, such that to every x only one y corresponds, $y = f(x)$. A function may depend on several variables, e.g. the components of a vector. A function is a relation between the elements of two or more sets, e.g., number sets. This relation is not always symmetrical. With each element of the first set $[x]$ corresponds only one element of the second set $[y]$. Conversely, each element of $[y]$ corresponds with zero, one or more elements of $[x]$. If the functional relation between $[x]$ and $[y]$ is symmetrical, the function is called one-to-one. This is important in particular in the case of a projection of a set onto itself. Sometimes such a projection is called a rotation. In a picture in which $[x]$ is represented on the horizontal axis and $[y]$ on the vertical axis, a graph represents the function spatially.

If the set $[x]$ is finite, then $[y]$ is finite as well and the prescription may be a table. More interesting are functions for which $[x]$ is a non-denumerable set of real or complex numbers within a certain interval. A function may be continuous or discontinuous. An example of a discontinuous function is the step function: $y = 0$ if $x < a$ and $y = 1$ if $x > a$. Here $[x]$ is the set of all real numbers, and $[y]$ is a subset of this set. The derivative of the step function is the characteristic delta function. The delta function equals zero

for all values of x , except for $x = a$. For $x = a$, the delta function is not defined. The integral of the delta function is 1. An approximate representation of the delta function is a rectangle having height h and width $1/h$. If h increases indefinitely, $1/h$ decreases, but the integral (the rectangle's area) is and remains equal to 1. The well-known Gauss-function approximates the delta function equally well.

Many a characteristic function defined by a specific *lawful* connection between two sets $[y]$ and $[x]$ has the disposition to be interlaced with spatial, kinetic, or physical characters. For instance, the quadratic function $y=ax^2+bx+c$ is interlaced with the spatial character of a parabola and with motion in free fall. Spatially defined, a parabola is a conic section. Of course, it can also be defined as the projection of the mentioned quadratic function. Contrary to laws, definitions are not very important.

The exponential function has the disposition to become interlaced with periodic motions and various physical processes. The exponential function with a real exponent ($\exp.at$) indicates positive or negative growth. If it has an imaginary exponent, the exponential function ($\exp.iat$) is periodical (i.e., $\exp.iat=\exp.i(at+n.2\pi)$ for each integral value of n), hence its character is interlaced with those of periodic motions like rotations, oscillations and waves.

Besides the above mentioned number vectors, mathematics knows of vectors which components are functions. Now a vector is an ordered set of n functions. (The dimension n may be finite or infinite, denumerable or non-denumerable). This is only possible if the scalar product $f.y$ is defined, including the magnitude of f (the square root of $f.f$), and if an orthonormal base of n unit functions f_1, f_2, \dots exists. 'Orthonormal' means that $f_i.f_j=d_{ij}$: the scalar product of each pair of basis functions equals 1 if $i=j$, it equals 0 if i differs from j . A function is an element of a complete addition group of functions if it is a linear combination of a set of basic functions. An n -dimensional linear combination of n basis functions is: $f=c_1f_1+c_2f_2+ \dots +c_nf_n$. In a complex function set, the components c_1, c_2, c_3, \dots are complex as well. The number of dimensions may be finite, denumerable infinite, or non-denumerable. In the latter case, the sum is an integral.

The basic functions being orthonormal, the group of functions is isomorphic with the group of number vectors having the same number of dimensions.

A function projects the elements of a number set onto another number set. Because many functions exist, sets of functions can be constructed. These too may be projected on each other, and such a projection is called an operator. Although the idea of an operator is developed and mostly applied in quantum physics, it is a mathematical concept. An operator A converses a function into another one, $y(x)=Af(x)$. This has the profile of an event. Having a quantitative character, a transition made by an operator is interlaced with the character of events qualified by a later relation frame. A spatial operation may be a translation or a rotation. A change of state is an example of a physical event. Quantum physics projects a physical change of state on the mathematical transition of a function by means of an operator.

If the converted function is proportional to the original one ($Af=af$, such that a is a real number), we call f an *eigenfunction* (proper function) of A , and a the corresponding *eigenvalue* (proper value). Trivial examples are the identity operator, for which any function is an eigenfunction (the eigenvalue being 1); or the operator multiplying a function by a real number (being its eigenvalue).

An operation playing an important part in kinematics, physics and chemistry is differentiating a function. (The reverse operation is called integrating). By differentiating a function is converted into its derivative. In mechanics, the derivative of the position function indicates the velocity of a moving body. Its acceleration is found by calculating the derivative of the velocity function.

For the operator (d/dx) , the real exponential function $f=b.exp.ax$ is an eigenfunction, for $(d/dx)f=ab.exp.ax=af$. The eigenvalue is the exponent a . The imaginary exponential function $y=b.exp.iat$ is an eigenfunction of the operator $(1/i)(d/dt)$, in quantum physics called the Hamilton-operator or Hamiltonian (after William Rowan Hamilton). Again, the eigenvalue is the exponent a . If $ABf = BAf$ for each f , A and B are called commutative. If two operators commute, they have the same eigenfunctions, but usually different eigenvalues.

Quantum physics calls a linear set of functions with complex components a *Hilbert space*. The quantum mechanical state space is called after David Hilbert, but was invented by John von Neumann, in 1927. This group is a representation of the ensemble of possible states of a physical system.

Consider an operator projecting a group onto itself. The operator A converts an element f of the group into another element Af of the same group. Such an operator is called *linear* if for all elements of the group $A(f+y)=Af+Ay$. If its eigenfunctions constitute an orthonormal basis for the group or a subgroup, the operator is called *hermitean*, after the mathematician Charles Hermite. The operation represented by a hermitean operator H is not a combination procedure for a group, but it projects a function on the eigenfunctions of H .

Besides hermitean operators, quantum physics applies unitary operators, which form a group representing the symmetry properties of a Hilbert space. To each operator A , an operator A^+ is conjugated such that the scalar product $y. Af = A^+f. y$. For a Hermitean operator H , $H^+ = H$, hence $f.Hy = Hf.y$. For a unitary operator U , $UU^+ = I$, the identity operator. Hence, $Uy.Uf=f.y=(y.f)^*$. This means that the probability of a state or a transition, being determined by a scalar product, is invariant under a unitary operation. Unitary operators are especially fit to describe symmetries and invariances.

2.4. Ensemble and probability

In our daily life as well as in science, we experience a thing first of all as a unit having specific properties. We know that an atom has the spatially founded character of a nucleus surrounded by a cloud of electrons. However, we also know it as a unit with a specific mass and chemical properties. A character determines a class of similar things. There are many hydrogen atoms having the same characteristic properties, even if deploying individual differences.

The arithmetic of characteristically equal individuals has a specific application in statistics. Statistics makes sense if it concerns the mutual variations of similar individuals. Statistics is only applicable to a specific set of individuals, a subset of a character class, a sample representative for the ensemble of possible variations. Both theoretically and empirically, we can apply statistics to the casting of dice, supposing all dice to have the same cubic symmetry, and assuming that the casting procedure is arbitrary.

I call an *ensemble* the set of all possible variations allowed by a character. Just like other sets, an ensemble has subsets, and sometimes the measure of a subset represents the relative probability of the possibilities. The concept of probability makes only sense if it concerns possibilities that can be realized by some physical interaction. Therefore, probability is a mathematical concept anticipating the physical one. I shall present a short summary of the classical theory of probability.

I discuss probability only in an ontological context, not in the epistemological meaning of the probability of a statement. Ontologically, probability does not refer to a lack of knowledge, but to the variation allowed by a character.

Consider the subsets A, B, \dots of the non-empty ensemble E of possibilities. Now A and B is the *union* of A and B , the subset of all elements belonging to A , to B , or to both. The *intersection* A or B is the subset of all elements belonging to A as well as to B . If A or $B = \mathbf{0}$ (the empty set) we call A and B *disjunct*, they have no elements in common. If A is a subset of B , then A and $B = B$ and A or $B = A$. Clearly, the intersection of A with E is A : A or $E = A$.

Formally, probability is defined as a quantitative measure $p(A)$ for any subset A of E . Observe that the theory ascribes a probability to the *subsets*, not to the elements of a set.

1. Probability is a non-negative measure: $p(A)$ is larger than or equal to 0.
2. Probability is normalized: $p(E)=1$.
3. Probability is an additive function for disjunct subsets of E : if A or $B = \mathbf{0}$, then $p(A$ and $B)=p(A)+p(B)$.

Starting from this definition, several theorems can be derived.

The *conditional probability*, the chance having A if B is given and if $p(B) > 0$, is defined as $p(A/B) = p(A \text{ or } B) / p(B)$. Because $p(A) = p(A/E)$, each probability is conditional. If A and B exclude each other, being disjunct ($A \text{ or } B = \mathbf{0}$), the conditional probability is zero. Now $p(A/B) = p(B/A) = 0$. If A is a subset of B then: $p(A \text{ and } B) = p(B)$; $p(A \text{ or } B) = p(A)$; $p(A/B) = p(A) / p(B)$.

A and B are called *statistically independent* if $p(A/B) = p(A)$ and $p(B/A) = p(B)$. Then $p(A \text{ or } B) = p(A)p(B)$ – for statistically independent subsets the chance of the combination is the product of their separate chances. Mark the distinction between disjunct and statistically independent subsets. In the first case probabilities are added, in the second case multiplied.

If an ensemble consists of n mutually statistically independent subsets, it can be projected onto an n -dimensional space. For instance, the possible outcomes of casting two dice simultaneously are represented on a 6×6 diagram. Genetics calls this a Punnett-square, after Reginald Punnett (1905). If E is a square with unit area, $p(A)$ is the area of a part of the square. Hence, so far the theory is not intrinsically a *probability* theory.

Finally, consider a set of disjunct subsets X of E , such that their sum equals E . Now the probability $p(X)$ is a function over the subsets X of E . We call $p(X)$ the *probability distribution* over the subsets X of E . Consider an arbitrary function $y(X)$ defined on this set. The average value of the function, also called its expectation value, is the sum over all X of the product $y(X)p(X)$, if the number of disjunct subsets is denumerable (otherwise it is the integral). In the form of a formula: $\langle y(X) \rangle = \sum_E y(X)p(X)$. In this sum, the probability expresses the 'weight' of each subset X .

This is called the *ensemble average* of the property. In statistical mechanics, it is an interesting question of whether this average is equal to the *time average* for the same property for a single system during a long time interval. This so-called *ergodic problem* is only solved for some very special cases, sometimes with a positive, sometimes with a negative result.[16] Besides the average of a property, it is often important to know how sharply peaked its probability distribution is. The 'standard deviation', the average difference from the average, is a measure of this peak. This is defined as $\langle y(X) - \langle y(X) \rangle \rangle$.

The formal theory is applicable to specific cases only if the value $p(A)$ can be theoretically or empirically established for the subsets A of E . Often this is only a *posteriori* possible by performing measurements with the help of a representative sample. Sometimes, symmetries allow of postulating an *a priori* probability distribution. Games of chance are the simplest, oldest, and best-known examples.

Although the above-summarized theory is not only relatively simple but almost universally valid as well, its application strongly depends on the situation. Quantum physics allows of interference of states, influencing probability in a way excluded by

classical probability theory (4.3). With respect to thing-like characters, the laws constituting the character determine the probability of possible variations. Another important field of application is formed by aggregates, for instance studied by statistical mechanics. For systems in or near equilibrium impressive results have been achieved, but for non-equilibrium situations (hence, for events and processes), the application of probability turns out to be fraught with problems.

Based on the characteristic similarity of the individuals concerned, statistical research is of eminent importance in all sciences. It is a means to research the character of individuals whose similarity is recognized or conjectured. It is also a means to study the properties of a homogeneous aggregate containing a multitude of individuals of the same character.

As early as 1860, James Clerk Maxwell applied statistics to an ideal gas, consisting of N molecules, each having mass m , in a container with volume V .^[17] He neglected the molecules' dimensions and mutual interactions. The vector \mathbf{r} gives the position of a molecule, and the vector \mathbf{v} represents its velocity. Maxwell assumed the probability for positions, $p_1(\mathbf{r})$, to be independent of the probability for velocities, $p_2(\mathbf{v})$. This means: $p(\mathbf{r},\mathbf{v})=p_1(\mathbf{r})p_2(\mathbf{v})$.

In equilibrium, the molecules are uniformly distributed over the available volume, hence the chance to find a molecule in a volume-element $d\mathbf{r}=dx.dy.dz$ equals $p_1(\mathbf{r})d\mathbf{r}=1/Vd\mathbf{r}$. Observe that $p_1(\mathbf{r})$ as well as $p_2(\mathbf{v})$ is a probability *density*. Maxwell based the velocity distribution on two kinds of symmetry. First, he assumed that the direction of motion is isotropic. This means that $p_2(\mathbf{v})$ only depends on the magnitude of the molecular speed. In this case, mathematically it does not matter to replace the speed by its square, hence $p_2(\mathbf{v})=p_2(v_x^2+v_y^2+v_z^2)$. Secondly, Maxwell assumed that the components of the velocity (v_x, v_y, v_z) are statistically independent. Only the exponential function satisfies these two requirements: $p_2(\mathbf{v})=p_2(v_x^2+v_y^2+v_z^2)=p_x(v_x)p_y(v_y)p_z(v_z)=a \cdot \exp. -\frac{1}{2}mb(v_x^2+v_y^2+v_z^2)$.

By calculating the pressure P exerted by the molecules on the walls of the container, and comparing the result with the law of Robert Boyle and Louis Gay-Lussac, Maxwell found that the exponent depends on temperature. From the law of Boyle and Gay-Lussac ($PV=NkT$, wherein T is the temperature and k is called Boltzmann's constant), it follows that $b=N/PV=1/kT$. The value of a follows from normalisation, i.e., the requirement that the total probability equals 1. Only in the twentieth century, experiments confirmed Maxwell's theoretical distribution function. The expression $\frac{1}{2}m(v_x^2+v_y^2+v_z^2)$ is recognizable as the kinetic energy of a molecule. The mean kinetic energy turns out to be equal to $(3/2)kT$. For all molecules together the energy is $(3/2)NkT$, hence, the specific heat is $(3/2)Nk$. This result was disputed in Maxwell's days, but it was later experimentally confirmed for mono-atomic gases. When Maxwell published his theory, it was not generally accepted that most known gases (hydrogen, oxygen, or nitrogen) consist of bi-atomic molecules.^[18] These gases have a different specific heat than mono-atomic gases like mercury vapour and the later discovered noble gases like helium and argon. Boltzmann explained this difference by observing that bi-atomic molecules have rotation and vibration kinetic energy besides

translational kinetic energy. An exact explanation became available only after the development of quantum physics.

Ludwig Boltzmann generalized Maxwell's distribution, by allowing other forms of energy besides kinetic energy. The Maxwell–Boltzmann distribution $p(\mathbf{r},\mathbf{v})=p_1(\mathbf{r})p_2(\mathbf{v})=(a/V)\cdot\exp.-E/kT$ turns out to be widely valid. The probabilities or relative occupation numbers of two-atomic, molecular, or nuclear states having energies E_1 and E_2 have a proportion according to the so-called Boltzmann-factor, determined by the *difference* between E_1 and E_2 . The Boltzmann-factor is: $p(E_1)/p(E_2)=(\exp.-E_1/kT)/(\exp.-E_2/kT)=\exp.-(E_1-E_2)/kT$. This means that a state having a high energy has a low probability.

The weakness of Maxwell's theory was neglecting the mutual interaction of the molecules, for without interaction equilibrium cannot be reached. Boltzmann corrected this by assuming that the molecules collide continuously with each other, exchanging energy. He arrived at the same result.

Maxwell and Boltzmann considered one system consisting of a large number of molecules, whereas Josiah Gibbs studied an ensemble of a large number of similar systems. Assuming that all microstates are equally probable, the probability of a macrostate can be calculated by determining the number of corresponding microstates. The logarithm of this number is proportional to the entropy of a macrostate.

Clausius and Boltzmann aimed at reducing the irreversibility expressed by the second law of thermodynamics to the reversible laws of mechanics. In how far they succeeded is still a matter of dispute. Anyhow, it could not be done without taking recourse to probability laws.[19] Boltzmann demonstrated the equilibrium state of a gas to have a much larger probability than a non-equilibrium state. He assumed that any system moves from a state with a low probability to a state with a larger one as a matter of course. This means that the irreversibility of the realization of possibilities is presupposed. In quantum mechanics, the combination of reversible equations of motion with probability leads to irreversible processes as well.[20]

Both in classical and in quantum statistics a character as a set of laws determines the ensemble of possibilities and the distribution of probabilities. It allows of individuality, the subject side of a character. Positivist philosophers defined probability as the limit of a frequency in an unlimited sequence of individual cases.[21] Later Karl Popper defended the 'propensity-interpretation' of probability: we have to '... interpret these weights of the possibilities (or of the possible cases) as measures of the propensity, or tendency, of a possibility to realize itself upon repetition'.[22] A propensity is a physical disposition or tendency '... to bring about the possible state of affairs ... to realize what is possible ... the relative strength of a tendency or propensity of this kind expresses itself in the relative frequency with which it succeeds in realizing the possibility in question.'[23] Besides subjectivist views, the frequency interpretation and the propensity interpretation, Lawrence Sklar calls 'probability' a theoretical term'[24] '... the meaning of probability attributions would be the rules of interference

that take us upward from assertions about observed frequencies and proportions to assertions of probabilities over kinds in the world, and downward from such assertions about probabilities to expectations about frequencies and proportions in observed samples. These rules of “inverse” and “direct” inference are the fundamental components of theories of statistical inference.’[25] This comes close to my interpretation of probability determined by a character.

Positivists tried to reduce the concept of probability to the subject side. Of course, the empirical measurement of a probability often has the form of a frequency determination. Each law statement demands testing, and that is only possible by taking a sample. This hypothesis must be regarded ‘as a postulate which can be ultimately justified only by the correspondence between the conclusions which it permits and the regularities in the behaviour of actual systems which are empirically found.’[26] This applies to all suppositions founding calculations of probabilities, but this does not justify the elimination of the law-side from probability theory.

An example of a frequency definition of probability is found in the study of radioactivity. A radioactive atom decays independent of other atoms, even if they belong to the same sample. During the course of time, the initial number of radioactive atoms (N_0) in a sample decreases exponentially to N_t at time t . When at a time t_0 , N_0 radioactive atoms of the same kind are left in a sample, then the expected number of remaining atoms at time t equals: $N_t = N_0 \exp.-(t-t_0)/t^*$, such that $N_t/N_0 = \exp.-(t-t_0)/t^*$. The characteristic constant t^* is proportional to the well-known half-life time. The law of decay is theoretically derivable from quantum field theory. This results in a slight deviation from the exponential function, too small to be experimentally verifiable,[27] Many scientists are content with this practical definition. However, a sample is a *collection* limited in time and space, it is not an ensemble of possibilities.

There are two limiting cases. In the one case, we extend the phenomenon of radioactivity to all similar atoms, increasing N_0 and N_t infinitely in order to get a *theoretical* ensemble. The ensemble has two possibilities, the initial state and the final state, and their distribution in the ensemble at time t after t_0 can be calculated, namely as the proportion $\exp.-(t-t_0)/t^* = [\exp.-t/t^*]/[\exp.-t_0/t^*]$. In the other limiting case we take $N_0=1$. Now $\exp.-(t-t_0)/t^*$ is the *chance* that a single atom decays after $t-t_0$ sec. This quotient depends on a time difference, not on a temporal instant. As long as the atom remains in its initial state, the probability of decay to the final state is unchanged.

Both limiting cases are theoretical. An ensemble is no more experimentally determinable than an individual chance. Only a collection of atoms can be subjected to experimental research. It makes no sense to consider one limiting case to be more fundamental than the other one. The first case concerns the law side, the second case the subject-side of the same phenomenon of radioactivity.

Statistics is not only applicable for the investigation of the ensemble of possibilities of a character. If two characters are interlaced, their ensembles are related as well. Sometimes, a one-to-one relation between the elements of both ensembles exists. Now the realization of a possibility in one ensemble reduces the number of

possibilities in the other ensemble to one. In other cases, several possibilities remain, with different probabilities.

Character interlacements are not always obvious. In a complex system, it is seldom easy to establish relations between structures, events and processes. Statistical research of correlations is a much applied expedient.

-
- [1] Zermelo in 1908, quoted by Quine 1963, 4; Putnam 1975, chapter 2.
 [2] Shapiro 1997, 98.
 [3] Quine 1963, 107-116.
 [4] Gödel 1962.
 [5] Putnam 1975, xi; Shapiro 1997, 109-112; Brown 1999, 182-191.
 [6] Barrow 1992, 137.
 [7] Amundson 1994, 102-106.
 [8] Brown 1999, 19, 22-23.
 [9] Quine 1963, 101.
 [10] Cassirer 1910, 49.
 [11] Grünbaum 1968, 13.
 [12] Quine 1963, chapter VI.
 [13] The relation between the elements CA and BA is $(CA)(BA)' = (CA)(A'B') = CB'$, the relation between C and B .
 [14] Slater 1975, 60-62.
 [15] Stafleu 1980, chapter 3.
 [16] Tolman 1938, 65-70; Khinchin 1949, Ch. III; Reichenbach 1956, 78-81; Prigogine 1980, 33-42, 64-65; Sklar 1993, 164-194.
 [17] Maxwell 1890, I, 377-409; Born 1949, 50ff; Achinstein 1991, 171-206.
 [18] Stafleu 2019, 10.3.
 [19] Bellone 1980, 91.
 [20] Belinfante 1975, chapter 2.
 [21] Von Mises 1939, 163-176, Reichenbach 1956, 96ff, Hempel 1965, 387, and initially Popper 1959, chapter VIII.
 [22] Popper 1967, 32.
 [23] Popper 1983, 286. See Settle 1974 discussing Popper's views; Margenau 1950, chapter 13; Nagel 1939, 23; Sklar 1993, 90-127.
 [24] Sklar 1993, 102-108
 [25] Sklar 1993, 103
 [26] Tolman 1938, 59.
 [27] Cartwright 1983, 118.

Chapter 3

Symmetry

3.1. Spatial magnitudes and vectors

3.2. Character, transformation and symmetry of spatial figures

3.3. Non-Euclidean space-time in the theory of relativity

Encyclopaedia of relations and characters. 3. Symmetry

3.1. Spatial magnitudes and vectors

The second relation frame for characters concerns their spatial relations. In 1899, David Hilbert formulated his foundations of projective geometry as relations between points, straight lines and planes, without defining these. 'Projective geometry' is since the beginning of the nineteenth century developed as a generalization of Euclidean geometry. Gottlob Frege thought that Hilbert referred to known subjects, but Hilbert denied this. He was only concerned with the *relations* between things, leaving aside their nature. According to Paul Bernays, geometry is not concerned with the nature of things, but with 'a system of conditions for what might be called a relational structure'.^[1] Inevitably, the later emphasis on structures was influenced by structuralism, for instance by Nicolas Bourbaki, pseudonym for a group of mainly French mathematicians.^[2]

Topological, projective, and affine geometries are no more metric than the theory of graphs. (A 'graph' is a two- or more-dimensional discrete set of points connected by line stretches.) This deals with spatial relations without considering the quantitative relation frame. I shall not discuss these non-metric geometries. The nineteenth- and twentieth-century views about metric spaces and mathematical structures turn out to be very important to modern physics.

This chapter is mainly concerned with the possibility to project a relation frame on a preceding one, and its relevance to characters. Section 3.1 discusses spatial magnitudes and vectors. The metric of space, being the law for the spatial relation frame, turns out to rest on symmetry properties. Symmetry plays an important part in the character and transformation of spatial figures that are the subject matter of section 3.2. Finally, section 3.3 deals with the metric of non-Euclidean kinetic space-time according to the theory of relativity.

Mathematics studies inter alia spatially qualified characters. Because these are interlaced with kinetic, physical, or biotic characters, spatial characters are equally important to science. This also applies to spatial relations concerning the position and posture of one figure with respect to another one. A characteristic point, like the centre of a circle or a triangle, represents the position of a figure objectively. The distance between these characteristic points objectifies the relative position of the circle and the triangle. It remains to stipulate the posture of the circle and the triangle, for instance with respect to the line connecting the two characteristic points. A co-ordinate system is an expedient to establish spatial positions by means of numbers.

Spatial relations are rendered quantitatively by means of magnitudes like distance, length, area, volume, and angle. These objective properties of spatial subjects and

their relations refer directly (as a subject) to numerical laws and indirectly (as an object) to spatial laws.

Science and technology prefer to define magnitudes that satisfy quantitative laws.

This is not the case with all applications of numbers. Numbers of houses project a spatial order on a numerical one, but hardly allow of calculations. Lacking a metric, neither Friedrich Mohs' scale for the hardness of minerals, nor Charles Richter's scale for the strength of earthquakes leads to meaningful calculations.

If we want to make calculations with a spatial magnitude, we have to project it on a suitable set of numbers (integral, rational, or real), such that *spatial* operations are isomorphic to *arithmetical* operations like addition or multiplication. This is only possible if a *metric* is available, a law to find magnitudes and their combinations.

For many magnitudes, the isomorphic projection on a group turns out to be possible. For magnitudes having only positive values (e.g., length, area, or volume), a multiplication group is suitable. For magnitudes having both positive and negative values (e.g., position), a combined addition and multiplication group is feasible. For a continuously variable magnitude, this concerns a group of real numbers. For a digital magnitude like electric charge, the addition group of integers may be preferred. It would express the fact that charge is an integral multiple of the electron's charge, functioning as a unit.

Every metric needs an arbitrarily chosen unit. Each magnitude has its own metric, but various metrics are interconnected. The metrics for area and volume are reducible to the metric for length. The metric for speed is composed from the metrics of length and time. Connected metrics form a metric system.

If a metric system is available, the government or the scientific community may decide to prescribe a metric to become a *norm*, for the benefit of technology, traffic, and commerce. Processing and communicating of experimental and theoretical results requires the use of a metric system.

A point has no dimensions and could have been considered a spatial object if extension were essential for spatial subjects. However, a relation frame is not characterized by any *essence* like continuous extension, but by *laws* for relations. Two points are spatially related by having a relative distance. The argument 'a point has no extension, hence it is not a subject' reminds of Aristotle and his adherents. They abhorred nothingness, including the vacuum and the number zero as a natural number. Roman numerals do not include a zero, and Europeans did not recognize it until the end of the Middle Ages. Galileo Galilei taught his Aristotelian contemporaries that there is no fundamental difference between a state of rest (the speed equals zero) and a state of motion (the speed is not zero).^[3]

It is correct that the property length does not apply to a point, any more than area can be ascribed to a line, or volume to a triangle. The difference between two line

segments is a segment having a certain length. The difference between two equal segments is a segment with zero length, but a zero segment is not a point. A line is a *set* having points as its *elements*, and each segment of the line is a *subset*. A subset with zero elements or only one element is still a subset, not an element. A segment has length, being zero if the segment contains only one point. A point has no length, not even zero length. Dimensionality implies that a part of a spatial figure has the same dimension as the figure itself. A three-dimensional figure has only three-dimensional parts. We can neither divide a line into points, nor a circle into its diameters. A spatial relation of a whole and its parts is not a subject-object relation, but a subject-subject relation. In a quantitative sense a triangle as well as a line segment is a set of points, and the side of a triangle is a subset of the triangle. But in a spatial sense, the side is not a part of the triangle.

Whether a point is a subject or an object depends on the nomic (*nomos* is Greek for law) context, on the laws we are considering. The relative position of the ends of a line segment determines in one context a subject-subject relation (to wit, the distance between two points), in another context a subject-object relation (the objective length of the segment). Likewise, the sides of a triangle, having length but not area, determine subjectively the triangle's circumference, and objectively its area.

The sequence of numbers can be projected on a line, ordering its points numerically. To order all points on a line or line segment the natural, integral, or even rational numbers are not sufficient. It requires the complete set of real numbers (2.2). The spatial order of equivalence or co-existence presents itself to full advantage only in a more-dimensional space. In a three-dimensional space, all points in a plane perpendicular to the *x*-axis correspond simultaneously to a single point on that axis. With respect to the numerical order on the *x*-axis, these points are equivalent. To lay down the position of a point completely, we need several numbers (*x,y,z,...*) simultaneously, as many as the number of dimensions. Such an ordered set of numbers constitutes a number vector (2.3).

For the character of a spatial figure too, the number of dimensions is a dominant characteristic. The number of dimensions belongs to the laws constituting the character. A plane figure has length and width. A three-dimensional figure has length, width, and height as mutually independent measures. The character of a two-dimensional figure like a triangle may be interlaced with the character of a three-dimensional figure like a tetrahedron. Hence, dimensionality leads to a hierarchy of spatial figures. At the base of the hierarchy, we find one-dimensional spatial vectors.

Contrary to a number vector, a spatial vector is localized and oriented in a metrical space. Localization and orientation are spatial concepts, irreducible to numerical ones. A spatial vector marks the relative position of two points. By means of vectors, each point is connected to all other points in space. Vectors having one point in common form an addition group. After the choice of a unit of length, this group is isomorphic to the group of number vectors having the same dimension. Besides spatial addition, a scalar product is defined (2.3). In an Euclidean space, the scalar product of two vectors **a** and **b** equals $\mathbf{a}\cdot\mathbf{b}=ab \cos \alpha$. Herein a is the square root of $\mathbf{a}\cdot\mathbf{a}$. It equals the

length of \mathbf{a} whereas α is the angle between \mathbf{a} and \mathbf{b} . If two vectors are perpendicular to each other, their scalar product is zero. The group's identity element is the vector with zero length. Its base is a set of orthonormal vectors, i.e., the mutually perpendicular unit vectors having a common origin. Each vector starting from that origin is a linear combination of the unit vectors. So far, there is not much difference with the number vectors.

However, whereas the base of a group of number vectors is rather unique, in a group of spatial vectors the base can be chosen arbitrarily. For instance, one can rotate a spatial base about the origin. It is both localized and oriented. The set of all bases with a common origin is a rotation group. The set of all bases having the same orientation but different origins is a translation group. It is isomorphic both to the addition group of spatial vectors having the same origin and to the addition group of number vectors.

Euclidean space is homogeneous (similar at all positions) and isotropic (similar in all directions). Combining spatial translations, rotations, reflections with respect to a line or a plane and inversions with respect to a point leads to the Euclidean group. It reflects the symmetry of Euclidean space. Symmetry points to a transformation keeping certain relations invariant.[4] At each operation of the Euclidean group, several quantities and relations remain invariant, for instance, the distance between two points, the angle between two lines, the shape and the area of a triangle, and the scalar product of two vectors.

Besides a relative position, a spatial vector represents a *displacement*, the result of a motion. This is a disposition, a tertiary characteristic of spatial vectors.

Each base in each point of space defines a co-ordinate system. In an Euclidean space, this is usually a Cartesian system of mutually perpendicular axes (named after, respectively, Euclid of Alexandria and René Descartes). Partly, the choice of the co-ordinate system is arbitrary. We are free to choose rectangular, oblique or polar axes. (Polar co-ordinates do not determine the position of a point by its projections on two or more axes, but by the distance r to the origin and by one or more angles. For example, think of the geographical determination of positions on the surface of the earth.)

If we have a reference system, we can replace it by translation, rotation, mirroring, or a combination of these. A co-ordinate system has to satisfy certain rules.

1. The number of axes and unit vectors equals the number of dimensions. With fewer co-ordinates, the system is underdetermined, with more it is overdetermined.

2. The unit vectors are mutually independent. Two vectors are mutually dependent if they have the same direction. An arbitrary vector is a linear combination of the unit vectors, and is said to depend on them. In two dimensions, $\mathbf{a}=(a_1,a_2)=a_1(1,0)+a_2(0,1)$.

3. Replacing a co-ordinate system should not affect the spatial relations between the subjects in the space. In particular the distance between two points should have the

same value in all co-ordinate systems. This rule warrants the *objectivity* of the co-ordinate systems. In a co-ordinate transformation, a magnitude that remains equal to itself is called 'invariant'. This applies e.g. to the magnitude of a vector and the angle between two vectors. 'Covariant' magnitudes change in analogy to the co-ordinates.

4. The choice of a *unit of length* is arbitrary, but should have the same value in all co-ordinate systems, as well as along all co-ordinate axes. That may seem obvious, but for a long time at sea, the units used for depth and height were different from those for horizontal dimensions and distances.

5. For calculating the distance between two points we need a law, called the spatial *metric*, see below.

6. The co-ordinate system should reflect the symmetry of the space. For an Euclidean space, a Cartesian co-ordinate system satisfies this requirement. Giving preference to one point, e.g. the source of an electric field, breaks the Euclidean symmetry. In that case, scientists often prefer a co-ordinate system that expresses the spherical symmetry of the field. In the presence of a homogeneous gravitational field, physicists usually choose one of the co-ordinate axes in the direction of the field. If the space is non-Euclidean, like the earth's surface, a Cartesian co-ordinate system is quite useless.

The fact that we are free to choose a co-ordinate system has generated the assumption that this choice rests on a *convention*, an agreement to keep life simple.[5] However, both the fact that a group of co-ordinate systems reflects the symmetry of the space and the requirement of objectivity make clear that these rules are *normative*. It is not imperative to follow these rules, but we *ought to* choose a system that reflects spatial relations objectively.

The metric depends on the symmetry of space. In an Euclidean space, Pythagoras' law determines the metric. If the co-ordinates of two points are given by (x_1, y_1, z_1) and (x_2, y_2, z_2) , and if we call $dx = x_2 - x_1$ etc., then the distance is the square root of $dr^2 = dx^2 + dy^2 + dz^2$. This is the Euclidean metric. Since the beginning of the nineteenth century, mathematics acknowledges non-Euclidean spaces as well. Non-Euclidean geometries were discovered independently by Nicolai Lobachevski (first publication, 1829-30), Janos Bolyai and Carl Friedrich Gauss, later supplemented by Felix Klein. Significant is to omit Euclides' fifth postulate, corresponding to the axiom that one and only one line parallel to a given line can be drawn through a point outside that line. (Long before, it was known that on a sphere the Euclidean metric is only applicable to distances small compared with the sphere's radius.)

Preceded by Gauss, in 1854 Bernhard Riemann formulated the general metric for an infinitesimal small distance in a multidimensional space. Riemann's metric is $dr^2 = g_{xx}dx^2 + g_{yy}dy^2 + g_{xy}dxdy + g_{yx}dydx + \dots$. Mark the occurrence of mixed terms besides quadratic terms. In the Euclidean metric $g_{xx} = g_{yy} = 1$, $g_{xy} = g_{yx} = 0$, and Δx and Δy are not necessarily infinitesimal. According to Riemann, a multiply extended magnitude allows of various metric relations, meaning that the theorems of geometry cannot be reduced to quantitative ones.[6]

For a non-Euclidean space, the coefficients in the metric depend on the position. If i and j indicate x or y , the g_{ij} 's, are components of a tensor. In the two-dimensional case g_{ij} is a second derivative (like $d^2r/dx dy$). For a more-dimensional space it is a *partial* derivative, meaning that other variables remain constant. To calculate a finite displacement requires the application of integral calculus. The result depends on the choice of the path of integration. The distance between two points is the smallest value of these paths. On the surface of a sphere, the distance between two points corresponds to the path along a circle whose centre coincides with the centre of the sphere.

The metric is determined by the structure and eventually the symmetry of the space. This space has the disposition to be interlaced with the character of kinetic space or with the physical character of a field. A well-known example is the general theory of relativity, being the relativistic theory of the gravitational field. In the general theory of relativity, the coefficients for the four-dimensional space-time manifold form a symmetrical tensor, i.e., $g_{ij}=g_{ji}$ for each combination of i and j . Hence, among the sixteen components of the tensor ten are independent.

An electromagnetic field is also described by a tensor having sixteen components. Its symmetry demands that $g_{ij}=-g_{ji}$ for each combination of i and j , hence the components of the quadratic terms are zero. This leaves six independent components, three for the electric vector and three for the magnetic pseudovector.

Gravity having a different symmetry than electromagnetism is related to the fact that mass is definitely positive and that gravity is an attractive force. In contrast, electric charge can be positive or negative and the electric force named after Charles-Augustin Coulomb may be attractive or repulsive. A positive charge attracts a negative one, two positive charges (as well as two negative charges) repel each other.

In general, a non-Euclidean space is less symmetrical than an Euclidean one having the same number of dimensions. Motion as well as physical interaction causes a break of symmetry in spatial relations.

3.2. Character, transformation and symmetry of spatial figures

This section discusses the *shape* of a spatial figure as an elementary example of a character. A spatial character has both a primary and a secondary characteristic. The tertiary characteristic plays an increasingly complex part in the path of a specific motion, the shape of a crystal, the morphology of a plant or the body structure of an animal. Besides, even the simplest figures display a spatial interlacement of their characters.

A spatial figure has the profile of a thing-like subject. Its shape determines its character. Consider a simple plane triangle in an Euclidean space. In a non-Euclidean space two figures only have the same shape if they have the same magnitude as well.^[7] Similarity (to be distinguished from congruence or displacement symmetry) is a characteristic of an Euclidean space. Many regular figures like squares or cubes only exist in an Euclidean space. The character of a triangle constitutes a set of widely different triangles, having different angles, linear dimensions, and relative positions. Because each triangle belonging to the character class is a possible triangle as well, the ensemble coincides with the character class. We distinguish this set easily from related sets of e.g., squares, ellipses, or pyramids. Clearly, the triangle's character is primarily spatially characterized and secondarily quantitatively founded. Thirdly, a triangle has the disposition to have an objective function in a three- or more-dimensional figure.

A triangle is a two-dimensional spatial thing, directly subject to spatial laws. The triangle is bounded by its sides and angular points, which have no two-dimensional extension but determine the triangle's objective magnitude. Quantitatively, we determine the triangle by the number of its angular points and sides, the magnitude of its angles, the length of its sides, and by its area.

With respect to the character of a triangle, its sides and angular points are objects, even if they are in another context subjects (3.1). Their character has the disposition to become interlaced with that of the triangle.

A triangle has a structure or character because its objective measures are *bound*, satisfying restricting laws or constraints. Partly this is a matter of definition, a triangle having three sides and three angular points. This definition is not entirely free, for a 'biangle' as a two-dimensional figure does not exist and a quadrangle may be considered a combination of two triangles. However, there are other lawlike relations not implied by the definition, for instance the law that the sum of the three triangles equals π , the sum of two right angles. This is a specific law, only valid for plane triangles.

A triangle is a whole with parts. As observed, the relation of a whole and its parts is not to be confused with a subject-object relation. It makes no sense to consider the sides and the angular points as parts of the triangle. With respect to a triangle, the whole-part relation has no structural meaning. In contrast, a polygon is a combination of triangles being parts of the polygon. Therefore, a polygon has not much more structure than it derives from its component triangles. The law that the sum of the angles of a polygon having n sides equals $(n-2)\pi$ is reducible to the corresponding law for triangles.

Two individual triangles can be distinguished in three ways, by their relative position, their relative magnitude, and their different shape. I shall consider two mirrored triangles to be alike.

Relative position is not relevant for the character of a triangle. We could just as well consider its relative position with respect to a circle or to a point as to another triangle. Relative position is the universal spatial subject-subject relation. It allows of the identification of any individual subject. Often, the position of a triangle will be objectified, e.g. by specifying the positions of the angular points with respect to a co-ordinate system.

Next, triangles having the same shape can be distinguished by their magnitude. This leads to the secondary variation in the quantitative *foundation* of the character.

Finally, two triangles may have different shapes, one being equilateral, the other rectangular, for example. This leads to the primary variation in the spatial *qualification* of the triangle's character. Triangles are spatially similar if they have equal angles. Their corresponding sides have an equal ratio, being proportional to the sines of the opposite angles.

For any polygon, the triangle can be considered the primitive form. It displays a primary spatial variability in its shape and a secondary quantitative variability in its magnitude. Another primitive form is the ellipse, with the circle as a specific variation.

There are irregular shapes as well, not subject to a specific law. These forms have a secondary variability in their quantitative foundation, but lack a lawlike primary variation regarding the qualifying relation frame.

Like two triangles can be different in three respects, a triangle can be changed in three ways: by displacement (translation, rotation, and/or mirroring); by making it larger or smaller; or by changing its shape, i.e., by transformation. A transformation means that the triangle becomes a triangle with different angles or it gets an entirely different shape. Displacement, enlargement or diminishment, and transformation are spatial expressions anticipating actual events.

An operator (2.3) describes a characteristic transformation, if co-ordinates and functions represent the position and the shape of the figure. The character of a spatial transformation preserving the shape of the figure is interlaced with the character of an operator having eigenfunctions and eigenvalues.

All displacements of a triangle in a plane form a group isomorphic to the addition group of two-dimensional vectors. All rotations, reflections and their combinations constitute groups as well. Enlargements of a given triangle form a group isomorphic to the multiplication group of positive real numbers. (A subgroup is isomorphic to the multiplication group of positive rational numbers).

A separate class of spatial figures is called symmetric, e.g., equilateral and isosceles triangles. Symmetry is a property related to a spatial transformation such that the figure remains the same in various respects. Without changing, an equilateral triangle can be reflected in three ways and rotated about two angles. An isosceles triangle has only one similar operation, reflection, and is therefore less symmetric. A circle is very

symmetric, because an infinite number of rotations and reflections transform it into itself.

The theory of groups renders good services to the study of these symmetries (2.3). In 1872, Felix Klein in his 'Erlangen Program' pointed out the relevance of the theory of groups for geometry, considered to be the study of properties invariant under transformations.[8]

Consider the group consisting of only three elements, I , A and B , such that $AB=I, AA=B, BB=A$. This is very abstract and only becomes transparent if an interpretation of the elements is given. This could be the rotation symmetry of an equilateral triangle, A being an anti-clockwise rotation of $\pi/3$, B of $2\pi/3$. The inverse is the same rotation clockwise. The combination AB is the rotation B followed by A , giving I , the identity. Clearly, the character of this group has the disposition of being interlaced with the character of the equilateral triangle. However, this triangle has more symmetry, such as reflections with respect to a perpendicular. This yields three more elements for the symmetry group, now consisting of six elements. The rotation group I, A, B is a subgroup, isomorphic to the group consisting of the numbers 0, 1 and 2 added modulo 3 (2.3). The group is not only interlaced with the character of an equilateral triangle, but with many other spatial figures having a threefold symmetry, as well as with the group of permutations of three objects. A permutation is a change in the order of a sequence; e.g., BAC is a permutation of ABC . A set of n objects allows of $n! = 1.2.3... n$ permutations.

In turn, the character of an equilateral triangle is interlaced with that of a regular tetrahedron. The symmetry group of this triangle is a subgroup of the symmetry group of the tetrahedron.

A group expresses spatial similarity as well. The combination procedure consists of the multiplication of all linear dimensions with the same positive real or rational number, leaving the shape invariant. The numerical multiplication group of either rational or real positive numbers is interlaced with a spatial multiplication group concerning the secondary foundation of figures.

