2 Complex Systems and the Evolution of Matter

How can order arise from complex, irregular, and chaotic states of matter? In classical antiquity philosophers tried to take the complexity of natural phenomena back to first principles. Astronomers suggested mathematical models in order to reduce the irregular and complex planetary orbits as they are experienced to regular and simple movements of spheres. Simplicity was understood, still for Copernicus, as a feature of truth (Sect. 2.1). With Newton and Leibniz something new was added to the theory of kinetic models. The calculus allows scientists to compute the instantaneous velocity of a body and to visualize it as the tangent vector of the body’s trajectory. The velocity vector field has become one of the basic concepts in dynamical systems theory. The cosmic theories of Newton and Einstein have been described by dynamical models which are completely deterministic (Sect. 2.2).

But Poincaré discovered that those models may be non-computable in the long run (the many-body-problem). Even in a fully deterministic world, the assumption of a Laplacean demon which can calculate the universe in the long run was exposed as an illusionary fiction. Chaos can arise not only in heaven, but also in the quantum world (as quantum chaos) (Sect. 2.3). From a methodological point of view, nonlinearity is a necessary but not sufficient condition of chaos. It also allows the emergence of order. In the framework of modern physics, the emergence of the structural variety in the universe from elementary particles to stars and living organisms is modeled by phase transitions and symmetry breaking of equilibrium states (Sect. 2.4). In the present state of superstring theories and M-theory, we do not have a complete theory explaining the evolution of matter with its increasing complexity. The presocratic wondering that “there is something and not nothing” is not dissolved. But the theory of complex systems opens new avenues of pattern formation in the nano world with applications for self-constructing materials in materials science (Sect. 2.5). From a methodological point of view, the question arises, how can we detect attractors of pattern formation in an immense variety of measured data? Time series analysis, fractals, and multifractals are challenges in the current theory of complex systems. The chapter closes with a survey of the degrees of complexity of different attractors in nonlinear dynamics (Sect. 2.6).
2.1 Aristotle’s Cosmos and the Logos of Heraclitus

Since the presocratics it has been a fundamental problem of natural philosophy to discover how order arises from complex, irregular, and chaotic states of matter [2.1]. What the presocratic philosophers did was to take the complexity of natural phenomena as it is experienced back to “first origins” (αρχή), “principles” or a certain order. Let us look at some examples. Thales of Miletus (625–545 B.C.), who is said to have proven the first geometric theorems, is also the first philosopher of nature to believe that only material primary causes could be the original causes of all things. Thales assumes water, or the wet, as the first cause. His argument points to the observation that nourishment and the seeds of all beings are wet and the natural substratum for wet things is water.

Anaximander (610–545 B.C.), who is characterized as Thales’ student and companion, extends Thales’ philosophy of nature. Why should water be the first cause of all this? It is only one of many forms of matter that exist in uninterrupted tensions and opposites: heat versus cold and wetness versus dryness . . . Therefore Anaximander assumes that the “origin and first cause of the existing things” is a “boundlessly indeterminable” original matter (ἀπειρόν) out of which the opposed forms of matter have arisen. Accordingly we have to imagine the “boundlessly indeterminable” as the primordial state in which matter was boundless, without opposites, and, therefore, everywhere of the same character. Consequently, it was an initial state of complete homogeneity and symmetry. The condition of symmetry is followed by symmetry breaking, from which the world arises with all its observable opposites and tensions:

The everlasting generative matter split apart in the creation of our world and out of it a sphere of flame grew around the air surrounding the earth like the bark around a tree; then, when it tore apart and bunched up into definite circles, the sun, moon and stars took its place. [2.2]

The ensuing states of matter that Anaximander described in his cosmogeny were therefore by no means chaotic; instead they were determined by new partial orders. The fascination with Anaximander increases when one reads his early ideas of biological evolution. He assumes that the first human beings were born from sea animals whose young are quickly able to sustain themselves, as he had observed in the case of certain kinds of sharks. A century later searches were already being made for fossils of sea animals as evidence of the rise of humans from the sea. The third famous Milesian philosopher of nature is Anaximenes (†525 B.C.), who is thought to have been a companion of Anaximander. He regards change as the effect of the external forces of condensation and rarefaction. In his view, every form of matter can serve as basic. He chooses air (ἀέρα):

And rarefied, it became fire; condensed, wind; then cloud; further, by still stronger condensation, water; then earth; then stones; but everything else originated by these. He, too, assumed eternal motion as the origin of transformation. – What contracts and condenses matter, he said is (the) cold; by contrast, what thins and slackens is (the) warm. [2.3]
Thus Anaximenes assumes external forces by which the various states of matter were produced out of a common original matter and were transformed into each other.

Heraclitus of Ephesus (ca. 500 B.C.), “the dark one”, as he was often called, is of towering significance for our theme. His language is indeed esoteric, more prophetic than soberly scientific, and full of penetrating metaphors. He took over from Anaximander the doctrine of struggle and the tension of opposites in nature. The original matter, the source of everything, is itself change and therefore is identified with fire:

The ray of lightning (i.e., fire) guides the All. – This world order which is the same for all was created neither by one of the gods nor by one of the humans, but it was always, and is, and will be eternally living fire, glimmering and extinguishing according to measures. [2.4]

Heraclitus elaborated further on how all states of matter can be understood as distinguishable forms of the original matter, fire. In our time the physicist Werner Heisenberg declared:

At this point we can interpose that in a certain way modern physics comes extraordinarily close to the teaching of Heraclitus. If one substitutes the word “fire”, one can view Heraclitus’ pronouncements almost word for word as an expression of our modern conception. Energy is indeed the material of which all the elementary particles, all atoms and therefore all things in general are made, and at the same time energy is also that which is moved . . . Energy can be transformed into movement, heat, light and tension. Energy can be regarded as the cause of all changes in the world. [2.5]

To be sure, the material world consists of opposite conditions and tendencies which, nevertheless, are held in unity by hidden harmony: “What is opposite strives toward union, out of the diverse there arises the most beautiful harmony (ἁρμονία), and the struggle makes everything come about in this way.” [2.6] The hidden harmony of opposites is thus Heraclitus’ cosmic law, which he called “logos” (λόγος).

What happens when the struggle of opposites comes to an end? According to Heraclitus, then the world comes to a final state of absolute equilibrium. Parmenides of Elea (ca. 500 B.C.) described this state of matter, in which there are no longer changes and motions in (empty) spaces. Matter is then distributed everywhere equally (homogeneously) and without any preferred direction for possible motion (isotropically). It is noteworthy that infinity is thought to be imperfection and therefore a finite distribution of matter is assumed. In this way Parmenides arrived at the image of a world that represents a solid, finite, uniform material sphere without time, motion or change. The Eleatic philosophy of unchanging being was, indeed, intended as a critique of the Heraclitean philosophy of constant change, which is put aside as mere illusion of the senses. And the later historical impact of the Eleatic philosophy in Plato appears in his critique of the deceptive changes that take place in sensory perception in contrast to the true world of unchangeable being of the Ideas. But from the point of view of philosophy of nature, the world Parmenides described was not necessarily opposite to the teaching of Heraclitus; in his cosmogeny it can be understood entirely as a singular end state of the highest symmetry.
After water, air, and fire were designated as original elements, it was easy to conceive of them as raw materials of the world. Empedocles (492–430 B.C.) took that step and added earth as the fourth element to fire, water, and air. These elements are free to mix and bind in varying proportions, and to dissolve and separate. What, now, according to Empedocles, were the enduring principles behind the constant changes and movements of nature? First there were the four elements, which he thought arose from nature and chance (τύχη), not from any conscious intention. Changes were caused by reciprocal effects among these elements, that is, mixing and separation: “I shall proclaim to you another thing: there is no birth with any of the material things, neither there is an ending in ruinous death. There is only one thing: mixture and exchange of what is mixed” [2.7]. Two basic energies were responsible for these reciprocal effects among the elements; he called them “love” (φιλία) for attraction and “hatred” (νείκος) for repulsion. There is an analogy in the yin–yang dualism of Chinese philosophy. Empedocles taught a constant process of transformation, i.e., combination and separation of the elements, in which the elements were preserved. He did not imagine these transformation processes to be at all mechanical (as the later atomists did), but rather physiological, in that he carried over processes of metabolism in organisms to inanimate nature.

In his medical theories, equilibrium is understood to be a genuinely proportional relationship. Thus, health means a particular balance between the opposite components and illness arises as soon as one of them gets the upper hand. If we think of modern bacteriology with its understanding of the antibodies in the human body, then this view of Empedocles proves to be surprisingly apt.

Anaxagoras (499–426 B.C.) advocated what was in many regards a refinement of his predecessors’ teaching. Like Empedocles he developed a mixing theory of matter. But he replaced Empedocles’ four elements with an unlimited number of substances that were composed of seed particles (σπέρματα) or equal-sized particles (ομόμερη). They were unlimited in their number and smallness, i.e., matter was assumed to be infinitely divisible. The idea of a granulated continuum comes forcefully to mind. Anaxagoras also tried to explain mixtures of colors in this way, when he said that snow is, to a certain degree, black, although the whiteness predominates. Everything was contained in each thing, and there were only predominances in the mixing relationships.