The translation operator, representing a displacement by a vector, formally represented by $T(\mathbf{a})\mathbf{r}=\mathbf{r}+\mathbf{a}$, is an element of various groups, e.g., the Euclidean group mentioned before. Solid-state physics applies translation groups to describe the regularity of crystals. This implies an interlacement of the quantitative character of a group with the spatial character of a lattice and with the physical character of a crystal. The translation group for this lattice is an addition group for spatial vectors. It is isomorphic to a *discrete* group of number vectors, which components are not real or rational but integral. The crystal's character has the disposition to be interlaced with the kinetic wave character of the X-rays diffracted by the crystal. As a consequence, this kind of diffraction is only possible for a *discrete* set of wave lengths.

The question of whether figures and kinetic subjects are *real* usually receives a negative answer, even in Protestant philosophy: 'No single real thing or event is

typically qualified or founded in an original mathematical aspect.'[9] 'If anything is to be actually real in the world of empirical existence, it must ultimately be founded in physical reality.' 'Existence is ordered so as to build on physical foundations.'[10] The view that only physical (material) things are real is a common form of physicalism or materialist naturalism.[11]

First, this is the view of natural experience, which appears to accept only tangible matters to be real. Nevertheless, without the help of any theory, everybody recognizes typical shapes like circles, triangles, or cubes. This applies to typical motions like walking, jumping, rolling, or gliding as well.

Second, reality is sometimes coupled to observability. Now shapes are very well observable, albeit that we always need a physical substrate for any actual observation. Moreover, it would be an impoverishment if we would restrict our experience to what is directly observable. Human imagination is capable of representing many things that are not directly observable. For instance, we are capable of interpreting drawings of two-dimensional figures as three-dimensional objects. Although a movie consists of a sequence of static pictures, we see people moving. We can even see things that have no material existence, like a rainbow.

Third, I observe that the view that shapes are not real is strongly influenced by Plato, Aristotle, and their medieval commentators. According to Plato, spatial forms are invisible, but more real than observable phenomena. In contrast, Aristotle held that forms determine the nature of the things, having a material basis as well. Moreover, the realization of an actual thing requires an operative cause. Hence, according to Aristotle, all actually existing things have a physical character.

In opposition, I maintain that in the cosmos everything is real that answers the laws of the cosmos. Then numbers, groups, spatial figures, and motions are no less real than atoms and stars.

But are these natural or cultural structures? It cannot be denied that the *concept* of a circle or a triangle is developed in the course of history, in human cultural activity. Yet I consider them to be natural characters, which existence humanity has *discovered*, like it discovered the characters of atoms and molecules.

Reality is a theoretical concept. It implies that the temporal horizon is much wider than the horizon of our individual experience, and in particular much wider than the horizon of natural experience. By scientific research, we enlarge our horizon, discovering characters that are hidden from natural experience. Nevertheless, such characters are no less real than those known to natural experience are.

We call the kinetic space for waves a *medium* (and sometimes a field), and we call the physical space for specific interactions a *field*. For the study of physical interactions, spatial symmetries are very important. For instance, in classical physics this is the case with respect to gravity (Newton's law), the electrostatic force (Coulomb's law) and the magnetostatic force. Each of these forces is subject to an 'inverse square law'.

This law expresses the isotropy of physical space. In all directions, the field is equally strong at equal distances from a point-like source, and the field strength is inversely proportional to the square of the distance. About 1830, Carl Friedrich Gauss developed a method allowing of calculating the field strength of combinations of point-like sources. He introduced the concept of 'flux' through a surface, popularly expressed, the number of field lines passing through the surface. (An infinitesimal surface is defined as a vector \mathbf{a} by its magnitude and the direction perpendicular to the surface. The flux is the scalar product of \mathbf{a} with the field strength \mathbf{E} at the same location and is maximal if \mathbf{a} is parallel to \mathbf{E} , minimal if their directions are opposite. If \mathbf{a} is perpendicular to \mathbf{E} the flux is zero. For a finite surface one finds the flux by integration.)

Gauss proved that the flux through a *closed* surface around one or more point-like sources is proportional to the total strength of the sources, independent of the shape of that surface and the position of the sources. The proportionality factor depends on the force law and is different in the three mentioned cases. This symmetry property has some important consequences.

Outside the sphere, a homogeneous spherical charge or mass causes a field that is equal to that of a point-like source concentrated in the centre of the sphere. Within the sphere, the field is proportional to the distance from the centre. Starting from the centre, the field initially increases linearly, but outside the sphere, it decreases quadratically. For gravity, Isaac Newton had derived this result by other means.

For magnetic interaction, physicists find empirically that the flux through a closed surface is always zero. This means that within the surface there are as many positive as negative magnetic poles. Magnetism only occurs in the form of dipoles or multipoles. There is no law excluding the existence of magnetic monopoles, but experimental physics has never found them.

In the electrical case, the combination of Gauss's law with the existence of conductors leads to the conclusion that in a conductor carrying no electric current the electric field is zero. All net charge is located on the surface and the resulting electric field outside the conductor is perpendicular to the surface. Therefore, inside a hollow conductor the electric field is zero, unless there is a net charge in the cavity. Experimentally, this has been tested with a large accuracy. Because this result depends on the inverse square law, it has been established that the exponent in Coulomb's law differs less than 10^{-20} from 2. If there is a net charge in the cavity, there is as much charge (with reversed sign) on the inside surface of the conductor. It is distributed such that in the conductor itself the field is zero. If the net charge on the conductor is zero, the charge at the outside surface equals the charge in the cavity. By connecting it with the 'earth', the outside can be discharged. Now outside the conductor the electric field is zero, and the charge within the cavity is undetectable. Conversely, a space surrounded by a conductor is screened from external electric fields.

Gauss's law depicts a purely spatial symmetry and is therefore only applicable in static or quasi-static situations. James Clerk Maxwell combined Gauss's law for electricity and magnetism with André-Marie Ampère's law and Michael Faraday's law for

changing fields. As a consequence, Maxwell found the laws for the electromagnetic field. These laws are not static, but relativistically covariant, as Albert Einstein established.

Spin is a well-known property of physical particles. It derives its name from the now as naive considered assumption that a particle spins around its axis. If the particle is subject to electromagnetic interaction, a magnetic moment accompanies the spin, even if the particle is not charged. A neutron has a magnetic moment, whereas a neutrino has not. Spin is an expression of the particle's rotation *symmetry*, and is similar to the angular momentum of an electron in its orbit in an atom. A pion has zero spin and transforms under rotation like a scalar. The spin of a photon is 1 and it transforms like a vector. The hypothetical graviton's spin is twice as large, behaving as a tensor at rotation. These particles, called bosons, have symmetrical wave functions. Having a half-integral spin (as is the case with, e.g., an electron or a proton), a fermion's wave function is antisymmetric. It changes of sign after a rotation of 2π (4.4). This phenomenon is unknown in classical physics, but plays an important part in quantum statistics.

3.3. Non-Euclidean space-time in the theory of relativity

Until the end of the nineteenth century, motion was considered as change of place, with time as the independent variable. Isaac Newton thought space to be absolute, the expression of God's omnipresence, a *sensorium Dei*.^[12] Newton's contemporaries Christiaan Huygens and Gottfried Wilhelm Leibniz were more impressed by the relativity of motion. They believed that anything only moves relative to something else, not relative to absolute space. As soon as Thomas Young, Augustin Fresnel and other physicists in the nineteenth century established that light is a moving wave, they started the search for the ether, the material medium for wave motion. They identified the ether with Newton's absolute space, now without the speculative reference to God's omnipresence. This search had little success, the models for the ether being inconsistent or contrary to observed facts. In 1865, James Clerk Maxwell formulated his electromagnetic theory, connecting magnetism with electricity, and interpreting light as an electromagnetic wave motion. Although Maxwell's theory did not require the ether, he persisted in believing its existence. In 1905, Albert Einstein suggested to abandon the ether.^[13] He did not prove that it does not exist, but showed it to be superfluous. Physicists intended the ether as a material substratum for electromagnetic waves. However, in Einstein's theory it would not be able to interact with anything else. Consequently, the ether lost its physical meaning. (The cosmic electromagnetic background radiation discovered by Arno Penzias and Robert Wilson in 1964 may be considered to be an ether.)

Until Einstein, kinetic time and space were considered independent frames of reference. In 1905, Albert Einstein shook the world by proving that the kinetic order implies a relativization of the quantitative and spatial orders. Two events being synchronous according to one observer turn out to be diachronous according to an observer moving at high speed with respect to the former one. This relativizing is unheard of in the common conception of time, and it surprised both physicists and philosophers.

Einstein based the special theory of relativity on two postulates or requirements for the theory. The first postulate is the principle of relativity. It requires each natural law to be formulated in the same way with respect to each inertial frame of reference. The second postulate demands that light have the same speed in every inertial system. From these two axioms, Einstein could derive the mentioned relativization of the quantitative and spatial orders. He also showed that the units of length and of time depend on the choice of the reference system. Moving rulers are shorter and moving clocks are slower than resting ones. In the theory of Hendrik Lorentz and others, time dilation and space contraction were explained as molecular properties of matter. Einstein explained them as kinetic effects. Only the speed of light is in all reference systems the same, acting as a *unit of motion*. Indeed, relativity theory often represents velocities in proportion to the speed of light.

An inertial system is a system of reference in which Newton's first law of motion, the law of inertia, is valid. Unless some unbalanced force is acting on it, a body moves with constant velocity (both in magnitude and in direction) with respect to an inertial system. This is a reference system for *motions*; hence, it includes clocks besides a spatial co-ordinate system. If we have one inertial system, we can find many others by shifting, rotating, reflecting, or inverting the spatial co-ordinates; or by moving the system at a constant speed; or by resetting the clock, as long as it displays kinetic time uniformly (4.1). These operations form a group, in classical physics called the Galileo group. Here time is treated as a variable parameter independent of the three-dimensional spatial co-ordinate system. Since Einstein proved this to be wrong, an inertial system is taken to be four-dimensional. The corresponding group of operations transforming one inertial system into another one is called the Lorentz group (sometimes called the Poincaré group, of which the Lorentz group (now without spatial and temporal translations) would be a subgroup.) The distinction between the classical Galileo group and the special relativistic Lorentz group concerns relatively moving systems. Both have an Euclidean subgroup of inertial systems not moving with respect to each other. The distinction concerns the combination of motions, objectified by velocities. Restricted to one direction, in the Galileo group velocities are combined by addition ($v+w$), in the Lorentz group by the formula $(v+w)/(1+vw/c^2)$, see section 2.3. The name 'Galileo group' dates from the twentieth century.

In a four-dimensional inertial system, a straight line represents a uniform motion. Each point on this line represents the position (x,y,z) of the moving subject at the time t . If the speed of light is the unit of velocity, a line at an angle of $\pi/4$ with respect to the t -axis represents the motion of a light signal. The relativistic metric concerns the spatio-temporal *interval* between two events. The metric of special relativity theory is $ds^2=dx^2+dy^2+dz^2-dt^2=dr^2-dt^2$. There are no mixed terms, and the interval is not

necessarily infinitesimal. This metric is pseudo-Euclidean because of the minus sign in front of dt^2 . If the speed of light is not taken as the unit of speed, this term becomes c^2dt^2 . The metric can be made *apparently* Euclidean by considering time an imaginary co-ordinate: $ds^2=dx^2+dy^2+dz^2+(idt)^2$. It is preferable to make visible that kinetic space is less symmetric than the Euclidean four-dimensional space, for lack of symmetry between the time axis and the three spatial axes. According to the formula, ds^2 can be positive or negative, and Δs real or imaginary. Therefore, one defines the interval as the absolute value of Δs .

The combination rule in the Lorentz group is formulated such that the interval is invariant at each transformation of one inertial system into another one. Only then, the speed of light (the unit of motion) is equal in all inertial systems. A flash of light expands spherically at the same speed in all directions, in any inertial reference system in which this phenomenon is registered. This system is called the block universe or Hermann Minkowski's space-time continuum.[14]

The magnitude of the interval is an objective representation of the relation between two events, combining a time difference with a spatial distance. For the same pair of events in another inertial system, both the time difference Δt and the spatial distance Δr may be different. Only the magnitude Δs of the interval is independent of the choice of the inertial system.

Whereas the Euclidean metric is always positive or zero, the pseudo-Euclidean metric, determining the *interval* between two events may be negative as well. For the motion of a light signal between two points, the interval is zero. For a light signal, $\Delta s=0$, for the covered distance Δr equals $c\Delta t$. If $\Delta r=0$, the two events have the same position and the interval is a time difference (Δt). If $\Delta t=0$, the interval is a spatial distance (Δr) and the two events are simultaneous. In other cases, an interval is called space-like if the distance $\Delta r > c\Delta t$, or time-like if the time difference $\Delta t > \Delta r/c$ (in absolute values). In the first case, light cannot bridge the distance within the mentioned time difference, in the second case it can.

For two events having a *space-like* interval, an inertial system exists such that the time difference is zero ($\Delta t=0$), hence the events are simultaneous. In another system, the time difference may be positive or negative. The distance between the two events is too large to be bridged even by a light signal, hence the two events cannot be causally related. Whether such a pair of events is diachronous or synchronous appears to depend on the choice of the inertial system.

Other pairs of events are diachronous in every inertial system, their interval being *time-like* ($\Delta s^2 < 0$). If in a given inertial system event A occurs before event B , this is the case in any other inertial system as well. Now A may be a cause of B , anticipating the physical relation frame. The causal relation is irreversible, the cause preceding the effect.[15]

The formula for the relativistic metric shows that space and time are not equivalent, as is often stated. By a rotation about the z-axis, the x-axis can be transformed into

the y -axis. In contrast, no physically meaningful transformation exists from the t -axis into one of the spatial axes or conversely.

In the four-dimensional space-time continuum, the spatial and temporal co-ordinates form a vector. Other vectors are four-dimensional as well, often by combining a classical three-dimensional vector with a scalar. This is meaningful if the vector field has the same or a comparable symmetry as the space-time continuum. For instance, the linear momentum and the energy of a particle are combined into the four-dimensional momentum-energy vector $(p_x, p_y, p_z, E/c)$. Its magnitude (the square root of $p_x^2 + p_y^2 + p_z^2 - E^2/c^2$) has in all inertial systems the same value. The theory of relativity distinguishes invariant, covariant, and contravariant magnitudes, vectors etc.

An unexpected consequence of the symmetry of physical space and time is that the laws of conservation of energy, linear and angular momentum turn out to be derivable from the principle of relativity. Emmy Noether first showed this in 1915.^[16] Because natural laws have the same symmetry as kinetic space, the conservation laws in classical mechanics differ from those in special relativity.

Considering the homogeneity and isotropy of a field-free space and the uniformity of kinetic time, theoretically the principle of relativity allows of two possibilities for the transformations of inertial systems.^[17] According to the classical Galileo group, the metric for time is independent of the metric for space. The units of length and time are invariant under all transformations. The speed of light is different in relatively moving inertial systems. In the relativistic Lorentz group, the metrics for space and time are interwoven into the metric for the interval between two events. The units of length and time are not invariant under all transformations. Instead, the unit of velocity (the speed of light) is invariant under all transformations. On empirical grounds, the speed of light being the same in all inertial systems, physicists accept the second possibility. Not the Galileo group but the Lorentz group turns out to be interlaced with kinetic space-time.

According to the principle of relativity, the natural laws can be formulated independent of the choice of an inertial system. Albert Einstein called this a postulate, a demand imposed on a theory. In contrast, Mario Bunge calls it a norm, a 'normative metanomological principle ...' constituting '...a necessary though insufficient condition for objectivity ...'^[18] I suggest that it rests on the irreducibility of physical interaction to spatial or kinetic relations. The principle of relativity is not merely a convention, an agreement to formulate natural laws as simple as possible. It is first of all a requirement of objectivity, to formulate the laws such that they have the same expression in every appropriate reference system.

Yet, physicists do not always stick to the principle of relativity. When standing on a revolving merry-go-round, anyone feels an outward centrifugal force. When trying to walk on the roundabout he or she experiences the force called after Gustave-Caspar Coriolis as well. These forces are not the physical cause of acceleration, but its effect. Both are inertial forces, only occurring in a reference system accelerating with respect to the inertial systems.

Although the centrifugal force and the Coriolis force do not exist with respect to inertial systems, they are real, being measurable and exerting influence. In particular, the earth is a rotating system. The centrifugal force causes the acceleration of a falling body to be larger at the poles than at the equator, partly directly, partly due to the flattening of the earth at the poles, another effect of the centrifugal force. The Coriolis force causes the rotation of the pendulum called after Léon Foucault, and it has a strong influence on the weather. The wind does not blow directly from a high- to a low-pressure area, but it is deflected by the Coriolis force to encircle such areas.

Another example of an inertial force occurs in a reference system having a constant acceleration with respect to inertial systems. This force experienced in an accelerating or braking lift or train is equal to the product of the acceleration and the mass of the subject on which the force is acting. It is a universal force, influencing the motion of all subjects that we wish to refer to the accelerated system of reference.

Often, physicists and philosophers point to that inertial force in order to argue that the choice of inertial systems is arbitrary and conventional. Only because of simplicity, we prefer inertial systems, because it is awkward to take into account these universal forces. A better reason to avoid such universal forces is that they do not represent subject-subject relations. Inertial forces do not satisfy Newton's third law, the law of equal action and reaction, for an inertial force has no reaction.^[19] The source of the force is not another subject. A Newtonian physicist would call such a force fictitious.^[20] The use of inertial forces is only acceptable for practical reasons. For instance, this applies to weather forecasting, because the rotation of the earth strongly influences the weather.

Another hallmark of inertial forces is to be proportional to the mass of the subject on which they act. In fact, it does not concern a force but an acceleration, i.e., the acceleration of the reference system with respect to inertial systems. We interpret it as a force, according to Isaac Newton's second law, but it does not satisfy his third law.

Gravity too happens to be proportional to the mass of the subject on which it acts. At any place, all freely falling subjects experience the same acceleration. Hence, gravity looks like an inertial force. This inspired Albert Einstein to develop the general theory of relativity, defining the metric of space and time such that gravity is eliminated. It leads to a curved space-time, having a strong curvature at places where - according to the classical view - the gravitational field is strong. Besides subjects having mass, massless things experience this field as well. Even light moves according to this metric, as confirmed by ingenious observations.

Yet, gravity is not an inertial force, because it satisfies Newton's third law. Contrary to the centrifugal and Coriolis forces, gravity expresses a subject-subject relation. The presence of heavy matter determines the curvature of space-time. In classical physics, gravity was the prototype of a physical subject-subject relation. One of the unexpected results of Isaac Newton's *Principia* was that the planets attract the sun, besides the

sun attracting the planets. It undermined Newton's Copernican view that the sun is at rest at the centre of the world.[\[21\]](#)

Einstein observed that a gravitational field in a classical inertial frame is equivalent with an accelerating reference system without gravity, like an earth satellite. The popular argument for this *principle of equivalence* is that locally one could not measure any difference.[\[22\]](#) I like to make four comments.

First, on a slightly larger scale the difference between a homogeneous acceleration and a non-homogeneous gravitational field is easily determined.[\[23\]](#) Even in an earth satellite, differential effects are measurable. Except for a homogeneous field, the principle of equivalence is only locally valid.[\[24\]](#)

Second, the curvature of space-time is determined by matter, hence it has a physical source. The gravity of the sun causes the deflection of starlight observed during a total eclipse. An inertial force lacks a physical source.

Third, in non-inertial systems of reference, the law of inertia is invalid. In contrast, the general theory of relativity maintains this law, taking into account the correct metric. A subject on which no force is acting – apart from gravity – moves uniformly with respect to the general relativistic metric. If considered from a classical inertial system, this means a curved and accelerated motion due to gravity. The general relativistic metric does not eliminate, but incorporates gravity.

Finally, in the general relativistic space-time, the speed of light remains the universal unit of velocity. Light moves along a 'straight' line (the shortest line according to Bernhard Riemann's definition). Accelerating reference systems still give rise to inertial forces. This means that Einstein's original intention to prove the equivalence of *all* moving reference systems has failed.

The metrics of special *and* general relativity theory presuppose that light moves at a constant speed everywhere. The empirically confirmed fact that light is subject to gravity necessitates an adaptation of the metric. In the general theory of relativity, kinetic space-time is less symmetric than in the special theory. Because gravity is quite weak compared to other interactions, this symmetry break is only observable at a large scale, at distances where other forces do not act or are neutralized. Where gravity can be neglected, the special theory of relativity is applicable.

The general relativistic space-time is not merely a kinetic, but foremost a physical manifold. The objection against the nineteenth-century ether was that it did not allow of interaction. This objection does not apply to the general relativistic space-time. This acts on matter and is determined by matter.[\[25\]](#)

The general theory of relativity presents models for the physical space-time, which models are testable. It leads to the insight that the physical cosmos is finite and expanding. It came into being about thirteen billions years ago, in a 'big bang'. According to the standard model to be discussed in section 5.1, the fundamental

forces initially formed a single universal interaction. Shortly after the big bang they fell apart by a symmetry break into the present electromagnetic, strong and weak nuclear interaction besides the even weaker gravity. Only then the characters to be discussed in the next two chapters were gradually realized in the astrophysical evolution of the universe.

-
- [1] Shapiro 1997, 158; Torretti 1999, 408-410.
 - [2] Barrow 1992, 129-134; Shapiro 1997, chapter 5.
 - [3] Galileo 1632, 20-22.
 - [4] Van Fraassen 1989, 262.
 - [5] Grünbaum 1973, chapter 1; Sklar 1974, 88-146.
 - [6] Jammer 1954, 150-166; Sklar 1974, 13-54; Torretti 1999, 157.
 - [7] Torretti 1999, 149.
 - [8] Torretti 1999, 155.
 - [9] Dooyeweerd 1953-1958, III, 99.
 - [10] Hart 1984, 156, 263.
 - [11] Stafleu 2019, chapter 11.
 - [12] Stafleu 2019, 4.4.
 - [13] Einstein 1905.
 - [14] Minkowski 1908.
 - [15] Bunge 1967a, 206.
 - [16] Stafleu 2019, 6.6.
 - [17] Rindler 1969, 24, 51-53.
 - [18] Bunge 1967a, 213, 214.
 - [19] French 1965, 494.
 - [20] French 1965, 511.
 - [21] Newton 1687, 419.
 - [22] Bunge 1967a, 207-210.
 - [23] Rindler 1969, 19; Sklar 1974, 70.
 - [24] Bunge 1967a, 210-212.
 - [25] Rindler 1969, 242.
-

Chapter 4

Periodic motion

-
- 4.1. Motion as a relation frame
 - 4.2. The character of oscillations and waves
 - 4.3. A wave packet as an aggregate
 - 4.4. Symmetric and antisymmetric wave functions
-

Encyclopaedia of relations and characters. 4. Periodic motion

4.1. Motion as a relation frame

Chapter 4 investigates characters primarily qualified by kinetic relations. In ancient and medieval philosophy, local motion was considered a kind of change. Classical mechanics emphasized uniform and accelerated motion of unchanging matter. In modern physics, the periodic motion of oscillations and waves is the main theme. In living nature and technology, rhythms play an important part as well.

Twentieth-century physics is characterized by the theory of relativity (chapter 3), by the investigation of the structure of matter (chapter 5), and by quantum physics. The latter is dominated by the duality of waves and particles. Section 4.1 discusses the kinetic relation frame and section 4.2 the kinetic character of oscillations and waves. Section 4.3 deals with the character of a wave packet with its anticipations on physical interaction. Section 4.4 concerns the meaning of symmetrical and antisymmetrical wave functions for physical aggregates.

Kinetically qualified characters are founded in the quantitative or the spatial relation frame and are interlaced with physical characters. Like numbers and spatial forms, periodic motions take part in our daily experience. And like irrational numbers and non-Euclidean space, some aspects of periodic phenomena collide with common sense. Chapter 4 aims to demonstrate that a realistic interpretation of quantum physics is feasible and even preferable to the standard non-realistic interpretations. This requires insight in the phenomenon of character interlacement.

In section 1.2, I proposed relative motion to be the third general type of relations between individual things and processes. Kinetic time is subject to the kinetic order of uniformity and is expressed in the periodicity of mechanical or electric clocks. Before starting the investigation of kinetic characters, I discuss some general features of kinetic time.

Like the rational and real numbers, points on a continuous line are ordered, yet no point has a unique successor (2.2). One cannot say that a point *A* is directly succeeded by a point *B*, because there are infinitely many other points between *A* and *B*. Yet, a uniformly or accelerating moving subject passes the points of its path *successively*.^[1] The succession of temporal moments cannot be reduced to quantitative and/or spatial relations. It presupposes the numerical order of earlier and later and the spatial order of simultaneity, being diachronic and synchronic aspects of kinetic time. Zeno of Elea recognized this long before the Christian era. Nevertheless, until the

seventeenth century, motion was not recognized as an independent principle of explanation.^[2] Later on, it was reinforced by Albert Einstein's theory of relativity (3.3).

The uniformity of kinetic time seems to rest on a convention.^[3] Sometimes it is even meaningful to construct a clock that is not uniform. For instance, the non-uniform physical order of radioactive decay is applied in the dating of archaeological and geological finds.^[4] However, the uniformity of kinetic time together with the periodicity of many kinds of natural motion yields a kinetic norm for clocks. A norm is more than a mere agreement or convention. If applied by human beings constructing clocks, the *law* of inertia becomes a *norm*. A clock does not function properly if it represents a uniform motion as non-uniform.

With increasing clarity, the law of inertia was formulated by Galileo Galilei, René Descartes and others, finding its ultimate form in Isaac Newton's first law of motion: 'Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it.'^[5]

Inertial motion is not in need of a physical cause. Classical and modern physics consider inertial motion to be a state, not a change. In this respect, modern kinematics differs from Aristotle's, who assumed that each change needs a cause, including local motion. Contrary to Aristotle (being the philosopher of common sense), the seventeenth-century physicists considered friction to be a force. Friction causes an actually moving subject to decelerate. In order to maintain a constant speed, another force is needed to compensate for friction. Aristotelians did not recognize friction as a force and interpreted the compensating force as the cause of uniform motion.

Uniformity of motion means that the subject covers equal distances in equal times. But how do we know which times are equal? The diachronous order of earlier and later allows of counting hours, days, months, and years. These units do not necessarily have a fixed duration. In fact, months are not equal to each other, and a leap year has an extra day. Until the end of the Middle Ages, an hour was not defined as 1/24th of a complete day, but as the 1/12th part of a day taken from sunrise to sunset. A day in winter being shorter than in summer, the duration of an hour varied with the seasons. Only after the introduction of mechanical clocks in the fifteenth century, it became customary to relate the length of an hour to the period from noon to noon, such that all hours are equal.

Mechanical clocks measure kinetic time. Time as measured by a clock is called *uniform* if the clock correctly shows that a subject on which no net force is acting moves uniformly.^[6] This appears to be circular reasoning. On the one side, the uniformity of motion means equal distances in equal times. On the other hand, the equality of temporal intervals is determined by a clock subject to the norm that it represents uniform motion correctly.^[7] This circularity is unavoidable, meaning that the uniformity of kinetic time is an unprovable axiom. However, this axiom is not a convention, but an expression of a fundamental and irreducible law.

The uniformity of time is sometimes derived from a *ceteris paribus* argument. If one repeats a process at different moments under exactly equal circumstances, there is no reason to suppose that the process would proceed differently. In particular the duration should be the same. This reasoning is applicable to periodic motions, like in clocks. But it betrays a deterministic vision and is not applicable to stochastic processes like radioactivity. Albert Einstein observed that the equality of covered distances provides a problem as well, because spatial relations are subject to the order of simultaneity, dependent on the state of motion of the clocks used for measuring uniform motion.

Uniformity is a *law* for kinetic time, not an intrinsic *property* of time. There is nothing like a stream of time, flowing independently of the rest of reality. Positivist philosophers denied the ontological status of uniform time. Mach states emphatically: 'The question of whether a motion is uniform *in itself* has no meaning at all. No more can we speak of an "absolute time" (independent of any change).' [8] In my view, the law of inertia determines the meaning of the uniformity of time. According to Hans Reichenbach, it is an 'empirical fact' that different definitions give rise to the same 'measure of the flow of time': natural, mechanical, electronic or atomic clocks, the laws of mechanics, and the fact that the speed of light is the same for all observers.[9] 'It is obvious, of course, that this method does not enable us to discover a "true" time, but that astronomers simply determine with the aid of the laws of mechanics that particular flow of time which the laws of physics implicitly define.' [10] However, if 'truth' means law conformity, 'true time' is the time subject to natural laws. It seems justified to generalize Reichenbach's 'empirical fact', to become the law concerning the uniformity of kinetic time. Rudolf Carnap poses that the choice of the metric of time rests on simplicity: the formulation of natural laws is simplest if one sticks to this convention.[11] But then it is quite remarkable that so many widely different systems confirm to this human agreement. More relevant is to observe that physicists are able to explain all kinds of periodic motions and processes based on laws that presuppose the uniformity of kinetic time. Such an explanation is completely lacking with respect to any alternative metric invented by philosophers. Time only exists in *relations* between events. The uniformity of kinetic time expressed by the law of inertia asserts the existence of motions being uniform *with respect to each other*.

Both classical and relativistic mechanics use this law to introduce inertial systems. An inertial system is a spatio-temporal reference system in which the law of inertia is valid. It can be used to measure accelerated motions as well. Starting with one inertial system, all others can be constructed by using either the Galileo group or the Lorentz group, reflecting the relativity of motion (3.3). Both start from the axiom that kinetic time is uniform.

The law of uniformity concerns all dimensions of kinetic space. Therefore, it is possible to project kinetic time on a linear scale, irrespective of the number of dimensions of kinetic space. Equally interesting is that kinetic time can be projected on a circular scale, as displayed on a traditional clock. The possibility of establishing the equality of temporal intervals is actualized in uniform circular motion, in oscillations, waves, and other periodic processes. Therefore, besides the *general* aspect of uniformity, the time measured by clocks has a *characteristic* component as

well, the periodic character of any clock. Mechanical clocks depend on the regularity of a pendulum or a balance. Electronic clocks apply the periodicity of oscillations in a quartz crystal. Periodicity has always been used for the measurement of time. The days, months, and years refer to periodic motions of celestial bodies. The modern definition of the second depends on atomic oscillations. The periodic character of clocks allows of digitalizing kinetic time, each cycle being a unit, whereas the cycles are countable.

The uniformity of kinetic time as a universal law for kinetic relations and the periodicity of all kinds of periodic processes reinforce each other. Without uniformity, periodicity cannot be understood, and vice versa.

The positivist idea that the uniformity of kinetic time is no more than a convention has the rather absurd consequence, that the periodicity of oscillations, waves and other natural rhythms would be a convention as well.

4.2. The character of oscillations and waves

Periodicity is the distinguishing mark of each primary kinetic character with a tertiary physical characteristic. The motion of a mechanical pendulum, for instance, is primarily characterized by its periodicity and tertiarily by gravitational acceleration. For such an oscillation, the period is constant if the metric for kinetic time is subject to the law of inertia. This follows from an analysis of pendulum motion. The character of a pendulum is applied in a clock. The dissipation of energy by friction is compensated such that the clock is periodic within a specified margin.

Kepler's laws determine the character of periodic planetary motion. Strictly speaking, these laws only apply to a system consisting of two subjects, a star with one planet or binary stars. Both Isaac Newton's law of gravity and the general theory of relativity allow of a more refined analysis. Hence, the periodic motions of the earth and other systems cannot be considered completely apart from physical interactions. However, in this section I shall abstract from physical interaction in order to concentrate on the primary and secondary characteristics of periodic motion.

The simplest case of a periodic motion appears to be uniform circular motion. Its velocity has a constant magnitude whereas its direction changes constantly. Ancient and medieval philosophy considered uniform circular motion to be the most perfect, only applicable to celestial bodies. Seventeenth-century classical mechanics discovered uniform *rectilinear* motion to be more fundamental, the velocity being constant in direction as well as in magnitude. Christiaan Huygens assumed that the

outward centrifugal acceleration is an *effect* of circular motion. Robert Hooke and Isaac Newton demonstrated the inward centripetal acceleration to be the *cause* needed to maintain a uniform circular motion. This force should be specified for any instance of uniform circular motion.

Not moving itself, the circular path of motion is simultaneously a kinetic object and a spatial subject. The position of the centre and the magnitude and direction of the circle's radius vector determine the spatial position of the moving subject on its path. The radius is connected to magnitudes like orbital or angular speed, acceleration, period and phase. The phase (φ) indicates a moment in the periodic motion, the kinetic time (t) in proportion to the period (T): $\varphi = t/T = ft$ modulo 1. If considered an angle, $\varphi = 2\pi ft$ modulo 2π . A phase difference of $\frac{1}{4}$ between two oscillations means that one oscillation reaches its maximum when the other passes its central position.

These quantitative properties allow of calculations and an objective representation of motion.

A uniform circular motion can be constructed as a composition of two mutually perpendicular linear harmonic motions, having the same period and amplitude and a phase difference of one quarter. But then circular uniform motion turns out to be merely a single instance of a large class of two-dimensional harmonic motions. A similar composition of two harmonics – having the same period but different amplitudes or a phase difference other than one quarter – does not produce a circle but an ellipse. If the force is inversely proportional to the square of the distance (like the gravitational force of the sun exerted on a planet), the result is a periodic elliptic motion as well, but this one cannot be constructed as a combination of only two harmonic oscillations. Observe that an ellipse can be defined primarily (spatially) as a conic section, secondarily (quantitatively) by means of a quadratic equation between the co-ordinates [e.g., $(x-x_0)^2/a^2 + (y-y_0)^2/b^2 = 1$], and tertiarily as a path of motion, either kinetically as a combination of periodic oscillations or physically as a planetary orbit.

We can also make a composition of two mutually perpendicular oscillations with *different* periods. Now according to Jules Lissajous, this constitutes a closed curve if and only if the two periods have a harmonic ratio, i.e., a rational number. If the proportion is an octave (1:2), then the resulting figure is a lemniscate (a figure eight). The Lissajous figures derive their specific regularity from periodic motions. Clearly, the two-dimensional Lissajous motions constitute a kinetic character. This character has a primary rational variation in the harmonic ratio of the composing oscillations, as well as a secondary variation in frequency, amplitude and phase. It is interlaced with the character of linear harmonic motion and several other characters. The structure of the path like the circle or the lemniscate is primarily spatially and secondarily quantitatively founded. A symmetry group is interlaced with the character of each Lissajous-figure, the circle being the most symmetrical of all.

In all mentioned characters, we find a typical subject-object relation determining an ensemble of possible variations. In the structure of the circle, the circumference has a fixed proportion to the diameter. This allows of an unbounded variation in diameter.

In the character of the harmonic motion, we find the period (or its inverse, the frequency) as a typical magnitude, allowing of an unlimited variability in period as well as a bounded variation of phase. Varying the typical harmonic ratio results in an infinite but denumerable ensemble of Lissajous-figures.

A linear harmonic oscillation is quantitatively represented by a harmonic function. This is a sine or cosine function or a complex exponential function, being a solution of a differential equation. This equation, the law for harmonic motion, states that the acceleration \mathbf{a} is proportional to the distance \mathbf{x} of the subject to the centre of oscillation \mathbf{x}_0 , according to: $\mathbf{a} = d^2\mathbf{x}/dt^2 = -(2\pi f)^2(\mathbf{x} - \mathbf{x}_0)$ wherein the frequency $f = 1/T$ is the inverse of the period T . The minus sign means that the acceleration is always directed to the centre.

This equation, the law for harmonic motion, concerns mechanical or electronic oscillations, for instance. Primarily, a harmonic oscillation has a specific kinetic character. It is a special kind of motion, characterized by its law and its period. An oscillation is secondarily characterized by magnitudes like its amplitude and phase, not determined by the law but by accidental initial conditions. Hence, the character of an oscillation is kinetically qualified and quantitatively founded.

The harmonic oscillation can be considered the *basic form* of any periodic motion, including the two-dimensional periodic motions discussed above. In 1822, Joseph Fourier demonstrated that each periodic function is the sum or integral of a finite or infinite number of harmonic functions. The decomposition of a non-harmonic periodic function into harmonics is called Fourier analysis.

A harmonic oscillator has a single natural frequency determined by some specific properties of the system. This applies, for instance, to the length of a pendulum; or to the mass of a subject suspended from a spring together with its spring constant; or to the capacity and the inductance in an electric oscillator consisting of a capacitor and a coil. This means that the kinetic character of a harmonic oscillation is interlaced with the physical character of an electric artefact.

Accounting for energy dissipation by adding a velocity-dependent term leads to the equation for a damped oscillator. Now the initial amplitude decreases exponentially. In the equation for a forced oscillation, an additional acceleration accounts for the action of an external periodic force. In the case of resonance, the response is maximal. Now the frequency of the driving force is approximately equal to the natural frequency. Applying a periodic force, pulse or signal to an unknown system and measuring its response is a widely used method of finding the system's natural frequency, revealing its characteristic properties.

An oscillation moving in space is called a *wave*. It has primarily a *kinetic* character, but contrary to an oscillation it is secondarily founded in the *spatial* relation frame. Whereas the source of the wave determines its period, the velocity of the wave, its wavelength and its wave number express the character of the wave itself.

In an isotropic medium, the wavelength λ is the distance covered by a wave with wave velocity v in a time equal to the period T : $\lambda=vT=v/f$. The inverse of the wavelength is the wave number (the number of waves per metre), $\sigma=1/\lambda=f/v$. In three dimensions, the wave number is replaced by the wave vector \mathbf{k} , which besides the number of waves per metre also indicates the direction of the wave motion. In a non-isotropic medium, the wave velocity depends on the direction. The wave velocity has a characteristic value independent of the motion of the source. It is a property of the medium, the kinetic space of a wave that specifically differs from the general kinetic space as described by the Galileo or Lorentz group.

Usually, the wave velocity depends on the frequency as well. This phenomenon is called dispersion. Only light moving in a vacuum is free of dispersion. (The medium of light in vacuum is the electromagnetic field.) The *observed* frequency of a source depends on the relative motions of source, observer and medium. This is the effect called after Christian Doppler.

A wave has a variability expressed by its frequency, phase, amplitude, and polarization. Polarization concerns the direction of oscillation. A sound wave in air is longitudinal, the direction of oscillation being parallel to the direction of motion. Light is transversal, the direction of oscillation being perpendicular to the direction of motion. Light is called unpolarized if it contains waves having all directions of polarization. Light may be partly or completely polarized. It may be linearly polarized (having a permanent direction of oscillation) or circularly polarized (the direction of oscillation itself rotating at a frequency independent of the frequency of the wave itself).

During the motion, the wave's amplitude may decrease. For instance, in a spherical wave the amplitude decreases in proportion to the distance from the centre.

Waves do not interact with each other, but are subject to *superposition*. This is a combination of waves taking into account amplitude as well as phase. Superposition occurs when two waves are crossing each other. Afterwards each wave proceeds as if the other had been absent. *Interference* is a special case of superposition. Now the waves concerned have exactly the same frequency as well as a fixed phase relation. If the phases are equal, interference means an increase of the net amplitude. If the phases are opposite, interference may result in the mutual extinction of the waves.

Just like an oscillation, each wave has a tertiary, usually physical disposition. This explains why waves and oscillations give a technical impression, because technology opens dispositions. During the seventeenth century, the periodic character of sound was discovered in musical instruments. The relevance of oscillations and waves in nature was only fully realized at the beginning of the nineteenth century. This happened after Thomas Young and Augustin Fresnel brought about a break-through in optics by discovering the wave character of light in quite technical experiments. Since the end of the same century, oscillations and waves dominate communication and information technology.

It will be clear that the characters of waves and oscillations are interlaced with each other. A sound wave is caused by a loudspeaker and strikes a microphone. Such an event has a physical character and can only occur if a number of physical conditions are satisfied. However, there is a kinetic condition as well. The frequency of the wave must be adapted to the oscillation frequency of the source or the detector. The wave and the oscillating system are *correlated*. This correlation concerns the property they have in common, i.e., their periodicity, their primary kinetic qualification.

Sometimes an oscillation and a wave are directly interlaced, for instance in a violin string. Here the oscillation corresponds to a standing wave, the result of interfering waves moving forward and backward between the two ends. The length of the string determines directly the wavelength and indirectly the frequency, dependent on the string's physical properties determining the wave velocity. Amplified by a sound box, this oscillation is the source of a sound wave in the surrounding air having the same frequency. In fact, all musical instruments perform according to this principle. The wave is always spatially determined by its wavelength. The length of the string fixes the fundamental tone (the keynote or first harmonic) and its overtones. The frequency of an overtone is an integral number times the frequency of the first harmonic.

A wave equation represents the law for a wave, and a real or complex wave function represents an individual wave. Whereas the equation for oscillations only contains derivatives with respect to time, the wave equation also involves differentiation with respect to spatial co-ordinates. Usually a linear wave equation provides a good approximation for a wave, for example, the equations for the propagation of light. Edwin Schrödinger's non-relativistic equation and Paul Dirac's relativistic equation describe the motion of material waves.

If j and f are solutions of a linear wave equation, then $aj+bf$ is a solution as well, for each pair of real (or complex) numbers a and b . Hence, a linear wave equation has an infinite number of solutions, an ensemble of possibilities. Whereas the equation for an oscillation determines its frequency, a wave equation allows of a broad spectrum of frequencies. The source determines the frequency, the initial amplitude and the phase. The medium determines the wave velocity, the wavelength and the decrease of the amplitude when the wave proceeds away from the source.

Events having their origin in relative motions may be characteristic or not. A solar or lunar eclipse depends on the relative motions of sun, moon and earth. It is accidental and probably unique that the moon and the sun are equally large as seen from the earth, such that the moon is able to cover the sun precisely. Such an event does not correspond to a character. However, wave motion gives rise to several characteristic events satisfying specific laws.

Willebrord Snell's and David Brewster's laws for the refraction and reflection of light at the boundary of two media only depend on the ratio of the wave velocities, the index of refraction. Because this index depends on the frequency, light passing a boundary usually displays dispersion, like in a prism. Dispersion gives rise to various special natural phenomena like a rainbow or a halo, or artificial ones, like Isaac Newton's rings.

If the boundary or the medium has a periodic character like the wave itself, a special form of reflection or refraction occurs if the wavelength fits the periodicity of the lattice. In optical technology, diffraction and reflection gratings are widely applied. Each crystal lattice forms a natural three-dimensional grating for X-rays, if their wavelength corresponds to the periodicity of the crystal lattice according to Lawrence and William Bragg's law.

These are characteristic kinetic phenomena, not because they lack a physical aspect, but because they can be explained satisfactorily by a kinetic theory of wave motion.

4.3. A wave packet as an aggregate

Many sounds are signals. A signal being a pattern of oscillations moves as an aggregate of waves from the source to the detector. This motion has a physical aspect as well, for the transfer of a signal requires energy. But the message is written in the oscillation pattern, being a signal if a human or an animal receives and recognizes it.

A signal composed from a set of periodic waves is called a wave packet. Although a wave packet is a kinetic subject, it achieves its foremost meaning if considered interlaced with a physical subject having a wave-particle character. The wave-particle duality has turned out to be equally fundamental and controversial. Neither experiments nor theories leave room for doubt about the existence of the wave-particle duality. However, it seems to contradict common sense, and its interpretation is the object of hot debates.

René Descartes and Christiaan Huygens assumed that space is completely filled up with matter, that space and matter coincide. They considered light to be a succession of mechanical pulses in space. Descartes believed that light does not move, but has a tendency to move. Huygens denied that wave motion is periodical.^[12] From the fact that planets move without friction, Isaac Newton inferred that interplanetary space is empty. He supposed that light consists of a stream of particles. In order to explain interference phenomena like the rings named after him, he ascribed the light particles (or the medium) properties that we now consider to apply to waves.^[13]

Between 1800 and 1825, Thomas Young in England and Augustin Fresnel in France developed the wave theory of light. Common sense dictated waves and particles to exclude each other, meaning that light is either one or the other. When the wave theory turned out to explain more phenomena than the particle model, the battle was over.^[14] Decisive was Léon Foucault's experimental confirmation in 1854 of the wave-theoretical prediction that light has a lower speed in water than in air. Isaac Newton's particle theory predicted the converse. Light is wave motion, as was later confirmed by James Clerk Maxwell's theory of electromagnetism. Nobody realized

that this conclusion was a *non sequitur*. At most, it could be said that light has wave properties, as follows from the interference experiments of Young and Fresnel, and that *Newton's* particle theory of light was refuted.[15]

Nineteenth-century physics discovered and investigated many other rays. Some looked like light, such as infrared and ultraviolet radiation (about 1800), radio waves (1887), X-rays and gamma rays (1895-96). These turned out to be electromagnetic waves. Other rays consist of particles. Electrons were discovered in cathode rays (1897), in the photoelectric effect and in beta-radioactivity. Canal rays consist of ions and alpha rays of helium nuclei. Cathode rays, canal rays and X-rays are generated in a cathode tube, a forerunner of our television tube, fluorescent lamp and computer screen.

At the end of the nineteenth century, this gave rise to a rather neat and rationally satisfactory worldview. Nature consists partly of particles, for the other part of waves, or of fields in which waves are moving. This *dualistic worldview* assumes that something is either a particle or a wave, but never both, *tertium non datur*.

It makes sense to distinguish a *dualism*, a partition of the world into two compartments, from a *duality*, a two-sidedness. The dualism of waves and particles rested on common sense, one could not imagine an alternative. However, twentieth-century physics had to abandon this dualism perforce and to replace it by the wave-particle *duality*. All elementary things have both a wave and a particle character.

Almost in passing, another phenomenon, called quantization, made its appearance. It turned out that some magnitudes are not continuously variable. The mass of an atom can only have a certain value. Atoms emit light at sharply defined frequencies. Electric charge is an integral multiple of the elementary charge. In 1905 Albert Einstein suggested that light consists of quanta of energy. Einstein never had problems with the duality of waves and particles, but he rejected its probability interpretation.[16] In Niels Bohr's atomic theory (1913), the angular momentum of an electron in its atomic orbit is an integer times Max Planck's reduced constant.[17] (Planck's reduced constant is $h/2\pi$. In Bohr's theory the angular momentum $L = nh/2\pi$, n being the orbit's number. For the hydrogen atom, the corresponding energy is $E_n = E_1/n^2$, with $E_1 = -13.6$ eV, the energy of the first orbit.)

Until Erwin Schrödinger and Werner Heisenberg in 1926 introduced modern quantum mechanics, repeatedly atomic scientists found new quantum numbers with corresponding rules.

The dualism of matter and field, of particles and waves, was productive as long as its components were studied separately. Problems arose when scientists started to work at the interaction between matter and field. The first problem concerned the specific emission and absorption of light restricted to spectral lines, characteristic for chemical elements and their compounds. Niels Bohr tentatively solved this problem in 1913. The spectral lines correspond to transitions between stationary energy states. The second question was under which circumstances light can be in equilibrium with

matter, for instance in an oven. This concerns the shape of the continuous spectrum of black radiation. After a half century of laborious experimental and theoretical work, this problem led to Max Planck's theory (1900) and Albert Einstein's photon hypothesis (1905). According to Planck, the interaction between matter and light of frequency f is in need of the exchange of energy packets of $E = hf$ (h being Planck's constant). Einstein suggested that light itself consists of quanta of energy. Later he added that these quanta have linear momentum as well, proportional to the wave number (s), $p = E/c = hs = h/\lambda$. The relation between energy and frequency ($E = hf$), applied by Bohr in his atomic theory of 1913, was experimentally confirmed by Robert Millikan in 1916, and the relation between momentum and wave number in 1922 by Arthur Compton. The particle character of electromagnetic radiation is easiest to demonstrate with high-energetic photons in gamma- or X-rays. The wave character is easiest proven with low-energetic radiation, with radio or microwaves.

Until 1920, Planck and Einstein did not have many adherents to their views. As late as 1924, Niels Bohr, Hendrik Kramers and John Slater published a theory of electromagnetic radiation, fighting the photon hypothesis at all cost. [18] They went as far as abandoning the laws of conservation of energy and momentum at the atomic level. That was after the publication of Arthur Compton's effect, describing the collision of a gamma-particle with an electron conserving energy and momentum. Within a year, experiments by Walther Bothe and Hans Geiger proved the 'BKS-theory' to be wrong. In 1924 Satyendra Bose and Albert Einstein derived Max Planck's law from the assumption that electromagnetic radiation in a cavity behaves like an ideal gas consisting of photons.

In 1923, Louis de Broglie published a mathematical paper about the wave-particle character of light. [19] Applying the theory of relativity, he predicted that electrons too would have a wave character. The motion of a particle or energy quantum does not correspond to a single monochromatic wave but to a group of waves, a wave packet. The speed of a particle cannot be related to the *wave velocity* (λ/T), being larger than the speed of light for a material particle. Instead, the particle speed corresponds to the speed of the wave packet, the *group velocity*. This is the derivative of frequency with respect to wave number rather than their quotient. Because of the relations of Planck and Einstein, this is the derivative of energy with respect to momentum as well (dE/dp). At most, the group velocity equals the speed of light. (The group velocity dE/dp equals approximately $E/p > c$ and $dE/dp < c$ follow from the relativistic relation between energy and momentum, $E = (E_0^2 + c^2 p^2)^{1/2}$, where E_0 is the particle's rest energy. Only if $E_0 = 0$, $E/p = dE/dp = c$. Observe that the word 'group' for a wave packet has a different meaning than in the mathematical theory of groups.)

In order to test these suggestions, physicists had to find out whether electrons show interference phenomena. Experiments by Clinton Davisson and Lester Germer in America and by George Thomson in England (1927) proved convincingly the wave character of electrons, thirty years after George's father Joseph Thomson established the particle character of electrons. As predicted by Louis De Broglie, the linear momentum turned out to be proportional to the wave number. Afterwards the wave character of atoms and nucleons was demonstrated experimentally.

We have seen that it took quite a long time before physicists accepted the particle character of light. Likewise, the wave character of electrons was not accepted immediately, but about 1930 no doubt was left among pre-eminent physicists.

This meant the end of the wave-particle (or matter-field) *dualism*, implying all phenomena to have either a wave character or a particle character, and the beginning of wave-particle *duality* being recognized as a universal property of matter. In 1927, Niels Bohr called the wave and particle properties *complementary*.^[20] Bohr's principle of complementarity presupposes that quantum phenomena only occur at an atomic level, which is refuted in solid state physics. According to Bohr, a measuring system is an indivisible whole, subject to the laws of classical physics, showing either particle or wave phenomena. In different measurement systems, these phenomena would give incompatible results. This view is out of date.

The concept of complementarity is not well-defined. Sometimes, non-commuting operators and the corresponding variables (like position and momentum) are called 'complementary' as well, at least if their 'commutator' is a number.

An interesting aspect of a wave is that it concerns a movement in motion, a propagating oscillation. Classical mechanics restricted itself to the motion of unchangeable pieces of matter. For macroscopic bodies like billiard balls, bullets, cars and planets, this is a fair approximation, but for microscopic particles it is not. Even in classical physics, the idea of a point-like particle is controversial. Both its mass density and charge density are infinite, and its intrinsic angular momentum cannot be defined.

The experimentally established fact of photons, electrons, and other microsystems having both wave and particle properties does not fit the still popular mechanistic worldview. However, the theory of characters accounts for this fact as follows.

The character of an electron consists of an interlacement of two characters, a *generic* kinetic wave character and an accompanying *specific* particle character that is physically qualified. The specific character (different for different physical kinds of particles) determines primarily how electrons interact with other physical subjects, and secondarily which magnitudes play a role in this interaction. These characteristics distinguish the electron from other particles, like protons and atoms being spatially founded, and like photons having a kinetic foundation (5.2-5.4).

Interlaced with the specific character is a generic pattern of motion having the kinetic character of a wave packet. Electrons share this generic character with all other particles. In experiments demonstrating the wave character, there is little difference between electrons, protons, neutrons, or photons. The generic wave character has primarily a kinetic qualification and secondarily a spatial foundation (4.2). The specific physical character determines the boundary conditions and the actual shape of the wave packet. Its wavelength is proportional to its linear momentum, its frequency to its energy. A free electron's wave packet looks different from that of an electron bound in a hydrogen atom.

The wave character representing the electron's motion has a tertiary characteristic as well, anticipating physical interaction. The wave function describing the composition of the wave packet determines the probability of the electron's performance as a particle in any kind of interaction.

A purely periodic wave is infinitely extended in both space and time. It is unfit to give an adequate description of a moving particle, being localized in space and time. A packet of waves having various amplitudes, frequencies, wavelengths, and phases delivers a pattern that is more or less localized. The waves are superposed such that the net amplitude is zero almost everywhere in space and time. Only in a relatively small interval (to be indicated by Δ) the net amplitude differs from zero.

Let us restrict the discussion to rectilinear motion of a wave packet at constant speed. Now the motion is described by four magnitudes. These are the position (x) of the packet at a certain instant of time (t), the wave number (s) and the frequency (f).

The packet is an aggregate of waves with frequencies varying within an interval Δf and wave numbers varying within an interval Δs . Generally, it is provable that the wave packet in the direction of motion has a minimum dimension Δx such that $\Delta x \cdot \Delta s > 1$. In order to pass a certain point, the packet needs a time Δt , for which $\Delta t \cdot \Delta f > 1$. If we want to compress the packet (Δx and Δt small), the packet consists of a wide spectrum of waves (Δs and Δf large). Conversely, a packet with a well defined frequency (Δs and Δf small) is extended in time and space (Δx and Δt large). It is impossible to produce a wave packet whose frequency (or wave number) has a precise value, and whose dimension is point-like simultaneously. If we make the variation Δs small, the length of the wave packet Δx is large. Or we try to localize the packet, but then the wave number shows a large variation.

Sometimes a wave packet is longer than one might believe. A photon emitted by an atom has a dimension of $\Delta x = c\Delta t$, Δt being equal to the mean duration of the atom's metastable state before the emission. Because Δt is of the order of 10^{-8} sec and $c = 3 \cdot 10^8$ m/sec, the photon's 'coherence length' in the direction of motion is several metres. This is confirmed by interference experiments, in which the photon is split into two parts, to be reunited after the parts have transversed different paths. If the path difference is less than a few metres, interference will occur, but this is not the case if the path difference is much longer. The coherence length of photons in a laser ray is many kilometres long, because in a laser, Δt has been made artificially long.

An oscillating system emits or absorbs a wave packet as a whole. During its motion, the *coherence* of the composing waves is not always spatial. A wave packet can split itself without losing its *kinetic* coherence. This coherence is expressed by phase relations, as can be demonstrated in interference experiments as described above. In general, two different wave packets do not interfere in this way, because their phases are not correlated. This means that a wave packet maintains its *kinetic identity* during its motion. The physical unity of the particle comes to the fore when it is involved in some kind of interaction, for instance if it is absorbed by an atom causing a black spot on a photographic plate or a pulse in a counter tube named after Hans Geiger and

Walther Müller. Emission and absorption are physically qualified events, in which an electron or a photon acts as an indivisible whole.

The identification of a particle with a wave packet seems to be problematic for various reasons. The first problem, the possible splitting and absorption of a wave packet, is mentioned above.

Second, the wave packet of a freely moving particle always expands, because of the composing waves having different velocities. (Light in vacuum is an exception.) Even if the wave packet is initially well localized, gradually it is smeared out over an increasing part of space and time. However, the assumption that the wave function satisfies a linear wave equation is a simplification of reality. Wave motion can be non-linearly represented by a 'soliton' that does not expand. Unfortunately, a non-linear wave equation is mathematically more difficult to treat than a linear one.

Third, in 1926 Werner Heisenberg observed that the wave packet is subject to a law known as indeterminacy relation, uncertainty relation or Heisenberg relation. As a matter of fact, there is as little agreement about its definition as about its name.

Combining the relations $\Delta x \cdot \Delta s > 1$ and $\Delta t \cdot \Delta f > 1$ with those of Max Planck ($E = hf$) and Albert Einstein ($p = hs$) leads to Heisenberg's relations for a wave packet: $\Delta x \cdot \Delta p > h$ and $\Delta t \cdot \Delta E > h$. (The values of '1' respectively 'h' in the mentioned relations indicate an order of magnitude. Sometimes other values are given).^[21] The meaning of Δx etc. is given above. In particular, Δt is the time the wave packet needs to pass a certain point. If $\Delta x \cdot \Delta s = \Delta t \cdot \Delta f = 1$, the wave packet's speed $v = \Delta x / \Delta t = \Delta f / \Delta s$ is approximately the group velocity df/ds , according to Louis De Broglie. This interpretation is the oldest one, for the indeterminacy relations – without Planck's constant – were applied in communication theory (where Δf is the band width) long before the birth of quantum mechanics.^[22] It is interesting to observe that the indeterminacy relations are not characteristic for quantum mechanics, but for wave motion. The relations are an unavoidable consequence of the wave character of particles and of signals. I shall discuss some alternative interpretations, in particular paying attention to Heisenberg's relation between energy and time.^[23]

Quantum mechanics connects any variable magnitude with a Hermitean operator having eigenfunctions and eigenvalues (2.3). The eigenvalues are the possible values for the magnitude in the system concerned. In a measurement, the scalar product of the system's state function with an eigenfunction of the operator is the square of the probability that the corresponding eigenvalue will be realized.

If two operators act successively on a function, the result may depend on their order. Heisenberg's relation $\Delta x \cdot \Delta p > h$ can be derived as a property of the non-commuting operators for position and linear momentum. In fact, each pair of non-commuting operators gives rise to a similar relation. This applies, e.g., to each pair out of the three components of angular momentum. Consequently, only one component of an electron's magnetic moment (usually along a magnetic field) can be measured. The

other two components are undetermined, as if the electron exerts a precessional motion about the direction of the magnetic field.

Remarkably, there is no operator for kinetic time. Therefore, some people deny the existence of a Heisenberg relation for time and energy.[24] On the other hand, the operator for energy, called Hamilton-operator or Hamiltonian after William Hamilton, is very important. Its eigenvalues are the energy levels characteristic for e.g. an atom or a molecule. Each operator commuting with the Hamiltonian represents a 'constant of the motion' subject to a conservation law.

From the wave function, the probability to find a particle in a certain state can be calculated. Now the indeterminacy is a measure of the mean standard deviation, the statistical inaccuracy of a probability calculation. The indeterminacy of time can be interpreted as the mean lifetime of a metastable state. If the lifetime is large (and the state is relatively stable), the energy of the state is well defined. The rest energy of a short living particle is only determined within the margin given by the Heisenberg relation for time and energy.

This interpretation is needed to understand why an atom is able to absorb a light quantum emitted by another atom in similar circumstances. Because the photon carries linear momentum, both atoms get momentum and kinetic energy. The photon's energy would fall short to excite the second atom. Usually this shortage is smaller than the uncertainty in the energy levels concerned. However, this is not always the case for atomic nuclei. Unless the two nuclei are moving towards each other, the process of emission followed by absorption would be impossible. Rudolf Mössbauer discovered this consequence of Heisenberg's relations in 1958. Since then, Mössbauer's effect became an effective instrument for investigating nuclear energy levels.

The position of a wave packet is measurable within a margin of Δx and its linear momentum within a margin of Δp . Both are as small as experimental circumstances permit, but their product has a minimum value determined by Heisenberg's relation. The accuracy of the measurement of position restricts that of momentum.

Initially the indeterminacy was interpreted as an *effect* of the measurement disturbing the system. The measurement of one magnitude disturbs the system such that another magnitude cannot be measured with an unlimited accuracy. Heisenberg explained this by imagining a microscope exploiting light to determine the position and the momentum of an electron.[25] Later, this has appeared to be an unfortunate view. It seems better to consider Heisenberg's relations to be the *cause* of the limited accuracy of measurement, rather than to be its effect.

The Heisenberg relation for energy and time has a comparable consequence for the measurement of energy. If a measurement has duration Δt , its accuracy cannot be better than $\Delta E > h/\Delta t$.

In quantum mechanics, the law of conservation of energy achieves a slightly different form. According to the classical formulation, the energy of a closed system is constant. In this statement, time does not occur explicitly. The system is assumed to be isolated for an indefinite time, and that is questionable. Heisenberg's relation suggests a new formulation. For a system isolated during a time interval Δt , the energy is constant within a margin of $\Delta E \approx h/\Delta t$. Within this margin, the system shows spontaneous energy fluctuations, only relevant if Δt is very small. In fact, the value of ΔE is less significant than the relative indeterminacy $\Delta E/E$. For a macroscopic system the energy E is so much larger than ΔE that the energy fluctuations can be neglected, and the law of conservation of energy remains valid.

According to quantum field theory, a physical vacuum is not an empty space. Spontaneous fluctuations may occur. A fluctuation leads to the creation and annihilation of a virtual photon or a virtual pair consisting of a particle and an antiparticle, having an energy of ΔE , within the interval $\Delta t < h/\Delta E$. Meanwhile the virtual particle or pair is able to exert an interaction, e.g. a collision between two real particles. (Such virtual processes are depicted in the diagrams called after Richard Feynman.) Virtual particles are not directly observable but play a part in several real processes.

The amplitude of waves in water, sound, and light corresponds to a measurable physical *real* magnitude. In water this is the height of its surface, in sound the pressure of air, in light the electromagnetic field strength. The energy of the wave is proportional to the square of the amplitude. This interpretation is not applicable to the waves for material particles like electrons. In this case the wave has a less concrete character, it has no direct physical meaning. Even in mathematical terms, the wave is not real, for the wave function has a complex value.

In 1926, Max Born offered a new interpretation, since then commonly accepted.^[26] He stated that a wave function (real or complex) is a probability function. In a footnote added in proof, Born observed that the probability is proportional to the *square* of the wave function.

The wave function we are talking about is prepared at an earlier interaction, for instance, the emission of the particle. It changes during its motion, and one of its possibilities is realized at the next interaction, like the particle's absorption. The wave function expresses the transition probability between the initial and the final state.^[27]

This probability may concern any measurable property that is variable. Hence, it does not concern natural constants like the speed of light or the charge of the electron. According to Born, the probability interpretation bridges the apparently incompatible wave and particle aspects: 'The true philosophical import of the statistical interpretation consists in the recognition that the wave-picture and the corpuscle-picture are not mutually exclusive, but are two complementary ways of considering the same process'.^[28] Wave properties determine the probability of position, momentum, etc., traditionally considered properties of particles.

Classical mechanics used statistics as a mathematical means, assuming that the particles behave deterministic in principle. In 1926, Born's probability interpretation put a definitive end to mechanist determinism, having lost its credibility before because of radioactivity. Waves and wave motion are still determined, e.g. by Schrödinger's equation, even if no experimental method exists to determine the phase of a wave. However, the wave function determines only the probability of future interactions. The fact that quantum physics is a stochastic theory has evoked widely differing reactions. Albert Einstein considered the theory incomplete. Max Born stressed that at least waves behave deterministically, only its interpretation having a statistical character. Niels Bohr accepted a fundamental stochastic element in his world-view. In quantum mechanics, the particles themselves behave stochastically.

Even more strange is that chance is subject to interference. In the traditional probability calculus (2.4) probabilities can be added or multiplied. Nobody ever imagined that probabilities could interfere. Interference of waves may result in an increase of probability, but to a decrease as well, even to the extinction of probability. Hence, besides a probability interpretation of waves, we have a wave interpretation of probability.[29]

Outside quantum mechanics, this is still unheard of, not only in daily life and the humanities, but in sciences like biology and ethology as well. The reason is that interference of probabilities only occurs *as long as there is no physical interaction* by which a chance realizes itself. Observe that an interference-*experiment* aims at *demonstrating* interference. This is only possible if the interference of waves is *followed* by an interaction of the particles concerned with, e.g., a screen. The absence of physical interaction is an exceptional situation. It only occurs if the system concerned has no internal interactions (or if these are frozen), as long as it moves freely. In macroscopic bodies, interactions occur continuously and interference of probabilities does not occur. Therefore, the phenomenon of interference of chances is unknown outside quantum physics.

The concept of probability or chance *anticipates* the physical relation frame, because only by means of a physical interaction a chance can be realized. An open-minded spectator observes an asymmetry in time. Probability always concerns future events. It draws a boundary line between a possibility in the present and a realization in the future. For this realization, a physical interaction is needed. The wave equation and the wave function describe probabilities, not their realization. The wave packet anticipates a physical interaction leading to the realization of a chance, but is itself a kinetic subject, not a physical subject. If the particle realizes one of its possibilities, it simultaneously destroys all alternative possibilities. In that respect, there is no difference between quantum mechanics and classical theories of probability.

As long as the position of an electron is not determined, its wave packet is extended in space and time. As soon as an atom absorbs the electron at a certain position, the probability to be elsewhere collapses to zero. Theoretically, this means the projection of a state vector on one of the eigenvectors of Hilbert space, representing all possible states of the system. 'No other permanent or transient principle of physics has ever

given rise to so many comments, criticisms, pleadings, deep remarks, and plain nonsense as the wave function collapse.’[30] In particular, the assumptions that probability is an expression of our limited knowledge of a system and that the observer causes the reduction of the wave packet, have led to a number of subjectivist and solipsist interpretations of quantum physics and related problems, of which I shall only briefly discuss that of Schrödinger’s cat. This so-called reduction of the wave packet requires a velocity far exceeding the speed of light. However, this reduction concerns the wave character, not the physical character of the particle. It does not counter the physical law that no material particle can move faster than light.

Likewise, Schrödinger’s equation describes the states of an atom or molecule and the transition probabilities between states. It does not account for the actual transition from a state to an eigenstate, when the system experiences a measurement or another kind of interaction. According to Nancy Cartwright, ‘This transition therefore does not belong to elementary quantum dynamics. But it is meant to express a physical interaction between the measured object and the measuring apparatus, which one would expect to be a direct consequence of dynamics’.[31] ‘Von Neumann claimed that the reduction of the wave packet occurs when a measurement is made. But it also occurs when a quantum system is prepared in an eigenstate, when one particle scatters from another, when a radioactive nucleus disintegrates, and in a large number of other transition processes as well ... There is nothing peculiar about measurement, and there is no special role for consciousness in quantum mechanics.’[32] But contrary to Cartwright stating: ‘... there are not two different kinds of evolution in quantum mechanics. There are evolutions that are correctly described by Schrödinger’s equation, and there are evolutions that are correctly described by something like von Neumann’s projection postulate. But these are not different kinds in any physically relevant sense’,[33] I believe that there is a difference. The first concerns a reversible motion, the second an irreversible physical process: ‘Indeterministically and irreversibly, without the intervention of any external observer, a system can change its state ... When such a situation occurs, the probabilities for these transitions can be computed; it is these probabilities that serve to interpret quantum mechanics.’[34]

Is the problem of the reduction of the wave packet relevant for macroscopic bodies as well? Historically, this question is concentrated in the problem of Edwin Schrödinger’s cat, hypothetically locked up alive in a non-transparent case. A mechanism releases a mortal poison at an unpredictable instant, for instance controlled by a radioactive process. As long as the case is not opened, one may wonder whether the cat is still alive. If quantum mechanics is applied consequently, the state of the cat is a mixture, a superposition of two eigenstates, dead and alive, respectively.

The principle of decoherence, developed at the end of the twentieth century, may provide a satisfactory answer. For a macroscopic body, a state being a combination of eigenstates will spontaneously change very fast into an eigenstate, because of the many interactions taking place in the macroscopic system itself. This solves the problem of Schrödinger’s cat, for each superposition of dead *and* alive transforms itself almost immediately into a state of dead *or* alive. The principle of decoherence is

in some cases provable, though it is not proved generally.[35] Decoherence even occurs in quite small molecules.[36] There are exceptions too, in systems without much internal energy dissipation, e.g. electromagnetic radiation in a transparent medium and superconductors (5.4)[37].

The principle of decoherence is part of a realistic interpretation of quantum physics. It does not idealize the 'reduction of the wave packet' to a projection in an abstract state space. It takes into account the character of the macroscopic system in which a possible state is realized by means of a physical interaction.

The so-called measurement problem constitutes the nucleus of what is usually called the interpretation of quantum mechanics. 'The interpretive challenge of quantum theory is often presented in terms of the measurement problem: i.e., that the formalism itself does not specify that only one outcome happens, nor does it explain why or how that particular outcome happens. This is the context in which it is often asserted that the theory is incomplete and is therefore in need of alteration in some way.'[38] It is foremost a philosophical problem, not a physical one, which is remarkable, because measurement is part of experimental physics, and the starting point of theoretical physics. After the development of quantum physics, both experimental and theoretical physicists have investigated the relevance of symmetry, and the structure of atoms and molecules, solids and stars, and subatomic structures like nuclei and elementary particles. Apparently, this has escaped the attention of many philosophers, who are still discussing the consequences of Heisenberg's indeterminacy relations.

4.4. Symmetric and antisymmetric wave functions

The concept of probability is applicable to a single particle as well as to a homogeneous set of similar particles, a gas consisting of molecules, electrons or photons. In order to study such systems, since circa 1860 statistical physics has developed various mathematical methods. A *distribution function* points out how the energy is distributed over the particles, how many particles have a certain energy value, and how the average energy depends on temperature. In any distribution function, the temperature is an important equilibrium parameter.

Classical physics assigned each particle its own state, but in quantum physics, this would lead to wrong results. It is better to design the possible states, and to calculate how many particles occupy a given state, without questioning which particle occupies which state. It turns out that there are two entirely different cases, referring to 'bosons', respectively 'fermions'. [39]

In the first case, the occupation number of particles in a well-defined state is unlimited. *Bosons* like photons are subject to a distribution function in 1924 derived by Satyendra Bose and published by Albert Einstein, hence called Bose-Einstein statistics. Bosons have an integral spin, the occupation number of each state may vary from zero to infinity. An *integral* spin means that the intrinsic angular momentum is an integer times Planck's reduced constant, 0 , $h/2\pi$, $2h/2\pi$, etc. A *half-integral* spin means that the intrinsic angular moment has values like $(1/2)h/2\pi$, $(3/2)h/2\pi$. I shall not discuss the connection of integral spin with bosons and half-integral spin with fermions.

In the other case, each well-defined state is occupied by at most one particle, according to Wolfgang Pauli's exclusion principle. The presence of a particle in a given state excludes the presence of another similar particle in the same state. *Fermions* like electrons, protons, and neutrons have a half-integral spin. They are subject to the distribution function that Enrico Fermi and Paul Dirac derived in 1926.

In both cases, the distribution approximates the classical Maxwell-Boltzmann distribution function, if the mean occupation of available states is much smaller than 1. This applies to molecules in a classical gas (2.4).

The distinction of fermions and bosons rests on permutation symmetry. In a finite set the elements can be ordered into a sequence and numbered using the natural numbers as indices. For n elements, this can be done in $n!=1.2.3.4\dots n$ different ways. The $n!$ permutations are symmetric if the elements are indistinguishable. Permutation symmetry is not spatial but quantitative.

In a system consisting of a number of similar particles, the state of the aggregate can be decomposed into a product of separate states for each particle apart. (It is by no means obvious that the state function of an electron or photon gas can be written as a product (or rather a sum of products) of state functions for each particle apart, but it turns out to be a quite close approximation.) A permutation of the order of similar particles should not have consequences for the state of the aggregate as a whole. However, in quantum physics only the *square* of a state is relevant to probability calculations. Hence, exchanging two particles allows of two possibilities: either the state is multiplied by $+1$ and does not change, or it is multiplied by -1 . In both cases, a repetition of the exchange produces the original state. In the first case, the state is called symmetric with respect to a permutation, in the second case antisymmetric.

In the antisymmetric case, if two particles would occupy the same state an exchange would simultaneously result in multiplying the state by $+1$ (because nothing changes) and by -1 (because of antisymmetry), leading to a contradiction. Therefore, two particles cannot simultaneously occupy the same state. This is Wolfgang Pauli's exclusion principle concerning fermions. No comparable principle applies to bosons, having symmetric wave functions with respect to permutation.

Both a distribution function like the Fermi-Dirac statistics and Pauli's exclusion principle are only applicable to a homogeneous aggregate of *similar* particles. In a

heterogeneous aggregate like a nucleus, they must be applied to the protons and neutrons separately.

The distinction of fermions and bosons, and the exclusion principle for fermions, have a fundamental significance for the understanding of the characters of material things containing several similar particles. To a large extent, it explains the orbital structure of atoms and the composition of nuclei from protons and neutrons.

When predicting the wave character of electrons, Louis de Broglie suggested that the stability of the electronic orbit in a hydrogen atom is explainable by assuming that the electron moves around the nucleus as a standing wave. This implies that the circumference of the orbit is an integral number times the wavelength. From the classical theory of circular motion, he derived that the orbital angular momentum should be an integral number times Max Planck's reduced constant ($h/2\pi$). This is precisely the quantum condition applied by Niels Bohr in 1913 in his first atomic theory. For a uniform circular motion with radius r , the angular momentum $L=rp$. The linear momentum $p = h/\lambda$ (λ for lambda, the wave length) according to Einstein. If the circumference $2\pi r = n\lambda$, n being a positive integer, then $L=n\lambda p/2\pi=nh/2\pi$. Quantum mechanics allows of the value $L=0$ for orbital angular momentum. This has no analogy as a standing wave on the circumference of a circle.

The atomic physicists at Copenhagen, Göttingen, and Munich considered de Broglie's idea rather absurd, but it received support from Albert Einstein, and it inspired Edwin Schrödinger to develop his wave equation.[40] In a stable system, Schrödinger's equation is independent of time and its solutions are stationary waves, comparable to the standing waves in a violin string or an organ pipe. Only a limited number of frequencies are possible, corresponding to the energy levels in atoms and molecules. In contrast, a time-dependent Schrödinger equation describes *transitions* between energy levels, giving rise to the discrete emission and absorption spectra characteristic for atoms and molecules. Although one often speaks of *the* Schrödinger equation, there are many variants, one for each physical character. Each variant specifies the system's boundary conditions and expresses the law for the possible motions of the particles concerned.

In the practice of solid-state physics, the exclusion principle is more important than Schrödinger's equation. This can be elucidated by discussing the model of particles confined to a rectangular box. Again, the wave functions look like standing waves.

In a good approximation the valence electrons in a metal or semiconductor are not bound to individual atoms but are free to move around. The mutual repulsive electric force of the electrons compensates for the attraction by the positive ions. The electron's energy consists almost entirely of kinetic energy, $E=p^2/2m$, if p is its linear momentum and m its mass.

Because the position of the electron is confined to the box, in Heisenberg's relation Δx equals the length of the box (analogous for y and z). Because Δx is relatively large, Δp is small and the momentum is well defined. Hence the momentum characterizes

the state of each electron and the energy states are easy to calculate. In a three-dimensional momentum space a state denoted by the vector \mathbf{p} occupies a volume $\Delta\mathbf{p}$. Momentum space is a three-dimensional diagram for the vector \mathbf{p} 's components, p_x , p_y and p_z . The volume of a state equals $\Delta\mathbf{p}=\Delta p_x\Delta p_y\Delta p_z$. In the described model, the states are mostly occupied up till the energy value E_F , the 'Fermi-energy', determining a sphere around the origin of momentum space. Outside the sphere, most states are empty. A relatively thin skin, its thickness being proportional to the temperature, separates the occupied and empty states.

According to the exclusion principle, a low energy state is occupied by two electrons (because there are two possible spin states), whereas high-energy states are empty. In a metal, this leads to a relatively sharp separation of occupied and empty states. The mean kinetic energy of the electrons is almost independent of temperature, and the specific heat is proportional to temperature, strikingly different from other aggregates of particles.

Mechanical oscillations or sound waves in a solid form wave packets. These bosons are called phonons or sound particles. Bose-Einstein statistics leads to Peter Debye's law for the specific heat of a solid. At a moderate temperature the specific heat is proportional to the third power of temperature. Except for very low temperatures, the electrons contribute far less to the specific heat of a solid than the phonons do. The number of electrons is independent of temperature, whereas the number of phonons in a solid or photons in an oven strongly depends on temperature.

A similar situation applies to an oven, in which electromagnetic radiation is in thermal equilibrium. According to Planck's law of radiation, the energy of this boson gas is proportional to the fourth power of temperature. For a gas satisfying the Maxwell-Boltzmann distribution, the energy is proportional to temperature. Some people who got stuck in classical mechanics *define* temperature as a measure of the mean energy of molecules. Which meaning such a definition should have for a fermion gas or boson gas is unclear.

Hence, the difference between fermion and boson aggregates comes quite dramatically to the fore in the temperature dependence of their energy. Amazingly, the *physical* character of the electrons, phonons, and photons plays a subordinate part compared to their *kinetic* character. Largely, the symmetry of the wave function determines the properties of an aggregate. Consequently, a neutron star has much in common with an electron gas in a metal.

The existence of antiparticles is a consequence of a symmetry of the relativistic wave equation. The quantum mechanics of Erwin Schrödinger and Werner Heisenberg in 1926 was not relativistic, but about 1927 Paul Dirac found a relativistic formulation.^[41] From his equation follows the electron's half-integral angular momentum, not as a spinning motion as conceived by its discoverers, Samuel Goudsmit and George Uhlenbeck, but as a symmetry property (still called spin).

Dirac's wave equation had an unexpected result, to wit the existence of negative energy eigenvalues for free electrons. According to relativity theory, the energy E and momentum p for a freely moving particle with rest energy $E_0 = m_0 c^2$ are related by the formula: $E^2 = E_0^2 + (cp)^2$. For a given value of the linear momentum p , this equation has both positive and negative solutions for the energy E . The positive values are minimally equal to the rest energy E_0 and the negative values are maximally $-E_0$. This leaves a gap of twice the rest energy, about 1 MeV for an electron, much more than the energy of visible light, being about 5 eV per photon. Classical physics could ignore negative solutions, but this is not allowed in quantum physics. Even if the energy difference between positive and negative energy levels is large, the transition probability is not zero. In fact, each electron should spontaneously jump to a negative energy level, releasing a gamma particle having an energy of at least 1 MeV.

Dirac took recourse to Pauli's exclusion principle. By assuming all negative energy levels to be occupied, he could explain why these are unobserved most of the time, and why many electrons have positive energy values. An electron in one of the highest negative energy levels may jump to one of the lowest positive levels, absorbing a gamma particle having an energy of at least 1 MeV. The reverse, a jump downwards, is only possible if in the nether world of negative energy levels, at least one level is unoccupied. Influenced by an electric or magnetic field, such a hole moves as if it were a positively charged particle. Initially, Dirac assumed protons to correspond to these holes, but it soon became clear that the rest mass of a hole should be the same as that of an electron.

After Carl Anderson in 1932 discovered the positron, a positively charged particle having the electron's rest mass, this particle was identified with a hole in Dirac's nether world. This identification took some time.^[42] The assumption of the existence of a positive electron besides the negative one was in 1928 much more difficult to accept than in 1932. In 1928, physics acknowledged only three elementary particles, the electron, the proton and the photon. In 1930, the existence of the neutrino was postulated and in 1932, Chadwick discovered the neutron. The completely occupied nether world of electrons is as inert as the nineteenth century ether. It neither moves nor interacts with any other system. That is why we do not observe it. For those who find this difficult to accept, alternative theories are available explaining the existence of antiparticles. Experiments pointed out that an electron is able to annihilate a positron, releasing at least two gamma particles. In the inertial system in which the centre of mass for the electron-positron pair is at rest, their total momentum is zero. Because of the law of conservation of momentum, the annihilation causes the emergence of at least two photons, having opposite momentum.

Meanwhile it is established that besides electrons all particles, bosons included, have antiparticles. Only a photon is identical to its antiparticle. The existence of antiparticles rests on several universally valid laws of symmetry. A particle and its antiparticle have the same mean lifetime, rest energy and spin, but opposite values for charge, baryon number, or lepton number (5.2).

However, if the antiparticles are symmetrical to particles, why are there so few? (Or why is Dirac's nether world nearly completely occupied?) Probably, this problem can only be solved within the framework of a theory about the early development of the cosmos.

The image of an infinite set of unobservable electrons having negative energy, strongly defeats common sense. However, it received unsolicited support from the so-called band theory in solid-state physics, being a refinement of the earlier discussed free-electron model. The influence of the ions is not completely compensated for by the electrons. An electric field remains having the same periodic structure as the crystal. Taking this field into account, Rudolf Peierls developed the band model. It explains various properties of solids quite well, both quantitatively and qualitatively.

A band is a set of neighbouring energy levels separated from other bands by an energy gap. (A band is comparable to an atomic shell but has a larger bandwidth.) It may be fully or partly occupied by electrons, or it is unoccupied. Both full and empty bands are physically inert. In a metal, at least one band is partly occupied, partly unoccupied by electrons. An isolator has only full (i.e., entirely occupied) bands besides empty bands. The same applies to semiconductors, but now a full band is separated from an empty band by a relatively small gap. According to Peierls in 1929, if energy is added in the form of heat or light (a phonon or a photon), an electron jumps from the lower band to the higher one, leaving a hole behind. This hole behaves like a positively charged particle. In many respects, an electron-hole pair in a semiconductor looks like an electron-positron pair. Only the energy needed for its formation is about a million times smaller. Dirac and Heisenberg corresponded with each other about both theories, initially without observing the analogy.^[43]

Another important difference should be mentioned. The set of electron states in Dirac's theory is an *ensemble*. In the class of possibilities independent of time and space, half is mostly occupied, the other half is mostly empty. There is only one nether world of negative energy values. In contrast, the set of electrons in a semiconductor is a spatially and temporally restricted *collection* of electrons, in which some electron states are occupied, others unoccupied. There are as many of these collections as there are semiconductors. To be sure, Peierls was interested in an ensemble as well. In his case, this is the ensemble of all semiconductors of a certain kind. This may be copper oxide, the standard example of a semiconductor in his days, or silicon, the base material of modern chips. But this only confirms the distinction from Dirac's ensemble of electrons.

Common sense did not turn out to be a reliable guide in the investigation of characters. At the end of the nineteenth century, classical mechanics was considered the paradigm of science. Yet, even then it was clear that daily experience was in the way of the development of electromagnetism, for instance. The many models of the ether were more an inconvenience than a stimulus for research.

When relativity theory and quantum physics unsettled classical mechanics, this led to uncertainty about the reliability of science. At first, the oncoming panic was warded

off by the reassuring thought that the new theories were only valid in extreme situations. These situations were, for example, a very high speed, a total eclipse, or a microscopic size. However, astronomy cannot cope without relativity theory, and chemistry fully depends on quantum physics. All macroscopic properties and phenomena of solid-state physics can only be explained in the framework of quantum physics.

Largely, daily experience rests on habituation. In hindsight, it is easy to show that classical mechanics collided with common sense in its starting phase with respect to the law of inertia. Action at a distance in Isaac Newton's *Principia* evoked the abhorrence of his contemporaries, but the nineteenth-century public did not experience any trouble with this concept. In the past, mathematical discoveries would cause heated discussions, but the rationality of irrational numbers or the reality of non-Euclidean spaces is now accepted almost as a matter of course.

This does not mean that common sense is always wrong in scientific affairs. The irreversibility of physical processes is part of daily experience. In the framework of the mechanist worldview of the nineteenth century, physicists and philosophers have stubbornly but in vain tried to reduce irreversible processes to reversible motion, and to save determinism. This is also discernible in attempts to find (mostly mathematical) interpretations of quantum mechanics that allow of temporal reversibility and of determinism, such as the so-called many-worlds interpretation, and the transaction interpretation.

Since the twentieth century, mathematics, science and technology dominate our society to such an extent, that new developments are easier to integrate in our daily experience than before. Science has taught common sense to accept that the characters of natural things and events are neither manifest nor evident. The hidden properties of matter and of living beings brought to light by the sciences are applicable in a technology that is accessible for anyone but understood by few. This technology has led to an unprecedented prosperity. Our daily experience adapts itself easily and eagerly to this development.

[1] Lucas 1973, 29.

[2] Stafleu 1987, 61; 2019, 2.1.

[3] Reichenbach 1957, 116-119; Grünbaum 1968, 19, 70; 1973, 22; Stafleu 2019, 4.4-4.5.

[4] Grünbaum 1973, 22-23.

[5] Newton 1687, 13.

[6] Margenau 1950, 139.

[7] Maxwell 1877, 29; Cassirer 1921, 364.

[8] Mach 1883, 217.

[9] Reichenbach 1957, 117.

[10] Reichenbach 1957, 118.

- [11] Carnap 1966, chapter 8.
 - [12] Huygens 1690, 15; Sabra 1967, 212; Stafleu 2019, 3.2.
 - [13] Newton 1704, 278-282; Sabra 1967, chapter 13.
 - [14] Achinstein 1991, 24.
 - [15] Hanson 1963, 13; Jammer 1966, 31.
 - [16] Klein 1964, Pais 1982, part IV.
 - [17] Pais 1991, 150.
 - [18] Bohr, Kramers, Slater 1924; cp. Slater 1975, 11; Pais 1982, chapter 22; 1991, 232-239.
 - [19] Darrigol 1986.
 - [20] Bohr 1934, chapter 2; Bohr 1949; Meyer-Abich 1965; Jammer 1966, chapter 7; 1974, chapter 4; Pais 1991, 309-316, 425-436.
 - [21] Messiah 1961, 133.
 - [22] Bunge 1967a, 265.
 - [23] Margenau 1950, chapter 18; Messiah 1961, 129-149; Jammer 1966, chapter 7; Jammer 1974, chapter 3; Omnès 1994, chapter 2.
 - [24] Bunge 1967a, 248, 267.
 - [25] Heisenberg 1930, 21-23.
 - [26] Jammer 1974, 38-44.
 - [27] Cartwright 1983, 179.
 - [28] M. Born, *Atomic physics*, Blackie 1944, quoted by Bastin (ed.) 1971, 5.
 - [29] Heisenberg 1958, 25.
 - [30] Omnès 1994, 509.
 - [31] Omnès 1994, 84:
 - [32] Cartwright 1983, 195.
 - [33] Cartwright 1983, 179.
 - [34] Cartwright 1983, 179.
 - [35] Omnès 1994, chapter 7, 484-488; Torretti 1999, 364-367.
 - [36] Omnès 1994, 299-302.
 - [37] Omnès 1994, 269.
 - [38] Kastner 2013, 202:
 - [39] Jammer 1966, 338-345.
 - [40] Klein 1964; Raman, Forman 1969.
 - [41] Kragh 1990, chapter 3, 5.
 - [42] Hanson 1963, chapter IX.
 - [43] Kragh 1990, 104-105.
-

Chapter 5

Physical characters

5.1. The unification of physical interactions

5.2. The character of electrons

5.3. The quantum ladder

5.4. Individualized currents

5.5. Aggregates and statistics

5.6. Coming into being, change and decay

Encyclopaedia of relations and characters. 5. Physical characters

5.1. The unification of physical interactions

The aim of chapter 5 is a philosophical analysis of physical characters. Their relevance can hardly be overestimated. The discovery of the electron in 1897 provided the study of the structure of matter with a strong impulse, both in physics and in chemistry. Our knowledge of atoms and molecules, of nuclei and sub-atomic particles, of stars and stellar systems, dates largely from the twentieth century. The significance of electrotechnology and electronics for the present society is overwhelming.