More distinctly than some of his predecessors, Anaxagoras tried in his philosophy of nature to give physical explanations for the celestial appearances and motions that were described only kinematically in the mathematical astronomy of the Greeks. So in his cosmology he proceeded from a singular initial state: a homogeneous mixture of matter. An immaterial original power, which Anaxagoras called “spirit” (νοῦς), set this mixture into a whirling motion which brought about a separation of the various things depending on the speed of each of them. Earth clumped together in the middle of the vortex, while heavier pieces of stone were hurled outward and formed the stars. Their light was explained by the glow of their masses, which was attributed to their fast speed. Anaxagoras’ vortex theory appears again in modern times with Descartes, and then in more refined form in the Kant–Laplace theory of the mechanical origin of the planetary system.
In modern natural sciences atomism has proved to be an extremely successful research program. In the history of philosophy the atomic theory of Democritus is often presented as a consequence of Heraclitus’ philosophy of change and Parmenides’ principle of unchanging being. The Democritean distinction between the “full” and the “empty”, the smallest indestructable atoms (ἄτομος) and empty space, corresponded to Parmenides’ distinction between “being” and “not-being”. Heraclitean complexity and change was derived from distinguishable reconfigurations of the atoms. Empty space was supposed to be homogeneous and isotropic.

Atoms differ in their form (μορφή), their position (θέσις), and their diverse configurations (τάξις) in material combinations. The configuration of the atoms for the purpose of designation is compared with the sequence of letters in words, which has led to the presumption that atomistic ideas were developed only in cultures with phonetic alphabets. In fact, in China, where there was no phonetic alphabet but instead ideographic characters, the particle idea was unknown and a field-and-wave conception of the natural processes prevailed. The Democritean atoms move according to necessity (νάγκη) in a constant whirl (ἐνός or ἑνη). Here, by contrast with later Aristotelian notions, motion means only change of location in empty space. All phenomena, all becoming and perishing, result from combination (σύνκρισις) and separation (διάκρισις). Aggregate states of matter, such as gaseous, liquid, or solid, are explained by the atoms’ differing densities and potentialities for motion. In view of today’s crystallography, the Democritean idea that even atoms in solid bodies carry out oscillations in place is noteworthy.

Plato, in his dialogue Timaeus, introduced the first mathematical model of atomism. The changes, mixings, and separations on earth of which the pre-socratics had spoken were to be traced to unchangeable mathematical regularities. In Empedocles’ four elements, namely fire, air, water and earth, a classification was at hand that was immediately accessible to experience. Theatetus made a complete determination of all the regular bodies that are possible in 3-dimensional (Euclidian) space: tetrahedra, octahedra, icosahedra, cubes and dodecahedra. Therefore what Plato proposed to do amounted to interpreting Empedocles’ four elements with these geometric building blocks.

Plato consciously avoided the Democritean designation “atom” for his elements, since they can be decomposed into separate plan figures. Thus tetrahedra, octahedra and icosahedra have faces consisting of equilateral triangles which, when they are bisected, yield right-angled triangles with sidelenghts 1, 2 and $\sqrt{3}$, while the square faces of cubes, when bisected, yield right-angled triangles with side lengths 1, 1 and $\sqrt{2}$. A consequence is that “fluids” like water, air and fire can combine with each other whereas a solid made of earth building blocks, because of its different triangles, can only be converted into another solid.

Then Plato developed a kind of elementary particle physics in which the specific elements are transformed into each other and “reciprocal effects” can take place with the “elementary particles” (i.e., the corresponding component triangles) according to geometric laws. Transformation of the elements results, for example, from their being cut open along the edges. Plato made this possibility dependent on the acuteness of the angles of the solid. The more acute plane angles can cleave polyhedra
which have a regular angle. Thus, in sequence, every tetrahedron, every cube, every octahedron, every icosahedron can, in each case, cleave the following polyhedron, but not the previous one or polyhedra of the same sort. The conclusion for the philosophy of nature is that fire can separate or dissolve all elements; earth, only air and water; air, only water.

Plato stated categorically that the elements are not all of the same size. For instance, in order to be able to explain that fire can cause water in solid form to change into water in liquid form, he maintained that in the liquid state the elements are smaller and more mobile, while in the solid state they are larger.

The escape from fire is called cooling and the state after eradicating fire, solidification. Fire and air can pass through the gaps in earth building blocks (cubes) without hindrance, without dissolution of the earth elements. Condensed air cannot be dissolved without destroying the element. Condensed air, namely, means an accumulation of octahedra under the best surface configurations possible. Even fire would not be able to penetrate into the necessarily remaining gaps, whose plane angles are smaller than those of all elements, without destroying the octahedra. In the case of water, only fire is capable of breaking the strongest condensation. The gaps between adjacent icosahedra form plane angles which do not admit penetration by either earth or air. Only fire (tetrahedra) can penetrate and dissolve the combination.

Indeed, Plato developed an internally consistent mathematical model by which various aggregate states and reciprocal effects of substances could be explained if one accepted his – albeit more or less arbitrary – initial conditions for interpretation of the elements. Naturally, a number of the consequences for the philosophy of nature are strange and ridiculous. And yet we have here the first attempt in the history of sciences to explain matter and its states by simple geometric laws. A high point up to now in this development is modern elementary particle physics. Heisenberg made this observation about it: “... The elementary particles have the form Plato ascribed to them because it is the mathematically most beautiful and simplest form. Therefore the ultimate root of phenomena is not matter, but instead mathematical law, symmetry, mathematical form” [2.8]. In Antiquity and the Middle Ages Plato’s mathematical atomism gained little support. The basic problem, for his successors, in his geometric theory of matter was already evident in the dialogue Timaeus. How are the functions of living organisms to be explained? The suggestion that certain corporeal forms are as they are in order to fulfill certain physiological purposes (e.g., the funnel shape of the gullet for assimilation of food) cannot, in any case, be derived from the theory of regular solids. In addition, the idea of explaining the changing and pulsating processes of life on the basis of the “rigid” and “dead” figures of geometry must have seemed thoroughly unnatural, speculative, and farfetched to the contemporaries of that time. Contemporaries of our time still have difficulties understanding the detour that today’s scientific explanations take through complicated and abstract mathematical theories. This is where the Aristotelian philosophy of nature begins.

Aristotle formulated his concept of a balance or “equilibrium” in nature chiefly on the basis of the ways in which living organisms such as plants and animals function. The process and courses of life are known to us from everyday experience.
What is more obvious than to compare and explain the rest of the world, which is unknown and strange, with the familiar? According to Aristotle, the task of science is to explain the principles and functions of nature’s complexity and changes. This was a criticism of those philosophers of nature who identified their principles with individual substances. The individual plant or the individual animal was not simply the sum of its material building blocks. Aristotle called the general, which made the individual being what it was, form (ἐἶδος). What was shaped by form was called matter (ὕλη). Yet form and matter did not exist in themselves, but were instead principles of nature derived by abstraction. Therefore Matter was also characterized as the potential (δύναμις) for being formed. Not until matter is formed does reality (ἐνέργεια) come into being.

The real living creatures that we observe undergo constant change. Here Heraclitus was right and Parmenides, for whom changes were illusory, was wrong. Changes are real. Yet according to Aristotle, Heraclitus was wrong in identifying changes with a particular substance (fire). Aristotle explained those changes by a third principle along with matter and form, namely, the lack of form (στέρησις), which was to be nullified by an adequate change. The young plant and the child are small, weak and immature. They grow because in accordance with their natural tendencies (form), they were meant to become big, strong, and mature. Therefore it was determined that movement (κίνησις) in general was change, transition from possibility to reality, “actualization of potential” (as people in the Middle Ages were to say). The task of physics was to investigate movement in nature in this comprehensive sense. Nature (φύσις) – in contrast to a work of art produced by man, or a technical tool – was understood to be everything that carried the principle of movement within itself. If the Aristotelian designations make us think, first of all, of the life processes of plants, animals, and people as they present themselves to us in everyday experience, these designations seem to us to be thoroughly plausible and apposite. Nature is not a stone quarry from which one can break loose individual pieces at will. Nature itself was imagined to be a rational organism whose movements were both necessary and purposeful. Aristotle distinguished two sorts of movement, namely quantitative change by increase or decrease in magnitude, qualitative change by alteration of characteristics, and spatial change by change of location. Aristotle designated four aspects of causality as the causes of changes. Why does a plant grow? It grows (1) because its material components make growth possible (causa materialis), (2) because its physiological functions determine growth (causa formalis), (3) because external circumstances (nutrients in the earth, water, sunlight, etc.) impel growth (causa efficiens), (4) because, in accordance with its final purpose, it is meant to ripen out into the perfect form (causa finalis).

Aristotle then employed these same principles, which are obviously derived from the life cycles of plants, animals, and humans, to explain matter in the narrower sense, that is, what was later called the inorganic part of nature. Here too Aristotle proceeded from immediate experience. What we meet with is not so and so many elements as isolated building blocks of nature. Instead we experience characteristics such as warmth and cold, wetness and dryness. Combination of these yield the following pairs of characteristics which determine the elements: warm–dry (fire),
warm–wet (air), cold–wet (water), cold–dry (earth). Warm–cold and wet–dry are excluded as simultaneous conditions. Therefore there are only four elements. This derivation was later criticized as arbitrary, but it shows the Aristotelian method, namely to proceed not from abstract mathematical models, but instead directly from experience. Fire, air, water, and earth are contained more or less, more intensively or less intensively, in real bodies and they are involved in constant transformation. According to Aristotle, eliminating the coldness of water by means of warmth results in air, and eliminating the wetness of the air results in fire. The changes of nature are interpreted as maturational and transformational processes.

How could such a predominantly organic philosophy of nature deliver physical explanations for mathematical natural science, insofar as it was extant at that time? There were only two elementary spatial motions – those that proceeded in a straight line and those that proceeded in a circle. Therefore there had to be certain elements to which these elementary motions come naturally. The motions of the other bodies were determined by these elements and their natural motions, depending on which motion predominated with each of them. The most perfect motion was circular motion. It alone could go on without end, which was why it had to be assigned to the imperishable element. This was the fifth element (quintessence), which made up the unchangeable celestial spheres and the stars. The continual changes within the earthly (sublunar) world were to be set off from the unchangeable regularity of the celestial (superlunar) world. These transformational processes were associated with the four elements to which straight-line motion is peculiar, and specifically the straight-line motion toward the center of the world, toward which the heavy elements earth and water strive as their natural locus, and the straight-line motion toward the periphery of the lunar sphere, toward which the light elements strive upwards as their natural locus.

Among the natural motions there was also free fall. But Aristotle did not start out from isolated motions in idealized experimental situations as Galilei did. A falling body is observed in its complex environment without abstraction of frictional ("dissipating") forces. During its free fall a body is sinking in the medium of air like a stone in water. Thus, Aristotle imagines free fall as a hydrodynamical process and not as an acceleration in vacuum. He assumes a constant speed of falling \( v \), which was directly proportional to the weight \( p \) of the body and inversely to the density \( d \) of the medium (e.g., air), thus in modern notation \( v \sim p/d \). This equation of proportionality at the same time provided Aristotle with an argument against the void of atomists. In a vacuum with the density \( d = 0 \), all bodies would have to fall infinitely fast, which obviously did not happen.

A typical example for a (humanly) forced motion is throwing, which, again, is regarded in its complex environment of "dissipative" forces. According to Aristotle a nonliving body can move only as a result of a constant external cause of motion. Think of a cart on a bumpy road in Greece, which comes to a stop when the donkey (or the slave) stops pulling or pushing. But why does a stone keep moving when the hand throwing it lets go? According to Aristotle, there could be no action at a distance in empty space. Therefore, said Aristotle, the thrower imparts a movement to the continuous medium of the stone’s surroundings, and this pro-
pels the stone farther. For the velocity $v$ of a pulling or pushing motion, Aristotle asserted the proportionality $v \sim K/p$ with the applied force $K$. Of course, these are not mathematical equations relating measured quantities, but instead proportionalities of qualitative determinants, which first emerged in this algebraic notation in the peripatetic physics of the Middle Ages. Thus in Aristotelian dynamics, in contrast to Galilean–Newtonian dynamics, every (straight-line) change of position required a cause of motion (force). The medieval theory of impetus altered Aristotelian dynamics by attributing the cause of motion to an “impetus” within the thrown body, rather than to transmission by an external medium.

How did peripatetic dynamics explain the cosmic laws of heaven? The central symmetry of the cosmological model was based on the (unforced) circular motion of the spheres, which was considered natural for the “celestial” element, and on the theory of the natural locus in the centerpoint of the cosmos. Ptolemy was still to account for the position of the earth on the basis of the isotropy of the model and by a kind of syllogism of sufficient reason. Given complete equivalence of all directions, there was no reason why the earth should move in one direction or another.

It was Aristotle’s teacher Plato who presented a centrally-symmetrical model with the earth in the center; around it the whole sky turns to the right around the celestial axis, which goes through the earth. Sun, Moon, and planets turn to the left on spheres that have different distances from the earth in the sequence Moon, Mercury, Venus, Sun, Mars, Jupiter, and Saturn. The outermost shell carries the sphere of the fixed stars. According to the Platonic–Pythagorean conception, the rotational periods are related to each other by whole numbers. There is a common multiple of all rotational times, at the end of which all the planets are exactly in the same place again. The motion of each one produces a sound, so that the tunes of the movements of the spheres jointly form a harmony of the spheres in the sense of a well-ordered musical scale. Geometry, arithmetic, and aesthetic symmetries of the cosmos ring through the universe in a harmonious music of the spheres. Soon this emphatically symmetrical model of the cosmos was called into question by exact observations. A difficult problem was presented by the irregular planetary orbits, especially their retrograde movements. The irregularities in the sky were disquieting, especially for philosophers in the Pythagorean tradition, who were accustomed to comprehending the heaven – in contrast to the earth – as the realm of eternal symmetry and harmony.

Plato posed a famous question in order to reduce the complexity of motions in the heaven: by means of what regular, ordered circular movements could the phenomena of the planets be “saved”, i.e., kinematically explained? An exact model of the observed curves was achieved when Apollonius of Perga (ca. 210 B.C.) recommended that the common center of the spheres be given up. But the spherical form of planetary movement and the equal speed of the spheres were to be retained. According to this proposal, the planets rotate uniformly on spheres (epicycles), whose imagined centers move uniformly on great circles (deferents) around the centerpoint (the earth). By appropriately proportioning the speed and diameter of the two circular motions and by varying their directions of motion, it was possible to produce an unanticipated potential for curves, and these found partial application in astron-
omy from Kepler to Ptolemy also. The spherical symmetry of the individual models was therefore preserved, even if they no longer had a common center, but various different centers.

The following examples of the epicycle-deferent technique show what a multiplicity of apparent forms of motion can be created by appropriately combining uniform circular motions \[2.9\]. This makes the Platonic philosophy more comprehensible in its view that behind the changes in phenomena there are the eternal and unchangeable forms. In Fig. 2.1 an elliptical orbit is produced by combining a deferent motion and an epicycle motion. Figure 2.2 shows a closed cycloid. In this way, changing distances between planets and the earth can also be represented. In principle, even angular figures can be produced. When the epicycle diameter approaches the deferent diameter, an exact straight line results. Even regular triangles and rectangles can be produced by means of appropriate combinations of an epicycle motion and a deferent motion, if one changes the speed of the east-west motion of a planet that is moving on an epicycle with a west-east motion.

Fig. 2.1. Deferent-epicycle model of an ellipse

Fig. 2.2. Deferent-epicycle model of a cycloid

If one lets the celestial body circle on a second epicycle whose midpoint moves on the first epicycle, one can produce a multiplicity of elliptical orbits, reflection-symmetric curves, periodic curves, and also non-periodic and asymmetric curves. From a purely mathematical and kinetic standpoint, Plato’s problem of “saving the phenomena” is completely solved. In principle, therefore, Plato’s reduction of complexity in the sense of uniform circular motion (modified by Apollonius and Ptolemy) could influence the sciences right up until today. In any case, it cannot be disproved by phenomenological description of curved paths. In particular, from this standpoint not only the reversed roles of the earth and the sun in the so-called Copernican revolution, but also Kepler’s change from circular to elliptical orbits, seem secondary, since both initiatives can be traced back to a combination of circular motions in accordance with the epicycle-deferent technique. This poses two questions: (1) How is the assertion mathematically substantiated? (2) If it is sub-
stantiated, why does it not play a role in modern scientific applications of curve theory? In order to answer the first question exactly and generally, it is necessary to go back to the modern structure of analytical geometry. But historically, Copernicus and Kepler also knew how the curves that they used (e.g., ellipses) could be reconstructed by the epicycle-deferent technique.

First of all, one must remember that points on the plane can be represented by complex numbers $z = x + iy = re^{i\theta}$ with the corresponding Cartesian coordinates $(x, y)$ or polar coordinates $(r, \theta)$. The addition of complex numbers then corresponds to vector addition [2.10]. A uniform circular motion with center $c$, radius $\varrho$ and period $T$ can be represented by

$$z = c + \varrho e^{i(2\pi t/T) + i\alpha} = c + \varrho e^{i(2\pi t/T) + i\alpha}$$

with time $t$ and initial phase $\alpha$ for the point. Now assume a point $A$ that is moving according to the equation $z = f(t)$. Let a point $B$ move relative to $A$ on a circle with radius $\varrho$, period $T$, and initial phase $\alpha$. The motion of $B$ is then described by the equation

$$z = f(t) + \varrho e^{i(2\pi t/T) + i\alpha}$$

Then it is possible to describe the movement of a point $B$ on an epicycle whose center moves around $A$. The addition of a new epicycle is described mathematically by the addition of a new term $\varrho e^{i(2\pi t/T) + i\alpha}$ to the expression for $z$. Clearly, $\varrho e^{i(2\pi t/T) + i\alpha} = \varrho e^{i\alpha} e^{i(2\pi t/T)} = ae^{ikt}$ with a complex number $a \neq 0$ and $k$ as a real number. In the case of a retrograde motion, $T$ or $k$, respectively, is taken to be negative. A motion that results from the superposition of $n$ epicycles is then expressed by the equation

$$z = a_1e^{ikt} + a_2e^{ikt} + \ldots + a_ne^{ikt}$$

Let us proceed first from a periodic motion on the plane $z = f(t)$ (e.g., with period $2\pi$). Mathematically, we assume $f$ continuous with limited variation. Then for $f$ there exists a representation with a uniformly converging series

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}$$

Therefore it can easily be proved mathematically that $f(t)$ can be approximated by means of sums

$$S_N(t) = \sum_{n=-N}^{N} c_n e^{int}$$

with any desired degree of exactitude for increasing $N$. Function $f$ is indeed uniformly convergent. Therefore for arbitrarily small $\varepsilon > 0$ one can choose an index $N_0$ so that for all $N \geq N_0$ and all $t$, it holds true that

$$|f(t) - S_N(t)| < \varepsilon$$

Astronomically, this result means that a constant-motion path (of limited variation) can be approximated to any desired degree of exactitude by means of finite superpositions of the epicycle motions.