The physical aspect of the cosmos is characterized by interactions between two or more subjects. Interaction is a relation different from the quantitative, spatial, or kinetic relations, on which it can be projected. It is subject to natural laws. Some laws are specific, like the electromagnetic ones, determining characters of physical kinds. Some laws are general, like the laws of thermodynamics and the laws of conservation of energy, linear and angular momentum. The general laws constitute the physical-chemical relation frame. Both for the generic and the specific laws, physics has reached a high level of unification.

Because of their relevance to study types of characters, this chapter starts with an analysis of the projections of the physical relation frame onto the three preceding ones (5.1). Next, I investigate the characters of physically stable things, consecutively quantitatively, spatially, and kinetically founded (5.2-5.4). Section 5.5 surveys aggregates and statistics. Finally, section 5.6 reviews processes of coming into being, change, and decay.

The existence of physically qualified things and events implies their interaction, the universal physical relation. If something could not interact with anything else it would be inert. It would not exist in a physical sense, and it would have no physical place in the cosmos. Groups, spatial figures, waves and oscillations do not interact, hence are not physical unless interlaced with physical characters. The noble gases are called inert because they hardly ever take part in chemical compounds, yet their atoms are able to collide with each other. The most inert things among subatomic particles are the neutrino's, capable of flying through the earth with a very small probability of colliding with a nucleus or an electron. Nevertheless, neutrinos are detectable and have been detected.

Wolfgang Pauli postulated the existence of neutrinos in 1930 in order to explain the phenomenon of β -radioactivity. Neutrino's were not detected experimentally before 1956. According to a physical criterion, neutrino's exist if they demonstrably interact with other particles. Sometimes it is said that the neutrino was 'observed' for the first time in 1956. Therefore one has to stretch the concept of 'observation' quite far. In no

experiment neutrino's can be seen, heard, smelled, tasted, or felt. Even their path of motion cannot be made visible in any experiment. But in several kinds of experiment, from observable phenomena the energy and momentum (both magnitude and direction) of individual neutrino's can be calculated. For a physicist, this provides sufficient proof for their physical existence as interacting particles.

The universality of the relation frames allows science of comparing characters with each other and to determine their specific relations. The projections of the physical relation frame onto the preceding frames allow us to measure these relations. Measurability is the base of the mathematization of the exact sciences. It allows of applying statistics and designing mathematical models for natural and artificial systems.

The simplest case of interaction concerns two isolated systems interacting only with each other. Thermodynamics characterizes an isolated or closed system by magnitudes like energy and entropy. The two systems have thermal, chemical, or electric potential differences, giving rise to currents creating entropy. According to the second law of thermodynamics, this interaction is irreversible. Here 'system' is a general expression for a bounded part of space inclusive of the enclosed matter and energy. A closed system does not exchange energy or matter with its environment. Entropy can only be defined properly if the system concerned is in internal equilibrium and isolated from its environment.

In kinematics, an interactive event may have the character of a collision, minimally leading to a change in the state of motion of the colliding subjects. Often, the internal state of the colliding subjects changes as well. Except for the boundary case of an elastic collision, these processes are subject to the physical order of irreversibility. Frictionless motion influenced by a force is the standard example of a reversible interaction. In fact, it is also a boundary case, for any kind of friction or energy dissipation causes motion to be irreversible.

The law of inertia expresses the independence of uniform motion from physical interaction. It confirms the existence of uniform and rectilinear motions having no physical cause. This is an abstraction, for concrete things experiencing forces have a physical aspect as well. In reality a uniform rectilinear motion only occurs if the forces acting on the moving body balance each other.

Kinetic time is symmetric with respect to past and future. If in the description of a motion the time parameter (t) is replaced by its reverse ($-t$), we achieve a valid description of a possible motion. In the absence of friction or any other kind of energy dissipation, motion is reversible. By distinguishing past and future we are able to discover cause-effect relations, assuming that an effect never precedes its cause. According to relativity theory, the order of events having a causal relation is in all inertial systems the same, provided that time is not reversed.

In our common understanding of time, the discrimination of past and future is a matter of course,^[1] but in the philosophy of science it is problematic. The existence of

irreversible processes cannot be denied. All motions with friction are irreversible. Apparently, the absorption of light by an atom or a molecule is the reverse of emission, but Albert Einstein demonstrated that the reverse of (stimulated) absorption is *stimulated* emission of light, making *spontaneous* emission a third process, having no reverse (5.6). This applies to radioactive processes as well. The phenomenon of decoherence makes most quantum processes irreversible.[2] Only wave motion subject to Edwin Schrödinger's equation is symmetric in time. Classical mechanics usually expresses interaction by a force between two subjects, this relation being symmetric according to Newton's third law of motion. However, this law is only applicable to spatially separated subjects if the time needed to establish the interaction is negligible, i.e., if the action at a distance is (almost) instantaneous. Einstein made clear that interaction always needs time, hence even interaction at a distance is asymmetric in time.

Irreversibility does not imply that the reverse process is impossible. It may be less probable, or requiring quite different initial conditions. The transport of heat from a cold to a hotter body (as occurs in a refrigerator) demands different circumstances from the reverse process, which occurs spontaneously if the two bodies are not thermally isolated from each other. A short living point-like source of light causes a flash expanding in space. It is not impossible but practically very difficult to reverse this wave motion, for instance applying a perfect spherical mirror with the light source at the centre. But even in this case, the reversed motion is only possible thanks to the first motion, such that the experiment as a whole is still irreversible.

Yet, irreversibility as a temporal order is philosophically controversial, for it does not fit into the reductionist worldview influenced by nineteenth-century mechanism.[3] This worldview assumes each process to be reducible to motions of as such unchangeable pieces of matter, interacting through Newtonian forces. Ludwig Boltzmann attempted to bridge reversible motion and irreversible processes by means of the concepts of probability and randomness. In order to achieve the intended results, he had to assume that the realization of chances is irreversible. According to Hans Reichenbach, 'the direction of time is supplied by the direction of entropy, because the latter direction is made manifest in the statistical behaviour of a large number of separate systems, generated individually in the general drive to more and more probable states.'^[4] But he also observes: 'The inference from time to entropy leads to the same result whether it is referred to the following or to preceding events'.^[5] One may conclude that '... the one great law of irreversibility (the Second Law) cannot be explained from the reversible laws of elementary particle mechanics...'^[6]

It is sometimes stated that all 'basic' laws of physics are symmetrical in time. This seems to be true as far as *kinetic* time is concerned, and if any law that belies temporal symmetry (like the second law of thermodynamics, or the law for spontaneous decay) is not considered 'basic'. Anyhow, all attempts to reduce irreversibility to the subject side of the physical aspect of reality have failed.

Interaction is first of all subject to general laws independent of the specific character of the things involved. Some conservation laws are derivable from Albert Einstein's principle of relativity, stating that the laws of physics are independent of the motion of inertial systems.

Being the *physical* subject-subject relation, interaction may be analysed with the help of *quantitative* magnitudes like energy, mass, and charge; *spatial* concepts like force, momentum, field strength, and potential difference; as well as *kinetic* expressions like currents of heat, matter, or electricity.

Like interaction, energy, force, and current are abstract concepts. Yet these are not merely covering concepts without physical content. They can be specified as projections of characteristic interactions like the electromagnetic one. Electric energy, gravitational force, and the flow of heat specify the abstract concepts of energy, force, and current.

For energy to be measurable, it is relevant that one concrete form of energy is convertible into another one. For instance, a generator transforms mechanical energy into electric energy. Similarly, a concrete force may balance another force, whereas a concrete current accompanies currents of a different kind. This means that characteristically different interactions are *comparable*, they can be *measured* with respect to each other. The physical subject-subject relation, the interaction projected as energy, force, and current, is the foundation of the whole system of measuring, characteristic for astronomy, biology, chemistry, physics, as well as technology. The concepts of energy, force, and current enable us to determine physical subject-subject relations objectively.

Measurement of a quantity requires several conditions to be fulfilled. First, a *unit* should be available. A measurement compares a quantity with an agreed unit. Secondly, a magnitude requires a law, a *metric*, determining how a magnitude is to be projected on a set of numbers, on a scale. The third requirement, being the availability of a *measuring instrument*, cannot always be directly satisfied. A magnitude like entropy can only be calculated from measurements of other magnitudes. Fourth, therefore, there must be a fixed relation between the various metrics and units, a *metrical system*. This allows of the application of measured properties in theories. Unification of units and scales such as the metric system is a necessary requirement for the communication of both measurements and theories.

I shall discuss the concepts of energy, force, and current in some more detail. It is by no means evident that these concepts are the most general projections of interaction. Rather, their development has been a long and tedious process, leading to a *general unification* of natural science, to be distinguished from a more specific unification to be discussed later on.

a. Since the middle of the nineteenth century, energy is the most important quantitative expression of physical, chemical, and biotic interactions.^[7] As such it has superseded mass, in particular since it is known that mass and energy are equivalent,

according to physics' most famous formula, $E=mc^2$. The formula means that mass and energy are *equivalent*, that each amount of energy corresponds with an amount of mass and conversely. It does not mean that mass is a form of energy, or can be converted into energy, as it is often misunderstood. Energy is specifiable as kinetic and potential energy, thermal energy, nuclear energy, or chemical energy. Affirming the total energy of a closed system to be constant, the law of conservation of energy implies that one kind of energy can be converted into another one, but not mass into energy. For this reason, energy forms a universal base for comparing various types of interaction.

Before energy, mass became a universal measure for the amount of matter, serving as a measure for gravity as well as for the amount of heat that a subject absorbs when heated by one degree. Energy and mass are general expressions of physical interaction. This applies to entropy and related thermodynamic concepts too. In contrast, the rest energy and the rest mass of a particle or an atom are characteristic magnitudes.

Velocity is a measure for motion, but if it concerns physically qualified things, linear momentum (quantity of motion, the product of mass and velocity) turns out to be more significant. The same applies to angular momentum (quantity of rotation, the product of moment of inertia and angular frequency; angular frequency equals 2π times the frequency. The moment of inertia is an expression of the distribution of matter about a body with respect to a rotation axis.) In the absence of external forces, linear and angular momentum are subject to conservation laws. Velocity, linear and angular momentum, and moment of inertia are not expressed by a single number (a scalar) but by vectors or tensors. Relativity theory combines energy (a scalar) with linear momentum (a vector with three components) into a single vector, having four components.

b. According to Isaac Newton's third law, the mechanical force is a subject-subject relation.^[8] If *A* exerts a force \mathbf{F} on *B*, then *B* exerts a force $-\mathbf{F}$ on *A*. The minus sign indicates that the two forces being equal in magnitude have opposite directions. The third law has exerted a strong influence on the development of physics during a quite long time. In certain circumstances, the law of conservation of linear momentum can be derived from it. However, nowadays physicists allot higher priority to this conservation law than to Newton's third law. In order to apply Newton's laws when more than one force is acting, we have to consider the forces simultaneously. This does not lead to problems in the case of two forces acting on the same body. But the third law is especially important for action at a distance, inherent in the Newtonian formulation of gravity, electricity, and magnetism. In Albert Einstein's theory of relativity, simultaneity at a distance turns out to depend on the motion of the reference system. The laws of conservation of linear momentum and energy turn out to be easier to amend to relativity theory than Newton's third law. Now one describes the interaction as an exchange of energy and momentum (mediated by a field particle like a photon). This exchange requires a certain span of time.

Newton's second law provides the relation between force and momentum: the net force equals the change of momentum per unit of time. The law of inertia seems to be deducible from Newton's second law. If the force is zero, momentum and hence velocity is constant, or so it is argued. However, if the first law would not be valid, there could be a different law, assuming that each body experiences a frictional force, dependent on speed, in a direction opposite to the velocity. (In its most simple form, $\mathbf{F} = -b\mathbf{v}$, $b > 0$.) Accordingly, if the total force on a body is zero, the body would be *at rest*. A unique reference system would exist in which all bodies on which no forces act would be at rest. This is the nucleus of Aristotle's mechanics, but it contradicts both the classical principle of relativity and the modern one. The principle of relativity is an alternative expression of the law of inertia, pointing out that absolute (non-relative) uniform motion does not exist. Just like spatial position on the one hand and interaction on the other side, motion is a universal *relation*.

Besides to a rigid body, a force is applicable to a fluid, usually in the form of a pressure (i.e., force per area). A pressure difference causes a change of volume or a current subject to Daniel Bernoulli's law, if the fluid is incompressible. Besides, there are non-mechanical forces causing currents. A temperature gradient causes a heat current, chemical potentials drive material flows (e.g., diffusion), and an electric potential difference directs an electric current.

To find a metric for a thermodynamic or an electric potential is not an easy task. On the basis of an analysis of the idealized cycles devised by Sadi Carnot, William Thomson (later Lord Kelvin) established the theoretical metric for the thermodynamic temperature scale.^[9] The practical definition of the temperature scale takes this theoretical 'absolute' scale as a norm. The definition of the metric of pressure is relatively easy, but finding the metric of electric potential caused almost as much trouble as the development of the thermodynamic temperature scale.

The Newtonian force can sometimes be written as the derivative of a potential energy (i.e., energy as a function of spatial position). Since the beginning of the nineteenth century, the concept of a force is incorporated in the concept of a field. At first a field was considered merely a mathematical device, until James Clerk Maxwell proved the electromagnetic field to have physical reality of its own. A field is a physical function projected on space. Usually one assumes the field to be continuous and differentiable almost everywhere. A field may be constant or variable. There are scalar fields (like the distribution of temperature in a gas), vector fields (like the electrostatic field), and tensor fields (like the electromagnetic field). A field of force is called 'conservative' if the forces are derivable from a space-dependent potential energy. This applies to the classical gravitational and electrostatic fields. It does not apply to the force derived by Hendrik Antoon Lorentz, because it depends on the velocity of a charged body with respect to a magnetic field. (The Lorentz force and Maxwell's equations for the electromagnetic field are derivable from a gauge-invariant vector potential. 'Gauge-invariance' is the relativistic successor to the static concept of a conservative field.)

c. A further analysis of thermodynamics and electricity makes clear that current is a third projection, now from the physical onto the kinetic relation frame. The concept

of *entropy* points to a general property of currents. In each current, entropy is created, making the current irreversible. (A current in a superconductor is a boundary case. In a closed superconducting circuit without a source, an electric current may persist indefinitely, whereas a normal current would die out very fast.)

In a system in which currents occur, entropy increases. Only if a system as a whole is in equilibrium, there are no net currents and the entropy is constant. Like several mechanical forces are able to balance each other, so do thermodynamic forces and currents. This leads to mutual relations like thermo-electricity, the phenomenon that a heat current causes an electric current (Thomas Seebeck's effect) or reverse (Charles Peltier's effect).[10] This is applied in the thermo-electric thermometer, measuring a temperature difference by an electric potential difference. Relations between various types of currents are subject to a symmetry relation discovered by William Kelvin and generalized by Lars Onsager.[11]

The laws of thermodynamics are generally valid, independent of the specific character of a physical thing or aggregate. For a limited set of *specific* systems (e.g., a gas consisting of similar molecules), statistical mechanics is able to derive the second law from mechanical interactions, starting from assumptions about their probability.[12] Whereas the thermodynamic law states that the entropy in a closed system is constant or increasing, the statistical law allows of fluctuations. The source of this difference is that thermodynamics supposes matter to be continuous, whereas statistical mechanics takes into account the molecular character of matter.

There are many different interactions, like electricity, magnetism, contact forces (e.g., friction), chemical forces (e.g., glue), or gravity. Some are reducible to others. The contact forces turn out to be of an electromagnetic nature, and chemical forces are reducible to electrical ones.

Besides the general unification discussed above allowing of the comparison of widely differing interactions, a characteristic unification can be discerned. James Clerk Maxwell's unification of electricity and magnetism implies these interactions to have the same character, being subject to the same specific cluster of laws and showing symmetry. The fact that they can still be distinguished points to an asymmetry, a break of symmetry. The study of characteristic symmetries and symmetry breaks supplies an important tool for achieving a characteristic unification of natural forces.

Since the middle of the twentieth century, physics discerns four fundamental specific interactions. These are gravity and electromagnetic interaction besides the strong and weak nuclear forces. Later on, the electromagnetic and weak forces were united into the electroweak interaction, whereas the strong force is reducible to the colour force between quarks. In the near future, physicists expect to be able to unite the colour force with the electroweak interaction. The ultimate goal, the unification of all four forces is still far away.

About 1900, the 'electromagnetic world view' supposed that all physical and chemical interactions could be reduced to electromagnetism.[13] Just like the modern standard

model, it aimed at deducing the (rest-) mass of elementary particles from this supposed fundamental interaction.^[14]

These characteristic interactions are distinguished in several ways, first by the particles between which they act. Gravity acts between all particles, the colour force only between quarks, and the strong force only between particles composed from quarks. A process involving a neutrino is weak, but the reverse is not always true.

Another difference is their relative strength. Gravity is weakest and only plays a part because it cannot be neutralized. It manifests itself only on a macroscopic scale. The other forces are so effectively neutralized, that the electrical interaction was largely unknown until the eighteenth century, and the nuclear forces were not discovered before the twentieth century. Gravity conditions the existence of stars and systems of stars.

Next, gravity and electromagnetic interaction have an infinite range, whereas the other forces do not act beyond the limits of an atomic nucleus. For gravity and electricity the inverse-square law is valid (the force is inversely proportional to the square of the distance from a point-like source). This law is classically expressed in Isaac Newton's law of gravity and Charles Coulomb's electrostatic law, with mass respectively charge acting as a measure of the strength of the source. A comparable law does not apply to the other forces, and the lepton and baryon numbers do not act as a measure for their sources. As a function of distance, the weak interaction decreases much faster than quadratically. The colour force is nearly constant over a short distance (of the order of the size of a nucleus), beyond which it decreases abruptly to zero.

The various interactions also differ because of the field particles involved. Each fundamental interaction corresponds to a field in which quantized currents occur. For gravity, this is an unconfirmed hypothesis. Field particles have an integral spin and they are bosons (3.2, 4.4). If the spin is even (0 or 2), it concerns an attractive force between equal particles and a repulsive force between opposite particles (if applicable). For an uneven spin it is the other way around. The larger the field particle's rest mass, the shorter is the range of the interaction. If the rest mass of the field particles is zero (as is the case with photons and gravitons), the range is infinite. Unless mentioned otherwise, the field particles are electrically neutral.

The mean lifetime of spontaneous decay differs widely. The stronger the interaction causing a transition, the faster the system changes. If a particle decays because of the colour force or strong force, it happens in a very short time (of the order of 10^{-23} to 10^{-19} sec). Particles decaying due to weak interaction have a relatively long lifetime (10^{-12} sec for a tauon up to 900 sec for a free neutron). Electromagnetic interaction is more or less between.

In high-energy physics, symmetry considerations and group theory play an important part in the analysis of collision processes. New properties like isospin and strangeness have led to the introduction of groups named SU(2) and SU(3) and the discovery of at first three, later six quarks. (SU(3) means special unitary group with

three variables. The particles in a representation of this group have the same spin and parity (together one variable), but different values for strangeness and one component of isospin.)

Quantum electrodynamics reached its summit shortly after the Second World War, but the other interactions are less manageable, being developed only after 1970. Now each field has a symmetry property called gauge invariance, related to the laws of conservation of electric charge, baryon number, and lepton number. Symmetry is as much an empirical property as any other one. After the discovery of antiparticles it was assumed that charge conjugation C (symmetry with respect to the interchange of a particle with its antiparticle), parity P (mirror symmetry), and time reversal T are properties of all fundamental interactions. Since 1956, it is experimentally established that β -decay has no mirror symmetry unless combined with charge conjugation (CP). In 1964 it turned out that weak interactions are only symmetrical with respect to the product CPT , such that even T alone is no longer universally valid.

The appropriate theory is called the *standard model*, since the discovery of the J/ψ particle in 1974 explaining successfully a number of properties and interactions of subatomic particles. Dating from the seventies of the twentieth century, it was tentatively confirmed in 2012 by the experimental discovery of Peter Higgs' particle, already predicted in 1964. Tentatively: the model does not include gravity, and some recently discovered properties of neutrinos do not quite fit into it. The general theory of relativity is still at variance with quantum electrodynamics, with the electroweak theory of Steven Weinberg and Abdus Salam, as well as with quantum chromodynamics.[15]

These fundamental interactions are specifications of the abstract concept of interaction being the universal physical and chemical relation. Their laws, like those of Maxwell for electromagnetism, form a specific set, which may be considered a character. But this character does not determine a class of things or events, but a class of relations.

5.2. The character of electrons

Ontology, the doctrine of *on* (or *ontos*, Greek for being), aims to answer the question of how matter is composed according to present-day insights. Since the beginning of the twentieth century, many kinds of particles received names ending with *on*, like electron, proton, neutron, and photon. At first sight, the relation with ontology seems to be obvious. Historically the suffix *-on* goes back to the electron. Whether the connection with ontology has really played a part is unclear.[16] The word electron comes from the Greek word for amber or fossilized resin, since antiquity known for its properties that we now recognize as static electricity. From 1874, Stoney used the word electron for the elementary amount of charge. Only in the twentieth century, electron became the name of the particle identified by Thomson in 1897. Rutherford introduced the names proton and neutron in 1920 (long before the actual discovery of

the neutron in 1932). Lewis baptized the photon in 1926, 21 years after Einstein proposed its existence. Yet, not many physicists would affirm that an electron is the essence of electricity, that the proton forms the primeval matter, that the neutron and its little brother, the neutrino, have the nature of being neutral, or that in the photon light comes into being, and in the phonon sound. In pion, muon, tauon, and kaon, *on* is no more than a suffix of the letters π , μ , τ and K , whereas Paul Dirac baptized fermion and boson after Enrico Fermi and Satyendra Bose. In 1833 Michael Faraday, advised by William Whewell, introduced the words ion, cation, and anion, referring to the Greek word for *to go*. In an electrolyte, an ion moves from or to an electrode, an anode or cathode (names proposed by Whewell as well). An intruder is the positive electron. Meant as positon, the positron received an additional *r*, possibly under the influence of electron or new words like magnetron and cyclotron, which however are machines, not particles.

Only after 1925 quantum physics and high-energy physics allowed of the study of the characters of elementary physical things. Most characters have been discovered after 1930. But the discovery of the electron (1897), of the internal structure of an atom, composed from a nucleus and a number of electrons (1911), and of the photon (1905) preceded the quantum era. These are typical examples of characters founded in the quantitative, spatial, and kinetic projections of physical interaction. In section 5.1, these projections were pointed out to be respectively energy; force or field; and current.

An electron is characterized by a specific amount of mass and charge and is therefore quantitatively founded. The foundation is not in the quantitative relation frame itself (because that is not physical), but in the most important quantitative projection of the physical relation frame. This is energy, expressing the quantity of interaction. Like other particles, an electron has a typical rest energy, besides specific values for its electric charge, magnetic moment, and lepton number.

In chapter 4, I argued that an electron has the character of a wave packet as well, kinetically qualified and spatially founded, anticipating physical interactions. An electron has a *specific* physical character and a *generic* kinetic character. The two characters are interlaced within the at first sight simple electron. The combined *dual* character is called the wave-particle duality. Electrons share it with all other elementary particles. As a consequence of the kinetic character and the inherent Heisenberg relations, the position of an electron cannot be determined much better than within 10^{-10} m (about the size of a hydrogen atom). But the physical character implies that the electron's collision diameter (being a measure of its physical size) is less than 10^{-17} m.

Except for quarks, all quantitatively founded particles are *leptons*, to be distinguished from field particles and baryons (5.3, 5.4). Leptons are not susceptible to the strong nuclear force or the colour force. They are subject to the weak force, sometimes to electromagnetic interaction, and like all matter to gravity. Each lepton has a positive or negative value for the lepton number (*L*), which significance appears in the occurrence or non-occurrence of certain processes. Each process is subject to the law of conservation of lepton number, i.e., the total lepton number cannot change. For

instance, a neutron ($L=0$) does not decay into a proton and an electron, but into a proton ($L=0$), an electron ($L=1$) and an antineutrino ($L=-1$). The lepton number is just as characteristic for a particle as its electric charge. For non-leptons the lepton number is 0, for leptons it is +1 or -1.

Leptons satisfy a number of characteristic laws. Each particle has an electric charge being an integral multiple (positive, negative or zero) of the elementary charge. Each particle corresponds with an antiparticle having exactly the same rest mass and lifetime, but opposite values for charge and lepton number. Having a half-integral spin, leptons are fermions satisfying the exclusion principle and the characteristic Fermi-Dirac statistics (4.3, 5.5).

Three *generations* of leptons are known, each consisting of a negatively charged particle, a neutrino, and their antiparticles. These generations are related to similar generations of quarks (5.3). A tauon decays spontaneously into a muon, and a muon into an electron. Both are weak processes, in which simultaneously a neutrino and an anti-neutrino are emitted.

The leptons display little diversity, their number is exactly 6. Like their diversity, the variation of leptons is restricted. It only concerns their external relations: their position, their linear and angular momentum, and the orientation of their magnetic moment or spin relative to an external magnetic field.

This description emphasizes the quantitative aspect of leptons. But leptons are first of all physically qualified. Their specific character determines how they interact by electroweak interaction with each other and with other physical subjects, influencing their coming into being, change and decay.

Electrons are by far the most important leptons, having the disposition to become part of systems like atoms, molecules, and solids. The other leptons only play a part in high-energy processes. In order to stress the distinction between a definition and a character as a set of laws, I shall dwell a little longer on hundred years of development of our knowledge of the electron.^[17]

Although more scientists were involved, it is generally accepted that Joseph J. Thomson in 1897 discovered the electron. He identified his cathode ray as a stream of particles and established roughly the ratio e/m of their charge e and mass m , by measuring how an electric and/or magnetic field deflects the cathode rays. In 1899 Thomson determined the value of e separately, allowing him to calculate the value of m . Since then, the values of m and e , which may be considered as defining the electron, are determined with increasing precision. In particular Robert Millikan did epoch-making work, between 1909 and 1916. Almost simultaneously with Thomson, Hendrik Lorentz observed that the Zeeman effect (1896) could be explained by the presence in atoms of charged particles having the same value for e/m as the electron. Shortly afterwards, the particles emerging from β -radioactivity and the photoelectric effect were identified as electrons.

The mass m depends on the electron's speed, as was first established experimentally by Walter Kaufmann, later theoretically by Albert Einstein. Since then, instead of the mass m the rest mass m_0 is characteristic for a particle. Between 1911 and 1913, Ernest Rutherford and Niels Bohr developed the atomic model in which electrons move around a much more massive nucleus. The orbital angular momentum turned out to be quantized. In 1923 Louis de Broglie made clear that an electron sometimes behaves like a wave, interpreted as the bearer of probability by Max Born in 1926 (4.3). In 1925, Samuel Goudsmit and George Uhlenbeck suggested a new property, half-integral spin, connected to the electron's intrinsic magnetic moment. In the same year, Wolfgang Pauli discovered the exclusion principle. Enrico Fermi and Paul Dirac derived the corresponding statistics in 1926. Since then, the electron is a fermion, playing a decisive part in all properties of matter (4.3, 5.3, 5.5). In 1930 it became clear that in β -radioactivity besides the electron a neutrino emerges from a nucleus. Neutrinos were later on recognized as members of the lepton family. β -radioactivity is not caused by electromagnetic interaction, but by the weak nuclear force. Electrons turned out not to be susceptible to strong nuclear forces. In 1931 the electron got a brother, the positron or anti-electron. This affirmed that an electron has no eternal life, but may be created or annihilated together with a positron. In β -radioactivity, too, an electron emerges or disappears (in a nucleus, an electron cannot exist as an independent particle), but apart from these processes, the electron is the most stable particle we know besides the proton. According to Paul Dirac, the positron is a hole in the nether world of an infinite number of electrons having a negative energy (4.3). In 1953, the law of conservation of lepton number was discovered. After the second world war, Richard Feynman, Julian Schwinger, and Shin'ichiro Tomonaga developed quantum electrodynamics. This is a field theory in which the physical vacuum is not empty, but is the stage of spontaneous creations and annihilations of virtual electron-positron pairs. Interaction with other (sometimes virtual) particles is partly responsible for the properties of each particle. As a top performance counts the theoretical calculation of the magnetic moment of the electron in eleven decimals, a precision only surpassed by the experimental measurement of the same quantity in twelve decimals. Moreover, the two values differ only in the eleventh decimal, within the theoretical margin of error. 'The agreement between experiment and theory shown by these examples, the highest point in precision reached anywhere in the domain of particles and fields, ranks among the highest achievements of twentieth-century physics.'^[18] Finally, the electron got two cousins, the muon and the tauon.

Besides these scientific developments, electronics revolutionized the world of communication, information, and control.

Since Joseph Thomson's discovery, the concept of an electron has been changed and expanded considerably. Besides being a particle having mass and charge, it is now a wave, a top, a magnet, and a fermion, half of a twin, and a lepton. Yet, few people doubt that we are still talking about the same electron.

What the essence of an electron is appears to be a hard question, if ever posed. It may very well be a meaningless question. But we achieve a growing insight into the *laws* constituting the electron's character, determining the electron's *relations* with other things and the *processes* in which it is involved. The electron's charge means that two

electrons exert a force on each other according to the laws of Charles Coulomb and Hendrik Lorentz. The mass follows from the electron's acceleration in an electric and/or magnetic field, according to James Clerk Maxwell's laws. The lepton number makes only sense because of the law of conservation of lepton number, allowing of some processes and prohibiting others. Electrons are fermions, satisfying the exclusion principle and the distribution law of Enrico Fermi and Paul Dirac.

The character of electrons is not *logically* given by a definition, but *physically* by a specific set of laws, which are successively discovered and systematically connected by experimental and theoretical research.

An electron is to be considered an individual satisfying the character described above. A much-heard objection to the assignment of individuality to electrons and other elementary particles is the impossibility to distinguish one electron from another. Electrons are characteristically equal to each other, having much less variability than plants or animals, even less than atoms.

This objection can be retraced to the still influential worldview of mechanism. This worldview assumed each particle to be identifiable by objective kinetic properties like its position and velocity at a certain time. Quantum physics observes that the identification of physically qualified things requires a physical interaction. In general, this interaction influences the particle's position and momentum (4.3). Therefore, the electron's position and momentum cannot be determined with unlimited accuracy, as follows from Werner Heisenberg's relations. This means that identification in a mechanistic sense is not always possible. Yet, in an interaction such as a measurement, an electron manifests itself as an individual.

If an electron is part of an atom, it can be identified by its state, because the exclusion principle precludes that two electrons would occupy the same state. The two electrons in the helium atom exchange their states continuously without changing the state of the atom as a whole. But it cannot be doubted that at any moment there are two electrons, each with its own mass, charge and magnetic moment. For instance, in the calculation of the energy levels the mutual repulsion of the two electrons plays an important part.

The individual existence of a bound electron depends on the binding energy being much smaller than its rest energy. Binding energy equals the energy needed to liberate an electron from an atom. It varies from a few eV (the outer electrons) to several tens of keV (the inner electrons in a heavy element like uranium). The electron's rest mass is about 0.5 MeV, much larger than its binding energy in an atom (13.6 eV). To keep an electron as an independent particle in a *nucleus* would require a binding energy of more than 100 MeV, much more than the electron's rest energy of 0,5 MeV. For this reason, physicists argue that electrons in a nucleus cannot exist as independent, individual particles, like they are in an atom's shell.

In contrast, protons and neutrons in a nucleus satisfy the criterion that an independent particle has a rest energy substantially larger than the binding energy. Their binding

energy is about 8 MeV, their rest energy is almost 1000 MeV. A nucleus is capable of emitting an electron (this is β -radioactivity). The electron's existence starts at the emission and eventually ends at the absorption by a nucleus. Because of the law of conservation of lepton number, the emission of an electron is accompanied by the emission of an anti-neutrino, and at the absorption of an electron a neutrino is emitted. This would not be the case if the electron could exist as an independent particle in the nucleus. Neutrinos are stable, their rest mass is zero or very small, and they are only susceptible to weak interaction. Neutrinos and anti-neutrinos differ by their parity, the one being left handed, and the other right handed. (This distinction is only possible for particles having zero rest mass. If neutrinos have a rest mass different from zero, as some experiments suggest, the theory has to be adapted with respect to parity). That the three neutrinos differ from each other is established by processes in which they are or are not involved, but in what respect they differ is less clear. For some time, physicists expected the existence of a fourth generation, but the standard model restricts itself to three, because astrophysical cosmology implies the existence of at most three different types of neutrinos with their antiparticles.

More than as free particles, the electrons display their characteristic properties as components of atoms, molecules and solids, as well as in processes. The half-integral spin of electrons was discovered in the investigation of atomic spectra. The electron's fermion character largely determines the shell structure of atoms. In 1930, Wolfgang Pauli suggested the existence of the neutrino because of the character of β -radioactivity. The lepton number is discovered by an analysis of specific nuclear reactions.

Electrons have the affinity or propensity of functioning as a component of atoms and molecules because electrons share electromagnetic interaction with nuclei. Protons and electrons have the same but opposite charge, allowing of the formation of neutral atoms, molecules and solids. Electric neutrality is of tremendous importance for the stability of these systems. This tertiary characteristic determines the meaning of electrons in the cosmos.

5.3. The quantum ladder

An important spatial manifestation of interaction is the *force* between two spatially separated bodies. An atom or molecule having a spatially founded character consists of a number of nuclei and electrons kept together by the electromagnetic force. More generally, any interaction is spatially projected on a *field*.

Sometimes a field can be described as the spatial derivative of the potential energy. A set of particles constitutes a stable system if the potential energy has an appropriate shape, characteristic for the spatially founded structure. In a spatially founded

structure, the relative spatial positions of the components are characteristic, even if their relative motions are taken care of. Atoms have a spherical symmetry restricting the motions of the electrons. In a molecule, the atoms or ions have characteristic relative positions, often with a specific symmetry. In each spatially founded character a number of quantitatively founded characters are interlaced.

It is a remarkable fact that in an atom the nucleus acts like a *quantitatively* founded character, whereas the nucleus itself is a spatial configuration of protons and neutrons kept together by forces. The nucleus itself has a spatially founded character, but in the atom it has the disposition to act as a whole, characterized by its mass, charge, and magnetic moment. Similarly, a molecule or a crystal is a system consisting of a number of atoms or ions and electrons, all acting like quantitatively founded particles. Externally, the nucleus in an atom and the atoms or ions in a molecule act as quantitatively founded wholes, as units, while preserving their own internal spatially founded structure.

However, an atom bound in a molecule is not completely the same as a free atom. In contrast to a nucleus, a free atom is electrically neutral and it has a spherical symmetry. Consequently, it cannot easily interact with other atoms or molecules, except in collisions. In order to become a part of a molecule, an atom has to open up its tertiary character. This can be done in various ways. The atom may absorb or eject an electron, becoming an ion. A common salt molecule does not consist of a neutral sodium atom and a neutral chlorine atom, but of a positive sodium ion and a negative chlorine ion, attracting each other by the Coulomb force. This is called heteropolar or ionic bonding. Any change of the spherical symmetry of the atom's electron cloud leads to the relatively weak Van der Waals interaction. A very strong bond results if two atoms share an electron pair. This homopolar or covalent bond occurs in diatomic molecules like hydrogen, oxygen, and nitrogen, in diamond, and in many carbon compounds. Finally, especially in organic chemistry, the hydrogen bond is important. It means the sharing of a proton by two atom groups.

The possibility of being bound into a larger configuration is a very significant tertiary characteristic of many physically qualified systems, determining their meaning in the cosmos.

The first stable system studied by physics is the solar system, in the seventeenth century investigated by Johannes Kepler, Galileo Galilei, Christiaan Huygens, and Isaac Newton. The law of gravity, mechanical laws of motion, and conservation laws determine the character of planetary motion. The solar system is not unique, there are more stars with planets, and the same character applies to a planet with its moons, or to a double star. Any model of the system presupposes its isolation from the rest of the world, which is only approximately the case. This approximation is pretty good for the solar system, less good for the system of the sun and each planet apart, and pretty bad for the system of earth and moon.

Spatially founded physical characters display a large disparity. Various specific subtypes appear. According to the standard model (5.1), these characters form a

hierarchy, sometimes called the quantum ladder.^[19] At the first rung there are six (or eighteen, see below) different quarks, with the antiquarks grouped into three generations related to those of leptons, as follows from analogous processes.

Like a lepton, a quark is quantitatively founded, it has no structure. But a quark cannot exist as a free particle. Quarks are *confined* as a duo in a meson (e.g., a pion) or as a trio in a baryon (e.g., a proton or a neutron) or an antibaryon. Confinement is a tertiary characteristic, but it does not stand apart from the secondary characteristics of quarks, their quantitative properties. Whereas quarks have a charge of $1/3$ or $2/3$ times the elementary charge, their combinations satisfy the law that the electric charge of a free particle can only be an integral multiple of the elementary charge. Likewise, in confinement the sum of the baryon numbers (for quarks $\pm 1/3$ or $\pm 2/3$) always yields an integral number. For a meson this number is 0, for a baryon it is +1, for an antibaryon it is -1.

Between quarks the colour force is acting, mediated by gluons. The colour force has no effect on leptons and is related to the strong force between baryons. In a meson the colour force between two quarks hardly depends on their mutual distance, meaning that they cannot be torn apart. If a meson breaks apart, the result is not two separate quarks but two quark-antiquark pairs.

Quarks are fermions, they satisfy the exclusion principle. In a meson or baryon, two identical quarks cannot occupy the same state. But an omega particle (sss) consists of three strange quarks having the same spin. This is possible because each quark exists in three variants, each indicated by a 'colour' besides six 'flavours'. For the antiquarks three complementary colours are available. The metaphor of 'colour' is chosen because the colours are able to neutralize each other, like ordinary colours can be combined to produce white. This can be done in two ways, in a duo by adding a colour to its anticolour, or in a trio by adding three different colours or anticolours. The law that mesons and baryons must be colourless yields an additional restriction on the number of possible combinations of quarks. A white particle is neutral with respect to the colour force, like an uncharged particle is neutral with respect to the Coulomb force. Nevertheless, an electrically neutral particle may exert electromagnetic interaction because of its magnetic moment. This applies e.g. to a neutron, but not to a neutrino. Similarly, by the exchange of mesons, the colour force manifests itself as the strong nuclear force acting between baryons, even if baryons are 'white'. Two quarks interact by exchanging gluons, thereby changing of colour.

The twentieth-century standard model has no solution to a number of problems. Why only three generations? If all matter above the level of hadrons consists of particles from the first generation, what is the tertiary disposition of the particles of the second and third generation? Should the particles of the second and third generation be considered excited states of those of the first generation? Why does each generation consist of two quarks and two leptons (with corresponding antiparticles)? What is the origin of the mass differences between various leptons and quarks?

The last question might be the only one to receive an answer in the twenty-first century, when the existence of Peter Higgs' particle and its mass were experimentally established (2012). For the other problems, at the end of the twentieth century no experiment is proposed providing sufficient information to suggest a solution.

The second level of the hierarchy consists of hadrons, baryons having half integral spin and mesons having integral spin. Although the combination of quarks is subject to severe restrictions, there are quite a few different hadrons. A proton consists of two up and one down quark (uud), and a neutron is composed of one up and two down quarks (udd). These two nucleons are the lightest baryons, all others being called hyperons. A pion consists of $d\bar{d}$, $u\bar{u}$ (charge 0), $d\bar{u}$ ($-e$) or $u\bar{d}$ ($+e$), where \bar{d} and \bar{u} are the antiquarks of d and u . As a free particle, only the proton is stable, whereas the neutron is stable within a nucleus. A free neutron decays into a proton, an electron and an antineutrino. The law of conservation of baryon number is responsible for the stability of the proton, being the baryon with the lowest rest energy. All other hadrons have a very short mean lifetime, a free neutron having the longest (900 sec). Their diversity is much larger than that of leptons and of quarks. Based on symmetry relations, group theory orders the hadrons into sets of e.g. eight baryons or ten mesons.

For a large part, the interaction of hadrons consists of rearranging quarks accompanied by the creation and annihilation of quark-antiquark pairs and lepton-antilepton pairs. The general laws of conservation of energy, linear and angular momentum, the specific laws of conservation of electric charge, lepton number and baryon number, and the laws restricting electric charge and baryon number to integral values, characterize the possible processes between hadrons in a quantitative sense. Besides, the fields described by quantum electrodynamics and quantum chromodynamics characterize these processes in a spatial sense, and the exchange of field particles in a kinetic way.

Atomic nuclei constitute the third layer in the hierarchy. With the exception of hydrogen, each nucleus consists of protons and neutrons, determining together the coherence, binding energy, stability, and lifetime of the nucleus. The mass of the nucleus is the sum of the masses of the nucleons less the mass equivalent to the binding energy. Decisive is the balance of the *repulsive* electric force between the protons and the *attractive* strong nuclear force binding the nucleons independent of their electric charge. In heavy nuclei, the surplus of neutrons compensates for the mutual repulsion of the protons. To a large extent, the exclusion principle applied to neutrons and protons separately determines the stability of the nucleus and its internal energy states.

The nuclear force is negligible for the external functioning of a nucleus in an atom or molecule. Only the mass of the nucleus, its electric charge and its magnetic moment are relevant for its external relations. Omitting the magnetic moment leads to two diversities in nuclei.

The first diversity concerns the number of protons. In a neutral atom it equals the number of electrons determining the atom's chemical propensities. The nuclear charge together with the exclusion principle dominates the energy states of the electrons, hence the position of the atom in the periodic system of elements.

The second diversity concerns the number of neutrons in the nucleus. Atoms having the same number of protons but differing in neutron number are called isotopes, because they have the same position (*topos*) in the periodic system. They have similar chemical propensities.