It is clear that so far we have used only superpositions with epicycle periods $\pm 2\pi, \pm \pi, \pm \frac{2}{3}\pi, \pm \frac{1}{2}\pi, \pm \frac{2}{5}\pi, \ldots$. In particular, therefore, only commensurable superpositions were employed, which can be expressed by means of integer number ratios in accordance with
the Pythagorean tradition. But in fact non-periodic curves can also be approximated by means of epicycloidal superpositions if we permit incommensurable periods. This result is mathematically supported by a proposition by Harald Bohr about almost-periodic functions (1932) [2.11]. The second question, why the epicycle-deferent technique for the explanation of the paths of motion was abandoned, cannot be answered by pointing to the observation of missing curves. Mathematically, observed curves – however exotic – could still be explained in principle (under the above, very broad mathematical frame conditions) by means of the Platonic–Apollonian use of this ancient strategy for reducing the complexity of motions.

The decisive question in this case, however, is which motions the planets “really” carry out, whether they are, in fact, combined, uniform, and unforced circular motions that seem to us on earth to be elliptical paths, or whether they are in fact compelled to follow elliptical paths by forces. This determination cannot be made geometrically and kinematically, but only dynamically, i.e., by means of a corresponding theory of forces, hence by means of physics.

Besides the epicycle-deferent-technique, Ptolemy employed imaginary balance points relative to which uniform circular motions were assumed that, relative to the earth as center, appear non-uniform. This technique proved to be useful for calculation, but constituted a violation of the central symmetry and therefore had the effect of an ad hoc assumption that was not very convincing from the standpoint of philosophy of nature, a criticism later made especially by Copernicus. The reasons that Copernicus exchanged the earth for the position of the sun were predominantly kinematic. Namely, a certain kinematic simplification of the description could be achieved in that way with a greater symmetry. Thus in the heliocentric model the retrograde planetary motions could be interpreted as effects of the annual motion of the earth, which according to Copernicus moved more slowly than the outer planets Mars, Jupiter and Saturn and faster than the inner planets Mercury and Venus. But Copernicus remained thoroughly conservative as a philosopher of nature since he considered greater simplicity in the sense of “natural” circular motion to be a sign of proximity to reality.

With Johannes Kepler, the first great mathematician of modern astronomy, the belief in simplicity was likewise unbroken. In his *Mysterium cosmographicum* of 1596, Kepler began by trying once more to base distance in the planetary system on the regular solids, alternatingly inscribed and circumscribed by spheres. The planets Saturn, Jupiter, Mars, Earth, Venus, and Mercury correspond to six spheres fitted inside each other and separated in this sequence by a cube, a tetrahedron, a dodecahedron, an icosahedron, and an octahedron. Kepler’s speculations could not, of course, be extended to accommodate the discovery of Uranus, Neptune, and Pluto in later centuries.

Yet Kepler was already too much of a natural scientist to lose himself for long in Platonic speculations. His *Astronomia Nova* of 1609 is a unique document for studying the step-by-step dissolution of the old Platonic concept of simplicity under the constant pressure of the results of precise measurement. In contrast to Copernicus, Kepler supplemented his kinematic investigations with original dynamic arguments. Here the sun is no longer regarded as being physically functionless at a kinemati-
cally eccentric point, as with Copernicus, but is seen as the dynamic cause for the motion of planets. The new task was to determine these forces mathematically as well. Kepler’s dynamic interpretation with magnetic fields was only a (false) initial venture. Success came later, in the Newtonian theory of gravity.

The simplicity of the celestial ("superlunar") world and the complexity of the earthly ("sublunar") are also popular themes in other cultures. Let us cast a glance at the Taoist philosophy of nature of ancient China. To be sure, it is edged with myth and less logically argued than the Greek philosophy of nature, and it also invokes more intuition and empathy; nevertheless, there are parallels between the two. Taoism describes nature as a great organism governed by cyclical motions and rhythms, such as the life cycles of the generations, dynasties, and individuals from birth to death; the food chains consisting of plant, animal, and human; the alternation of the seasons; day and night; the rising and setting of the stars; etc. Everything is related to everything else. Rhythms follow upon each other like waves in the water. What forces are the ultimate cause of this pattern in nature? As with Empedocles, in Taoism two opposite forces are distinguished, namely \textit{yin} and \textit{yang}, whose rhythmic increase and decrease govern the world. In the book \textit{Kuei Ku Tzu} (4th century B.C.) it says: “Yang returns cyclically to its origin. Yin reaches its maximum and makes way for yang.” [2.12] While according to Aristotle all individuals carry their natural purposes and movements in themselves, the Tao of yin and yang determines the internal rhythms of individuals, and those energies always return to their origins. The cyclical rotational model of the Tao provides explanations for making calendars in astronomy, for water cycles in meteorology, for the food chain, and for the circulatory system in physiology. It draws its great persuasiveness from the rhythms of life in nature, which people experience every day and can apply in orienting themselves to life. Nature appears as a goal-directed organism.

It is noteworthy that the Chinese philosophy of nature had no notions of atomistic particles and therefore did not develop mathematical mechanics in the sense of the occidental Renaissance. Instead, at its center there was a harmonious model of nature with rhythmic waves and fields that cause everything to be connected to everything. The preference for questions of acoustics and the early preoccupation with magnetic and electrostatic effects is understandable given this philosophy of nature. The view of the Taoists bear more resemblance to the philosophy of nature of the Stoics than to Aristotle. Here too the discussion centers on effects that spread out in a great continuum like waves on water. This continuum is the Stoics’ \textit{pneuma}, whose tensions and vibrations are said to determine the various states of nature. The multifarious forms of nature are only transitory patterns that are formed by varied tensions of the pneuma. Modern thinking leaps, of course, to the patterns of standing water waves or sound waves or the patterns of magnetic fields. Nevertheless, neither the Stoic nor the Taoist heuristic background led to the development of a physical theory of acoustic or magnetic fields comparable to Galilean mechanics with its background of an atomistic philosophy of nature. The emergence of order from complex, irregular, and chaotic states of matter was only qualitatively described, using different models for earth and for heaven.
2.2 Newton’s and Einstein’s Universe and the Demon of Laplace

Since antiquity, astronomers and philosophers have believed that the celestial motions are governed by simple geometric laws. Simplicity was not only understood as the demand for an economical methodology, but, still for Copernicus, as a feature of truth. Thus, the astronomical doctrine from Plato to Copernicus proclaimed: reduce the apparent complexity of the celestial system to the simple scheme of some true motions! The simple building blocks were given by the basic concepts of Euclidean geometry: circle (compass) and straight line (ruler). In contrast to the simplicity of the superlunar world, the sublunar earthly world seemed to be really complex. Thus its dynamics could not be mathematized at least in the framework of Euclidean geometry. That was the reason why Plato’s mathematical atomism was soon forgotten, and Aristotle’s belief in a complex qualitative dynamics of nature which cannot be mathematized in principle overshadowed scientific research until the Renaissance.

Early physicists like Galileo overcame the boundary of a superlunar (“simple”) and sublunar (“complex”) world. They were convinced that the dynamics of nature is governed by the same simple mathematical laws in heaven and on earth. Technically, Galileo simplified the dynamics of, e.g., free fall by selecting some observable and measurable quantities and neglecting other constraints. In short, he made a simplified mathematical model of an idealized experimental situation. Of course, even the astronomical models of antiquity only considered a few parameters, such as angular velocity and position of the planets, and neglected the complex diversity of other constraints (e.g., density, mass, friction of the celestial spheres). From a modern point of view, even the presocratic philosophers suggested qualitative “models” of a complex dynamics in nature by selecting some dominant “parameters” (e.g., water, fire, air, and earth).

In general, a system which may be physical, biological, or social, is observed in different states. The strategies for making models of observed phenomena may have changed since ancient times, but the target of the modeling activity is in some sense the same: the dynamics of the changing states in the observed systems. Obviously, the real states cannot be described by only a few observable parameters, but it is assumed that they can. In the case of early astronomy and mechanics, this was the first step of mathematical idealization and led to a geometric model for the set of idealized states which is nowadays called the state space of the model. The presocratic “models” of nature differ from modern ones not only because of their mathematization and measurability, but also because the relationship between the actual states of a real system and the points of the geometric model was believed to be ontologically necessary, while in modern systems it is a fiction maintained for the sake of theory, prediction, and so on.

The simplest scheme is the one-parameter model. Early medical experience with mammals shows that the state of health or illness can be correlated with the parameter of temperature. Many animals correlate observable features with the emotional states of other animals: the ear attitude of a dog corresponds to its state of fear, while its fang exposure is a qualitative “parameter” for its degree of rage. A combination of both attitudes represents a more adequate characterization of the dog’s
emotional state. The state of a planet in medieval cosmology can be defined by its angular velocity and localization. States of other systems may need more than two features (e.g., temperature, blood pressure, and pulse rate for the healthy state of a mammal).

In any case, if these parameters are numerical, then the corresponding state spaces can be represented by geometric spaces. Thus the values of two numerical parameters may be represented by a single point in a two-dimensional state space visualized by the plane of Euclidean geometry. Changes in the actual state of the system are observed and may be represented as a curve in the state space. If each point of this curve carries a label recording the time of observation, then we get a trajectory of the model. Sometimes it is useful to introduce another coordinate of time and to represent the changing parameters of states by its time series. This kind of representation is called the graph of a trajectory.