The diversity of atomic nuclei is represented in a two-dimensional diagram, a configuration space. The horizontal axis represents the number of protons ($Z =$ atomic number), the vertical axis the number of neutrons (N). In this diagram the isotopes (same Z , different N) are positioned above each other. The configuration space is mostly empty, because only a restricted number of combinations of Z and N lead to stable or metastable (radioactive) nuclei. The periodic system of elements is a two-dimensional diagram as well. Dmitri Mendeleev ordered the elements in a sequence according to a secondary property (the atomic mass) and below each other according to tertiary propensities (the affinity of atoms to form molecules, in particular compounds with hydrogen and oxygen). Later on, the atomic mass was replaced by the atomic number Z . However, quantum physics made clear that the atomic chemical properties are not due to the nuclei, but to the electrons subject to the exclusion principle. The vertical ordering in the periodic system concerns the configuration of the electronic shells. In particular the electrons in the outer shells determine the tertiary chemical propensities.

This is not an ordering according to a definition in terms of necessary and sufficient properties distinguishing one element from the other, but according to their characters. The properties do not define a character, as essentialism assumes, but the character (a set of laws) determines the properties and propensities of the atoms.

In the hierarchical order, we find globally an *increase* of spatial dimensions, diversity of characters and variation within a character, besides a *decrease* of the binding energy per particle and the significance of strong and weak nuclear forces. For the characters of atoms, molecules, and crystals, only the electromagnetic interaction is relevant.

The internal variation of a spatially founded character is very large. Quantum physics describes the internal states with the help of David Hilbert's space, having the eigenvectors of William Hamilton's operator as a base (2.3). A Hilbert space describes the ensemble of possibilities (in particular the energy eigenvalues) determined by the system's character. In turn, the atom or molecule's character itself is represented by Edwin Schrödinger's time-independent equation. This equation is exactly solvable only in the case of two interacting particles, like the hydrogen atom, the helium ion, the lithium ion, and positronium. (Positronium is a short living composite of an electron and a positron, the only spatially founded structure entirely consisting of leptons.) In

other cases, the equation serves as a starting point for approximate solutions, usually only manageable with the help of a large computer.

The hierarchical connection implies that the spatially founded characters are successively interlaced, for example nucleons in a nucleus, or the nucleus in an atom, or atoms in a molecule. Besides, these characters are interlaced with kinetically, spatially, and quantitatively qualified characters, and often with biotically qualified characters as well.

The characters described depend strongly on a number of *natural constants*, which value can be established only experimentally, not theoretically. Among others, this concerns the gravitational constant G , the speed of light c , Planck's constant h and the elementary electric charge e , or combinations like the fine structure constant ($2\pi e^2/hc=1/137.036$) and the mass ratio of the proton and the electron (1836.104). If the constants of nature would be slightly different, both nuclear properties and chemical properties would change drastically.[\[20\]](#)

The quantum ladder is of a physical and chemical nature. As an ordering principle, the ladder has a few flaws from a logical point of view. For instance, the proton occurs on three different levels, as a baryon, as a nucleus, and as an ion. The atoms of the noble gases are their molecules as well. This is irrelevant for their character. The character of a proton consists of the specific laws to which it is subjected. The classification of baryons, nuclei, or ions is not a characterization, and a proton is not 'essentially' a baryon and 'accidentally' a nucleus or an ion.

The number of molecular characters is enormous and no universal classification of molecules exists. In particular the characters in which carbon is an important element show a large diversity.

The molecular formula indicates the number of atoms of each element in a molecule. Besides, the characteristic spatial structure of a molecule determines its chemical properties. The composition of a methane molecule is given by the formula CH_4 , but it is no less significant that the methane molecule has the symmetrical shape of a regular tetrahedron, with the carbon atom at the centre and the four hydrogen atoms at the vertices. The V-like shape of a water molecule (the three atoms do not lie on a straight line, but form a characteristic angle of 105°) causes the molecule to have a permanent electric dipole moment, explaining many of the exceptional properties of water. Isomers are materials having the same molecular formula but different spatial orderings, hence different chemical properties. Like the symmetry between a left and a right glove, the spatial symmetry property of mirroring leads to the distinction of dextro- and laevo-molecules.

The symmetry characteristic for the generic (physical) character is an emergent property, in general irreducible to the characters of the composing systems. Conversely, the original symmetry of the composing systems is broken. In methane, the outer shells of the carbon atom have exchanged their spherical symmetry for the tetrahedron symmetry of the molecule. Symmetry breaks also occur in fields. The

symmetry of strong nuclear interaction is broken by electroweak interaction. For the strong interaction, the proton and the neutron are symmetrical particles having the same rest energy, but the electroweak interaction causes the neutron to have a slightly larger rest energy and to be metastable as a free particle.

From quantum field theory, in principle it should be possible to derive successively the emergent properties of particles and their spatially founded composites. This is the synthetic, reductionist or fundamentalist trend, constructing complicated structures from simpler ones. It cannot explain symmetry breaks.^[21] For practical reasons too, a synthetic approach is usually impossible. The alternative is the analytical or holistic method, in which the symmetry break is explained from the empirically established symmetry of the original character. Symmetries and other structural properties are usually *a posteriori* explained, and hardly ever *a priori* derived. However, analysis and synthesis are not contrary but complementary methods.

Climbing the quantum ladder, *complexity* seems to increase. On second thoughts, complexity is not a clear concept. An atom would be more complex than a nucleus and a molecule even more. However, in the character of a hydrogen atom or a hydrogen molecule, weak and strong interactions are negligible, and the complex spatially founded nuclear structure is reduced to the far simpler quantitatively founded character of a particle having mass, charge, and magnetic moment. Moreover, a uranium nucleus consisting of 92 protons and 146 neutrons has a much more complicated character than a hydrogen molecule consisting of two protons and two electrons, having a position two levels higher on the quantum ladder.

Inward a system is more complex than outward. An atom consists of a nucleus and a number of electrons, grouped into shells. If a shell is completely filled in conformity with the exclusion principle, it is chemically inert, serving mostly to reduce the effective nuclear charge. A small number of electrons in partially occupied shells determines the atom's chemical propensities. Consequently, an atom of a noble gas, having only completely occupied shells, is less complicated than an atom having one or two electrons less. The complexity of molecules increases if the number of atoms increases. But some very large organic molecules consist of a repetition of similar atomic groups and are not particularly complex.

In fact, there does not exist an unequivocal criterion for complexity.

An important property of hierarchically ordered characters is that for the explanation of a character it is sufficient to descend to the next lower level. For the understanding of molecules, chemists need the atomic theory, but they do not need to know much about nuclear physics. A molecular biologist is acquainted with the chemical molecular theory, but his knowledge of atomic theory may be rather superficial. This is possible because of the phenomenon that a physical character interlaced in another one both keeps its properties and hides them.

Each system derives its stability from an internal equilibrium that is hardly observable from without. The nuclear forces do not range outside the nucleus. Strong electric

forces bind an atom or a molecule, but as a whole it is electrically neutral. The strong internal equilibrium and the weak remaining external action are together characteristic for a stable physical system. If a system exerts a force on another one, it experiences an equal external force. This external force should be much smaller than the internal forces keeping the system intact, otherwise it will be torn apart. In a collision between two molecules, the external interaction may be strong enough to disturb the internal equilibrium, such that the molecules fall apart. Eventually, a new molecule with a different character emerges. Because the mean collision energy is proportional to the temperature, the stability of molecules and crystals depend on this parameter. In the sun's atmosphere no molecules exist and in its centre no atoms occur. In a very hot star like a neutron star, even nuclei cannot exist.

Hence, a stable physical or chemical system is relatively inactive. It looks like an isolated system. This is radically different from plants and animals that can never be isolated from their environment. The internal physical equilibrium of a plant or an animal is maintained by metabolism, the continuous flow of energy and matter through the organism.

5.4. Individualized currents

I consider the primarily physical character of a photon to be secondarily kinetically founded. A photon is a field particle in the electromagnetic interaction, transporting energy, as well as linear and angular momentum from one spatially founded system to another. Besides photons, nuclear physics recognizes gluons being field particles for the colour force, mesons for the strong nuclear force, and three types of vector bosons for the weak interaction. The existence of the graviton, the field particle for gravity, has not been experimentally confirmed. All these interaction particles have an integral spin and are bosons. Hence, these are not subject to the exclusion principle. Field particles are not quantitatively or spatially founded things, but individualized characteristic currents, hence kinetically founded 'quasiparticles'. Bosons *carry* forces, whereas fermions *feel* or *experience* forces.

By absorbing a photon, an atom comes into an excited state, i.e. a metastable state at a higher energy than the ground state. Whereas an atom in its ground state can be considered an isolated system, an excited atom is always surrounded by the electromagnetic field.

A photon is a wave packet, like an electron it has a dual character. Yet there is a difference. Whereas the electron's *motion* has a wave character, a photon is a *current* in an electromagnetic field, a current being a kinetic projection of physical interaction. With respect to electrons, the wave motion only determines the probability of what will happen in a future interaction. In a photon, besides determining a similar

probability, the wave consists of periodically changing electric and magnetic fields. A real particle's wave motion lacks a substratum, there is no characteristic medium in which it moves, and its velocity is variable. Moving quasiparticles have a substratum, and their wave velocity is a property of the medium. The medium for light in empty space is the electromagnetic field, all photons having the same speed independent of any reference system.

Each inorganic solid consists of crystals, sometimes microscopically small. Amorphous solid matter does not exist or is very rare. The ground state of a crystal is the hypothetical state at zero temperature. At higher temperatures, each solid is in an excited state, determined by the presence of quasiparticles.

The crystal symmetry, adequately described by the theory of groups, has two or three levels. First, each crystal is composed of space filling unit cells. All unit cells of a crystal are equal to each other, containing the same number of atoms, ions or molecules in the same configuration. A characteristic lattice point indicates the position of a unit cell. The lattice points constitute a Bravais lattice (called after Auguste Bravais), representing the crystal's translation symmetry. Only fourteen types of Bravais lattices are mathematically possible and realized in nature. Each lattice allows of some variation, for instance with respect to the mutual distance of the lattice points, as is seen when the crystal expands on heating. Because each crystal is finite, the translation symmetry is restricted and the surface structure of a crystal may be quite different from the crystal structure.

Second, the unit cell has a symmetry of its own, superposed on the translation symmetry of the Bravais lattice. The cell may be symmetrical with respect to reflection, rotation, or inversion. The combined symmetry determines how the crystal scatters X-rays or neutrons, presenting a means to investigate the crystalline structure empirically. Hence, the long distance *spatial* order of a crystal evokes a long time *kinetic* order of specific waves.

Third, in some materials we find an additional ordering, for instance that of the magnetic moments of electrons or atoms in a ferromagnet. Like the first one, this is a long-distance ordering. It involves an interaction that is not restricted to nearest neighbours. It may extend over many millions of atomic distances.

The atoms in a crystal oscillate around their equilibrium positions. These elastic oscillations are transferred from one atom to the next like a sound wave, and because the crystal has a finite volume, this is a stationary wave, a collective oscillation. The crystal as a whole is in an elastic oscillation, having a kinetically founded character. These waves have a broad spectrum of frequencies and wavelengths, being sampled into wave packets. In analogy with light, these field particles are called sound quanta or *phonons*.

Like the electrons in a metal, the phonons act like particles in a box (4.4). Otherwise they differ widely. The number of electrons is constant, but the number of phonons increases strongly at increasing temperature. Like all quasiparticles, the phonons are

bosons, not being subject to the exclusion principle. The mean kinetic energy of the electrons hardly depends on temperature, and their specific heat is only measurable at a low temperature. In contrast, the mean kinetic energy of phonons strongly depends on temperature, and the phonon gas dominates the specific heat of solids. At a low temperature this increases proportional to T^3 to become constant at a higher temperature. Peter Debye's theory (originally 1912, later adapted) explains this from the wave and boson character of phonons and the periodic character of the crystalline structure.

In a solid or liquid, besides phonons many other quantized excitations occur, corresponding, for instance, with magnetization waves or spin waves. The interactions of quasiparticles and electrons cause the photoelectric effect and transport phenomena like electric resistance and thermo-electricity.

The specific properties of some superconductors can be described with the help of quasiparticles. (This applies to the superconducting metals and alloys known before 1986. For the ceramic superconductors, discovered since 1986, this explanation is not sufficient.) In a superconductor two electrons constitute a pair called after Leon Cooper. This is a pair of electrons in a bound state, such that both the total linear momentum and the total angular momentum are zero. The two electrons are not necessarily close to each other. Superconductivity is a phenomenon with many variants, and the theory is far from complete.

Superconductivity is a collective phenomenon in which the wave functions of several particles are macroscopically coherent. There is no internal dissipation of energy. It appears that on a macroscopic scale the existence of kinetically founded characters is only possible if there is no decoherence (4.3). Therefore, kinetically founded physical characters on a macroscopic scale are quite exceptional.

5.5. Aggregates and statistics

We have now discussed three types of physically qualified characters, but this does not exhaust the theory of matter. The inorganic sciences acknowledge many kinds of mixtures, aggregates, alloys or solutions. In nature, these are more abundant than pure matter. Often, the possibility to form a mixture is restricted and some substances do not mix at all. In order to form a stable aggregate, the components must be tuned to each other. Typical for an aggregate is that the characteristic magnitudes (like pressure, volume, and temperature for a gas) are variable within a considerable margin, even if there is a lawful connection between these magnitudes.

Continuous variability provides quantum physics with a criterion to distinguish a composite thing (with a character of its own) from an aggregate. Consider the

interaction between an electron and a proton. In the most extreme case this leads to the absorption of the electron and the transformation of the proton into a neutron (releasing a neutrino). At a lower energy, the interaction may lead to a bound state having the character of a hydrogen atom if the total energy (kinetic and potential) is negative. Finally, if the total energy is positive, we have an unbound state, an aggregate. In the bound state the energy can only have discrete values, it is quantized, whereas in the unbound state the energy is continuously variable.

Hence, if the rest energy has a characteristic value and internal energy states are lacking, we have an elementary particle (a lepton or a quark). If there are internal discrete energy states we have a composite character, whereas we have an aggregate if the internal energy is continuously variable.

With aggregates it is easier to abstract from specific properties than in the case of the characters of composite systems discussed in section 5.3. Studying the properties of macroscopic physical bodies, thermodynamics starts from four general laws, for historical reasons numbered 0 to 3 and written with capitals.

The Zeroth Law states that two or more bodies (or parts of a single body) can be in mutual equilibrium. Now the temperature of the interacting bodies is the same, and in a body as a whole the temperature is uniform. Depending on the nature of the interaction, this applies to other intensive magnitudes as well, for instance the pressure of a gas, or the electric or chemical potential. In this context bodies are not necessarily spatially separated. The thermodynamic laws apply to the components of a mixture as well. Equilibrium is an equivalence relation (2.1). An intensive magnitude like temperature is an equilibrium parameter, to be distinguished from an extensive magnitude like energy, which is additive. If two unequal bodies are in thermal equilibrium with each other, their temperature is the same, but their energy is different and the total energy is the sum of the energies of the two bodies apart. An additive magnitude refers to the quantitative relation frame, whereas an equilibrium parameter is a projection on the spatial frame.

According to the First Law of thermodynamics, the total energy is constant, if the interacting bodies are isolated from the rest of the world. The thermodynamic law of conservation of energy forbids all processes in which energy would be created or annihilated. The First Law does not follow from the fact that energy is additional. Volume, entropy, and the mass of each chemical component are additive as well, but not always constant in an interaction.

The Second Law states that interacting systems proceed towards an equilibrium state. The entropy decreases if a body loses energy and increases if a body gains energy, but always in such a way that the total entropy increases as long as equilibrium is not reached. Based on this law only entropy *differences* can be calculated.

According to the Third Law the absolute zero of temperature cannot be reached. At this temperature all systems would have the same entropy, to be considered the zero point on the entropy scale.

From these axioms other laws are derivable, such as Joshua Gibbs's phase rule (see below). As long as the interacting systems are not in equilibrium, the gradient of each equilibrium parameter acts as the driving force for the corresponding current causing equilibrium. A temperature gradient drives a heat current, a potential difference drives an electric current, and a chemical potential difference drives a material current. Any current (except a superconducting flow) creates entropy.

The thermodynamic axioms describe the natural laws correctly in the case of interacting systems being close to equilibrium. Otherwise, the currents are turbulent and a concept like entropy cannot be defined. Another restriction follows from the individuality of the particles composing the system. In the equilibrium state, the entropy is not exactly constant, but it fluctuates spontaneously around the equilibrium value. Quantum physics shows energy to be subject to Werner Heisenberg's relations (4.3). In fact, the classical thermodynamic axioms refer to a continuum, not to the actually coarse matter. Thermodynamics is a general theory of matter, whereas statistical physics studies matter starting from the specific properties of the particles composing a system. This means that thermodynamics and statistical physics complement each other.

An equilibrium state is sometimes called an 'attractor', attracting a system from any instable state toward a stable state. Occasionally, a system has several attractors, now called *local* equilibrium states. If there is a strong energy barrier between the local equilibrium states, it is accidental which state is realized. By an external influence, a sudden and apparently drastic transition may occur from one attractor to another one. In quantum physics a similar phenomenon is called 'tunnelling', to which I shall return in section 5.6.

a. A homogeneous set of particles having the same character may be considered a quantitatively founded aggregate, if the set does not constitute a structural whole with a spatially founded character of its own (like the electrons in an atom). In a gas the particles are not bound to each other. Usually, an external force or a container is needed to keep the particles together. In a fluid, the surface tension is a connective force that does not give rise to a characteristic whole. The composing particles' structural similarity is a condition for the applicability of statistics. Therefore I call a homogeneous aggregate quantitatively founded.

It is not sufficient to know that the particles are structurally similar. At least it should be specified whether the particles are fermions or bosons (4.4). Consider, for instance, liquid helium, having two varieties. In the most common isotope, a helium nucleus is composed of two protons and two neutrons. The net spin is zero, hence the nucleus is a boson. In a less common isotope, the helium nucleus has only one neutron besides two protons. Now the nucleus' net spin is $\frac{1}{2}$ and it is a fermion. This distinction (having no chemical consequences) accounts for the strongly diverging physical properties of the two fluids.

Each homogeneous gas is subjected to a specific law, called the *statistics* or distribution function. It determines how the particles are distributed over the available

states, taking into account parameters like volume, temperature, and total energy. The distribution function does not specify which states are available. Before the statistics is applicable, the energy of each state must be calculated separately.

The Fermi-Dirac statistics based on Wolfgang Pauli's exclusion principle applies to all homogeneous aggregates of fermions, i.e., particles having half-integral spin. For field particles and other particles having an integral spin, the Bose-Einstein statistics applies, without an exclusion principle. If the mean occupation number of available energy states is low, both statistics may be approximated by the classical Maxwell-Boltzmann distribution function. Except at very low temperatures, this applies to every dilute gas consisting of similar atoms or molecules. The law of Robert Boyle and Louis Gay-Lussac follows from this statistics. It determines the relation between volume, pressure and temperature for a dilute gas, if the interaction between the molecules is restricted to elastic collisions and if the molecular dimensions are negligible. Without these two restrictions, the state equation of Johannes Van der Waals counts as a good approximation. Contrary to the law of Boyle and Gay-Lussac, Van der Waals' equation contains two constants characteristic for the gas concerned. It describes the condensation of a gas to a fluid as well as the phenomena occurring at the critical point, the highest temperature at which the substance is liquid.

b. It is not possible to apply statistics directly to a mixture of subjects having different characters. Sometimes, it can be done with respect to the components of a mixture apart. For a mixture of gases like air, the pressure exerted by the mixture equals the sum of the partial pressures exerted by each component apart in the same volume at the same temperature (John Dalton's law). The chemical potential is a parameter distinguishing the components of a heterogeneous mixture.

I consider a heterogeneous mixture like a solution to have a spatial foundation, because the solvent is the *physical environment* of the dissolved substance. Solubility is a characteristic disposition of a substance dependent on the character of the solvent as the potential environment.

Stable characters in one environment may be unstable in another one. When solved in water, common salt molecules fall apart into sodium and chlorine ions. In the environment of water, the dielectric constant is much higher than in air. Now Charles Coulomb's force between the ions is proportionally smaller, too small to keep the ions together. (A more detailed explanation depends on the property of a water molecule to have a permanent electric dipole moment (5.3). Each sodium or chlorine ion is surrounded by a number of water molecules, decreasing their net electric charge. This causes the binding energy to be less than the mean kinetic energy of the molecules.)

The composition of a mixture, the number of grams of solved substance in one litre water, is accidental. It is not determined by any character but by its history. This does not mean that two substances can be mixed in any proportion whatsoever. However, within certain limits dependent on the temperature and the characters of the substances concerned, the proportion is almost continuously variable.

c. Even if a system only consists of particles of the same character, it may not appear homogeneous. It exists in two or more different 'phases' simultaneously, for example, the solid, liquid, and vaporous states. A glass of water with melting ice is in internal equilibrium at 0 °C. If heat is supplied, the temperature remains the same until all ice is melted. Only chemically pure substances have a characteristic melting point. In contrast, a heterogeneous mixture has a melting trajectory, meaning that during the melting process, the temperature increases. A similar characteristic transition temperature applies to other phase transitions in a homogeneous substance, like vaporizing, the transition from a paramagnetic to a ferromagnetic state, or the transition from a normal to a superconducting state. Addition of heat or change of external pressure shifts the equilibrium. A condition for equilibrium is that the particles concerned move continuously from one phase to the other. Therefore I call it a homogeneous *kinetically* founded aggregate.

An important example of a *heterogeneous* kinetic equilibrium concerns chemical reactions. Water consists mostly of water molecules, but a small part (10^{-7} at 25°C) is dissociated into positive H-ions and negative OH-ions. In the equilibrium state, equal amounts of molecules are dissociated and associated. By adding other substances (acids or bases), the equilibrium is shifted.

Both phase transitions and chemical reactions are subject to characteristic laws and to general thermodynamic laws, for instance Joshua Gibbs's phase rule.[\[22\]](#)

5.6. Coming into being, change and decay

I call an event physically qualified if it is primarily characterized by an interaction between two or more subjects. A process is a characteristic set of events, partly simultaneously, partly successively. Therefore, physically qualified events and processes often occur in an aggregate, sometimes under strictly determined circumstances, among which the temperature. In a mixture, physical, chemical, and astrophysical reactions lead to the realization of characters. Whereas in physical *things* properties like stability and life time are most relevant, physical and chemical *processes* concern the coming into being, change, and decay of those things.

In each characteristic event a thing changes of character (it emerges or decays) or of state (preserving its identity). With respect to the thing's character considered as a law, the first case concerns a *subjective* event (because the subject changes). The second case concerns an *objective* event (for the objective state changes). Both have secondary characteristics. I shall briefly mention some examples.

Annihilation or creation of particles is a subjective numerically founded event. Like any other event, it is subject to conservation laws. An electron and a positron emerge

simultaneously from the collision of a γ -particle with some other particle, if the photon's energy is at least twice the electron's rest energy. The presence of another particle, like an atomic nucleus, is required in order to satisfy the law of conservation of linear momentum. For the same reason, at least two photons emerge when an electron and a positron destroy each other.

By emitting or absorbing a photon, a nucleus, atom or molecule changes its state. This is a spatially founded *objective* transformation. In contrast, in a nuclear or chemical reaction one or more characters are transformed, constituting a *subjective* spatially founded event. In alpha- or beta-radioactivity, a nucleus changes subjectively its character, in gamma-activity it only changes objectively its state.

An elastic collision is an event in which the kinetic state of a particle is changed without consequences for its character or its internal state. Hence, this concerns an objective kinetically founded event. In a non-elastic collision a subjective change of character or an objective change of state occurs. Quantum physics describes such events with the help of operators determining the transition probability.

A process is an aggregate of events. In a homogeneous aggregate, phase transitions may occur. In a heterogeneous aggregate chemical reactions occur (5.5). Both are kinetically founded. This also applies to transport phenomena like electric, thermal or material currents, thermo-electric phenomena, osmosis and diffusion.

Conservation laws are 'constraints' restricting the possibility of processes. For instance, a process in which the total electric charge would change is impossible. In atomic and nuclear physics, transitions are known to be forbidden or improbable because of selection rules for quantum numbers characterizing the states concerned.

Physicists and chemists take for granted that each process that is not forbidden is possible and therefore experimentally realizable. In fact, several laws of conservation like those of lepton number and baryon number were discovered because certain reactions turned out to be impossible. Conversely, in 1930 Wolfgang Pauli postulated the existence of neutrinos, because otherwise the laws of conservation of energy and momentum would not apply to beta-radioactivity. Experimentally, the existence of neutrinos was not confirmed until 1956.

In common parlance, a collision is a rather dramatic event, but in physics and chemistry a collision is just an interaction between two or more subjects moving towards each other, starting from a large distance, where their interaction is negligible. In classical mechanics, their interaction means an attractive or repelling force. In modern physics, it implies the exchange of real or virtual particles like photons.

In each collision, at least the state of motion of the interacting particles changes. If that is all, we speak of an *elastic collision*, in which only the distribution of kinetic energy, linear and angular momentum over the colliding particles changes. A photon can collide elastically with an electron (Arthur Compton's effect), but an electron

cannot absorb a photon. Only a composite thing like a nucleus or an atom is able to absorb a particle.

Collisions are used to investigate the character of the particles concerned. A famous example is the scattering of alpha-particles (helium nuclei) by gold atoms (1911). For the physical process, it is sufficient to assume that the particles have mass and charge and are point-like. It does not matter whether the particles are positively or negatively charged. The character of this collision is statistically expressed in a mathematical formula derived by Ernest Rutherford. The fact that the experimental results (by Hans Geiger and Ernest Marsden) agreed with the formula indicated that the nucleus is much smaller than the atom, and that the mass of the atom is almost completely concentrated in the nucleus. A slight deviation between the experimental results and the theoretical formula allowed of an estimate of the size of the nucleus, its diameter being about 10^4 times smaller than the atom's. The dimension of a microscopic invisible particle is calculable from similar collision processes, and is therefore called its collision diameter. Its value depends on the projectiles used. The collision diameter of a proton differs if determined from collisions with electrons or with neutrons.

In a non-elastic collision the internal structure of one or more colliding subjects changes in some respect. With billiard balls only the temperature increases, kinetic energy being transformed into heat, causing the motion to decelerate.

In a non-elastic collision between atoms or molecules, the state of at least one of them changes into an excited state, sooner or later followed by the emission of a photon. This is an objective characteristic process.

The character of the colliding subjects may change subjectively as well, for instance, if an atom loses an electron and becomes an ion, or if a molecule is dissociated or associated.

Collisions as a means to investigate the characters of subatomic particles have become a sophisticated art in high-energy physics.

Spontaneous decay became first known at the end of the nineteenth century from radioactive processes. It involves strong, weak or electromagnetic interactions, respectively in α -, β -, and γ -radiation. The decay law of Ernest Rutherford and Frederick Soddy (1902) approximately represents the character of a single radioactive process. This statistical law is only explainable by assuming that each atom decays independently of all other atoms. It is a random process. Besides, radioactivity is almost independent of circumstances like temperature, pressure and the chemical compound in which the radioactive atom is bound. Such decay processes occur in nuclei and sub-atomic particles, as well as in atoms and molecules being in a metastable state. The decay time is the *mean duration of existence* of the system or the state.

Besides spontaneous ones, stimulated transformations occur. Albert Einstein first investigated this phenomenon in 1916, with respect to transitions between two energy levels of an atom or molecule, emitting or absorbing a photon. He found that (stimulated) absorption and stimulated emission are equally probable, whereas spontaneous emission has a different probability.^[23] In stimulated emission, an incoming photon causes the emission of another photon such that there are two photons after the event, mutually coherent, i.e., having the same phase and frequency. Stimulated emission plays an important part in lasers and masers, in which coherent light respectively microwave radiation is produced. Absorption is always stimulated.

Stimulated emission is symmetrical with stimulated absorption, but spontaneous emission is asymmetric and irreversible.

A stable system or a stable state may be separated from other systems or states by an energy barrier. It may be imagined that a particle is confined in an energy well, for instance an alpha-particle in a nucleus. According to classical mechanics, such a barrier is insurmountable if it has a larger value than the kinetic energy of the particle in the well, but quantum physics proves that there is some probability that the particle leaves the well. This is called 'tunnelling', for it looks like the particle digging a tunnel through the energy mountain.

Consider a chemical reaction in which two molecules A and B associate to AB and conversely, AB dissociates into A and B . The energy of AB is lower than the energy of $A+B$ apart, the difference being the binding energy. A barrier called the activation energy separates the two states. In an equilibrium situation, the binding energy and the temperature determine the proportion of the numbers of molecules ($N_A \cdot N_B / N_{AB}$). It is independent of the activation energy. At a low temperature, if the total number of A 's equals the total number of B 's, only molecules AB will be present. In an equilibrium situation at increasing temperatures, the number of molecules A and B increases, and that of AB decreases. In contrast, the *speed* of the reaction depends on the activation energy (and again on temperature). Whereas the binding energy is a characteristic magnitude for AB , the activation energy partly depends on the environment. In particular the presence of a catalyst may lower the activation energy and stimulate tunnelling, increasing the speed of the reaction.

The possibility to overcome energy barriers explains the possibility of transitions from one stable system to another one. It is the basis of theories about radioactivity and other spontaneous transitions, chemical reaction kinetics, the emergence of chemical elements and of phase transitions, without affecting theories explaining the existence of stable or quasi-stable systems.

In such transition processes the characters do not change, but a system may change of character. The laws do not change, but their subjects do.

The chemical elements have arisen in a chain of nuclear processes, to be distinguished as fusion and fission. The chain starts with the fusion of hydrogen nuclei (protons) into helium nuclei, which are so stable that in many stars the next steps do

not occur. Further processes lead to the formation of all known natural isotopes up to uranium. Besides helium with 4 nucleons, beryllium (8), carbon (12), oxygen (16), and iron (56) are relatively stable. In all these cases, both the number of protons and the number of neutrons is even.

The elements only arise in specific circumstances. In particular, the temperature and the density are relevant. The transition from hydrogen to helium occurs at 10 to 15 million Kelvin and at a density of 0.1 kg/cm^3 . The transition of helium into carbon, oxygen and neon occurs at 100 to 300 million Kelvin and 100 kg/cm^3 .^[24] Only after a considerable cooling down, these nuclei form with electrons the atoms and molecules to be found on the earth.

Once upon a time the chemical elements were absent. This does not mean that the laws determining the existence of the elements did not apply. The laws constituting the characters of stable and metastable isotopes are universally valid, independent of time and place. But the realization of the characters into actual individual nuclei does not depend on the characters only, but on circumstances like temperature as well. On the other hand, the available subjects and their relations determine these circumstances. Like initial and boundary conditions, characters are conditions for the existence of individual nuclei. Mutatis mutandis, this applies to electrons, atoms and molecules as well.

In the preceding chapters, I discussed quantitative, spatial, and kinetic characters. About the corresponding subjects, like groups of numbers, spatial figures, or wave packets, it cannot be said that they come into being or decay, except in relation to physical subjects. Only interacting things emerge and disappear. Therefore there is no quantitative, spatial or kinetic evolution comparable to the astrophysical one, even if the latter is expressed in numerical proportions, spatial relations, and characteristic rhythms.

Although stars have a lifetime far exceeding the human scale, it is difficult to consider them stable. Each star is a reactor in which continuously processes take place. Stars are subject to evolution. There are young and old stars, each with their own character. Novae and supernovae, neutron stars, and pulsars represent various phases in the evolution of a star. The simplest stellar object may be the black hole, behaving like a thermodynamic black body subject to the laws of thermodynamics.^[25]

These processes play a part in the theory about the astrophysical evolution, strongly connected to the standard model (5.1). It correctly explains the relative abundance of the chemical elements.^[26] After the start of the development of the physical cosmos, about thirteen billion years ago, it has expanded. As a result all galaxies move away from each other, the larger the distance, the higher their speed. Because light needs time to travel, the picture we get from galaxies far away concerns states from era's long past. The most remote systems are at the spatio-temporal horizon of the physical cosmos. In this case, astronomers observe events that occurred shortly after the big bang, the start of the astrophysical evolution.

Its real start remains forever behind the horizon of our experience. Astrophysicists are aware that their theories based on observations may approach the big bang without ever reaching it. The astrophysical theory describes what has happened since the beginning - not the start itself - according to laws discovered in our era. The extrapolation towards the past is based on the supposition that these laws are universally valid and constant. This agrees with the realistic view that the cosmos can only be investigated from within. It is not uncommon to consider our universe as one realized possibility taken from an ensemble of possible worlds.^[27] However, there is no way to investigate these alternative worlds empirically.

-
- [1] Lucas 1973, 43-56.
 [2] Omnès 1994, 193-198, 315-319.
 [3] Dijksterhuis 1950; Reichenbach 1956; Gold (ed.) 1967; Grünbaum 1973; 1974; Sklar 1974, chapter V; Sklar 1993; Prigogine 1980; Coveney, Highfield 1990; Stafleu 2019, chapter 9.
 [4] Reichenbach 1956, 135.
 [5] Reichenbach 1956, 115.
 [6] Putnam 1975, 88.
 [7] von Laue 1949; Jammer 1961; Elkana 1974a; Harman 1982.
 [8] Jammer 1957; Cohen, Smith (eds.) 2002.
 [9] Morse 1964, 53-58; Callen 1960, 79-81; Stafleu 1980, 70-73.
 [10] Callen 1960, 293-308.
 [11] Morse 1964, 106-118; Callen 1960, 288-292; Prigogine 1980, 84-88.
 [12] Sklar 1993, chapters 5-7.
 [13] McCormach 1970a; Kragh 1999, chapter 8.
 [14] Jammer 1961, chapter 11.
 [15] Pickering 1984, chapter 9-11; Pais 1986, 603-611.
 [16] See Walker, Slack 1970, who do not mention Faraday's ion.
 [17] See Millikan 1917; Anderson 1964; Thomson 1964; Pais 1986; Galison 1987; Kragh 1990; 1999.
 [18] Pais 1986, 466; Pickering 1984, 67.
 [19] Weisskopf 1972, 41-51.
 [20] See Barrow, Tipler 1986, 5, 252-254.
 [21] Cat 1998, 288.
 [22] Callen 1960, 206-207.
 [23] Einstein 1916.
 [24] Mason 1991, 50.
 [25] Hawking 1988, chapter 6, 7.
 [26] Mason 1991, chapter 4.
 [27] Barrow, Tipler 1986, 6-9.
-

Chapter 6

Organic characters

- 6.1. The biotic relation frame
 - 6.2. The organization of biochemical processes
 - 6.3. The character of biotic processes
 - 6.4. The secondary characteristic of organisms
 - 6.5. Populations
 - 6.6. The gene pool
 - 6.7. Does a species correspond with a character?
-

Encyclopaedia of relations and characters. 6. Organic characters

6.1. The biotic relation frame

No doubt, 1859 was the birth year of modern biology. Charles Darwin and Alfred Wallace were neither the first nor the only evolutionists, and their path was paved by geologists in the preceding century establishing that the earth is much older than was previously perceived, and that many animals and plants living in prehistoric times are now extinct.[1] The publication of Darwin's *On the origin of species by means of natural selection* drew much attention, criticism as well as approval. In contrast, Gregor Mendel's discovery in 1865 of the laws called after him, which would become the basis of genetics, was ignored for 35 years. The synthesis of Darwin's idea of natural selection with genetics, microbiology, and molecular biology (circa 1930) constitutes the foundation of modern biology.

This chapter applies the relational character theory, introduced in chapter 1, to living beings and life processes. The genetic relation, leading to renewal and ageing, is the primary characteristic of living subjects (6.1). I investigate successively the characters of organized and of biotic processes (6.2, 6.3), of individual organisms (6.4) and of populations and their dynamic evolution (6.5, 6.6). For the time being, I shall take for granted that a species corresponds to a character. Section 6.7 deals with the question of whether this assumption is warranted.

Life presupposes the existence of inorganic matter, including the characters typified by the relation frames of number, space, motion, and interaction. Organisms do not consist of other atoms than those occurring in the periodic system of chemical elements. All physical and chemical laws are unrestrictedly valid for living beings and life processes, taking into account that an organism is not a physically or chemically closed system.

Both in living organisms and in laboratory situations, the existence of organized and controlled chemical processes indicates that biotic processes are not completely reducible to physical and chemical ones. In particular, the genetic laws for reproduction make no sense in a restricted physical or chemical context. Rather, they transcend the physical and chemical laws without denying these:

'Except for the twilight zone of the origin of life, the possession of a genetic program provides for an absolute difference between organisms and inanimate matter.'^[2] '... the existence of a genetic program ... constitutes the most fundamental difference between living organisms and the world of inanimate objects, and there is no biological phenomenon in which the genetic program is not involved ...'^[3] 'Everything in a living being is centred on reproduction'.^[4] '... "life" is not so much defined by certain single characters but by their combination into individualized, purposefully

functioning systems showing a specific activity, limited to a certain life span, but capable of reproduction, and undergoing gradually hereditary alterations over long periods.'[5]

For the biotic relation frame the *genetic law* is appropriate. Each living organism descends from another one, and all living organisms are genetically related. This also applies to cells, tissues, and organs of a multicellular plant or animal. Its descent determines the function of a cell, a tissue, or an organ in an organism, as well as the position of an organism in taxonomy. The genetic law constitutes the universal relation frame for all living beings. Empirically, it is amply confirmed, and it is the starting point of major branches of biological research, like genetics, evolution theory, and taxonomy. However, in physical and chemical research, the genetic law only plays a part in biochemistry and biophysics.

The genetic order is more than a static relationship. It has the dynamics of innovation and ageing. Renewal is a characteristic of life, strongly related to sexual or asexual cell division, to growth and differentiation. The individual life cycle of fertilization, germination, growth, reproduction, ageing, and dying is irreversible. Rejuvenation occurs in a series from one generation to the next, and between cells in a multicellular organism. (This does not exclude neoteny and other forms of heterochrony.[6]) A population goes through periods of rise, blooming, regress, and extinction. Speciation implies innovation as well.

Each living being descends from another living being. The law statement, *omne vivum e vivo*, is relatively recent. Even in the nineteenth century, *generatio spontanea* was accepted as a possibility. Empirical and theoretical research have led to the conviction that life can only spring from life.[7] The theory of evolution does not exclude spontaneous generation entirely, for that would constitute the beginning of the biotic evolution. It might even be possible that the two kingdoms of prokaryotes arose independently. In contrast, there are good reasons to assume that eukaryotic cells have evolved from the prokaryotes, and multicellular plants, fungi, and animals from unicellular eukaryotes.

Most biologists accept a stronger law than *omne vivum e vivo*. It states that all living beings are genetically related, having a common ancestry. This law, to be called *the genetic law*, is hard to prove. Paleontological research alone does not suffice to demonstrate that all organisms have the same ancestors,[8] but it achieves support from other quarters. The argument that all living beings depend on the same set of four or five nucleic acids and twenty amino acids is not strong. Perhaps no other building blocks are available. But in eukaryotes these molecules only occur in the laevo variant, excluding the mirror-symmetric dextro variant. These two are energetically equivalent, and chemical reactions (as far as applicable) always produce molecules of the two variants in equal quantities. In the production of amino acids, similar DNA and RNA molecules are involved. In widely differing organisms, many other processes proceed identically: '... from the bacterium to man the chemical machinery is essentially the same ... First, in its structure: all living beings ... are made up of ... proteins and nucleic acids ... constituted by the assembling of the same

residues ... Second, in its functioning: the same reactions, or rather sequences of reactions, are used in all organisms for the essential chemical operations ...'[9] Moreover, all plants, animals, and fungi consist of cells, although there are large differences between prokaryotic and eukaryotic cells, as well as between plant and animal cells. Prokaryotic cells are more primitive and much smaller than eukaryotic cells, and the cell wall is in plants thicker and more rigid than in animals.

The fundamental laws of the universal relation frames cannot be logically derived from empirical evidence, even if this is abundantly available. The laws of thermodynamics, the mechanical conservation laws, and the law of inertia are no more provable than the genetic law. Such fundamental laws function as axioms in a theory, providing the framework for scientific research of characters. In this sense, the genetic law has proved to be as fruitful as the generally valid physical and kinetic laws. This does not mean that such laws are not debatable, or void of empirical content. On the contrary, the law of inertia was accepted in the seventeenth century only after a long struggle with the Aristotelian philosophy of nature, from which science had to be emancipated. The law of conservation of energy and the Second Law of thermodynamics were accepted only about 1850. Similarly, only in the twentieth century the genetic law was recognized after laborious investigations. In all these cases, empirically sustained arguments ultimately turned the scale.

With respect to the biotic relation frame, the theory of evolution is as general as thermodynamics is with respect to physical and chemical relations. Both theories concern aggregates, but they are nevertheless indispensable for understanding the characters of individual things and processes. The main axioms of evolution theory are the genetic law and laws for natural selection with respect to populations.[10] In general terms, the theory of evolution explains why certain species can maintain themselves in their environment and others cannot, pointing out the appropriate conditions. In specific cases, the evolution theory needs additional data and characteristic laws, in order to explain why a certain species is viable in certain circumstances. Also in this respect, evolution theory is comparable to thermodynamics. At the end of the nineteenth century, energeticists like Friedrich Ostwald assumed that thermodynamics should be able to explain all physical and chemical processes. Atomic theory and quantum physics made clear that thermodynamics is too general for that. Likewise, in my view, evolution theory is not specific enough to explain biotic characters.

The genetic law lies at the basis of biological taxonomy. Like plants and fungi, as well as protists and prokaryotes, animals are subject to biotic laws, but I shall assume that they are primarily characterized by another relation frame, to be called *psychical* (chapter 7). Within their generic psychic character, a specific organic character is interlaced. Genetic relations primarily characterize all other living beings and life processes. Each biotic process is involved with replication (6.3), and the nature of each living being is genetically determined (6.4). Within an organism, physical and chemical processes have the tertiary disposition to function in biotic processes (6.2). Living beings support symbiotic relations leading to evolution (6.5).

The genetic law is a leading principle of explanation for taxonomy and the modern species concept. The universal relation frames allow us of identifying any thing or event, to establish its existence and change, and to find its temporal relations to other things and events. In principle, the genetic law allows of the possibility to order all organisms into a biological taxonomy. The empirical taxonomy does not originate from human thought but from living nature. Its leading principle is not logical but biological. A logical, i.e., deductive classification is based on a division of sets into subsets, considering similarities and differences. It descends *logically* from general (the kingdoms and phyla) to specific (the species). In contrast, the biological ordering depends on *genetic* descent, ascending inductively from species to the higher categories.

Genetic relations can be projected on the relation frames preceding the biological one. On the different levels of taxonomy, a species, and a multicellular organism, these mappings can be distinguished as follows.

a. A *lineage* is a serial projection of the genetic order on the quantitative relation frame. Within a species one finds the linear relation of parent to offspring. Within a multicellular organism the serial order concerns each line of replicating cells. 'Biological entity', 'parent of', and 'ancestor' are primitive, undefinable concepts in the following two axioms: 'No biological entity is a parent of itself. If *a* is an ancestor of *b*, then *b* is not an ancestor of *a*.'^[11] However, if the mentioned terms are undefined, the natural numbers satisfy these axioms as well (2.1). By counting the intermediary specimens, it is possible to establish the genetic relation between two individuals, organs, or cells, that are serially connected.

b. Parallel lineages are mutually connected by common ancestry. Therefore species, organs, or cells having no serial relation may be related by *kinship*, the genetic relation between siblings, cousins, etc. Kinship of parallel lineages is to be considered a spatial expression of the genetic relation. Each branching means a new species, a new individual, a new organ, or a new cell. In taxonomy, biologists establish kinship between species on the basis of similarities and differences. These concern shape (morphology), way of life (physiology), development of an organism (in particular embryology), the manner of reproduction, and nowadays especially comparing DNA, RNA, or the proteins they produce.^[12] Kindred lineages are connected in a *cladogram*, a diagram showing the degree of kinship between species. If an organism has several descendants, the lineage branches within a species. In sexual reproduction lineages are connected and each organism has two parents, four grandparents, etc. Within an organism cell division causes branching. In a plant, fungus, or animal, recently branched cells lie close to each other. The larger the genetic distance between two cells, the smaller is their kinship.

c. *Genetic development* may be considered the kinetic projection of the order of renewal and ageing. Temporal relations are recognizable in the generation difference as a biotic measure mapped on kinetic time. It is the time between two successive bifurcations of a species, between the germination of a plant and that of its seeds, or between two successive cell divisions. If timing is taken into account, a cladogram

becomes a phylogenetic tree. Between two splits a population evolves. From germination to death an organism develops, and cells differentiate and integrate into tissues and organs.

d. The dynamic force of evolution within a species and the splitting of species consist of *competition* and *natural selection*. These may be considered projections of the genetic relation on the physical. Between plants, the competition concerns physical and chemical resources for existence, between fungi and animals organic ones as well. Competition is a repulsive force, to use a physical term. Besides natural selection, accidental processes lead to genetic changes, mostly in small isolated populations. This phenomenon is called 'random genetic drift' or 'inbreeding' in common parlance. Breeders use it to achieve desirable plant or cattle variations. There are attractive forces as well. Only within a species, sexual reproduction is the most innovative form of replication. Sexual interaction may be considered a specific physical expression of the genetic relation. Within an organism, neighbouring cells influence each other during their differentiation and integration.

These projections give rise to four types, each of organized chemical processes (6.2), biotic processes (6.3), biotically qualified thing-like characters (6.4), and their aggregates (6.5, 6.6).

6.2. The organization of biochemical processes

In each living being, many organized biochemical processes take place, having a function in the life of a cell, a tissue, an organ, or an organism. The term organism for an individual living being points to its character as an organized and organizing unit. The organism has a temporal existence. It emerges when the plant germinates, it increases in largeness and complexity during its development, it ages, and after its death it falls apart.

An organized unit is not necessarily a living being. A machine does not live, but it is an organized whole, made after a design. A machine does not reproduce itself and is not genetically related to other machines. Because human persons design a machine, its design cannot be found in the machine itself. In a living organism, the natural design is laid down in the genome, the ordered set of genes based on one of more DNA molecules. 'If you find something, anywhere in the universe, whose structure is complex and gives the strong appearance of having been designed for a purpose, then that something either is alive, or was once alive, or is an artefact created by something alive.'^[13] 'Entities have functions when they are designed to do something, and their function is what they are designed to do. Design can stem from the intentions of a cognitive agent or from the operation of selection ...'^[14]

The organism transfers the design from cell to cell and from generation to generation. The natural design changes because of mutation at the level of a single cell, because of sexual interaction at the level of organisms, or caused by natural selection at the level of a population. It is bound to general and specific laws determining the conditions under which the design is executable or viable. A design is the objective prescription for a biotic character. It is a chemical character having a tertiary biotic characteristic.

The processes to be discussed in the present section are primarily physically qualified, and some of them can be organized in a laboratory or factory. Their disposition to have a function in biotic processes is a tertiary characteristic (6.3).

a. Molecules are assembled according to a design. Although the concept of a lineage points to a relation between living beings, there is an analogy on the molecular level. This refers to the assemblage of molecules according to a genetic design as laid down in the DNA molecules. The DNA composition is partly species specific, partly it is unique for each individual living being.

The natural design for an organism is laid down in its genome, the genetic constellation of the genes in a specific sequence. The DNA molecules are the objective bearers of the genetic design, which is the genotype determining the phenotype, that is the appearance of a living being. Each organism has its own genome, being the objective expression of the species to which it belongs as well as of its individuality. Like the DNA molecules, the genome is mostly species specific.

A DNA molecule consists of a characteristic sequence of bases (nucleotides) of nucleic acids indicated by the letters A (adenine), C (cytosine), G (guanine) and T (thymine). In RNA, uracil (U) replaces thymine. The production of uracil costs less energy than that of thymine, which is more stable.^[15] Stability is for DNA more important than for RNA that is assembled repeatedly. Hence, mistakes in the transfer of design are easy to correct. The double helix structure enhances the stability of DNA as well. RNA consists of only one series of nucleotides. DNA is the start of the assembly lines of the molecules having a function in life processes. Three nucleotides form the design for one of the twenty possible amino acids. An RNA molecule is a replica of the part of the DNA molecule corresponding to a single gene. Mediated by an RNA molecule each gene designs a polypeptide or protein consisting of a characteristic sequence of amino acids. A protein is a large polypeptide. Sometimes the same gene assembles more than one protein. Often a gene occurs more than once in the DNA, its locus determining how the gene co-operates with other genes. Hence, similar genes may have different functions. A direct relation between a gene and a phenotypic characteristic is rare.^[16] Some proteins are enzymes acting as catalysts in these and other material transformations.

According to the 'central dogma of molecular biology', formulated by Francis Crick, the transfer of information from DNA via RNA ('transcription' by mRNA) to the polypeptides ('translation' by tRNA) is irreversible. With respect to the first step, the dogma does not apply entirely to viruses, and there are important differences between

prokaryotes and eukaryotes. The intervention of RNA is necessary in eukaryotes, because DNA is positioned in the cell nucleus, whereas the assembly of the polypeptides occurs elsewhere (in ribosomes). In prokaryotes, the translation may start before the transcription is finished. In transcription, a third form is produced, rRNA, concentrated in the ribosome, the organelle where the assembly of polypeptides takes place. Because RNA has mostly a transport function, its tertiary characteristic may be called kinetic.

Although a great deal of the assembly of molecules takes place according to a genetically determined pattern, interaction with the surroundings takes place as well. The environment is first of all the cell plasma, the more or less independent specialized organelles and the cell wall, in which many specific biochemical processes occur. Second, via the cell wall a cell is in contact with the physical and chemical environment. Third, in multicellular organisms the environment includes other cells in the immediate neighbourhood. Finally, only animal cells exert some kind of action at a distance (7.1).

To a large extent, the environment determines which genes or combinations of genes are active, being selectively switched on or off. The activity of genes in a multicellular organism depends on the phase of development. The genome acts in the germination phase differently than in an adult plant, in a root otherwise than in a flower. The genetic constellation determines the growth of an organism. Conversely, the genetic action depends on the development of the plant and the differentiation of its cells. *Epigenesis* is the name of the process, in which each phase in the development of a plant or animal is determined by preceding phases, genes and environment.[17]

Therefore, DNA is not comparable to a code, a blueprint, a map or diagram in which the phenotype is represented on a small scale. Rather, it is an extensive prescription, a detailed set of instructions for biochemical processes.[18] The conception of the composition of DNA as a code is a metaphor, inspired by the discovery that the structure of DNA can be written in only four symbols.

The enormous variation of molecules is possible because of the equality of the atoms and the uniformity of chemical bonding. This is comparable with the construction of machines. It is easy to vary machines if and as far as the parts are standardized and hence exchangeable. This applies to the disparity of organisms as well. The organization of a plant or an animal consists partly of standardized modules, some of which are homologous in widely different organisms. Such modules exist on the level of molecules (there are only twenty different amino acids, with an enormous variation in combinations), genes (standardized combinations of genes), cells (the number of cell types is restricted to several hundreds), tissues and organs. For evolutionary innovations, the existence of exchangeable parts having a different function in different combinations and circumstances is indispensable.[19] 'If each new species required the reinvention of control elements, there would not be time enough for much evolution at all, let alone the spectacularly rapid evolution of novel features observed in the phylogenetic record. There is a kind of tinkering at work, in which the same regulatory elements are recombined into new developmental machines. Evolution

requires the dissociability of developmental processes. Dissociability of processes requires the dissociability of molecular components and their reassembly.'[20]

b. The biotic functions of molecules depend on their shape. Although the macromolecules occurring in living beings have an enormous diversity, they have much in common as well. Polymers are chains of monomers connected by strong covalent bonds (6.3). Polysaccharides consist of carbohydrates (sugars); polypeptides are constructed from amino acids; and nucleic acids consist of nucleotides. The lipids (fats, oils and vitamins) constitute a fourth important group of large molecules. Lipids are not soluble in water. Lipids are not characterized by covalent bonds but by the weaker Van der Waals bonding. Phospholipids are the most important molecules in biotic membranes. In the double cell wall the molecules are at one end hydrophilic (attracting water), at the other end hydrophobic (repelling water). In the assembly of polymers, water is liberated, whereas polymers break down by hydrolysis (absorption of water).

All organisms apply the same monomers as building blocks of polymers. In contrast, the polymers, in particular the polypeptides and nucleic acids, are species specific. The twenty different amino acids can be connected to each other in each order and in large amounts. As a consequence, the diversity of proteins and their functions is enormous.

Polymers do not only differ because of their serial composition, but in particular by their spatial shape. Like all molecules, they are primarily physically qualified and secondarily spatially founded. DNA's double helix structure plays a part in its replication in cell division. Also other macromolecules display several spatial structures simultaneously. For the functioning of a protein as an enzyme, its spatial structure is decisive.

Each biochemical process has to overcome an energy barrier (6.6). Increasing the temperature is not suitable, because it accelerates each chemical process and is therefore not selective enough. Catalysis by specialized proteins (enzymes) or RNA molecules (ribozymes) is found in all organisms. In plants, the enzyme rubisco is indispensable for photosynthesis.

The polymers have various functions in an organism, like energy storage, structural support, safety, catalysis, transport, growth, defence, control, or motion. Only nucleic acids have a function in the reproduction of cells and organisms.

c. The genetic development of a living being depends on metabolism, a transport process. A cell can only live and replicate because of a constant stream of matter and energy through various membranes. A unicellular organism has direct contact with its environment, in which it finds its food and deposits its waste. This also applies to a multicellular organism consisting of a colony of independently operating cells, like many algae. These organisms' ideal environment is salt water, followed by fresh water and moist situations like mud or the intestines of animals. To colonial organisms, this imposes the constraint that a tissue cannot be thicker than two cells.

Multicellular fungi, plants, or animals need internal transport of food, energy, and waste, requiring cell differentiation, in which, for instance, the photosynthetic cells lie at the periphery of plants. Metabolism is an organized stream of matter through the organism. It allows of life outside water. In the atmosphere, oxygen is better accessible than in water, other materials are less accessible.

The cell wall is not merely a boundary of the cell. Nor is it a passive membrane that would transmit some kinds of matter better than others. Rather, it is actively involved in the transport of all kinds of matter from one cell to another. Membranes have an important function in the organization of biochemical processes, the assemblage of molecules, the transformation of energy, the transport of matter, the transfer of information, and the processing of signals. Hence, the presence of membranes may be considered a condition for life.

Plant cells are close together, and transport takes place directly from one cell to the other one. A plant cell has at least one intracellular cavity enclosed by a membrane. This is a vacuole, mostly filled with water, acting as a buffer storage and waste disposal. Animals have intercellular cavities between their cells. Animal cells are connected by proteins regulating the exchange of molecules and information. These proteins play an important part in the development of the embryo as well.

Passive transport is distinguished from active transport. Passive transport lacks an external source of energy and is caused by diffusion in a chemical solution or by osmosis if the solution passes a membrane. Osmosis occurs if a membrane lets pass a solvent (usually water) but not the solved matter. The solvent moves through the membrane in the direction of the highest concentration of the solved matter. This induces a pressure difference across the membrane that counteracts the transport. In equilibrium, the osmotic pressure in some desert plants can be up to a hundred times the atmospheric pressure.

Some substances pass a membrane together with proteins acting as carriers. The concentration gradient is the driving force of diffusion. The size and the electric charge of the molecules concerned and the distance to be travelled also influence the diffusion speed. In particular the distance is a constraint, such that diffusion is only significant within a cell and between two neighbouring cells. To cover larger distances other means of transport are needed.

Active transport requires a source of energy, like adenosine triphosphate (ATP). This transport is coupled to carriers and proceeds against a concentration difference like a pump. Endo- or exocytosis in eukaryotic cells is the process in which the cell wall encapsulates the substance to be transported. After the capsule has passed the wall it releases the transported substance. Animal cells have receptors in their wall sensitive for specific macromolecules. Besides, animals have organs specifically designed for transport, for instance by the circulation of blood.

No organism can live without energy. Nearly all organisms derive their energy directly or indirectly from the sun, by photosynthesis. This process transforms water, carbon

dioxide, and light into sugar and oxygen. This apparently simple chemical reaction is in fact a complicated and well organized process, only occurring in photosynthetic bacteria and in green plants. The product is glucose (a sugar with six carbon atoms in its molecule), yielding energy rich food for plants and all organisms that feed on plants.

The transformation of energy is a redox reaction. Some molecules oxidize by donating electrons, whereas other molecules reduce by accepting electrons. The first step is glycolysis (transformation of glucose in pyruvate), which does not require oxygen. Most organisms use oxygen for the next steps (cellular respiration). Other organisms are anaerobic, transforming energy by fermentation, which is less efficient. In the absence of oxygen, many aerobic cells switch to fermentation. Because nerve cells are unable to do so, they become easily damaged at a shortage of oxygen. Glycolysis, cellular respiration and fermentation are organized processes with many intermediate steps. The end product consists of ATP and other energy carriers that after transport cede their energy in other chemical reactions.

d. Self-replication of DNA molecules has a function in reproduction. Serving as a starting point for the assemblage of polypeptides, the DNA molecule has a specific spatial structure. It consists of a double helix of two sequences of nucleotides being each other's complement, because each adenine (A) in one sequence connects to a thymine (T) in the other one and each cytosine (C) in one sequence to a guanine (G) in the other. If the DNA molecule consists of two such strings it is called *diploid*. In *haploid* cells, the cell nucleus contains a single string of chromosomes, in *diploid* cells the chromosomes are paired. Each diploid gene occurs in a pair, except the sex chromosomes, being different in males (XY), equal in females (XX). Each chromosome is a single DNA molecule and consists of a large number of genes. The position of the genes on a chromosome is of decisive significance. On each position (locus) in a chromosome pair, there is at most one pair of genes, being *homozygote* (equal) or *heterozygote* (unequal). If in different individuals different genes can occupy the same locus, these genes are called *alleles*. The two halves are not identical, even if they look alike. This structure makes the DNA molecule very stable. A RNA molecule, acting as an intermediary between a gene on the DNA molecule and the assemblage of a polypeptide, is *haploid*. Consisting of a single helix, it is less stable than DNA. DNA is not always diploid. Many fungi consist of haploid cells. Only during sexual reproduction, their sex cells are diploid.

The DNA molecule itself is not assembled by another molecule. It has a unique way of self-duplication. Preceding a cell division, the diploid helix unfolds itself and the two haploid halves separate. In sexual cell division (meiosis) the next steps differ from those in the far more frequent asexual cell division (mitosis).

Mitosis is the asexual form of reproduction for unicellular organisms. It also occurs in the growth of all multicellular organisms. After the first division of the diploid DNA molecule, each half doubles itself by separating the two sequences and connecting a new complementary base to each existing base. Hence, two new diploid DNA molecules arise, after which the cell splits as well. The daughter cells are genetically identical to the mother cell.

Meiosis, the sexual cell division, is more varied. In a much occurring variant, the DNA remains haploid after the first splitting. As a rule, after the second division four daughter cells arise, each with half the DNA. Either all four are sperm cells or one is an egg cell, whereas the other three die or become organelles in the egg cell. Only after the egg cell merges with a sperm cell of another individual, a new diploid cell arises. This cell has a different composition of DNA, hence a new identity. This is only possible if the two merging DNA halves fit to each other. In most cases this means that the individuals concerned belong to the same species. In prokaryotes meiosis is often a more complicated process than in eukaryotes.

Cell division is not restricted to DNA replication. The membranes are not formed *de novo* but grow from the existing ones. In particular, the cell wall of the original cell is divided among the daughter cells. Life builds on life.

6.3. The character of biotic processes

Besides the organized biochemical processes there are processes that are typically biotically qualified. In section 6.5, I shall discuss the genetic changes occurring in a population. Important processes for the dynamic development of an individual organism, to be discussed in the present section, are cell division, spatial shaping, growth, and reproduction.

The genetic identity of a living organism as a whole is determined by its genetic contents. Its heredity is expressed in the genes and their material basis, the DNA molecules. All cells of a multicellular organism have the same DNA configuration and every two living beings have different DNA molecules, except in the case of asexual reproduction. The genes organize the biochemical processes discussed in section 6.2 as well as the biotic processes to be discussed in section 6.3. The genetic identity as an *organizing principle* of a living being determines its temporal unity. This unity disappears when the organism dies and disintegrates.

a. Cell division is a biotically qualified process that is quantitatively founded. The cell as a subjective unit multiplies itself. Sexual cell division (preceded by sexual interaction, see below) is distinguished from the more frequent asexual cell division (6.2). In the case of an eukaryotic cell, the nucleus too is divided into two halves. The other cell bodies, the cell plasma, and the cell wall, are shared out to the daughter cells and supplied by new ones.

Many organisms reproduce asexually. The prokaryotes and protists (mostly unicellular eukaryotes) reproduce by cell division. Many plants do so by self-

pollination. Now the daughter has the same genetic identity as the parent. In this respect they could be considered two spatially separated parts of the same plant. On the one hand, nothing is wrong with this view. Alaska is an integral part of the United States, though it is spatially separated from the mainland. The primary character determines the temporal unity of an individual, and in the case of a bacterium, a fungus, or a plant, this is its genetic identity. Only after a sexual reproduction the daughter plant is a real new individual, genetically different from its parents and any other individual. On the other hand, this view counters natural experience, accepting a plant as an individual only if it is coherent. Moreover, in asexual reproduction not only the spatial coherence is lost, but all kinds of biochemical and biotic interactions as well. This seems to be sufficient to suppose that asexual reproduction gives rise to a new individual.[\[21\]](#)

No single organism is subject to genetic change. Hardly anything can be found that is more stable than the genetic character and the identity of a living being. From its germination to its death, a plant remains identical to itself in a genetic sense. Only in sexual replication genetic change occurs. Of course, a plant is subject to other changes, both cyclic (seasonal) and during its development in its successive phases of life.

b. The genetic relation is not the only factor determining the biotic position of a living being. For each plant and every animal, its relation to the environment (the biotope or ecosystem) is a condition of life. First, the environment concerns subject-subject relations. Symbiosis should be considered a spatially founded way of living together. It is found on all levels of life. Within an eukaryotic cell symbiosis occurs between the cell nucleus and the organelles having their own DNA. In multicellular organisms cells form tissues or organs. In an ecosystem, unicellular and multicellular organisms live together, mutually dependent, competitive or parasitic.

Second, each organism has a subject-object relation to the physical and chemical surroundings of air, water, and soil. Just like the organized matter in the plant, the physical environment has a dynamic function in life processes.

Third, the character of plants anticipates the behaviour of animals and human beings. This constitutes an object-subject relation. By their specific shape, taste, colour, and flavour plants are observable and recognizable by animals as food, poison, or places suited for nesting, hunting, and hiding.

c. The dynamic development of a plant from its germination to adulthood may be considered a kinetically founded biotic process. It is accompanied by differentiation of cells and pattern formation in tissues, and by relative motion of cells in animals. The growth of a plant is strongly determined, programmed by the genome. In the cell division the DNA does not change, but the genes are differentially switched on and off. During the growth, cells differentiate into various types, influenced by neighbouring cells.[\[22\]](#)

There are other influences from the environment, for a plant only grows if the circumstances permit it. Most seeds never start developing, because the external factors are not favourable. Even for a developing plant, the genotype does not determine the phenotype entirely. The development of the plant occurs in phases from cell to cell, in which the phenotype of the next cell is both determined by its genotype and by the phenotype of the preceding cell and the surrounding cells, as well as by the physical, chemical, and organic environment.

The dynamic development of a plant or animal belongs to the least understood processes in biology. During the twentieth century, the attention of biologists was so much directed to evolution and natural selection, that the investigation of individual development processes (in which natural selection does not play a part) receded to the background. The complexity of these processes yields an alternative or additional explanation for the fact that relatively little is known about them. Some creationists take for granted 'natural' processes in the development of a human being from its conception, during and after pregnancy, while considering similar processes incomprehensible in evolution. A standard objection is that one cannot understand how by natural selection such a complicated organ like the human eye could evolve even in five hundred million years. However, who can explain the development of the human eyesight in nine months, starting from a single fertilized cell? In both cases, biologists have a broad understanding of the process, without being able to explain all details. (I am not suggesting here that the evolution and the development of the visual faculty are analogous processes.) It starts from a single point, fertilization, and expands into a series of parallel but related pathways. Sometimes one pathway may be changed without affecting others, leading to a developmental dissociation. Usually such dissociation is lethal, but if it is viable, it may serve as a starting point for evolutionary renewal.[23]

d. Sexual reproduction may be considered a primarily biotically qualified process that is secondarily physically founded, like a biotic interaction. Two genetically different cells unite, and the DNA splits before forming a new combination of genes (6.2). By sexual reproduction a new individual comes into being, with a new genetic identity.

Contrary to the growth of a plant, reproduction is to a large extent accidental. Which cells unite sexually is mostly incidental. Usually only sex cells from plants of the same species may pair, although hybridization occurs frequently in the plant kingdom. By their mating behaviour, animals sometimes limit accidents, increasing their chances. Even if a viable combination is available, the probability is small that the seed germinates, reaches adulthood and becomes a fruit bearing plant. Because the ultimate chance of success is small, a plant produces during its life an enormous amount of gametes. On the average and in equilibrium circumstances, only one fruit bearing descendant survives. The accidental nature and abundance of reproduction, supplied with incidents like mutation, is a condition for natural selection. But if it would occur in a similar way during the growth of a plant, no plant would ever reach the adult stage. Dynamic development is a programmed and reproducible process. Sexual reproduction (as well as evolution according to Charles Darwin) is neither.[24]

Fertilization is a biotically qualified process, interlaced with biochemical processes having a biotic function. Moreover, in animals fertilization is interlaced with the psychically qualified mating behaviour that is biotically founded.

6.4. The secondary characteristic of organisms

Because four relation frames precede the biotic one, we should expect four secondary types of biotically qualified thing-like characters. These are, respectively, quantitatively, spatially, kinetically, or physically founded. Each type is interlaced with the corresponding type of biotic processes mentioned in section 6.3. Moreover, the characters of different types are interlaced with each other as well.

a. It seems obvious to consider the *cell* to be the smallest unit of life. Each living being (viruses excepted) is either a cell or a composite of cells. However, this conceals the distinction between prokaryotes (bacteria and some algae) and eukaryotes. According to many biologists, this difference is more significant than that between plants and animals.[25] The oldest known fossils are prokaryotes, and during three-quarters of the history of the terrestrial biosphere, eukaryotes were absent. Prokaryotic cells are more primitive and usually smaller than eukaryotic cells. Most prokaryotes like bacteria are unicellular, although some colonial prokaryotes like algae exist. The protists, fungi, plants, and animals consist of eukaryotic cells. A bacterium has only one membrane, the cell wall. An eukaryotic cell has several compartments each enclosed by a membrane. Besides vacuoles these are particles like the cell nucleus, ribosomes (where RNA molecules assemble polypeptides), mitochondria (the power stations of a cell), and chloroplasts (responsible for photosynthesis). Prokaryotes have only one chromosome, eukaryotes more than one. Therefore, biologists consider the prokaryotes to belong to a separate kingdom, or even two kingdoms, the (eu)bacteria and the much smaller group of archaeobacteria (archaea).

It appears that the chromosomes in an eukaryotic cell have a prokaryotic character, as well as the genetically more or less independent mitochondria and chloroplasts. Having their own DNA, the latter organelles' composition is genetically related to that of the prokaryotes. These organelles are about as large as prokaryotic cells. RNA research indicates that mitochondria are genetically related to the purple group and chloroplasts to the cyanobacteria, both belonging to the eubacteria. The most primitive eukaryotes, the archaezoa, do not contain mitochondria or other organelles besides their nucleus. The similarity between prokaryotes and the organelles in eukaryotic cells was first pointed out by Lynn Margulis, in 1965.

Therefore, I consider the character of prokaryotes to be primarily biotic and secondarily quantitative. This may also apply to the characters of the mitochondria, chloroplasts, and chromosomes in an eukaryotic cell, and to the character of viruses

as well. None of these can exist as a living being outside a cell, but each has its own character and a recognizable genetic identity.

Contrary to bacteria, viruses are not capable of independently assembling DNA, RNA, and polypeptides, and they can only reproduce parasitically in a cell. Some viruses can be isolated forming a substance that is only physically and chemically active. Only if a virus enters a cell, it comes to life and starts reproducing. Outside the cell, a virus is primarily physically qualified, having a biotic disposition, to be actualized within a cell. Because a virus mainly *transports* DNA, its character may be considered to have a (tertiary) kinetic disposition (like RNA). A virus has a characteristic shape differing from the shape of a cell. Their character has the tertiary disposition to become interlaced in that of an eukaryotic cell. In eukaryotic organisms, reproduction starts in the prokaryotic chromosomes.

b. A spatially founded biotic character is characterized by *symbiosis* (6.3). The symbiosis of prokaryotes in an eukaryotic cell is called endosymbiosis. In the character of an eukaryotic cell several quantitatively founded prokaryotic characters are encapsulated. In turn, eukaryotic cells are the characteristic units of a multicellular fungus, plant, or animal. Likewise, an atomic nucleus (having a spatially founded character) acts like a quantitative unit in the character of an atom (5.3). Each cell has a spatial (morphological) shape, determined by the functions performed in and by the cell.

In colonial plants (thallophytes like some kinds of algae), the cells are undifferentiated. As in colonial prokaryotes, metabolism takes place in each cell independent of the other cells. In higher organisms, eukaryotic cells have the disposition to differentiate and to integrate into tissues and organs. Both in cell division and in growth, cells, tissues, or organs emerge having a specific shape. The spatial expression of an organism is found in its morphology, of old a striking characteristic of living beings. Since the invention of the optical microscope in the seventeenth century and the electron microscope in the twentieth, the structure of a cell is well known.

c. Differentiated organisms and organs have a kinetically founded character. Except for unicellular and colonial organisms, each living being is characterized by its dynamic development from the embryonic to the adult phase. Now the replication of cells leads to morphological and functional differentiation. In a succession of cell divisions, changes in morphology and physiology of cells occur. Their tertiary character takes distance from that of the gametes. This gives rise to differentiated tissues and organs like fibres, the stem and its bark, roots, and leaves. These have different morphological shapes and various physiological functions. In a differentiated plant, metabolism is an organized process, involving many cells in mutually dependent various ways (6.2). Growth is a biotic process (6.3). Differentiation enhances the plant's stability, fitness, and adaptive power.

Differentiation concerns in particular the various functions that we find in a plant. The biological concept of a *function* represents a subject-object relation as well as a disposition. Something is a biotic object if it has a function with respect to a biotic

subject (6.2). Cells, tissues, and organs are biotic subjects themselves. A cell has the disposition to be a part of a spatially founded tissue, in which it has a function of its own. A tissue has an objective function in a differentiated organ. By differentiation the functions are divided between cells and concentrated in tissues. In a differentiated plant, chlorophyll is only found in leaves, but it is indispensable for the whole plant. The leaves have a position such that they catch a maximum amount of light.

Differentiation leads to the natural development from germination to death. The variety in the successive life phases of fertilization, germination, growth, maturing, reproduction, ageing, and natural death is typical for differentiated fungi, plants, and animals.

Although the cells of various tissues display remarkable differences, their kinship is large. This follows from the fact that many plants are able to reproduce asexually by the formation of buds, bulbs, stolons, tubers, or rhizomes. In these processes, new individuals emerge from differentiated tissues of plants. Grafting and slipping of plants are agricultural applications of this regenerative power.

d. Sexual reproduction appears to be an important specific projection of the genetic relation on the physical and chemical relation frame. This biotic interaction between two living beings is the most important instrument of biotic renewal. All eukaryotic organisms reproduce by sexual cell division (even if some species reproduce by other means most of the time). In prokaryotes, the exchange of genetic matter does not occur by sexual interaction, but by the merger of two individuals. Reproduction is a biotic process (6.3). The part played by DNA replication is discussed in section 6.2.

In the highest developed plants, sexuality is specialized in typical sexual organs, like flowers, pistils, and stamens. Some plant species have separate male and female specimens. In sexually differentiated plants, the sexual relation determines the genetic cycle, including the formation of seeds. Fertilized seeds can exist some time independent of the parent plant without germinating, for instance in order to survive the winter. Sometimes they are provided with a hard indigestible wall, surrounded by pulp being attractive food for animals. The animal excretes the indigestible kernel, co-operating in the dispersal of the seeds.

In particular, sexual reproduction is relevant for the genetic variation within a population. This variation enhances the population's adaptability considerably. The genetic kinship between individuals in a population is much less than the genetic relation between cells within an individual organism.

The characteristic distinctions between an egg cell and pollen, between male and female sex organs in bisexual plants, and between male and female specimens in unisexual plants, have a function to prevent the merger of sex cells from the same individual. In bisexual plants self-pollination does occur, but sometimes the genetic cycle is arranged such as to preclude this. Fungi are not sexually differentiated but have other means to prevent self-fertilization. Within each fungus species several types occur, such that only individuals of different types can fertilize each other.

The above presented distinction of four biotic types of thing-like characters is only the start of their analysis. Real characters almost always consist of an interlacement of differently typed characters.

First, one recognizes the interlacement of equally biotically qualified but differently founded characters. In eukaryotic cells, an interlacement occurs with various organelles having a prokaryotic character. Because the organelles have various functions, this interlacement leads to a certain amount of differentiation. In all multicellular plants, the character of the cells is interlaced into that of a tissue. In differentiated plants, the character of organs is interlaced with those of tissues. This concerns both their morphological structure and their physiological functions. The highest developed plants display an interlacement of cells, tissues, leaves, roots, flowers, and seeds. Together they constitute the *organism*, the plant's primary biotic character. The differentiation of male and female organs or individuals is striking.

Second, the biotic organism is interlaced with characters that are not biotically qualified. First of all, these concern the physically qualified characters of the molecules composing the plant (6.2). Besides we find in a plant kinetic characters, typical motions of the plant as a whole or of its parts. An example is the daily opening and closing of flowers, or the transport of water from roots to leaves. Each plant and each of its cells, tissues, and organs have typical shapes. These characters are by no means purely physical, chemical, kinetic, or spatial. They are opened up by the biotic organism in which their characters are encapsulated. Their tertiary biotic disposition is more obvious than their primary qualifying or secondary founding relation frames. They have a function determined by the organism. Unlike cells and tissues, they do not form *parts* of the organism, as follows from the fact that they often persist some time after the death of the organism. Everybody recognizes the typical structure of a piece of wood to be of organic origin, even if the plant concerned is dead for a long time. Wood is not alive, but its physical properties and spatial structure cannot be explained from physical laws only. Wood is a *product* of a living being, which organism orders the physically qualified molecules in a typically biotic way.

Third, we encounter the interlacement of the organism with many kinds of biochemical and biotic processes (6.2, 6.3). Whereas physical systems always proceed to an equilibrium state, an organism is almost never at rest. (A boundary case is a seed in a quasi-stable state). Metabolism is a condition for life. Reproduction, development, and growth of a multicellular organism, and the seasonal metamorphosis of perennial plants, are examples of biotic processes. Each has its own character, interlaced with that of the organism.

The typology of characters differs from the biotic taxonomy. A relatively recent taxonomy of living beings still distinguished five kingdoms: *monera* (prokaryotes); *protocista* or *protista* (unicellular and colonial eukaryotic organisms); *fungi*; *animalia*; and *plantae*.^[26] Up till the sixties, besides the animal kingdom only one kingdom of plants was recognized, including the monera, protista, and fungi besides the 'true' plants.^[27] Nowadays the prokaryotes are divided into the kingdoms of (*eu*)*bacteria* and *archaebacteria* or *archaea*, differing from each other as much as they differ from

the eukaryotes. The protists form a set of mutually hardly related unicellular or colonial eukaryotes. Fungi are distinguished from plants by having haploid cells most of the time. Being unable to assimilate carbon, they depend on dead organic matter, or they parasitize plants or animals. DNA research reveals that fungi are more related to animals than to plants.

It cannot be expected that the typology discussed in this section would correspond to the biological taxonomy of species. Taxonomy is based on specific similarities and differences and on empirically found or theoretically assumed lineages and kinship. If the biotic kingdoms in the taxonomy would correspond to the division according to their secondary characteristic, this would mean that the four character types would have developed successively in a single line. In fact, many lineages evolve simultaneously. In each kingdom the actualization of animal phyla or plant divisions, classes, orders, etc. proceeds in the order of the four secondary character types and their interlacements. However, their disparity cannot be reduced to the typology based on the general relation frames.

The biological taxonomy, the division of species into genera, families, orders, classes, phyla or divisions, and kingdoms, is not based on the *general* typification of characters according to their primary, secondary, and tertiary characteristics. Rather, it is a *specific* typification, based on specific similarities and differences between species.

6.5. Populations

Sections 6.2 and 6.3 investigated physical, chemical, and biotic processes based on projections of the biotic relation frame on the preceding frames. Section 6.4, too, was mainly directed to secondary characteristics of biotic subjects. Now a tertiary characteristic will be considered, the disposition of organisms to adapt to their environment. Organisms do not evolve individually, but as a population in a biotope or ecosystem. Section 6.5 discusses the laws for populations and aggregates of populations, whereas section 6.6 treats the genome and the gene pool as objective aggregates.

a. A population is a homogeneous aggregate, a spatio-temporally bounded and genetically coherent set of living beings of the same species. Hence, a population is not a class but a collection. It is a spatial cross section of a lineage, which in turn is a temporally extended population.[28] Besides being genetically homogeneous, a population is also genetically varied, see below.

Two sets of organisms of the same species are considered different populations, if they are spatially isolated and the exchange of genetic matter is precluded. A

population as a whole evolves and isolated populations evolve independently from each other.

A population is a quantitatively founded biotic aggregate, having a number of objective properties open to statistical research, like number, dispersion, density, birth rate, and death rate. These numbers are subject to the law of abundance. Each population produces much more offspring than could reach maturity. The principle of abundance is a necessary condition for survival and evolutionary change. Competition, the struggle for life, sets a limit to abundance.[29]

Being threatened by extinction, small populations are more vulnerable than larger ones. Nevertheless, they are better fit to adaptation. Important evolutionary changes only occur in relatively small populations that are reproductively isolated from populations of the same species. As a 'founder population', a small population is able to start a new species. Large, widely dispersed populations are evolutionary inert.[30]

b. A biotope or ecosystem is a heterogeneous aggregate. It is a spatially more or less bounded collection of organisms of different species, living together and being more or less interdependent. The biotic environment or habitat of a population consists of other populations of various species.

A biotope is characterized by the symbiosis of prokaryotes and eukaryotes, of unicellular and multicellular organisms, of fungi and plants. Most biotopes are opened up because animals take part in them, and sometimes because they are organized by human interference. Biotopes like deserts, woods, meadows, or gardens are easily recognizable.

A population occupies a *niche* in a biotope. A *niche* or *adaptive zone* indicates the living room of a population. Both physical and biotic circumstances determine a niche, in particular predator-prey relations and the competition about space and food. Each niche is both possible and constrained because of the presence of other populations in the same area. In general, the geographic boundaries of the habitats of different species will not coincide. Therefore the boundary of a biotope is quite arbitrary.

Each niche is occupied by at most one population. This competitive exclusion principle is comparable to Pauli's exclusion principle for fermions (6.2, 6.4). If a population that would fit an occupied niche invades an ecosystem, the result is a conflict ending with the defeat of one of the two populations. Sooner or later, some population will occupy an empty niche.

If the physical or biotic environment changes, a population can adapt by genetically evolving or by finding another niche. If it fails it becomes extinct.

c. In each biotope, the populations depend on each other. Each biotope has its food chains and cycles of inorganic material. Fungi living mainly off dead plants form a kingdom of recyclers.[31] Many bacteria parasitize living plants or animals, which,

conversely, often depend on bacteria. Sometimes the relation is very specific. For instance, a lichen is a characteristic symbiosis of a fungus and a green or blue alga.

The biotic equilibrium in an ecosystem may change by physical causes like climatic circumstances, by biotic causes like the invasion of a new species, or by human intervention. Like a physical equilibrium, the biotic balance has a dynamic character. If an ecosystem gets out of balance, processes start having the disposition to restore equilibrium or to establish a new equilibrium.

Sometimes the ecological equilibrium has a specific character, if two populations are more or less exclusively dependent on each other, for instance in a predator-prey relation. If the prey increases its number, the number of predators will grow as well. But then the number of prey will decrease, causing a decrease of predators. In such an oscillating bistable system, two 'attractors' appear to be active (5.5).

d. Individual organisms are not susceptible to genetic change, but populations are subject to evolutionary change. Besides competition, the driving force of this dynamic development is natural selection, the engine of evolution. With each genotype a phenotype corresponds, the external shape and the functioning of the individual plant. Rather than the genotype, the phenotype determines whether a plant is fit to survive in its environment and to reproduce. Fitness depends on the survival value of an individual plant at short term, and on its reproduction capability and the viability of its offspring.[32] Fitness is a long-term measure for the ability of a population to maintain and reproduce itself.

Natural selection concerns a population and acts on the phenotype. It has the effect that 'the fittest survives', as Herbert Spencer would have it. 'Survival of the fittest' is sometimes called circular.[33] 'That which is fit survives, and that which survives is fit'.[34] This circularity may be caused by the fact that fitness is a primitive, undefinable concept in the theory of evolution.[35] Fitness is not definable, but it is measurable as reproductive success. The circularity may be removed by relating survival to an individual and fitness to its offspring: 'the fit are those who fit their existing environments and whose descendants will fit future environments ... in defining fitness, we are looking for a quantity that will reflect the probability that, after a given lapse of time, the animal will have left descendants'.[36]

Fitness is a quantitatively founded magnitude, lacking a metric. Fitness depends on the reproduction of an individual, and on that of its next of kin. The latter is called 'inclusive fitness', explaining the 'altruistic' behaviour of bees, for instance.

The struggle for life is a process taking place mostly *within* a population, much less *between* related populations (if occupying overlapping niches), and hardly ever between populations of different species.[37]

However, the evolution of a population depends on the environment, including the evolution of other populations. The phenomenon of co-evolution means that several lineages evolve simultaneously and mutually dependently. An example is the evolution

of seed eating birds and seed carrying plants. The plant depends on the birds for the dispersion of its seeds, whereas the birds depend on the plants for their food. Sometimes, the relation is very specific.

Besides co-evolution, biologists distinguish divergent and convergent evolution of homologous respectively analogous properties.[38] *Homology* concerns a characteristic having a common origin. In related species, its function evolved in diverging directions. *Analogy* concerns a characteristic having a corresponding function but a different origin. The emergence of analogous properties is called convergent or parallel evolution. The stings of a cactus are homologous to the leaves of an oak, but analogous to the spines of a hedgehog. The wings of a bird and a bat are homologous to each other, but analogous to the wings of an insect. Light sensitivity or visual power emerged at least forty times independently, hence analogously, but the underlying photoreceptors may have arisen only once, they appear to be homologous.[39]

6.6. The gene pool

The insight that populations are the units of evolution is due to Charles Darwin and Alfred Wallace. It is striking that they could develop their theory of evolution without knowledge of genetics. Besides populations being *subjective* aggregates of living beings, in the biotic evolution *objective* aggregates play a part. These objective aggregates consist of genes. Six years after Darwin's publication of *The origin of species* (1859), Gregor Mendel discovered the laws of heredity. These remained unnoticed until 1900, and only some time later they turned out to be the necessary supplements to the laws for populations.

Some populations reproduce only or mostly asexually (6.7). In section 6.6, I restrict myself to populations forming a reproductive community, a set of organisms reproducing sexually. Within and through a population, genes are transported, increasing and decreasing in number.

a. The genetic identity of each living being is laid down in its genome, the ordered set of genes (6.2). The genes do not operate independent of each other. Usually, a combination of genes determines a characteristic of the organism. In different phases of development, combinations of genes are simultaneously switched on or off. The linear order of the genes is very important. The number of genes is different in different species and may be very large. They are grouped into a relatively small number of chromosomes, each chromosome corresponding to a DNA molecule. The human genome consists of 23 chromosome pairs and about 30,000 genes. The genes take only 5% of the human DNA, the rest is non-coding 'junk-DNA', which (possibly stabilizing) function was not very clear at the end of the twentieth century. A

prokaryote cell has only one chromosome. In eukaryotes, genes occur in the cell nucleus as well as in several organelles, such as the mitochondria. The organelles are considered encapsulated prokaryotes (6.4).

Genes are not subjectively living individuals like organisms, organs, tissues, cells, or even organelles. Richard Dawkins assumes that the 'selfish genes' are the subjects to evolution,[40] but Ernst Mayr objects: 'The geneticists, almost from 1900 on, in a rather reductionist spirit preferred to consider the gene the target of evolution. In the past 25 years [of the twentieth century], however, they have largely returned to the Darwinian view that the individual is the principal target.'[41] Genes have an objective function in the character of a living cell. A genome should nor be identified with the DNA molecules forming its material basis, neither with a gene with a sequence of bases. 'The claim that genetics has been reduced to chemistry after the discovery of DNA, RNA, and certain enzymes cannot be justified ... The essential concepts of genetics, like gene, genotype ... are not chemical concepts at all ...'[42]

Confusion arises from using the same word for a sequence of nucleotides in a DNA molecule and its character, the pattern. In all cells of a plant the DNA molecules have the same pattern, the same character, which is called the plant's genome. Likewise, a gene is not a sequence of nucleotides, nor a particle in a physical or chemical sense, but a pattern of design. The same gene, the same pattern can be found at different positions in a genome, and at the same locus one finds in all cells of a plant the same pair of genes. Each gene is the design for a polypeptide, and the genome is the design of the organism.