The dynamical concepts of the Middle Ages included both kinds of representation. In the 1350s, the Parisian scholastic Nicole Oresme introduced the concept of graphical representations or geometrical configurations of intensities of qualities. Oresme mainly discussed the case of a linear quality whose extension is measured by an interval or line segment of space or time (“longitude of the quality”). He proposed to measure the intensity of the quality at each point of the interval by a perpendicular ordinate (“latitude of the quality”) at that point. The quantity of a linear quality is visualized by the configuration of both parameters. In the case of a uniformly accelerated motion during a time interval corresponding to the longitude $AB$ in Fig. 2.3, the latitude at each point $P$ of $AB$ is an ordinate $PQ$ whose length is the velocity at the corresponding instant [2.13]. The straight line $DC$ of the configuration is the graph of a trajectory representing the state of velocity. The so-called Merton Rule is immediately derived with a geometrical verification of Fig. 2.3: namely, it follows from the formula for the area of the trapezoid in Fig. 2.3 that the total distance traveled is $s = \frac{1}{2} (v_0 + v_f) t$.

Perhaps this interpretation was found on the basis of regarding this area as made up of very many vertical segments (“indivisibles”), each representing a velocity con-

Fig. 2.3. Oresme’s coordinates of a linear quality
tinued for a very short ("infinitesimal") time. The Merton Rule shows that even in the very early beginning of state space approaches a good geometric representation is not only a useful visualization, but gives new insight into the concepts of dynamics. Of course, Oresme and the Merton scholars at first only wanted to mathematize an Aristotelean-like physics of qualities. But their work was widely disseminated in Europe and led to the work of Galileo. In his famous Discorsi (1638), he introduced the basic concepts of modern mechanics and proceeded to the well-known distance formula $s = \frac{1}{2}gt^2$ for uniformly accelerated motion from rest (free fall) with a proof and an accompanying geometric diagram that are similar to Oresme’s ideas.

With Newton and Leibniz, something new was added to the theory of dynamical systems. The calculus allows one to compute the instantaneous velocity as the derivative of a velocity function and to visualize it as the tangent vector of the corresponding curve (Fig. 2.4a). The velocity vector field has become one of the basic concepts in dynamical systems theory (Fig. 2.4b). Trajectories determine velocity

![Fig. 2.4a–c. Geometric representation of a dynamical system: (a) Instantaneous velocity as tangent vector, (b) velocity vector field, (c) phase portrait](image)
vectors by the differentiation procedure of the calculus. Conversely, the integration procedure of the calculus allows one to determine trajectories from velocity vectors.

The strategy of modeling a dynamical system begins with the choice of a state space in which observations can be represented by several parameters. Continuing the observations leads to many trajectories within the state space. In the sense of Newton’s and Leibniz’ calculus, a velocity vector may be derived at any point of these curves, in order to describe their inherent dynamical tendency at any point. A velocity vector field is defined by prescribing a velocity vector at each point in the state space. The state space filled with trajectories is called the “phase portrait” of the dynamical system (Fig. 2.4c). This basic concept of dynamical system theory was originally introduced by Henri Poincaré. The velocity vector field was derived from the phase portrait by differentiation [2.14].

Of course, the velocity vector field visualizes the dynamics of the particular system being modeled. Actually, extensive observations over a long period of time are necessary to reveal the dynamical tendencies of the system which is represented by the corresponding velocity vector field. The modeling procedure is only adequate if we assume that (a) the velocity vector of an observed trajectory is at each point exactly the same as the vector specified by the dynamical system and (b) the vector field of the model is smooth. The word “smooth” means intuitively that there are no jumps and no sharp corners. In the case of a one-dimensional state space, the vector field is specified by a graph in the plane. Thus, the graph is smooth if it is continuous and its derivative is continuous as well. Historically, condition (b) corresponds to Leibniz’ famous principle of continuity playing a dominating role in the framework of classical physics.

In general, we summarize the modeling process as follows. A dynamical model is prepared for some experimental situation. We may imagine the laboratory devices of physicists like Galileo and Newton or biologists observing some organisms or even sociologists working on some social groups. The dynamical model consists of the state space and a vector field. The state space is a geometrical space (e.g., the Euclidean plane or in general a topological manifold) of the experimental situation. The vector field represents the habitual tendencies of the changing states and is called the dynamics of the model. How can we find the trajectories, thus the behaviour of the system? Technically, this problem is solved by creating the phase portrait of the system. That means we have to construct the trajectories of the dynamical system. Given a state space and a (“smooth”) vector field, a curve in the state space is a trajectory of the dynamical system if its velocity vector agrees with the vector field in the sense of tangent vectors (Fig. 2.5). The point corresponding to time zero is called the initial state of the trajectory. These trajectories are supposed to describe the behaviour of systems as observed over an interval of time. Moreover, physicists have been ambitious enough to aim at making predictions indefinitely far into the future and to calculate the course of nature as if it were a huge clock.

Let us have a short glance at Newton’s cosmos, which seemed to be a successful application of dynamical system theory evolving by the mathematical tools of Newton, Leibniz, Euler, etc. Newton gave three laws governing the behavior of material bodies. The first law (“lex inertiae”) says that a body continues to move uniformly
in a straight line if no force acts on it. If a force does act on it, then its mass times its acceleration is equal to that force (second law). The basic framework is completed by a third law: to every action there is always opposed an equal reaction. The Newtonian cosmos consists of particles moving around in a space which obeys the laws of Euclidean geometry. The accelerations of these particles are determined by the forces acting upon them. The force on each particle is obtained by adding together all the forces of other particles in the sense of the vector addition law. If the force is a gravitational one, then it acts attractively between two bodies and its strength is proportional to the product of the two masses and the inverse square of the distance between them. But, of course, there may be other types of forces.

Actually, Newton’s second law was understood as a universal scheme for all forces of nature in the macrocosmos and microcosmos. With a specific law of force the Newtonian scheme translates into a precise system of dynamical equations. If the positions, velocities, and masses of the various particles are known at one time, then their positions and velocities are mathematically determined for all later times. In short, the state of a body in Newton’s cosmos is specified by the two parameters of position and velocity. The Newtonian trajectories are determined by the dynamical equations of motion. If the initial states were known, then the behavior of Newton’s cosmos seemed to be determined completely. This form of determinism had a great influence on the philosophy of the 18th and 19th centuries. Newton’s dynamics was understood as basic science for modeling nature. But, of course, the mechanistic models are valid only in the limiting case of vanishing friction and are never fully achieved experimentally. Nature is so complex that physicists preferred to observe unnatural (“artificial”) limiting cases. Later on we shall see that the physicists’ belief in simple laws completely neglected the complexity of initial conditions and constraints and, thus, created an illusory model of a deterministic as well as computable nature.
According to Newton, there is only one real world of matter in one absolute framework of space-time, in which we may choose relative reference systems. This means that for any two events it is regarded as objectively decidable whether they are simultaneous and also whether they occur at the same place. Mathematically, Newton’s absolute space is represented by a 3-dimensional Euclidean space the metric of which is measurable by means of rulers, while time is taken to be a 1-dimensional Euclidean space with coordinate $t$ which is measured by standard clocks.

Because of its absolute simultaneity the Newtonian 4-dimensional space-time is stratified by maximal subsets of simultaneous events. Each stratum is a possible 3-dimensional hyperplane $t = t(e)$ of an event $e$ which separates its causal future, with strata $t > t(e)$, from its causal past, with strata $t < t(e)$. In Fig. 2.6a the third spatial coordinate is neglected, in order to visualize each stratum as 2-dimensional plane. This causal structure includes the Newtonian assumption that there are arbitrarily fast signals by means of instantaneous action at a distance [2.15].

Newton’s relative spaces are made precise by Lange as inertial systems designating reference systems for a force-free body moving in a straight line with a steady velocity. It is not stipulated which of the many possible inertial systems is used. Particular transformations (Galilean transformations) from one inertial system to another give the corresponding coordinates. Mechanical laws are preserved (invariant) with respect to these transformations. As every Galilean transformation has ten continuous parameters (one parameter for time and three times three parameters for rotation, steady velocity and translation), we can derive ten laws of conservation. Thus, e.g., the Galilean invariance of the time coordinate implicates the law of conservation of energy. Reference systems which are not inertial systems have typical effects. A disk rotating relative to the fixed stars has radial forces which cannot be eliminated by Galilean transformations. In short, in Newtonian space-time, uniform motions are considered as absolutely preferred over accelerated motions. Its structure is defined by the group of Galilean transformations.

![Fig. 2.6a. Newtonian space-time model with spatial strata of simultaneous events and trajectories of uniform inertial movements (straight lines) and accelerations (curve)](image-url)
At the beginning of this century, Einstein proved that Newton’s model of spacetime is restricted to mechanical motions with slow speed relative to the speed \( c \) of light. The constancy of \( c \) independently of any moving reference system is a fact of Maxwell’s electrodynamics. Thus, Newton’s addition law of velocities and Galilean invariance cannot hold true in electrodynamics. In his theory of special relativity (1905), Einstein assumed the constancy of the speed of light and the invariance of physical laws with respect to all inertial systems (“principle of special relativity”) and derived a common framework of space-time for electrodynamics and mechanics. Einstein’s special relativistic space-time was modeled by Minkowski’s four-dimensional geometry. The four-dimensionality should not surprise us, because Newton’s space-time has three (Cartesian) space and one time coordinate, too.

For the sake of simplicity, the units are chosen in a way that the speed of light is equal to one, and, thus, the units of length and time can be exchanged. Each point in this space-time represents an event, which means a point in space at a single moment. As a particle persists in time, it is not represented by a point, but by a line which is called the world-line of the particle. In order to visualize the Minkowskian model, we depict a space-time system with a standard time coordinate, measured in the vertical direction, and two space coordinates, measured in the horizontal direction (Fig. 2.6b) [2.16].

Uniformly moving particles are represented by straight lines, accelerated particles by curved lines. As particles of light (photons) uniformly travel with the fundamental speed \( c \), their world-lines are straight lines at an angle of 45° to the vertical. They form a light cone centred at the common origin 0. The system of light cones

![Fig. 2.6b,c. Minkowskian space-time cone in special relativity (b), and the twin brother paradox of special relativity: the Minkowskian distance \( RQ \) is greater than the length of \( RS \) and \( SQ \) together (c) ](image-url)
at all space-time points is regarded as the Minkowskian model of relativistic space-time.

Whereas the world-line of a photon is always along the light cone at each point, the world-line of any accelerated or uniformly moved material particle with a speed slower than c must always be inside the light cone at each point. As material particles or photons cannot travel faster than light, only the world-lines along and inside the light cone are determined physically. An event is called later than 0, if it is in the future cone above 0; it is called earlier than 0, if it is in the past cone below 0. Thus, the light cones determine the causal structure of relativistic space-time.

An essential difference between the Minkowskian model and ordinary Euclidean representations is the fact that the length of world-lines is interpreted as the time measured by physical clocks. Thus, time measurement becomes path-dependent, contrary to Newton’s assumption of an absolute time. The so-called “twin paradox” visualizes this effect dramatically. In Fig. 2.6c, one twin brother remains on the earth R moving uniformly and very slowly, while the other makes a journey to a nearby star S at great speed, nearly that of light. Minkowskian geometry forecasts that the travelling brother is still young upon his return at Q, while the stay-at-home brother is an old man. This is not science fiction, of course, but a consequence of the time-measuring length of Minkowskian world-lines: the Minkowskian distance RQ is greater than the length of the distance RS and SQ together, contrary to the usual Euclidean interpretation. Today, these effects are experimentally well confirmed for elementary particles at high speeds near c.

In the framework of Minkowskian space-time, the invariance of physical laws with respect to particular inertial systems is realized by the Lorentz transformation. Newtonian space-time with Galilean invariance remains a limiting case for reference systems like celestial motions of planets or earthly motions of billiard balls with slow speed relative to the constant c. In this sense, Einstein’s space-time is the culmination of classical physics rather than a revolutionary break with Newton.

An important concept which was first introduced into classical physics by Leibniz is energy, consisting of the kinetic energy $T$ and the potential energy $U$ of a system. The mechanical work done on a point mass which is displaced from a position 1 to a position 2 corresponds to the difference between the kinetic energy at position 1 and that of position 2. If this mechanical work is independent of the path followed from 1 to 2, then the corresponding force field is called conservative. Frictional forces are not conservative. In one dimension all forces must be conservative, since there is a unique path from one point to another point in a straight line, ignoring friction. The total energy $T + U$ is constant in a conservative field of force.

An important application of Newton’s mechanics is the harmonic oscillator, such as the small amplitude pendulum, or the weight oscillating up and down on a spring. The harmonic oscillator appears as a model through all parts of physics and even chemistry and biology. For example, remember electromagnetic light waves, where the electric and magnetic field energies oscillate. Harmonic oscillations are also well known in technology, for example as oscillating electrical currents in a coil and a condenser, with friction corresponding to the electrical resistance. In the philosophy of the 18th and 19th centuries the pendulum was a symbol of the mech-
anistic universe which seemed to be completely determined and calculable by the Newtonian equations of motion.

Thus, the pendulum may be considered as a classical example of the dynamical modeling procedure. This model assumes that the rod is very light, but rigid. The hinge at the top is perfectly frictionless. The weight at the lower end is heavy, but very small. The force of gravity always pulls it straight down. In Fig. 2.7a, the pendulum is drawn in a two-dimensional Euclidean plane with the angle $\alpha$ of elevation, the force $F$ of gravity, the pull $F \cos \alpha$ along the rod, and the force $F \sin \alpha$ turning it. In order to visualize the dynamical behavior of the pendulum we have to develop a dynamical model with a state space and a phase portrait. The state of the pendulum is fully determined by the angular variable $\alpha$ (with $\alpha = 0$ and $\alpha = 2\pi$ denoting the same angle) and the angular velocity $v$. Thus, we get a two-dimensional state space which can be visualized by the circular cylinder in Fig. 2.7b. The vertical circle in the center of this cylinder denotes the states of zero angular velocity $v = 0$. The straight line from front to back, at the bottom of the cylinder, is the axis of zero inclination with $\alpha = 0$, where the pendulum is lowest. At the origin with $(\alpha, v) = (0, 0)$, the pendulum is at rest at its lowest position [2.17].

As there is no friction and no air in the way, moving the pendulum a little to the left causes it to swing back and forth indefinitely. The full trajectory in the state space, corresponding to this oscillating motion, is a cycle, or closed loop. In the next case, the pendulum is balanced at the top, in unstable equilibrium. A tiny touch on the left causes it to fall to the right and pick up speed. The angular velocity reaches its maximum when the pendulum passes the bottom of the swing. On the way back up to the top again, the pendulum slows down. Then the pendulum balances at the top again. But when the pendulum at the beginning of its rotation shoved hard to the right, then its rate of angular velocity is rather large. Moving back up again, it

![Fig. 2.7a,b. Dynamical system (pendulum) (a) with 2-dimensional state space (circular cylinder) (b) [2.17]](image-url)
slows down, but not enough to come to rest at the top. Thus, the pendulum rotates clockwise indefinitely. The corresponding trajectory in the cylindrical state space is a cycle. Unlike the slow oscillation, the fast cycle goes around the cylinder. Performing many experiments would reveal the phase portrait of this dynamical model (Fig. 2.8a). There are two equilibrium points. At the top, there is a saddle point. At the origin, there is a vortex point which is not a limit point of the nearby trajectories. The phase portrait is easier to see when the cylinder is cut open along the straight line from front to back through the saddle point at the top (Fig. 2.8b).

If the system is not closed and the effects of friction are included as in physical reality, then the equilibrium point at the origin is no longer a vortex point (Fig. 2.8c). It has become a spiraling type of point attractor. As any motion of the pendulum will come to rest because of friction, any trajectory representing a slow motion of the pendulum near the bottom approaches this limit point asymptotically.

In two dimensions or more, other types of trajectories and limit sets may occur. For example, a cycle may be the asymptotic limit set for a trajectory (Fig. 2.9), or in a three-dimensional system a torus or even other more or less strange limit sets may occur.

Fig. 2.8a,b. Phase portrait of the pendulum on the cylindrical state space (a) and cut open into a plane (b). (c) Phase portrait of the pendulum with friction
Fig. 2.9. Cycle as asymptotic limit set for a trajectory

Limit sets enable us to model a system’s evolution to its equilibrium states. The key concepts are limit sets called “attractors” [2.18]. Mathematically, a limit set (limit point, cycle, torus, etc.) is called an attractor if the set of all trajectories approaching this limit set asymptotically is open. Roughly speaking, attractors receive most of the trajectories in the neighborhood of the limit set. Of all limit sets which represent possible dynamical equilibria of the system, the attractors are the most prominent. In the case of a limit point, an attractor represents a static equilibrium, while a limit cycle as attractor designates the periodic equilibrium of an oscillation. Vibrations on a pendulum, spring, or musical instrument are only a few of the mechanical applications. As we will see later on, periodic equilibria of oscillating dynamical systems play an important role in physics, chemistry, biology, and social sciences.

In a typical phase portrait, there will be more than one attractor. The phase portrait will be divided into their different regions of approaching trajectories. The dividing boundaries or regions are called separatrices. In Fig. 2.10, there are two point attractors with their two open sets of approaching trajectories and their separatrix.

In reality, a dynamical system cannot be considered as isolated from other dynamical systems. In order to get more adequate models, we will study two coupled systems. A simple example is provided by coupling two clocks. Historically, this particular system was observed by Christian Huygens in the 17th century. He noticed that two clocks hanging on the same wall tend to synchronize. This phenomenon is caused by nonlinear coupling through the elasticity of the wall. Indeed, any two dynamical systems can be combined into a single system by constructing the Cartesian product of the two corresponding state spaces. A small perturbation of this combined system is called a coupling of the two systems. The geometric model for the states of this combined system is formed as follows [2.19].

Each clock $A$ and $B$ is a kind of oscillator. For the sake of visualizing the asymptotic behaviour of both oscillators, the transient behavior is ignored and the two-
dimensional state model of the Euclidean plane with a limit cycle around the origin for the two parameters of displacement and velocity is replaced by the limit cycle alone. A state of oscillator $A$ is specified by an angle $\alpha$ corresponding to its phase (Fig. 2.11a), a state of oscillator $B$ by an angle $\beta$ (Fig. 2.11b).

Fig. 2.10. Phase portrait with two point attractors, two open sets of approaching trajectories, and a separatrix

Fig. 2.11a,b. Two clocks as oscillators with two cycles as their corresponding state spaces
In order to construct the state space for the combined system of both oscillators, we consider the limit cycle of clock $A$ in a horizontal plane. Each point of this horizontal cycle represents a phase state of $A$. We consider such a point as the center of the limit cycle of clock $B$ erected perpendicular to the horizontal plane of clock $A$ (Fig. 2.11c). Each point of this vertical cycle represents a phase state of $B$. The pair $(\alpha, \beta)$ of phases represents the state of the combined system [2.20].