The biotic character of the genome is interlaced with the chemical character of DNA molecules. The genome or genotype determines the organism's hereditary constitution. The phenotype is developed according to the design expressed in the genome. Both phenotype and genotype refer to the same individual organism.[43]

Nevertheless, genes have their own objective individuality. In asexual cell division, the genome remains the same. The parent cell transfers its genetic individuality to the daughter cells. In sexual reproduction, objective individual genes are exchanged and a new subjective individual organism emerges.

b. A population is characterized by the possibility to exchange genes and is therefore the carrier of a gene pool. Although the members of the population belong to the same species, they are genetically differentiated. In a diploid cell, a DNA molecule consists of a double helix. At each position or locus there are two genes. These genes may be identical (homozygote) or different (heterozygote). Different genes that can occupy the same locus in different organisms in a population are called *alleles*. Some alleles dominate others. The distribution of the alleles over a population determines their genetic variation, satisfying Gregor Mendel's laws in simple cases. In sexual reproduction, the pairs of genes separate, in order to form new combinations in the new cell (6.3).

At any time, the gene pool is the collection of all genes present in the population. The exchange of alleles in sexual reproduction leads to changes in the frequencies within the gene pool, but does not change the genes themselves. For changing genes, several other mechanisms are known, such as mutation, crossing-over, and polyploidy.

Mutations may have a physical cause (e.g., radioactivity), or a biotic one (e.g., a virus). Mutations are usually indifferent or even lethal, but sometimes enriching. For every gene, they are very rare, but because there are many genes in an individual and even more in a gene pool, they contribute significantly to the variation within a species. Crossing-over means a regrouping of genes over the chromosomes. Polyploidy means that a DNA molecule consists of more than two strings, on each or some loci there are three genes instead of two. Usually, the location of the genes does not change. It is a specific property of the species. Hence, the way genes co-operate is also specific for a species.

A population in which sexual reproduction occurs without constraints is subject to the statistical law of Godfrey Hardy and Wilhelm Weinberg (1908): on a certain locus in the genome the frequency of the alleles in the gene pool in a stable population is constant, generation after generation. Only selective factors and hybridization with another population may disturb the equilibrium.^[44] Populations are hardly ever in equilibrium. The relevance of the law of Hardy and Weinberg is that deviations point to equilibrium disturbing factors. In small populations 'genetic drift' occurs, changes in the gene pool caused by accidental circumstances.

Hybridization between related species or different populations of the same species give rise to a new species or race if three conditions are met. First, the hybrids are fertile. Second, there is a niche available in which the hybrids are better adapted than the original population. Third, the new combination of genes becomes isolated and sufficiently stabilized to survive.^[45] Usually, hybridization is impossible, because the offspring is not viable, or because the offspring is not fertile, or because the offspring has a decreasing fertility in later generations.

Observe that the organisms determine the frequency of the genes in the pool. The character of each gene is realized in DNA. Still, it makes no sense to count the number of DNA molecules in a population, because DNA is found in each cell and most cells have no significance for the gene pool. Even the number of gametes is irrelevant for calculating the gene frequency. The frequency of genes in the pool is the weighed frequency of the *organisms* in the population, being the carriers of the gene concerned. 'A community of interbreeding organisms is, in population genetic terms, a gene pool.'^[46] For instance, if at a certain locus a gene occurs once in 10% of the organisms and twice in 10%, the gene has a frequency of 15% in the gene pool, because each locus contains two alleles. (The complication that on different loci the same gene may occur is left out of consideration in this example.) By natural selection, the frequency of a gene may increase or decrease, depending on the fitness of the organisms in the population.

c. Because of external circumstances, the gene pool may change very fast. Within a few generations, the distribution of a gene pair AB may change from 90% A , 10% B into 10% A , 90% B . This means that a population is able to adapt itself to changes in its habitat, and to increase its chances of survival and reproduction. In a radical environmental change (in particular if a part of the population is isolated), hereditary variation within a species may give rise to the realization of a new species. Hence, adaptation and survival as concepts in the theory of evolution do not concern individual organisms (being genetically stable), but populations. Only populations are capable of genetic change.

Natural selection as such is not a random process,^[47] but it is based on at least two random processes, to wit mutation and sexual reproduction. Which alleles combine in mating rests on chance. The enormous amounts of cells involved in reproduction compensate for the relatively small chance of progress.

e. The phenotype (not the genotype) determines the chance of survival of an organism in its environment. The phenotype is the coherent set of the functions of all parts of the organism, its morphology, physiology, and its ability to reproduce. The genotype generates the phenotype, whereby development and environment factors play an additional but important part. Natural selection advances some phenotypes at the cost of others, leading to changes in the gene pool. Together with changes in the genes themselves, natural selection induces small changes in each generation to accumulate to large changes after a large number of generations.

The received theory of evolution emerged shortly after 1930 from a merger of Charles Darwin's theory of natural selection with genetics and molecular biology. It presupposes that evolution occurs in small steps. Major changes consist of a sequence of small changes. In many cases, this is an acceptable theory. Nevertheless, it would be honest to admit that there is no biological explanation available for the emergence of prokaryotes (about three billion years ago); of eukaryotes (circa one billion years ago); of multicellular organisms (in the Cambrian, circa 550 million years ago); of sexual reproduction; of animals; and of the main animal phyla, plant divisions, classes, and orders. At the end of the twentieth century, the empirical evidence available from fossils and DNA sequencing is not sufficient to arrive at theories withstanding scientific critique.

6.7. Does a species correspond with a character?

In this encyclopaedia, a natural character is defined as a set of laws determining an *ensemble* of possibilities besides a *class* of individuals (1.2). A class and an ensemble are not restricted in number, space, and time. They do not change in the course of time and do not differ at different places. A population is not a class but a

collection. Hence, it does not correspond to a character. The question of whether a *species* corresponds to a character is more difficult to answer. 'There is probably no other concept in biology that has remained so consistently controversial as the species concept.'^[48] Philosophers interpreting the concept of a natural kind in an essentialist way rightly observe that a biotic species does not conform to that concept. However, the idea that a character is not an essence but a set of laws sheds a different light on the concept of a species. The main problem appears to be that insufficient knowledge is available of the laws determining species. Instead, one investigates the much better accessible subject and object side of these unknown laws.

Generally speaking, biologists have a realist view on sets, considering a species to be a natural set. Each living being belongs to a species, classified according to a variety of practical criteria, which do not always yield identical results. Besides, there are quite a few theoretical definitions of a species.^[49] The distinction between operational criteria used in practice and theoretical definitions is not always sharp. Practice and theory are mutually dependent. However, not distinguishing them gives rise to many misunderstandings. '... the species problem results from confusing the concept of a species itself with the operations and evidence that are used to put that concept in practice.'^[50]

Criteria to distinguish species from each other are grouped into genealogical (or phylogenetic), structural, and ecological criteria. This corresponds more or less to a division according to primary, secondary, and tertiary characteristics as defined in chapter 1.

Species can be distinguished because they show distinctive, specific properties. These are regular, therefore lawlike. This is not merely interesting for biologists. In particular in sexual relationships, animals are able to distinguish other living beings from those of their own kind.

Primary criteria to distinguish species are genealogical. The biological taxonomy is based on empirical or theoretically established lineages. A population is a segment of a lineage. A taxon (for instance, a species, genus, family, order, or phylum) is defined as a set of organisms having a common ancestry. A monophyletic taxon or clade comprises all and only organisms having a common ancestry. Birds and crocodiles are monophyletic, both apart and together. A set of organisms lacking a common ancestry is called polyphyletic. Such a set, like that of all winged animals, is not suited for taxonomy. A taxon consisting of some but not all descendants of a common ancestor is called paraphyletic. For instance, reptiles have a common ancestry, but they share it with the birds, which are not reptiles. Opinions differ about the usefulness of paraphyletic taxa.

The biological taxonomy clearly presupposes genetic relations to constitute a general biotic relation frame. '... the general lineage concept is a quintessential biological species concept: inanimate objects don't form lineages.'^[51] Descent providing the primary, genealogical criterion for a species has two important consequences.

The first consequence is seldom explicitly mentioned, but always accepted. It is the assumption that an individual living being belongs to the same species throughout its life. (It may change of population, e.g., by migration.) This means that species characteristics cannot be exclusively morphological. In particular the shape of multicellular fungi, plants, and animals changes dramatically during various phases of life. The metamorphosis of a caterpillar into a butterfly is by no means an exception. The application of similarities and differences in taxonomy has to take into account the various phases of life of developing individuals.

Second, as a rule each living being belongs to the same species as its direct descendants and parents. Therefore the dimorphism of male and female specimens does not lead to a classification into different species. A very rare exception to this rule occurs at the realization of a new species. A minimal theoretical definition says that a species necessarily corresponds to a lineage, starting at the moment it splits off from an earlier existing species, and ending at its extinction.^[52]

If this minimal definition would be sufficient as well as necessary, a species would be a collection, like a population bounded in number, space, and time. But this definition cannot be sufficient, because it leaves utterly unclear what the splitting of a species means. Branching alone is not a sufficient criterion, because each lineage branches (an organism has various descendants, and in sexual replication each organism has two parents, four grandparents, etc.). According to the primary criterion alone, the assumption that all organisms are genetically related would mean that either all organisms belong to the same species, or each sexual reproduction leads to a new species. Hence, additional secondary and perhaps tertiary criteria are needed to make clear, which kind of branching leads to a new species.^[53] 'In effect, the alternative species definitions are conjunctive definitions. All definitions have a common primary necessary property – being a segment of a population-level lineage – but each has a different secondary property – reproductive isolation, occupation of a distinct adaptive zone, monophyly, and so on.'^[54]

The most practical criteria are structural. It concerns similarities and differences based on the DNA-structure (the genotype), besides the shape (morphology), and processes (physiology, development), making up the phenotype. In DNA and RNA research, biologists look at similarities and differences with respect to various genes and their sequences, taking into account the locus where they occur. The comparison of genes at different loci does not always give the same results. Hence people should be cautious with drawing conclusions. It should be observed that DNA and RNA research is usually only possible with respect to living or well-conserved cells and only establishes more or less contemporary relations (in a geological sense: 'contemporary' may concern millions of years). This also applies to other characteristics that cannot be fossilized, like behaviour. Non-contemporary similarities and differences are mostly restricted to morphological ones. For the agreement between various related species, homologies are very important (6.6).

Many biologists accept as a decisive distinction between species the existence of a reproductive gap between populations. 'A species is a reproductive community of

populations (reproductively isolated from others) that occupies a specific niche in nature.'[55] Within a species, individuals can mate fruitfully with each other, whereas individuals of different species cannot. Ernst Mayr mentions three aspects of a biotic species. 'The first is to envision species not as types but as populations (or groups of populations), that is, to shift from essentialism to population thinking. The second is to define species not in terms of degree of difference but by distinctness, that is, by the reproductive gap. And third, to define species not by intrinsic properties but by their relation to other co-existing species, a relation expressed both behaviourally (non-interbreeding) and ecologically (not fatally competing).'[56] 'The word "species" ... designates a relational concept'.[57] According to this definition, horses and donkeys belong to different species. A horse and a donkey are able to mate, but their offspring, the mules, are not fertile. The mentioning of populations is relevant. The reproduction gap does not concern individuals but populations.

Sometimes, a population *A* belongs to the same species as population *B*, *B* ditto with *C*, but *C* does not with respect to *A*. [58] Hence, the concept of a species according to this criterion is not always transitive. The possibility to mate and having fertile descendants is only relevant for *simultaneously* living members of a population. Hence it serves as a secondary addition to the primary genealogical criterion, stating that organisms living long after each other (and therefore unable to mate) may belong to the same species. Taking this into account, the mentioned lack of transitivity can be explained by assuming that one of the populations concerned is in the process of branching off. After some time, either *A* or *C* may become to belong to an independent species.

The reproduction gap is in many cases a suitable criterion, but not always. First, some species only reproduce asexually. This is not an exception, for they include the prokaryotes (the only organisms during three-quarters of the history of life on earth). [59] Second, many organisms that experts rank to different species are able to fertilize each other. Hybrid populations are more frequent in plants than in animals. The reproductive gap is in animals more pronounced than in plants, because of the animals' mating behaviour and the corresponding sexual dimorphism. Mating behaviour leads to the 'recognition species concept'. [60]

A tertiary criterion concerns the disposition of a species to find a suitable niche or adaptive zone (6.5). How organisms adapt to their environment leads to the formulation of ecological criteria to distinguish species. This is a relational criterion too, for adaptation does not only concern physical (e.g., climatic) circumstances, but in particular the competition with individuals of the same or of a different species.

Biologists and monist biophilosophers look after a universal concept of a species. According to Hull, the concept of a species ought to be universal (applicable to all organisms), practical in use, and theoretically significant. [61] He observes that monists are usually realists, pluralists being nominalists. [62] Supposing that a species corresponds to a character, it should be primarily biotically qualified. No difference of opinion is to be expected on that account. But what should be its secondary characteristic? Considering the analysis in section 6.4, for prokaryotes the quantitative

relation frame comes to mind (cell division); for unicellular or colonial eukaryotes the spatial frame (morphological shape and coherence); for differentiated plants the kinetic frame (physiology and development); finally, for sexually specialized plants and animals the physical relation frame (the reproductive gap). A species can only be a universal biotic character if the concept of a species is differentiated with respect to secondary and tertiary characteristics. For instance, the secondary criterion based on the reproductive gap is only applicable to sexually reproducing organisms. The pluralistic concept of a species finds its origin in the fact that all secondary and tertiary criteria are restrictively applicable, whereas the universal primary criterion is necessary but not sufficient.[63] Likewise, the physical concept of natural kinds is not universal. For quantitatively, spatially, and kinetically founded characters, different secondary criteria apply (chapter 5).

Some philosophers assume that species are comparable with organisms and they consider a species to be a biotic individual.[64] 'When species are supposed to be the things that evolve, they fit more naturally in the category *individual* (or historical entity) than the category *class* (or *kind*).'[65] Hull assumes a duality: 'Classes are spatio-temporally unrestricted, whereas individuals are spatio-temporally localized and connected. Given this fairly traditional distinction, we argued that species are more like individuals than classes'[66] Clearly, Hull does not distinguish between aggregates and individuals.

A species comes into being by branching off from another species, and it decays at extinction. Species change during their existence. It is true that these processes depend entirely on the replication of the organisms that are part of the species, but that applies to multicellular organisms as well, whose development and growth depend on the reproduction of their cells.

Organisms belonging more or less simultaneously to the same species form a *population*. Usually a population is a geographically isolated subset of a *lineage*, a set of organisms having the same ancestry. Both populations and lineages are temporal collections of individuals, not timeless classes. They are aggregates as well, because their members are genetically related. However, an aggregate is not always an individual, and it is always a set of individuals. If considered as a lineage or population (or a set of populations), a species is a temporal collection of individual organisms, subject to biotic laws. I shall not contest this vision that stresses the subject side of a species. But it does not answer the question of whether a species has a law side as well, corresponding with a character.

Both lineages and populations are products of a biotic dynamic evolution. Natural selection, genetic drift, and ecological circumstances explain *how* lineages emerge, change, and disappear. Geographic isolation explains the existence of various populations belonging to the same species. But natural selection, genetic drift, or geographic isolation does not explain *why* a group of living beings is viable in the circumstances constituting an adaptive zone. Unavoidably, such an explanation takes its starting point from law statements, whether hypothetical or confirmed. These propositions may very well include the supposition that a lineage and its populations

are spatio-temporal subsets of a timeless class, without violating the received facts and theories of evolution and genetics. The character of this class determines an ensemble of possibilities, partly realized in the individual variation occurring in a population.

No more than species, the chemical elements have been realized from the beginning of the universe. Only after the universe was cooled down sufficiently, protons and electrons could form hydrogen and helium. Only after the formation of stars, hydrogen and helium nuclei fused to become heavier nuclei. Nuclear physics provides a quite reliable picture of this chemical evolution. Doubtless, each isotope satisfies a set of laws constituting a character. I believe that the same applies to biotic species, although the complexity of organisms makes it far more difficult to state in any detail which laws constitute a biotic character. Based on an essentialist interpretation, Ernst Mayr turns down the analogy of the species concept in biology with that of mineralogy or chemistry: 'For a species name in mineralogy is on the whole a class name, defined in terms of a set of properties essential for membership in the class.'[\[67\]](#)

The crossing of a barrier between two species has an analogy in the well-known phenomenon of *tunnelling* in quantum physics (5.7). An energy barrier usually separates a radioactive nucleus from a more stable nucleus. This barrier is higher than the energy available to cross it. According to classical mechanics, a nucleus could never cross such a barrier, but quantum physics proves that there is a finite (even if small) probability that the nucleus overtakes the barrier, like a car passes a mountain through a tunnel. A similar event occurs in the formation of molecules in a chemical reaction. In this case, the possibility to overtake the energy barrier depends on external circumstances like the temperature. The presence of a catalyst may lower the energy barrier. In biochemical processes enzymes have a comparable function. The possibility that an individual physical or chemical thing changes of character is therefore a fact, both theoretically and experimentally established. Moreover, in all chemical reactions molecules change of character, dependent on circumstances like temperature.

Similarly, at the realization of a new species, circumstances like climate changes may enhance or diminish the probability of overcoming one or more constraints. A small, geographically isolated population will do that easier than a large, widely dispersed population. Since 1972, biology knows the theory of 'punctuated equilibrium'. From paleontological sources, Niles Eldredge and Stephen Gould derived that in a relatively short time (compared to much longer periods of stable equilibrium), a major transition from one species to another may occur.[\[68\]](#) Such a transition would take 50,000 years or more, whereas a stable period may last millions of years.[\[69\]](#)

Quantum physics explains the *transition* from one character to the other by tunnelling, but tunneling does not explain the *existence* of the characters concerned. Natural selection explains why constraints can be overcome, not why there are constraints, or which types of constraints are operative. Natural selection explains changes within species and from one species to the other, but not why there are species, and which species exist. On the contrary, the existence of species is a condition for the action of

natural selection. Populations change within a species, and sometimes they migrate from one species to another one, and its motor, its dynamic force, is natural selection. However, natural selection does not explain everything. The success of natural selection is only explicable by assuming that a population after adaptation is in a more stable equilibrium with its environment than before. What is stable or better adapted, why the chances of survival of an organism increase by a change induced by natural selection, cannot be explained by natural selection itself. Natural selection explains why a population changes its gene pool, but it does not explain why the new situation is more viable. To explain this requires research into the circumstances in which the populations live and into the characters that determine the species.

On the one side, the standard theories about evolution, genetics, ecology, and molecular biology do not exclude the possibility that each species corresponds to a character, a set of laws defining an ensemble of possibilities, sometimes (and never exhaustively) realized by a population of organisms. After all, 'by far the commonest fate of every species is either to persist through time without evolving further or to become extinct.'^[70]

On the other hand, it cannot be proved that a species corresponds to a character. That would only be empirically demonstrable. The idea that an empirical species is a subset of a class subject to a specific set of laws can only be confirmed by pointing out such laws. Evolutionists have a tendency to deny the existence of biotic laws.^[71] In contrast, Paul Griffiths asserts that there are laws valid for taxonomy.^[72] Michael Ruse stresses that biology needs laws no less than the inorganic sciences. He mentions Gregor Mendel's laws as an example.^[73] And Marc Ereshefsky observes that '... there may be universal generalizations whose predicates are the names of *types* of basal taxonomic units ... So though no laws exist about particular species taxa, there may very well be laws about types of species taxa.'^[74] For instance, both genetics and development biology look for lawful conditions concerning the constitution of genes and chromosomes determining the phenotype of a viable organism belonging to a species. That is because the biotic expression of a character is a natural design, the genome, objectively laid down in the species-specific DNA.

Natural selection may be considered a random push for the dynamic development of populations of living beings. This development also requires the specific lawful pull of the species concerned.

Should we not consider the ascription of an unchangeable and lawful character to species a relapse into essentialism?^[75] Ernst Mayr observes that in Carl Linnaeus' taxonomy the genera are defined in an essentialist way.^[76] He quotes from Linnaeus' *Philosophia Botanica* (1751): 'The 'character' is the definition of the genus, it is threefold: the *factitious*, the *essential*, and the *natural*. The *generic character* is the same as the *definition of the genus* ... The *essential definition* attributes to the genus to which it applies a characteristic which is very particularly restricted to it, and which is special. The essential definition [character] distinguishes, by means of a unique idea, each genus from its neighbours in the same natural order.'^[77]

Essentialism is a theory ascribing *a priori* an autonomous existence to plants, animals, and other organisms. Their essence is established on rational grounds, preceding empirical experience. Essentialism presupposes the possibility to formulate necessary and sufficient conditions for the existence of each species. The conditions for any species should be independent of the conditions for any other species.[78] This view differs widely from the idea of a character being a specific set of laws. With respect to the subject side, as far as essentialism excludes evolution, the theory of characters is by no means essentialist.

According to Aristotelian essentialism, each species would be autonomous. Biologists and philosophers seem to assume that this paradigm is still applicable to physics and chemistry. But physical things can only exist in interaction with other things, and the actual realization of physically qualified characters is only possible if circumstances permit. For instance, in the centre of the sun no molecules can exist, but we can only say that assuming that the laws which determine the possible existence of molecules are valid within the sun as much as elsewhere. The astrophysical and chemical theories of evolution assume that physical things emerged gradually, analogous to organisms in the biotic evolution. Nevertheless, it is generally accepted that particles, atoms, molecules, and crystals are subject to laws that are everywhere and always valid.

Physical and chemical things can only exist in interaction with each other in suitable circumstances. Similarly, living organisms can only exist in genetic relations with other organisms, permitting the circumstances. Each living organism would perish in absence of other living beings, and no organism can survive in an environment that does not provide a suitable niche.

My reasons to consider a species to be a character are *a posteriori*, based on scientific arguments open to empirical research. It is a hypothesis, like any other scientific assumption open to discussion. And it is a hypothesis leaving room for the evolution of a population within a species as well as from one species to another one. It is a hypothesis fully acknowledging Charles Darwin's great discovery of natural selection. Moreover, this hypothesis recognizes the importance of environmental circumstances both determining possibilities and their realization. The laws are not the only conditions for existence. Physical and ecological circumstances are conditions as well. The realization of species can only occur in a certain order, with relatively small transitions. In this respect, too, the evolution of species does not differ from the evolution of chemical elements.

Although essentialists are able to take circumstances into account, the theory of characters moves ahead. The possibilities offered by a character are not merely realizable if the circumstances permit, but the ecological laws are partly the same as the laws constituting the character of a species. The laws forming a character for one species are not separated from the laws forming the character of another species, or from the laws determining biotic processes. Essentialism supposes that each species can be defined independent of any other species.

It is undeniable that my hypothesis runs counter to the kind of evolutionism that denies the existence of constant laws. From the above discussion it will be clear that I do not criticize Darwin's *theory* and its synthesis with genetics and molecular biology. By natural selection, the theory of evolution explains the actual dynamic process of becoming and the evolution of populations. I believe that this theory does not contradict the view that species correspond to unchangeable characters and their ensembles. On the contrary, I believe that the facts corroborate the proposed model better than a radical evolutionism denying the existence of laws. The hypothesis that unchangeable laws dominate the species can be investigated on *empirical* grounds. This discussion belongs to the competence of empirical science.

The answer to the question of whether a species corresponds to a character does not depend on the acceptance or rejection of the belief that characters – not only biotic species – consist of laws given by God. The empirical approach that I advocate is at variance with the creationist view assuming *a priori* that the species are unchangeable, rejecting any theory of evolution. Creationism uses the Bible as a source of scientific knowledge *preceding* and superseding scientific research. It contradicts the view that the problem of whether species correspond to constant characters can only be solved *a posteriori*, based on scientific research.

For the time being, I am inclined to conclude that a species at the law side corresponds with a biotically qualified character, an unchangeable set of laws. The least one can say is that the recognition of a species or a higher taxonomical unit requires an insight into the regularities which make an organism to belong to that category. At the subject side, a species is realized by a lineage, an aggregate of individual organisms, hence with a collection, bounded in number, space, and time.

Evolution means the subjective realization of species. Natural selection is its motor and explains *how* species are realized. Whether a species is *realizable* at a certain place and time depends on the character of the species; on the preceding realization of a related species (on which natural selection acts); on the presence of other species (the ecological environment); and on physical circumstances like the climate (the physical environment).

I have no intention to suggest that the biotic evolution is comparable to the astrophysical and chemical evolution in all respects. I conceive of each evolution as a realization of possibilities and dispositions. But the way by which this occurs is strongly different. For physical and chemical things and events, interaction is decisive, including circumstances like temperature and the availability of matter and energy. The biotic evolution depends on sexual and asexual reproduction, with the possibility of variation and natural selection.

Another difference concerns the reproducibility of evolution. The physical evolution of the chemical elements and of molecules repeats itself in each star and each stellar system. In contrast, it is often stated that the biotic evolution is unique and cannot be repeated. It may be better to assert that the actual course of the biotic evolution is far more improbable than that of the physical and chemical ones. Comparable

circumstances – a condition for recapitulation – never or hardly ever occur in living nature. In particle accelerators the astrophysical evolution is copied, the chemical industry produces artificial materials, agriculture improves races, in laboratories new species are cultivated, and the bio-industry manipulates genes. All this would be difficult to explain if one loses sight of the distinction between law and subject.

As a character, a biotic design is a set of laws, but for a scientist this does no longer imply a divine designer.[79] Whereas this does not solve the question of the origin of the natural laws, natural science became liberated from too naive views about the observability of divine interventions in empirical reality.

Essentialism survived longest in the plant and animal taxonomy. Until the middle of the twentieth century, this considered the system of species, genera, families, classes, orders, and phyla or divisions to be a logical classification. In this classification, each category was characterized by one or more essential properties. 'The essentialist species concept ... postulated four species characteristics: (1) species consist of similar individuals sharing in the same essence; (2) each species is separated from all others by a sharp discontinuity; (3) each species is constant through time; and (4) there are severe limitations to the possible variation of any one species.'[80]

Biological essentialism is not a remains of the Middle Ages, but a fruit of the Renaissance. From John Ray to Carl Linnaeus, many realistic naturalists accepted the existence of unchangeable species. Besides, other biologists had a nominalist view of species.[81]

Ray and Linnaeus were more (Aristotelian) realists than (Platonic) idealists. Ernst Mayr ascribes the influence of essentialism to Plato.[82] 'Without questioning the importance of Plato for the history of philosophy, I must say that for biology he was a disaster.'[83] Mayr shows more respect for Aristotle, who indeed has done epoch-making work for biology.[84] However, Aristotle was an essentialist no less than Plato was.

The difficulty that some philosophers have with the modern concept of a species can be reduced to a conscious or subconscious allegiance to an essentialist view. The difficulty that some biologists have with the idea of natural law is their abhorrence of essentialism.[85] Therefore, it is important to distinguish essence from lawfulness. The 'essential' (necessary and sufficient) properties do not determine a character. Rather, the laws constituting a character determine the objective properties of the things or processes concerned. 'Essentialism with respect to species is the claim that for each species there is a nontrivial set of properties of individual organisms that is central to and distinctive of them or even individually necessary and jointly sufficient for membership in that species.'[86] Paul Griffiths contests the view that there are no natural laws (in the form of generalizations allowing of counterfactuals) concerning taxonomy.[87] Definition of a natural kind by properties may have a place in natural history, but not in a modern scientific analysis based on theories, in which laws dominate, not properties. The identification of a class by necessary and sufficient conditions is a remnant of rationalistic essentialism, These properties, represented in

an ensemble, may display such a large statistical variation that necessary and sufficient properties are hard to find.[88] Moreover, the laws and properties do not determine essences but relations.

A second reason why some biologists are wary of the idea of natural law is that they (like many philosophers) have a physicalist view of laws. Rightly, they observe that the physical and chemical model of a natural law is not applicable to biology.[89] To the nineteenth-century physicalist idea of law belonged determinism and causality. However, determinism is past, and causality is no longer identified with law conformity but is considered a physical relation.

The theory of evolution is considered a *narrative* about the history of life, rather than a theory about processes governed by natural laws. 'Laws and experiments are inappropriate techniques for the explication of such events. Instead, one constructs a historical narrative, consisting of a tentative construction of the particular scenario that led to the events one is trying to explain.'[90] But probably biologists will not deny that their work consists of finding order in living nature.[91] 'Biology is not characterized by the absence of laws; it has generalizations of the strength, universality, and scope of Newton's laws: the principles of the theory of natural selection, for instance.'[92] About M.B. Williams' axiomatization of the theory of evolution,[93] Alexander Rosenberg observes: 'None of the axioms is expressed in terms that restrict it to any particular spatio-temporal region. If the theory is true, it is true everywhere and always. If there ever were, or are now, or ever will be biological entities that satisfy the parent-of relation, anywhere in the universe, then they will evolve in accordance with this theory (or else the theory is false).'types of basal taxonomic units ... So though no laws exist about particular species taxa, there may very well be laws about types of species taxa.'[99]

The theory of evolution would not exist without the supposition that the laws for life, that are now empirically discovered, held millions of years ago as well. The question of whether other planets host living organisms can only arise if it is assumed that these laws hold there, too.[100]

A third reason may be the assumption that a law only deserves the status of natural law, if it holds universally and is expressible in a mathematical formula. A mathematical formulation may enlarge the scope of a law statement. Yet the idea of natural law does not imply that it has necessarily a mathematical form. Neither should a law apply to all physical things, plants, and animals. Every regularity, every recurrent

design or pattern, and every invariant property is to be considered lawful. In particular each character expresses its own specific law conformity. In the theory of evolution biologists apply whatever patterns they discover in the present to events in the past. Hence they implicitly acknowledge the persistence of natural laws, also in the field of biology.

Anyhow, Charles Darwin was not wary of natural laws. At the end of his *On the origin of species* he wrote:

'It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us. These laws, taken in the largest sense, being Growth with Reproduction; Inheritance which is almost implied by reproduction; Variability from the indirect and direct action of the external conditions of life, and from use and disuse; a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing Divergence of Character and the Extinction of less-improved forms.'^[101]

[1] Rudwick 2005; 2008.

[2] Mayr 1982, 56.

[3] Mayr 1982, 629.

[4] Jacob 1970, 4.

[5] Rensch 1968, 35.

[6] Raff 1996, chapter 8.

[7] Farley 1974; Bowler 1989.

[8] Ruse 1973, 118-121.

[9] Monod 1970, 102-103.

[10] Rosenberg 1985, 136-152.

[11] Rosenberg 1985, 137-138.

[12] Panchen 1992, chapter 9.

[13] Dawkins 1983, 16.

[14] Kitcher 1993, 270.

[15] Rosenberg 1985, 38-43.

[16] Hull 1974, 15-19.

[17] McFarland 1999, 27-29.

[18] Dawkins 1986, 295-296; McFarland 1999, 27.

[19] Raff 1996, chapter 10.

[20] Raff 1996, 27.

[21] Griffiths, Gray 1994.

[22] Griffiths, Gray 1994.

[23] Raff 1996, 260.

[24] Raff 1996, 23.

- [25] Mayr 1982, 140, 244; Margulis, Schwartz 1982, 5-11; Ruse 1982, 169-171.
- [26] Margulis, Schwartz 1982.
- [27] Greulach, Adams 1962, 28.
- [28] de Queiroz 1999, 53-54.
- [29] Darwin 1859, chapter 3.
- [30] Mayr 1982, 602.
- [31] Purves et al. 1998, chapter 28: 'Fungi: A kingdom of recyclers.'
- [32] McFarland 1999, 72.
- [33] Popper 1974, 137.
- [34] Dampier 1929, 319.
- [35] Rosenberg 1985, chapter 6; Sober 1993, 69-73.
- [36] McFarland 1999, 78.
- [37] Darwin 1859, chapter 4.
- [38] Panchen 1992, chapter 4.
- [39] Mayr 1982, 611; Raff 1996, 375-382.
- [40] Dawkins 1976
- [41] Mayr 2000, 68-69; Sober 1993, chapter 4.
- [42] Mayr 1982, 62.
- [43] Ruse 1982, 21, 30, 200-207.
- [44] Hull 1974, 57-58; Ridley 1993, 87-92, 131-132.
- [45] Ridley 1993, 42-43.
- [46] Ridley 1993, 387.
- [47] Dawkins 1986, 43, 62.
- [48] Mayr 1982, 251. On the biological species concept, see Mayr 1982, chapter 6; Rosenberg 1985, chapter 7; Ereshefsky 1992; Ridley 1993, chapter 15; Wilson (ed.) 1999.
- [49] Panchen 1992, 337-338 mentions seven species concepts, others count more than twenty, see Hull 1999.
- [50] de Queiroz, 1999, 64.
- [51] de Queiroz 1999, 77.
- [52] de Queiroz 1999; Mishler, Brandon, 1987, 310.
- [53] Ereshefsky 1992, 350.
- [54] de Queiroz 1999, 60, 63.
- [55] Mayr 1982, 273.
- [56] Mayr 1982, 272.
- [57] Mayr 1982, 286.
- [58] Ridley 1993, 40-42.
- [59] Nanney 1999.
- [60] Ridley 1993, 392-393.
- [61] Hull 1999, 38-39
- [62] Hull 1999, 25.
- [63] Dupré 1999.
- [64] Rosenberg 1985, 204-212; Ridley 1993, 403-404.:
- [65] Hull 1999, 32. For a criticism, see Mishler, Brandon, 1987; de Queiroz, Donoghue, 1988; Sober 1993, 149-159; de Queiroz 1999, 67-68.
- [66] Hull 1999, 32-33.
- [67] Mayr 1982, 251.
- [68] Gould, Vrba 1982; Ridley 1993, chapter 19; Strauss 2009, 487-496.
- [69] Stebbins 1982, 16-21.
- [70] Stebbins 1982, 23.
- [71] Dawkins 1986, 10-15.
- [72] Griffiths 1966.
- [73] Ruse 1973, 24-31.

- [74] Ereshefsky 1992, 360
 - [75] Toulmin, Goodfield 1965.
 - [76] Mayr 1982, 175-177.
 - [77] Mayr 1982, 176.
 - [78] Sober 1993, 145-149; Hull 1999, 33; Wilson 1999, 188.
 - [79] Dawkins 1986, chapter 1.
 - [80] Mayr 1982, 260:
 - [81] Toulmin, Goodfield 1965, chapter 8; Panchen 1992, chapter 6.
 - [82] Mayr 1982, 38, 87, 304-305.
 - [83] Mayr 1982, 87.
 - [84] Mayr 1982, 87-91, 149-154.
 - [85] Stafleu 2019, 6.8.
 - [86] Rosenberg 1985, 188; Hull 1999, 33; Wilson 1999, 188.
 - [87] Griffiths 1999.
 - [88] Hull 1974, 47; Rosenberg 1985, 190-191.
 - [89] Hull 1974, 49; Mayr 1982, 37-43, 846.
 - [90] Mayr 2000, 68:
 - [91] Rosenberg 1985, 122-126, 211, 219.
 - [92] Rosenberg 1985, 211.
 - [93] Rosenberg 1985, 136-152, Hull 1974, 64-66.
 - [94] Rosenberg 1985, 152.
 - [95] Rosenberg 1985, 219.
 - [96] Dawkins 1986, 10-15.
 - [97] Ruse 1973, 24-31.
 - [98] Rensch 1968; Griffiths 1999.
 - [99] Ereshefsky 1992, 360; Hull 1974, chapter 3.
 - [100] Dawkins 1983.
 - [101] Darwin 1859, 459.
-

-

Chapter 7

Characteristics of animal behaviour

7.1. The primary characteristic of animals

7.2. Secondary characteristics of animals

7.3. Control processes

7.4. Controlled processes

7.5. Goal-directed behaviour

7.1. The primary characteristic of animals

The sixth and final relation frame for characters of natural things and processes concerns animals and their behaviour. This, too, is a typical twentieth-century subject. In the United States and the Soviet Union, especially positivistically oriented behaviourists were concerned with laboratory research of the behaviour of animals, in particular with their learning ability. Later on, in Europe ethology emerged, investigating animal behaviour in natural circumstances. I shall not discuss human psychology, which witnessed important developments during the twentieth century as well. Besides ethology and animal psychology, neurology is an important source of information for chapter 7.

Section 7.1 argues that animals are characterized by goal directed behaviour, implying the establishment of informative connections and control. Section 7.2 discusses the secondary characteristic of animals. Section 7.3 deals with the psychical processing of information, section 7.4 with controlled processes and section 7.5 with their goals.

A psychical character is a pattern of behaviour or a program, a lawful prescription. This is a scheme of fixed processes laid down in detail, with their causal connections leading to a specified goal. Behaviour has an organic basis in the nervous system and in the endocrine system (7.2), and a physical and chemical basis in signals and their processing (7.3).

No more than the preceding chapter, this inventory of animal behaviour contains anything new from a scientific point of view. Only the *ordering* is uncommon, as it derives from the philosophy of dynamic development. The proposed ordering intends to demonstrate that the characters studied in mathematics and science do not merely show similarities. Rather, the natural characters are mutually interlaced and tuned to each other.

For the psychic subject-subject relation, I suggest the ability to make informative and goal directed connections. Psychic control influences organic, physical, chemical, kinetic, spatial, and quantitative relations, but it does not mean their abolishment. On the contrary, each new order means an enrichment of the preceding ones. Physical interactions allow of more (and more varied) motions than the kinetic relation frame alone does. Ever more kinds of motion are possible in the organic and animal worlds. The number of organic compounds of atoms and molecules is much larger than the number of inorganic ones. Organic variation, integration, and differentiation are in the animal kingdom more evolved than in the kingdom of plants. Each new order opens and enriches the preceding ones. By making informative connections, an animal functions organically better than a plant. For this purpose, an animal applies internally

its nervous system and its hormones, and externally its behaviour, sustained by its senses and motor organs.

Animals differ in important respects from plants, fungi, and bacteria. No doubt, they constitute a separate kingdom. The theory of evolution assumes that animals did not evolve from differentiated multicellular plants, but from unicellular protozoans.

According to a modern definition, *animalia* are multicellular: 'An organism is an animal if it is a multicellular heterotroph with ingestive metabolism, passes through an embryonic stage called a blastula, and has an extracellular matrix containing collagen.'^[1] Within the kingdom of the *protista* (the set consisting of all eukaryotes that do not belong to the *animalia*, *plantae*, or *fungi*), the unicellular protozoans like flagellates and amoebas do not form a well-defined group. The *animalia* probably form a monophyletic lineage, which would not be the case if the protozoans were included. Therefore, some biologists do not consider the protozoans to be animals, but others do.

In the evolutionary order, the plants may have emerged after the animals. The first fossils of multicellular animals occur in older layers than those of differentiated plants. Fungi are genetically more related to animals than to plants. Possibly, the plants branched off from the line that became the animal kingdom. If so, this branching is characterized by the encapsulation of prokaryotes evolving into chloroplasts. The distinctive property of green plants is their ability of photosynthesis, which is completely absent in animals and fungi. Another difference is the mobility of most animals in contrast to the sedentary nature of most plants. Animals lack the open growth system of many plants, the presence of growth points of undifferentiated cells, from which periodically new organs like roots or leaves grow. After a juvenile period of development, an animal becomes an adult and does not form new organs. Animal organs are much more specialized than plant organs.

If asked to state the difference, a biologist may answer that plants are autotroph and animals heterotroph. Plants achieve their food directly from their physical and chemical environment, whereas animals depend partly on plants for their food supply. David MacFarland divides organisms into producers, consumers, and decomposers. Plants produce chemical energy from solar energy. Animals consume plants or plant eaters. Fungi and bacteria decompose plant and animal remains to materials useful for plants.^[2] However, fungi too depend on plants or their remains, and some plants need bacteria for the assimilation of nitrogen. Apart from that, this criterion is not very satisfactory, because it does not touch the primary, qualifying relation frames of plants and animals. It seems to be inspired by a world view reducing everything biological to physical and chemical processes. This view stresses the energy balance, metabolism, and the production of enzymes out of proportion. I believe the distinction between autotroph and heterotroph to be secondary.

Animals are primarily distinguished by their behaviour. A relational philosophy does not look for reductionist or essentialist definitions, but for qualifying relations. The most typical biotic property of all living beings, whether bacteria, fungi, plants, or

animals, is the genetic relation, between organisms and between their parts, as discussed in chapter 6. Superposed on this relation, animals have psychic relations between their organs by means of their nervous system, and mutually by means of their observable behaviour. In part, this behaviour is genetically determined; in part, it is adaptable. Obviously, in particular species specific behaviour is genetically determined, because species are biotically qualified characters, if not aggregates (6.7). Different animal species can be distinguished because of their genetically determined behaviour. More differentiated animals have a complex nervous system with a larger capacity for learning and more freedom of action, than simpler animals have.

The taxonomy of the animal kingdom is mostly based on descent and on morphological and physiological similarities and differences. Its methodology hardly differs from that of the plant taxonomy. But there are examples of species that can only be distinguished because of their behaviour. When a new animal species is realized, a change of behaviour precedes changes in morphology or physiology.[3] This means that controlled behaviour plays a leading part in the formation of a new animal species. Because of the multiformity of species specific behaviour, there are far more animal species than plant species, and much less hybrids.

However, animals have a lot in common with plants and fungi, too, because their psychic character is interlaced with biotic characters. Conversely, as a tertiary characteristic some plants are tuned to animal behaviour. Flowering and fruit bearing plants have a symbiotic relation with insects transferring pollen, or with birds and mammals eating fruits and distributing indigestible seeds.

The psychically qualified character of an animal comes to the fore in its body plan (morphology) and body functions (physiology), being predisposed for its behaviour. For this purpose, animals have organs like the nervous system, hormonal glands, and sense organs, that plants and fungi lack. Animals differ from plants because of their sensitivity for each other, their aptitude to observe the environment, and their ability to learn. They are sensitive to internal stimuli and external signals. Sometimes, also plants react to external influences like sunlight. But they lack special organs for this purpose and they are not sensitive to each other or to signals. In a multicellular plant, a combination of such reactions may give rise to organized motions, for instance flowers turning to the sun. Animal movements are not primarily organized but controlled. However, control does not replace organization, but superposes it.

Each plant cell reacts to its direct surroundings, to neighbouring cells or the physical and biotic environment. A plant cell only exerts action by contact, through its membranes. Neighbouring animal cells are less rigidly connected than plant cells. There are more intercellular cavities. Animal cells and organs are informatively linked by neurons, capable of bridging quite long distances. An animal exerts action at a distance within its environment as well, by means of its sense organs, mobility, and activity.

A physical system is stable if its internal interactions are stronger than its external interactions. An organism derives its stability from maintaining its genetic identity during its lifetime (6.3). Only sexual reproduction leads to a new genetic identity. For the stability of an animal, internal control by the nervous and hormonal systems is more important than the animal's external behaviour.