If oscillator $A$ is stuck at phase $\alpha$ and oscillator $B$ moves through a full cycle, then the combined phase point traverses the vertical cycle in Fig. 2.11c. If oscillator $A$ also moves through a full cycle, then the vertical cycle in Fig. 2.11c is pushed around the horizontal cycle, sweeping out the torus in Fig. 2.11d. Thus, the state space for the combined system of two oscillators is the torus, which is the Cartesian product of the two cycles. Of course, the actual state model for two oscillators is four-dimensional and not only two-dimensional as in our reduced figures.

In order to get the phase portrait of the dynamical behavior for the combined system, we have to study the vector field and the trajectories on the state space of the torus. Let us first assume that each clock is totally indifferent to the state of the other. In this case, the clocks are uncoupled. The trajectory of a point on the torus corresponding to the time phase of each clock winds around the torus. If the rate of each clock is constant, then on the flat rectangular model of the torus, the trajectory is a straight line (Fig. 2.12). The slope of this line is the ratio of the rate of clock $B$ to the rate of clock $A$. If the two clocks run at the same rates, the ratio is one. Telling the same time means that both clocks have identical phases. Then the trajectory on the flat torus is the diagonal line in Fig. 2.12a.

A slight change in the system results in a slight change in the ratio of the rates or frequencies of the oscillators. Then the trajectory on the torus changes from a periodic trajectory to an almost periodic trajectory or to a periodic trajectory winding many times around, instead of just once (Fig. 2.12b). If two oscillators are coupled (for instance by Huygens’ shared wall for the two clocks), then a small vector field must be added to the dynamical model representing the uncoupled system. It is a noteworthy theorem of geometric analysis that the braid of trajectories on the torus is structurally stable in the sense that a small perturbation does not result in a significant change in the phase portrait. Experimentally, this result was already confirmed by Huygens’ observation of the synchronizing phenomenon of two clocks on the same wall.
Oscillators are a central dynamical paradigm for the modeling procedure of nature. They are by no means restricted to mechanical applications. In the 19th century, Hermann von Helmholtz invented an electrical vibrator and Lord Rayleigh studied coupled systems of vacuum tube oscillators used in the first radio transmitters. In this century, it was van der Pol who used the further development of radio frequency electronics in understanding coupled oscillators.

In the Newtonian universe, coupled oscillators provide examples of many-body problems. What can in general be said about the mechanics of point mass systems with several moving point masses, exerting forces upon each other? Systems with two point masses have simple and exact solutions. In a two-body problem with two point masses with isotropic central forces, the (twelve) unknowns are determined by the (ten) laws of conserved quantities and Newton’s laws of motion for the two particles. The problem of two-point masses can be successfully reduced to the already solved problem of a single point mass by considering Newton’s law of motion for the difference vector $r$ and reduced mass $\mu = m_1 m_2 / (m_1 + m_2)$ of both point masses $m_1$ and $m_2$. Historically, Galileo assumed that the earth moves around the sun, which is at rest. Thus he reduced the celestial motions to the simple case of a two-body problem. As we all know, the sun is actually moving around the combined centre of mass of the sun–earth system, which lies inside the surface of the sun. But this assumption is still inaccurate, of course, since many planets are simultaneously moving around the sun and all of them are exerting forces upon each other.

Another example of such a many-body-problem is given by a triple collision of three billiard balls. Provided that the balls collide only in pairs, and no triple or higher-order collisions occur, then the situation is reduced to two-body problems. The outcome depends in a continuous way on the initial state. Sufficiently tiny changes in the initial state lead only to small changes in the outcome. If three balls come together at once, the resulting behavior depends critically upon which balls come together first. Thus, the outcome depends discontinuously on the input, contrary to Leibniz’ principle of continuity, which he used basically to criticize
Descartes’ inquiries into percussion. In the Newtonian universe, many-body problems of billiard balls and planets can be described in deterministic models in the sense that physical behavior is mathematically completely determined for all times in the future and past by the positions and velocities of the balls or planets. But the models may be non-computable in practice and in the long run. In the case of planetary theory, numerical simulations on computers for many millions of years can produce very large errors, because the initial positions and velocities are not known exactly. A very tiny change in the initial data may rapidly give rise to an enormous change in the outcome. Such instabilities in behavior are typical for many-body problems. Even in a fully deterministic world, the assumption of a Laplacean demon which can calculate the Newtonian universe in the long run will eventually be exposed as an illusory fiction.

2.3 Hamiltonian Systems and the Chaos of Heaven and the Quantum World

In the 18th and 19th centuries, Newtonian mechanics seemed to reveal an eternal order of nature. From a modern point of view, Newtonian systems are only a useful kind of dynamical system for modeling reality. In order to specify the initial state of a Newtonian system, the positions and the velocities of all its particles must be known. Around the middle of the 19th century, a very elegant and efficient formalism was introduced by the mathematician William Hamilton [2.21]. His fruitful idea was to characterize a conservative system by a so-called Hamiltonian function $H$ which is the expression for the total energy (= sum of kinetic and potential energy) of the system in terms of all the position and momentum variables. While the velocity of a particle is simply the rate of change of its position with respect to time, its momentum is its velocity multiplied by its mass. Newtonian systems are described with Newton’s second law of motion in terms of accelerations, which are rates of change of rates of change of position. Thus, mathematically, they are defined by second-order equations. In the Hamiltonian formulation, there are two sets of equations. One set of equations describes how the momenta of particles are changing with time, and the other describes how the positions are changing with time. Obviously, Hamiltonian equations describe the rates of change of quantities (i.e., position or momentum). Thus, we get a reduction of mathematical description with first-order equations which are, of course, deterministic. For dynamical systems of $n$ unconstrained particles with three independent directions of space, there are $3n$ position coordinates and $3n$ momentum coordinates.

With suitable choices of the Hamiltonian function $H$, Hamiltonian equations can be used to characterize any classical dynamical system, not just Newtonian systems. Even in Maxwell’s electrodynamics, Hamiltonian-like equations deliver the rate of change with time of the electric and magnetic fields in terms of what their values are at any given time. The only difference is that Maxwell’s equations are field equations rather than particle equations, needing an infinite number of parameters to describe the state of the system, with field vectors at every single point in
space, rather than the finite number of parameters with three coordinates of position and three of momentum for each particle. Hamiltonian equations also hold true for special relativity and (in somewhat modified form) for general relativity. Even, the crucial step from classical mechanics to quantum mechanics is made by Bohr’s correspondence principle in the framework of the Hamiltonian formalism. These applications will be explained later on. Just now, it is sufficient to recall that Hamiltonian equations deliver a universal formalism for modeling dynamical systems in physics.

The corresponding state spaces allow us to visualize the evolution of the dynamical systems in each “phase”. Thus they are called phase spaces. For systems with \( n \) particles, phase spaces have \( 3n + 3n = 6n \) dimensions. A single point of a phase space represents the entire state of a perhaps complex system with \( n \) particles. The Hamiltonian equations determine the trajectory of a phase point in a phase space. Globally, they describe the rates of change at every phase point, and therefore define a vector field on the phase space, determining the whole dynamics of the corresponding system.

It is a well-known fact from empirical applications that states of dynamical models cannot be measured with arbitrary exactness. The measured values of a quantity may differ by tiny intervals which are caused by the measuring apparatus, constraints of the environment, and so on. The corresponding phase points are concentrated in some small regions of a neighborhood. Now, the crucial question arises if trajectories starting with neighboring initial states are locally stable in the sense that they have neighboring final states. In Fig. 2.13a, a phase state region \( R_0 \) of initial states of time zero is dragged along by the dynamics of the vector field to a region \( R_t \) at later time \( t \) (of course, the actual large number of coordinates must be neglected in such a visualization of a phase space) [2.22].

**Fig. 2.13.** (a) A phase state region \( R_0 \) at time 0 is dragged along by a Hamiltonian dynamics to a region \( R_t \) at later time \( t \) [2.22]. (b) According to Liouville’s theorem, the volume of an initial phase state region is conserved under a Hamiltonian dynamics, although its shape may be distorted, stretched, and spread outwards [2.22]
In this case, similar initial states lead to similar final states. This assumption is nothing else than a classical principle of causality in the language of Hamiltonian dynamics: similar causes lead to similar effects. Historically, philosophers and physicists from Leibniz to Maxwell believed in this causal principle, which seemed to secure the stability of the measuring process and the possibility of forecasts despite an appreciable interval of inaccuracy.

It is noteworthy that the representation in the Hamiltonian formalism allows a general statement about the causality of classical dynamical systems. Due to a famous theorem of the mathematician Liouville, the volume of any region of the phase space must remain constant under any Hamiltonian dynamics, and thus for any conservative dynamical system. Consequently, the size of the initial region $R_0$ in Fig. 2.13a cannot grow by any Hamiltonian dynamics if we understand “size” in the right manner as phase-space volume. But its conservation does not exclude that the shape of the initial region is distorted and stretched out to great distances in the phase space (Fig. 2.13b) [2.22].

We may imagine a drop of ink spreading through a large volume of water in a container. That possible spreading effect in phase spaces means that the local stability of trajectories is by no means secured by Liouville’s theorem. A very tiny change in the initial data may still give rise to a large change in the outcome. Many-body problems of celestial mechanics and billiard balls remain non-computable in the long run, although their dynamics are deterministic. Nevertheless, Liouville’s theorem implies some general consequences concerning the final regions which can be displayed by Hamiltonian dynamics, and thus by conservative dynamical systems. Remember the phase portrait Fig. 2.8c of a pendulum with friction (which is not a conservative system) with a different equilibrium point at the origin. While the non-conservative system has a spiraling type of point attractor (Fig. 2.14a), the conservative system has a vortex point (Fig. 2.14b) which is not an attractor [2.23].