Informative goal-directed connections express the universal psychic subject-subject relation. Animals receive information from their environment, in particular from other animals, and they react upon it. *Mutatis mutandis*, this also applies to animal organs. Both internally and externally, an animal may be characterized as an information processor. Provisionally, I propose the following projections on the five relation frames preceding the psychic one.

a. As units of information, signals or stimuli quantitatively express the amount of information. A signal has an external source, causing a stimulus in a sensor, or an impression on a sense organ. A stimulus may have an internal or an external source. In communication technology, the unit of information is called a bit. A neuron has an input for information and an output for instructions, both in the quantitative form of one or more stimuli. The nerve cell itself processes the information.

b. A behaviour program integrates stimuli into information and instruction patterns. Neurons make connections and distribute information. By their sense organs, higher animals make observations and transfer signals bridging short or large distances. The animal's body posture provides a spatially founded signal.

c. A net of neurons transports and amplifies information, with application of feedback. Communication between animals could be a kinetic expression of the psychic subject-subject relation.

d. Behaviour requires an irreversible causal chain from input to output, intermitted by programmed information processing. Interpretation, the mutual interaction and processing of complex information, requires a memory, the possibility to store information for a lapse of time.

e. The animal's ability to learn, to generate new informative links, to adapt behaviour programs, may be considered a projection on the biotic subject-subject relation. Learning is an innovative process, unlearning is a consequence of ageing. In the nervous system, learning implies both making new connections between neurons and developing programs.

The psychic subject-subject relation and its five projections should be recognizable in all psychic characters. They are simulated in computers and automatized systems.

7.2. Secondary characteristics of animals

Animal behaviour has an organic basis in the nervous system. Its character has a genetic foundation. 'The study of behavior is the study of the functioning of the nervous system and must be carried out at the behavioral level, by using behavioral concepts ... the *output* of the nervous system, manifested as perceptions, thoughts, and actions.'^[4] The sense organs are specialized parts of the nervous system, from which they emerge during the development of the embryo. The nervous system controls the animal body and requires observation, registration, and processing of external and internal information. The processing of stimuli, coming from inside or outside the animal body, occurs according to a certain program. This program is partly fixed, partly adaptable because of experience. Consequently, animals react to changes in their environment much faster and more flexibly than plants do. Besides the nervous system, the whole body and its functioning are disposed to behaviour.

a. The basic element of the nervous system is the nerve cell or neuron, passing on stimuli derived from a sensor to an effector. A unicellular animal (a protozoan) has no nerve cells. Rather, it is a nerve cell, equipped with one or more sensors and effectors. 'Protozoa, being single-cell systems ... seem to be organized along principles similar to those governing the physiology of neurons ... the protozoan is like a receptor cell equipped with effector organelles.'^[5] An effector may be a cilium by which the animal moves. The simplest multicellular animals like sponges consist only of such cells.^[6] A sponge (to the phylum Porifera belong about 10.000 species) has no nervous system, no mouth, muscles, or other organs. The cells are grouped around a channel system allowing of streaming water. Each cell is in direct contact with water. A sponge has at least 34 different cell types. The cells are organically but not psychically connected. The even more primitive Placozoa (of which only two species are known) too lack a nervous system. A nerve cell in a more differentiated animal is a psychic subject with a character of its own, spatially and functionally interlaced with the nervous system and the rest of the body. The protozoans and the sponges as well as the neurons in higher animals may be considered to be primarily psychically and secondarily quantitatively characterized thing-like subjects. For all multicellular animals, the neurons and their functioning (inclusive their neurochemistry) are strikingly similar, with only subordinate differences between vertebrates and invertebrates.^[7]

b. In a multicellular nervous system, a neuron usually consists of a number of dendrites, the cell body and the axon ending in a number of synapses. Each synapse connects to a dendrite or the cell body of another cell. (There are two kinds of nerve cells, neurons that are connected to each other besides glial cells, supporting the activity of the neurons. In the human brain, glial cells are more numerous than neurons, but I shall only discuss neurons. Glia perform a number of functions in the nervous system, including physically supporting the brain; assisting in nervous system development, repair, and maintenance; insulating neurons; and providing metabolic functions for neurons.)

A dendrite collects information from a sensor or from another neuron. After processing, the cell body transfers the information via the axon and the synapses to other neurons, or to a muscle or a gland. The dendrites collect the *input* of information that is *processed* in the cell body. The axon *transports* the output, the processed information that the synapses *transfer* to other cells. In this way, neurons form a network, an organ typical for all animals except the most primitive ones like protozoans and sponges. Neurons are distinct from other cells. The other cells may be sensitive for instructions derived from neurons, but they are unable to generate or process stimuli themselves. The neurons make psychic connections between each other and to other cells, sometimes bridging a long distance. The network's character is primarily psychically qualified, secondarily spatially founded. One or more neurons contain a program that integrates simultaneously received stimuli and processes them into a co-ordinated instruction.

Jellyfish, sea anemones, corals, and hydrozoans belong to the phylum of cnidarians (now about 10,000 species, but in the early Cambrian much more numerous; whether the pre-Cambrian Ediacaran fauna mostly consisted of cnidarians is disputed[8]). They have a net of neurons but not a central nervous system. The net functions mostly as a connecting system of more or less independent neurons. The neurons inform each other about food and danger, but they do not constitute a common behaviour program. The body plan of cnidarians is more specialized than that of sponges. Whereas the sponges are asymmetrical, the cnidarians have an axial symmetry. They cannot move themselves. Sea anemones and corals are sedentary, whereas jellyfish are moved by sea currents. The nerve net of cnidarians can only detect food and danger. It leads to activating or contracting of tentacles, and to contracting or relaxing of the body. However, even if a jellyfish is a primitive animal, it appears to be more complex than many plants.

c. In the nervous system, signals follow different pathways. Each signal has one or more addresses, corresponding to differentiated functions.

The behaviour of animals displays several levels of complexity. Sensorial, central, and motor mechanisms are distinguished as basic units of behaviour. Often these units correspond with structures in the nervous system and sometimes they are even recognizable in a single neuron. 'There may often be a close correspondence between systems defined in structural and functional terms, but this is by no means always the case, and it is very easy for confusion to arise.'[9] Only in a net, neurons can differentiate and integrate. Now the three mentioned functions are localized respectively in the sense organs, the central nervous system, and in specialized muscles.

The simplest differentiated net consists of two neurons, one specialized as a sensor, the other as a motor neuron. The synapses of a motor neuron stimulate a muscle to contract. In between, several inter-neurons may be operative, in charge of the transport, distribution, or amplification of stimuli. In the knee reflex, two circuits are operational, because two muscles counteract, the one stretching, the other bending the knee. The two circuits have a sensor neuron in common, sensitive to a pat on the

knee. In the first circuit, the sensor neuron sends a stimulus to the motor neuron instructing the first muscle to contract. In the second circuit, a stimulus first travels to the inter-neuron, blocking the motor neuron of the other muscle such that it relaxes.

A differentiated nervous system displays a typical left-right symmetry, with many consequences for the body plan of any animal having a head and a tail. In contrast with the asymmetric sponges and axially symmetric cnidarians, bilateral animals can move independently, usually with the head in front. The bilateral nervous system allowing of information transport is needed to control the motion. The more differentiated the nervous system is, the faster and more variable an animal is able to move. In the head, the mouth and the most important junction (ganglion) of the nerve net are located, in the tail the anus. From the head to the tail stretches a longitudinal chain of neurons, branching out in a net. Sometimes there is a connected pair of such chains, like a ladder. Apparently, these animals are primarily psychically and secondarily kinetically characterized. At this level, real sense organs and a central brain are not present yet, but there are specialized sensors, sensitive for light, touch, temperature, etc.

The simplest bilateral animals are flatworms, having hardly more than a net of neurons without ganglions. A flatworm has two light sensitive sensors in its head enabling it to orient itself with respect to a source of light. Round worms and snails have ganglions co-ordinating information derived from different cells into a common behaviour program. Their reaction speed is larger and their behaviour repertoire is more elaborate than those of flatworms, but considerably less than those of arthropods, e.g.

Progressing differentiation of the nervous system leads to an increasing diversity of animal species in the parallel-evolved phyla of invertebrates, arthropods, and vertebrates. Besides the nervous system, the behaviour, the body plan, and the body functions display an increasing complexity, integration, and differentiation. In various phyla, the evolution of the body plan and the body functions, that of the nervous system and the behaviour, have influenced each other strongly.

Remarkable is an increasing internalization, starting with a stomach.[10] After the conception, every multicellular animal starts its development by forming a *blastula*, a hollow ball of cells. A sponge is not much more than such a ball. Sponges and cnidarians have only one cavity, with an opening that is mouth and anus simultaneously. The cavity wall is at most two cells thick, such that each cell has direct contact with the environment. Animals with a differentiated nervous system have an internal environment, in cavities which walls are several cells thick. Between neighbouring cells, there are intercellular cavities.

In differentiated animals, biologists distinguish four kinds of tissues (with their functions): epithelium (the external surface of the body and its organs, taking care of lining, transport, secretion, and absorption); connective tissue (support, strength, and elasticity); muscle tissue (movement and transport); and nervous tissue (information, synthesis, communication, and control).[11] Vertebrates have an internal skeleton and

internal organs like blood vessels, kidneys, liver, and lungs. These may be distinguished according to their ethological functions: information and control (nervous system and endocrine system); protection, support, and motion (skin, skeleton, and muscles); reproduction (reproductive organs); food (digestion and respiration organs); transport and defence (blood, the heart, the blood-vessels, the lymph nodes, the immune system); secretion, water and salts balance (kidneys, bladder, and guts).[12] As far as a plant has differentiated organs (leaves, flowers, roots, the bark), these are typically peripheral, with an outward direction to the acquisition of food and reproduction. Animal organs are internalized. This is compensated for by the formation of specific behaviour organs directed outward. These are the movement organs like feet or fins; catching organs like a beak or the hands; fighting organs like horns or nails; and in particular the sense organs.

d. Manipulation of the environment requires a central nervous system and sense organs. The most interesting capacities of the nervous system emerge from the mutual interaction of neurons. The storage and processing of information requires a central nervous system. Reflexes are usually controlled outside this centre. The peripheral nervous system takes care of the transport of information to the centre and of instructions from the centre to muscles and glands. It is therefore secondarily kinetically characterized. The physically founded storage and processing of information requires specialization of groups of cells in the centre, each with its own programs.

In particular the sensors are grouped into specialized sense organs allowing of the formation of images. The best known example is the eye that in many kinds of animals uses light sensitivity to produce an image of the surroundings. In 1604, Johann Kepler demonstrated how in the human eye image formation as a physical process proceeds, thanks to the presence of a lens. In all vertebrates and octopuses, it works in the same way. Other animals, e.g., insects, do not have an eye lens. In vertebrates, the image formation occurs at the backside of the retina, in squids at the front side. The visual image formation does not end at the retina. An important part of the brain is involved in the psychic part of imaging. Besides visual, an image may be tactile or auditive, but now there is no preceding physical image formation comparable to the visual one in the eye.

On this level, chains of successive acts occur, in which different organs and organ systems co-operate, such as in food gathering, reproduction, movement, or fighting. Animals have manipulative organs, like teeth and claws. Animals with a central nervous system are primarily psychically and secondarily physically characterized.

e. In the highest animals' neocortex, the brain is superposed on the autonomous nervous system. In the latter, the same processes occur as in the entire nervous system of lower animals. With respect to the construction of their nervous system and their behaviour, octopuses, birds, and mammals are comparable. Within the nervous system, a division appears between the routine control of the body and less routine tasks. The neocortex can be distinguished from the 'old brain', including the limbic system, that controls processes also occurring in lower animals. In primates, there is

a further division of labour between the global, spatio-temporal right half and the more concentrated, analytical and serial left half, which in human beings harbours the speech centre. In the neocortex, the learning capacity of animals is concentrated. The difference between old and new brains, or between left and right half, is not rigorous. It points to the phenomenon that new programs always make use of existing ones.

Animals are capable of learning. Learned behaviour called habituation, i.e. an adaptive change in the program caused by experience, occurs both in higher and in lower animals. During habituation a new program emerges that the animal applies in a stimulus-reflex relation. The reverse is dehabituation. A stronger form is sensitivation, learning to be alert for new stimuli.

Instrumental learning, based on trial and error, is biotically founded. It requires imagination besides a sense for cause-effect relations. Only the highest animals are able to learn by experiment (experimental trial and error), in which the animal's attention is directed to the effect of its activities, to the problem to be solved. Sometimes an *AH-Erlebnis* occurs. Whether this should be considered insightful learning is controversial.[13]

Sometimes animals learn tricks from each other. Singing birds learn the details of their songs from their parents, sometimes prenatal. Some groups display initiation behaviour. In the laboratory, imitation learning is the imitation of a new or improbable activity or expression for which no instinctive disposition exists. It is a consequence of observing that another animal of the same or a different species performs an act for which it is rewarded.

Mammals, birds, and octopuses have programs that require to make choices. They apply these programs in the exploration of their environment and in playing. Initially, the animal makes an arbitrary choice, but it remembers its choices and their effects. By changing its programs, the animal influences its later choices. The new circumstances need not be the same as in the earlier case, but there must be some recognizable similarity.

Starting from the lowest level, each psychic character has dispositions to interlace with characters at a higher level. Neurons have the disposition to become interlaced into a net that allows of differentiation. The differentiated net may form a central nervous system, at the highest level divided into an autonomous system and a brain. These levels constitute a hierarchy, comparable to the quantum ladder (5.3).

On the one hand, the phenomenon of character interlacement means that the characters having different secondary foundations remain recognizable, on the other hand, it implies an amount of adaptation. A neuron in a net is not the same as a unicellular animal, but it displays sufficient similarities to assume that they belong primarily and secondarily to the same character type. Only the tertiary characteristic is different, because a unicellular protozoan cannot become part of a net of neurons, and because it has sensors and flagellates instead of dendrites and an axon.

The relation between 'old' and 'new' brains can be understood as a character interlacement as well. In particular, instinctive processes and states like emotions that mammals and birds share with fish, amphibians, and reptiles are located in the limbic system, the 'reptilian brain'. Hence, the difference between the limbic system in the higher animals and the central nervous system in the lower animals is tertiary, whereas the difference of both with the neocortex is secondary. This character interlacement is not only apparent in the structure of the nervous system. Both the programming and the psychic functioning of the nervous system display an analogous characteristic hierarchy.

7.3. Control processes

Animals are sensitive for their own body, for each other and for their physical and biotic environment. By observing, an animal as a subject establishes relations with its environment, being the object of its observation. In the subjective observation space (which is not necessarily Euclidean), an animal observes a number of objects in their mutual relationships, dependent on the animal's needs. Motion of an object is observed against the background of the observation space. Between some changes (as far as of the animal's interest), the animal makes a causal connection. Together with its own position and its memory, the observation space constitutes the subjective world of experience of an animal, to be distinguished from its objective environment.

Organically, sensors or sense organs bring about observation. The gathering of information is followed by co-ordination, transfer, and processing into instructions for behaviour, via the nervous and endocrine systems. Together, this constitutes a chain of information processing.

I shall distinguish between control processes (7.3), controlled processes (7.4), and psychically qualified behaviour (7.5), each having their specific characters. For information processing, projections on the quantitative up to the biotic relation frames can be indicated, as follows. Although control on the level of genes is very important for animal development, I shall not discuss it.

a. The simplest form of control is to switch on or off a programmed pattern of behaviour, like an electric appliance is put into operation by an on/off switch. Psychology calls this the release after the reception of an appropriate signal. Each signal and each stimulus must surpass a threshold value in order to have effect. Mathematically, a step function represents the transition from one state to the other. Its derivative is the delta function describing a pulse, the physical expression of a stimulus or signal, kinetically represented by a wave packet (chapter 4). In a neuron, a stimulus has the character of a biotically organized chemical process, called an action potential, in which specific molecules (neurotransmitters) play a part. Hence, the

objective psychological character of a signal or a stimulus is interlaced with various other characters.

The simplest form of behaviour consists of a direct relation between a stimulus and a response (e.g., a reflex). It depends on a specific stimulus that switches the program on or off. (The program itself may be quite complex). Often, only the output is called behaviour, but there is an unbreakable connection between input, program, and output. Hence it appears better to consider the whole as a kind of behaviour.

Sometimes a program as a whole is out of operation, such that it is insensitive for a stimulus or signal that should activate it. Hormonal action has the effect that animals are sexually active only during certain periods. Hormones determine the difference between the behaviour of male and female specimens of the same species. Sometimes, female animals display male behaviour (and conversely), if treated with inappropriate hormones. Being switched on or off by hormones, sexual behaviour programs appear to be available to both genders.

b. A spatially founded system of connected neurons receives simultaneous stimuli from various directions and co-ordinates instructions at different positions. The integration of stimuli and reflexes does not require a real memory. 'Immediate memory' is almost photographic and it lasts only a few seconds. It allows of the recognition of patterns and the surroundings. The reaction speed is low. Recognition of a spatial pattern requires contrast, the segregation of the observed figure from its background.

Often, a program requires more information than provided by a single signal. The observation of a partner stimulates mating behaviour, whereas the presence of a rival inhibits it. Moreover, internal motivation is required. Aggressive behaviour against a rival only occurs if both animals are in a reproductive phase. Besides stimulating, a stimulus may act relaxing, blocking, numbing, or paralysing.

Via the dendrites, several incoming pulses activate simultaneously the psychic program localized in a single cell body or a group of co-operating neurons. Some pulses act stimulating, others inhibiting. In this case, only the integration of stimuli into a pattern produces an instruction that may be a co-ordinated pattern of mutually related activities as well. Each neuron in a net co-ordinates the information received in the form of stimuli through its dendrites. It distributes the processed information via the axon and synapses to various addresses. Various mechanisms can be combined into more complex behaviour systems, like hunting, eating, sexual or aggressive behaviour. A behaviour system describes the organization of sensorial, central, and motor mechanisms being displayed as a whole in certain situations. In electronics, such a system is called an integrated circuit, in computers it is an independent program.

c. Each neuron transports information via its axon to other cells. In a differentiated nervous system, transport and amplification of information occurs in steps, mediating between the reception of signals and the exertion of instructions. As discussed so far,

information exists as a single pulse or a co-ordinated set of pulses. However, the information may consist of a *succession* of pulses as well. The short-term memory (having duration of 10-15 minutes) allows the animal to observe signals arriving successively instead of simultaneously. The stored information is deleted as soon as the activity concerned is completed.

If an observed object moves, it changes its position with respect to its background, enhancing its contrast. Hence, with respect to its background, a moving object is easier to be observed than a stationary object. Likewise, an animal enhances the visibility of an object by moving its eyes.

Amplification of stimuli makes negative feedback possible. This control process requires a sensor detecting a deviation from a prescribed value (the set point) for a magnitude like temperature. Transformed into a signal, the deviation is amplified and starts a process that counters the detected deviation. For a feedback process, no memory is required.

d. Psychologists distinguish sensation from perception. Sensations are the basic elements of experience, representing information. Perception is the interpretation process of sensorial information. 'Sensations are the basic data of the senses ... Perception is a process of interpretation of sensory information in the light of experience and of unconscious inference'.^[14] Perception is a new phase between the reception of signals and the exertion of instructions. It allows the animal to observe changes in its environment, other than motions for which a short-term memory is sufficient.

A physically differentiated nervous system may include chemical, mechanical, thermal, optic, acoustic, electric, and magnetic sensors, besides sensors sensitive for gravity or moisture. The sense organs distinguish signals of a specific nature and integrate these into an image, that may be visual, tactile, or auditive, or a combination of these.

An animal having sense organs is capable of forming an image of its environment and storing it in its memory. It is able to make a perceptive connection between cause and effect.^[15] This does not mean a conceptual insight into the abstract phenomenon of cause and effect - that is reserved to human beings. It concerns concrete causal relations, with respect to the satisfaction of the animal's needs of food, safety, or sex. For instance, an animal learns fast to avoid sick making food. An animal is able to foresee the effects of its behaviour, for the best predictor of an event is its cause.

Imaging allows an animal to get an impression of its changing environment in relation to the state of its body. The animal stores the image during some time in its memory, in order to compare it with an image formed at an earlier or later time. This is no one-way traffic. Observation occurs according to a program that is partly genetically determined, partly shaped from earlier experiences, and partly adapts itself to the situation of the moment. Observation is selective: an animal only sees what it needs in order to function adequately.

In observation, recollection, and recognition, comparison with past situations as well as knowledge and expectations play a part. If an animal recognizes or remembers an object, this gives rise to a latent or active readiness to react adequately. Not every circuit reacts to a single stimulus switching it on or off. Stimuli derived from a higher program may control a circuit in more detail. This is only possible in a nervous system having differentiation and perception besides co-ordination, and allowing of transport and storage of information. The long-term memory is located in the central nervous system, requiring specialized areas coupled to the corresponding sense organs. (The distinction between immediate, short, and long term memory does not concern their duration, but their function - on which the duration depends - as described in this section.)

Recognition based on image formation does not occur according to the (logical) distinction of similarities and differences, but holistic, as a totality, in the form of a *Gestalt*. Recollection, recognition, and expectation, respectively concerning the past, the present, and the future, give rise to emotions like joy, sorrow, anger, or fear. Probably, in animals it concerns always 'object bound' emotions, e.g., the fear of an other animal. In human beings one finds anxiety in the I-self relation as well. Images psychically interact with each other or with inborn programs. Emotions act like forces in psychic processes, in which both the nervous and the endocrine system play their parts. Sometimes the cause of behaviour is an internal urge or driving force (the original meaning of 'instinct'). This waits to become operative as a whole until the animal arrives at the appropriate physiological state and the suitable environment to utter its instinct.

Imaging allows an animal to control its behaviour by its expectations, by anticipations, by 'feedforward'. ('The term feedforward ... is used for situations in which the feedback consequences of behaviour are anticipated and appropriate action is taken to forestall deviations in physiological state.'^[16]) The intended goal controls the process. Animals drink not only to lessen their thirst, but also to prevent thirst. Taking into account observations and expectations, animals adapt the set point in a feedback system. (In a thermostat, the desired temperature is called the 'set point'. In homeostasis, the set point is constant, in an active control process the set point is adaptable.)

e. Fantasy or imagination is more than processing of information. It is innovative generation of information about situations which are not realized yet. It allows higher animals to anticipate on expected situations, to make choices, to solve problems, and to learn from these. It requires a rather strongly developed brain able to generate information, in order to allow of choosing between various possibilities. At this level, emotions like satisfaction and disappointment occur, because of agreement or disagreement between expectation and reality. In particular young mammals express curiosity and display playful behaviour.

Animals control their learning activity by directing their attention. Attention for aspects of the outer world depends on the environment and on the animal's internal state. A well-known form of learning in a new born animal is imprinting, for instance the

identification of its parents. Sometimes, comparing of experience leads to the adaptation of behaviour programs, to learning based on recognized experiences. Associative learning means the changing of behaviour programs by connecting various experiences. In a conditioned reflex, an animal makes connections between various signals. Repetition of similar or comparable signals gives a learning effect known as *reinforcement* (amplification by repetition).[17]

7.4. Controlled processes

All controlled processes are organized processes as well, and subject to physical and chemical laws. In an organized process (6.2), enzymes are operative, by lowering or heightening energy barriers. Hormones play a comparable stimulating or inhibiting part. Technologists speak of control if a process having its own source of energy influences another process, having a different energy source.

a. Like an electron, a stimulus corresponds to a kinetic wave packet, whereas the transport and processing of a pulse have a physical nature. Transport of information occurs by means of an electric current or a chemical process in a nerve. The distribution of hormones from the producing gland to some organ, too, constitutes information transport. In invertebrates, the stimulus has often the form of an electric pulse, in vertebrates it is a chemical pulse (an action potential). Whereas the neurons produce most stimuli, external signals induce stimuli as well. The induction and transport of stimuli happens in a characteristic way only occurring in animals. However, the accompanying characters of a physical pulse and a kinetic wave packet are fairly well recognizable.

b. The body plan of an animal is designed for its behaviour. Complex behaviour requires co-ordinated control by an integrated circuit in the nervous system, usually combined with the endocrine system. A special form of co-ordinated behaviour follows an alarm. This brings the whole body in a state of alertness, sometimes after the production of a shot of adrenalin. The animal's body posture expresses its emotive state.

c. Controlled motions recognizable as walking, swimming, or flying, are evidently different from physical or biotic movements, even without specifying their goal. Psychically qualified behaviour is recognizable because of its goal, like hunting or grazing. One of the most important forms of animal behaviour is motion. For a long time, the possibility to move itself was considered the decisive characteristic of animals. Crawling, walking, springing, swimming, and flying are characteristic movements that would be impossible without control by feedback. The animal body is predisposed to controlled motion, such that from fossils it can be established how now extinct animals were moving. Not all movements are intended to displacement,

they may have other functions. Catching has a function in the achievement of food, chewing in processing it. Animal motions are possible because animal cells are not rigidly but flexibly connected, having intercellular cavities (unlike plant cells). Muscular tissues are characteristically developed to make moving possible.

Many of the mentioned movements are periodic, consisting of a rhythmic repetition of separate movements.^[18] Many animals have an internal psychic clock regulating movements like the heartbeats or the respiration. The circadian clock (*circa die* = about a day) tunes organic processes to the cycle of day and night. Other clocks are tuned to the seasons (e.g., fertility), and some coastal animals have a rhythm corresponding to the tides.

The more complicated an animal is, the more important the control of its internal environment. Homeostasis is a characteristic process controlled by feedback. Many animals keep their temperature constant within narrow limits. The same applies to other physical and chemical parameters.

Animals with a central nervous system and specialized sense organs control their external behaviour by means of feedback. They are able to react fast and adequately to changes in their environment.

d. In particular in higher animals, the nervous system controls almost all processes in some way. Respiration, blood circulation, metabolism and the operation of the glands would not operate without control. The animal controls its internal environment by its nervous system, that also controls the transport of gases in respiration and of more or less dissolved materials in the guts and the blood vessels. Whereas in plants metabolism is an organized process, in animals it is controlled as well.

Internal processes are usually automatically controlled, but in specific actions, an animal can accelerate or decelerate them or influence them in other ways. Animals with sense organs also control external processes like the acquisition of food.

The development of a differentiated animal from embryo to the adult form is a controlled biotic process. The growth of an animal starting from its conception is influenced by the simultaneously developing nervous system. In mammals before the birth, there is an interaction with the mother, via the placenta. Emotions induced by the observation of a partner or a rival control mating behaviour.

e. Many forms of behaviour, such as mating, are genetically programmed. Through the genes, they are transferred from generation to generation. They are stereotype, progressing according to a fixed action pattern. The programming of other forms of behaviour occurs during the individual's development after its conception. Earlier, I observed that the genome should not be considered a blueprint (6.2). Even in multicellular differentiated plants, the realization of the natural design during the growth is not exclusively determined by the genome, but by the environment of the dividing cell as well. The tissue to which the cell belongs determines in part the phenotype of the new cells. Besides, in animal development during the growth the

nervous and endocrine systems play a controlling part. While the nervous system grows, it controls the simultaneous development of the sense organs and of other typically animal organs like the heart or the liver.

Besides the animal body including the nervous system, the programs in the nervous system are genetically determined, at least in part. Partly they are developed during growth. Moreover, animals are capable of changing their programs, to learn from information acquired from their environment. Finally, the exertion of a program depends on information received by the program from elsewhere.

Behaviour programs consist of these four components. Hence, there is no dualism of genetically determined and learned behaviour. 'We cannot dichotomize mammalian behaviour into learned and unlearned ...' Daniel Lehrman and others criticize Konrad Lorenz's definition of instinctive behaviour to be genetically determined (in contrast to learned behaviour).[19] Each kind of behaviour has inherited, learned and environmental components.[20] 'The innate/learnt type of dichotomy can lead to the ignoring of important environmental influences on development.' [21] Behaviour emerges as a relation of the animal with its environment, as adaptation in a short or a long time. First, by natural selection a population adapts the genetic component to a suitable niche. Next, an individual animal actualizes this adaptation during its development from embryo to adult. Third, its learning capacity enables the individual to adapt its behaviour to its environment much faster than would be possible by natural selection or growth. Fourth, the input of data in the program allows the animal to adapt its behaviour to the situation of the moment.

7.5. Goal-directed behaviour

Behaviour consists of psychically qualified events and processes. It emerges as a chain from stimulus or observation via information processing to response. It is always goal-directed, but it is not goal-conscious, intentional, or deliberate, these concepts being applicable to human behaviour only. Since the eighteenth century, physics has expelled goal-directedness, but the psychic order is no more reducible to the physical order than the biotic one. Behaviour is goal-directed and its goal is the object of subjective behaviour.

Since Aristotle, there is a dualism of causal and teleological explanations ('proximate' versus 'ultimate' causes). By 'teleology' is understood both the (biotic) function and (psychic) goal.[22] I restrict my discussion of goal-directedness to behaviour. Goal-directed behaviour always has a function, but a biotic function is not always goal-directed. Function and purpose presuppose (physical) causality, but cannot be considered causes themselves. Ernest Nagel associates teleological explanations with '... the doctrine that goals or ends of activity are dynamic agents in their own

realizations ... they are assumed to invoke purposes or end-in-views as causal factors in natural processes.'[23] In order to prevent this association, I shall avoid the term teleology (or teleonomy[24]). The goal being the *object* of animal behaviour cannot be a 'dynamic agent'. Only the animal itself as a psychic *subject* pursuing a goal is an agent of behaviour. This is in no way at variance with physical laws.

Often an animal's behaviour is directed to that of another animal. In that case, besides a subject-object relation, a subject-subject relation is involved. Animal behaviour is observable, both to people and to animals. By hiding, an animal tries to withdraw from being observed. Threatening and courting have the function to be observed. This occurs selectively, animal behaviour is always directed to a specific goal. Courting only impresses members of the same species.

According to the theory of characters various types of behaviour are to be expected, based on projections of the psychic relation frame onto the preceding ones. It has been established that many animals are able to recognize general relations in a restricted sense. These relations concern small numbers (up to 5), spatial dimensions and patterns in the animals' environment, motions and changes in their niche, causality with respect to their own behaviour and biotic relations within their own population.

For human beings, activity is not merely goal-directed, but goal-conscious as well. In the following overview, I shall compare animal with human behaviour.

a. A neuron transforms stimuli coming from a sensor into an instruction for an effector, e.g. a muscle or a gland. Muscles enable the animal's internal and external movements. The glands secrete materials protecting the body's health or alerting the animal or serving its communication with other animals. The direct stimulus-response relation occurs already in protozoans and sponges. The reflex, being the direct reaction of a single cell, organ, or organ system to a stimulus, is the simplest form of behaviour. It may be considered the unit of behaviour. Reflexes are always direct, goal-directed, and adapted to the immediate needs of the animal. Whereas complex behaviour is a psychically qualified *process*, a reflex may be considered a psychically qualified *event*.

Often, a higher animal releases its genetically determined behaviour (fixed action pattern) after a single specific stimulus, a *sign stimulus* or *releaser*. If there is a direct relation between stimulus and response, the goal of a fixed action pattern is the response itself, for instance the evasion of immediate danger.

Humans, too, display many kinds of reflexes. More than animals, they are able to learn certain action patterns, exerting them more or less 'automatically'. For instance, while cycling or driving a car, people react in a reflexive way to changes in their environment.

Human beings and animals are sensitive for internal and external states like hunger, thirst, cold, or tiredness. Such psychically experienced states are quantitatively determined. An animal can be more or less hungry, thirsty, or tired, feeling more or

less stimulated or motivated to acting. The satisfaction of needs is accomplished by complex behaviour. Taken together, animals apply a broad scale of food sources. Animals of a certain species restrict themselves to a specific source of food, characterizing their behaviour. In contrast, human beings produce, prepare, and vary their food. People do not have a genetically determined ecological niche. Far more than animals, they can adapt themselves to circumstances, and change circumstances according to their needs.

Contrary to the animals themselves, scientists analyse the quantitative aspect of behaviour by a balance of costs and benefits.[25] A positive cost-benefit relation is appropriate behaviour and favours natural selection. Behaviour always costs energy and sometimes gains energy. Behaviour involves taking risks. Some kinds of behaviour exclude others. The alternation of characteristic behaviour like hunting, eating, drinking, resting, and secreting depends on a trade-off of the effects of various forms of behaviour.[26] People, too, deliberate in this way, conscious or subconscious.

Animals of the same species may form a homogeneous aggregate like a breeding colony, an ants' or bees' nest, a herd of mammals, a shoal of fish, or a swarm of birds. Such an aggregate is a psychically qualified and biotically founded community, if the animals stay together by communicating with each other, or if the group reacts collectively to signals. (A population of animals as a gene pool is biotically qualified, but mating behaviour is a characteristic psychical subject-subject relation.) Human beings form numerous communities qualified by relation frames other than the psychic one.

The study of animals living in groups is called 'sociobiology'[27]. For quite some time, sociobiology has been controversial as far as its results were extrapolated to human behaviour.[28] Sociobiology was accused of 'genetic determinism', i.e. the view that human behaviour is mostly or entirely genetically determined.[29]

b. An ecosystem is a biotically qualified heterogeneous aggregate of organisms (6.5). The environment of a population of animals, its *Umwelt*, is psychically determined by the presence of other animals, biotically by the presence of plants, fungi, and bacteria, and by physical and chemical conditions as well. Each animal treats its environment in a characteristic way. In a biotope, animals of different species recognize, attract or avoid each other. The predator-prey relation and parasitism are characteristic examples. The posture of an animal is a spatial expression of its state controlled by its emotions, but it has a goal as well, e.g. to court, to threaten, to warn, or to hide. Characteristic spatially founded types of behaviour are orientation, acclimatization, and defending a territory.

The *Umwelt* and the horizon of experience of a population of animals are restricted by their direct needs of food, safety, and reproduction. Animals do not transcend their *Umwelt*. Only human beings are aware of the cosmos, the coherence of reality transcending the biotic and psychic world of animals.

c. The movements of animals are often very characteristic: resting, sleeping, breathing, displacing, cleaning, flying, reconnoitring, pursuing, or hunting. On a large scale, the migration of birds, fish, and turtles are typical motions. Usually the goal is easily recognizable. An animal does not move aimlessly. Many animal movements are only explainable by assuming that the animals observe each other. In particular animals recognize each other's characteristic movements. Human motions are far less stereotype than those of animals, and do not always concern biotic and psychic needs.

Communication is behaviour of an individual (the sender) influencing the behaviour of another individual (the receiver).[30] In communication, structuralists recognize the following six elements: the transmitter, the receiver, the message from transmitter to receiver, the shared code that makes the message understandable, the medium, and the context (the environment) to which the message refers. It consists of a recognizable signal, whether electric or chemical (by pheromones), visual, auditive, or tactile. It is a detail of something that a receiver may observe and it functions as a trigger for the behaviour of the receiver. Communication is most important if it concerns mating and reproduction, but it occurs also in situations of danger. Ants, bees, and other animals are capable of informing each other about the presence of food. Higher animals communicate their feelings by their body posture and body motions ('body language').

A signal has an objective function in the communication between animals if the sender's aim is to influence the behaviour of the receiver. A signal is a striking detail (a specific sound or a red spot, the smell of urine or a posture), meant to draw attention. It should surpass the noise generated by the environment. Many signals are exclusively directed to members of the same species, in mating behaviour or care for the offspring, in territory defence and scaring of rivals. Animal communication is species specific and stereotype. It is restricted to at most several tens of signals. In particular between predators and prey, one finds deceptive communication. As a warning for danger, sound is better suited than visual signals. Smelling plays an important part in territory behaviour. Impressive visual sex characteristics like the antlers of an elk or the tails of a peacock have mostly a signal value.

A signal in animal communication is a concrete striking detail. Only human communication makes use of *symbols*, having meaning each apart or in combination. Whereas animal signals always directly refer to reality, human symbols also (even mainly) refer to each other. A grammar consists of rules for the inter-human use of language, determining largely the character of a language. 'Animal communication signals are not true language because animals do not use signals as symbols that can take the place of their referent and because they do not string signals together to form novel sentences.'[31]

d. Often, animal behaviour can be projected on cause-effect relations. Higher animals are sensitive for these relations, whereas human beings have insight in them. Sensory observation, image formation, manipulations, emotions, and conflicts are related forms of behaviour.

The senses allow an animal of forming an image of its environment in order to compare it with images stored in its memory. This enables an animal having the appropriate organs to manipulate its environment, e.g. by burrowing a hole. Characteristic is the building of nests by birds, ants, and bees, and the building of dams by beavers. These activities are genetically determined, hardly adaptable to the environment

The formative activity of animals often results in the production of individual objects like a bird's nest. Plants are producers as well, e.g. of wood displaying its typical cell structure even after the death of the plant. The atmosphere consisting of nearly 20% oxygen is a product of ages of organic activity. In addition, animals produce manure. From the viewpoint of the producing plant or animal, these are by-products, achieving a relatively independent existence after secretion by some plant or animal. In this respect, wood and manure differ obviously from an individual object like a bird's nest.

A nest has primarily a physical character and is secondarily spatially founded, but its tertiary biotic and psychic dispositions are more relevant. It is produced with a purpose. Its structure is recognizable as belonging to a certain species. The nest of a blackbird differs characteristically from the nest of a robin. However, the nest itself does not live or behave. It is not a subject in biotic and psychic relations, but an object. It is a subject in physical relations, but these do not determine its character. It is an individual object, characteristic for the animals that produce it, fish, birds, mammals, insects, and spiders. The construction follows from a pattern that is inborn to the animal. Usually, the animal's behaviour during the construction of its nest is very stereotype. Only higher animals are sometimes capable of adapting it to the circumstances. The tertiary psychic characteristic of a nest, its purpose, dominates its primary physical character and its secondary spatial shape.

Manipulating the environment concerns a subject-object relation. The mutual competition, in particular the trial of strength between rivals, may be considered a physically founded subject-subject relation. Both are species-specific and stereotype. Stereotype animal behaviour contrasts with the freedom of human activity, for which human beings are consequently responsible.

e. Much animal behaviour has a biotic function, like reproduction and survival of the species. Animals are sensitive for genetic relations. Whether protozoans experience each other is difficult to establish, but their mating behaviour makes it likely. The courting and mating behaviour of higher animals is sometimes strongly ritualized and stereotype. It is both observable and meant to be observed. It has an important function in the natural selection based on sexual preferences.^[32] The body plan, in particular the sexual dimorphism, is tuned to this behaviour.

Mating behaviour and care for the offspring are psychically qualified and biotically founded types of behaviour. Animals are sensitive to the members of their species, distinguishing between the sexes, rivals, and offspring. For biotically founded behaviour, the mutual communication between animals is important. Sexually mature animals excrete recognizable scents. In herds, families, or colonies, a rank order with

corresponding behaviour is observable. An animal's rank determines its chance of reproduction.

Human mating behaviour is cultivated, increasing its importance. People distinguish themselves from animals by their sense of shame, one reason to cover themselves with clothing. The primary and secondary sex characteristics are both hidden and shown, in a playful ritual that is culturally determined, having many variations (11.1). Human sexuality is not exclusively directed to biotic and psychic needs and inborn sexual differences. It is expressed in many kinds of human behaviour.

The ability to learn is genetically determined and differs characteristically from species to species. Every animal is the smartest for the ecological niche in which it lives. Its ability to learn changes during its development. In birds and mammals, learning takes place already during the prenatal phase. In the juvenile phase, animals display curiosity, a tendency to reconnoitre the environment and their own capacities, e.g. by playing (acting as if). Usually, a young animal has more learning capability than an adult specimen.

The capacity of learning is hereditary and species specific, but what an animal learns is not heritable. The content of the animal's learning belongs to its individual experience. Sometimes, an animal is able to transfer its experiences to members of its population.

The genetic identity of a plant or animal is primarily determined by the individual configuration of its genes. The identity is objectively laid down in the configuration of the DNA molecule, equal in all cells of the organism. Only sexual reproduction changes the genetic configuration, but then a new individual comes into existence. In contrast, the identity of an animal is not exclusively laid down in its genetic identity. An animal changes because of its individual experience, because of what it learns. By changing its experience (by memorizing as well as forgetting), the animal itself changes, developing its identity. Even if two animals have the same genetic identity (think of clones or monozygotic twins), they will develop divergent psychic identities, having different experiences. In the nervous system, learning increases the number of connections between neurons and between programs.

The individual variation in the behaviour of animals of the same species or of a specified population can often be statistically expressed. The statistical spread is caused by the variation in their individual possibilities (inborn, learned, or determined by circumstances), as far as it is not caused by measurement inaccuracies. When the statistics displays a maximum (for instance, in the case of a Gauss or Poincaré distribution), the behaviour corresponding to the maximum is called 'normal'. Behaviour that deviates strongly from the maximum value is called 'abnormal'. This use of the word normal is not related to norms. However, these statistics can be helpful in finding law conformities, in particular if comparison between various species reveals corresponding statistics.

Their learning capacity implies that animals are able to recognize signals or patterns, and to react by adapting their behaviour programs. This means that animals in concrete situations have a sense of regularity. This sense is not comparable to the knowledge of and insight into the universal law conformity that humanity has achieved laboriously. Still, it should not be underestimated. The sense of regularity shared by human beings and animals is a condition for the insight into lawfulness that is exclusively human.

The learning capacity of an animal is restricted to behaviour serving the animal's biotic and psychic needs. It is an example of the capacity of animals (and plants) to adapt themselves to differing circumstances. In this respect, animals differ from human beings, whose behaviour is not exclusively directed to the satisfaction of biotic and psychic needs.

Besides animal psychology studying *general* properties of behaviour, ethology is concerned with the *characteristic* behaviour of various animal species. This does not imply a sharp boundary between animal psychology and ethology. In this chapter, I discussed the general relations constituting the psychic relation frame together with the characters that it qualifies.

Human psychology and psychiatry too are concerned with behaviour, but human behaviour is usually not psychically qualified. Hence, it is not always possible to compare animal with human behaviour. In animals, goal-directed behaviour and transfer of information always concerns psychic and biotic needs like food, reproduction, safety, and survival of the species. In human persons, behaviour may serve other purposes, for instance practicing science.

[1] Purves et al. 1998, 553-554.

[2] McFarland 1999, 62-63

[3] Wallace 1979, 23.

[4] Hogan 1994, 300-301.

[5] McFarland 1999, 174.

[6] Purves et al. 1998, 632-633.

[7] Churchland 1986, 36, 76-77.

[8] Raff 1996, 72.

[9] Hogan 1994, 300-301.

[10] Margulis, Schwartz 1982, 161.

[11] Purves et al. 1998, 810.

[12] Purves et al. 1998, 809-814.

[13] McFarland 1999, 343-346.

[14] McFarland 1999, 204.

[15] McFarland 1999, 340.

[16] McFarland 1999, 278:

[17] About various forms of learning, see Eibl-Eibesfeldt 1970, 251-302; Hinde 1966, chapters 23, 24; Wallace 1979, 151-174; Goodenough et al. 1993, 145; McFarland 1999, part 2.3.

[18] Hinde 1970, chapter 14.

- [19] Lehrman 1953.
 - [20] Cp. Hebb 1953, 108.
 - [21] Hinde 1966, 426.
 - [22] Nagel 1977.
 - [23] Nagel 1961, 402; Ayala 1970, 38.
 - [24] Mayr 1982, 47-51.
 - [25] Houston, McNamara 1999.
 - [26] McFarland 1999, 125-130.
 - [27] Wilson 1975.
 - [28] Segerstråle 2000.
 - [29] Midgley 1985.
 - [30] Goodenough et al. 1993, chapter 17.
 - [31] Goodenough et al. 1993, 596.
 - [32] Darwin 1859, 136-138.
-