In Fig. 2.14a, trajectories are attracted to a field point, and the volume of an initial area shrinks. In Fig. 2.14b, the trajectories rotate around a vortex point, and the volume of an initial area is conserved. Thus, due to Liouville’s theorem, we can generally conclude that in any conservative system attracting points must be excluded.

![Fig. 2.14a,b. Point attractor of a non-conservative system without conservation (a), vortex point of a conservative system with conservation (b)](image)
The effect of shrinking initial areas can easily be visualized for the trajectories of limit cycles, too. So, limit cycles as attractors are also not possible in conservative systems for the same mathematical (a priori) reasons.

These results are derived a priori by a far-reaching mathematical theorem of Hamiltonian systems. We must be aware that conservative physical systems like planetary systems, pendula, free fall, etc., are only some of the empirical applications of Hamiltonian systems. Hamiltonian systems are defined by a particular kind of mathematical equation (Hamiltonian equations). Features of Hamiltonian systems are derived from the mathematical theory of the corresponding equations. Consequently, modeling reality by Hamiltonian systems means that we can epistemically forecast some a priori features, e.g., that no static equilibrium of a limit point attractor and no periodic equilibrium of a limit cycle attractor can be expected.

Philosophically, this point of view obviously conforms to Kant’s epistemology in some modified sense. If we assume the mathematical framework of some dynamical systems, then, of course, we can assert some a priori statements about our empirical models, independently of their empirical applications in several sciences. But Kant’s epistemology and the dynamical system approach differ in the following sense: not only is there one categorial framework (e.g., Newtonian systems), but there are many kinds of systems modeling reality with more or less success. So, it will not be physicalist or reductionist to apply conservative systems even in cognitive and economical science, later on.

A further a priori result of Hamiltonian (conservative) systems says that there are irregular and chaotic trajectories. In the 18th and 19th centuries, physicists and philosophers were convinced that nature is determined by Newtonian- or Hamiltonian-like equations of motion, and thus future and past states of the universe can be calculated at least in principle if the initial states of present events are well known. Philosophically, this belief was visualized by Laplace’s demon, which like a huge computer without physical limitations can store and calculate all necessary states. Mathematically, the belief in Laplace’s demon must presume that systems in classical mechanics are integrable, and, thus are solvable. In 1892, Poincaré was already aware that the non-integrable three-body problem of classical mechanics can lead to completely chaotic trajectories [2.24]. About sixty years later, Kolmogorov (1954), Arnold (1963) and Moser (1967) proved with their famous KAM theorem that motion in the phase space of classical mechanics is neither completely regular nor completely irregular, but that the type of trajectory depends sensitively on the chosen initial conditions [2.25].

As celestial mechanics is an empirically well confirmed dynamical model of a Hamiltonian system, the KAM theorem refutes some traditional opinions about the “superlunar” world. Heaven is not a world of eternal regularity, either in the sense of Aristotle’s cosmos or in the sense of Laplace’s demon. Obviously, it is not the seat of the Gods. Nevertheless, it is not completely chaotic. Heaven, as far as recognized by Hamiltonian systems, is more or less regular and irregular. It seems to have more similarity with our human everyday life than our forefathers believed. This may be a motivation for writers to be curious about Hamiltonian systems. But now let us see some mathematical facts.
One of the simplest examples of an integrable system is a harmonic oscillator. Practically, the equations of motion of any integrable system with \( n \) degrees of freedom are the same as those of a set of \( n \) uncoupled harmonic oscillators. The corresponding phase space has \( 2n \) dimensions with \( n \) position coordinates and \( n \) momentum coordinates. For a harmonic oscillator with \( n = 1 \) we get a circle, and for two harmonic oscillators with \( n = 2 \) a torus (compare Fig. 2.11d). Thus, the existence of \( n \) integrals of motion confines the trajectories in the \( 2n \)-dimensional phase space of an integrable system to an \( n \)-dimensional manifold which has the topology of an \( n \)-torus. For an integrable system with two degrees of freedom and a four-dimensional phase space, the trajectories can be visualized on a torus. Closed orbits of trajectories occur only if the frequency ratios of the two corresponding oscillators are rational (Fig. 2.15). For irrational frequency ratios, the orbit of a trajectory never repeats itself, but approaches every point on the torus infinitesimally closely [2.26].

A nonintegrable system of celestial mechanics was studied by Hénon and Heiles in 1964. The dynamical model consists of an integrable pair of harmonic oscillators coupled by nonintegrable cubic terms of coordinates. If the initial state of the model with two position coordinates \( q_1, q_2 \) and two momentum coordinates \( p_1, p_2 \) is known, then its total energy \( E \) is determined by the corresponding Hamiltonian function \( H \) depending on these position and momentum coordinates. The trajectories of the system move in the four-dimensional phase space on a three-dimensional hyperplane which is defined by \( H(q_1, q_2, p_1, p_2) = E \).

The values of \( E \) can be used to study the coexistence of regular and irregular motions which was forecast by the KAM theorem. For small values of \( E \), the dynamical system has regular behavior, while for large values it becomes chaotic. In order to visualize this changing behavior, we consider the intersections of the trajectories with the two-dimensional plane of coordinates \( q_1 \) and \( q_2 \) (Poincaré maps). For \( E = \frac{1}{24} \) (Fig. 2.16a) and \( E = \frac{1}{12} \) (Fig. 2.16b), the Poincaré maps show the intersections of somewhat deformed tori which signal regular motion. Above a critical value of \( E = \frac{1}{3} \), most, but not all, tori are destroyed, and spots of irregular points appear to be random. For \( E = \frac{1}{8} \) (Fig. 2.16c), the Poincaré map illustrates a state of transition with the coexistence of regular and irregular motions. For \( E = \frac{1}{6} \) (Fig. 2.16d), the motion seems to be almost completely irregular and chaotic [2.27].
An empirical application is given in the following three-body problem of celestial mechanics, which is nonintegrable. Consider the motion of Jupiter perturbing the motion of an asteroid around the sun (Fig. 2.17).

Jupiter and the asteroid are interpreted as two oscillators with certain frequencies. According to the KAM theorem, stable and unstable motions of the asteroid can be distinguished, depending on the frequency ratio.

In general, we must be aware that stable as well as unstable trajectories are mathematically well defined. Consequently, even nonintegrable many-body problems describe deterministic models of the world. Metaphorically, we may say that the God of Leibniz and Newton would have no difficulty in forecasting regular and irregular trajectories *sub specie aeternitatis* and does not need to calculate their de-
development step by step. The observed chaotic behavior is neither due to a large number of degrees of freedom (a celestial three-body problem has rather few degrees of freedom) nor to the uncertainty of human knowledge. The irregularity is caused by the nonlinearity of Hamiltonian equations which let initially close trajectories separate exponentially fast in a bounded region of phase. As their initial conditions can only be measured with finite accuracy, and errors increase exponentially fast, the long-term behavior of these systems cannot be predicted. Mathematically, initial conditions are characterized by real values which may be irrational numbers with infinite sequences of digits. Thus, computer-assisted calculations will drive the errors faster and faster with improved measurement of more and more digits.

The macrocosmos of celestial mechanics, the world of asteroids, planets, stars, and galaxies, is determined by the coexistence of regular and irregular behavior. Deterministic chaos in the heavens is not everywhere, but locally possible, and thus may cause cosmic catastrophes which cannot be excluded in principle. What about the microcosmos of quantum mechanics, the quantum world of photons, electrons, atoms, and molecules? Is there chaos in the quantum world? In order to answer this question, we first must remind the reader of some basic concepts of Hamiltonian systems and phase spaces corresponding to objects in the quantum world [2.28].

In 1900, Max Planck proposed that electromagnetic oscillations occur only in quanta, whose energy $E$ bears the definite relation $E = h\nu$ to the frequency $\nu$ depending on the constant $h$ (“Planck’s quantum”). Besides Einstein’s huge constant $c$ of light’s speed, Planck’s tiny constant of quanta is the second fundamental constant of nature, according to 20th century physics. Planck’s relation was experimentally supported by, e.g., the radiation of black bodies. In 1923, Louis de Broglie proposed that even the particles of matter should sometimes behave as waves. De Broglie’s wave-frequency $\nu$ for a particle of mass $m$ satisfies the Planck relation. Combined with Einstein’s famous theorem $E = mc^2$ of Special Relativity (“mass is a particular state of energy and can therefore be transformed into energy by radiation”), we get a relation telling us that $\nu$ is related to $m$ by $h\nu = mc^2$. It follows that in the quantum world, fields oscillating with some frequency can occur only in discrete units of mass, depending on Planck’s and Einstein’s constants. Obviously, in the
quantum world, phenomena can be considered as waves as well as particles. This so-called particle-wave duality was well confirmed by many experiments which reveal features of waves or particles for quantum systems like photons or electrons, depending on the preparation of experimental conditions.

In 1913, Niels Bohr introduced his “planetary” model for the atom which could explain the observed and measured discrete stable energy levels and spectral frequencies with surprising accuracy. Bohr’s rules required that the angular momentum of electrons in orbit about the nucleus can occur only in integer multiples of \( h = h/2\pi \). His successful, albeit somewhat ad hoc rules only delivered an approximate geometric model which must be derived from a dynamical theory of the quantum world, corresponding to Newtonian or Hamiltonian classical mechanics which can explain Kepler’s planetary laws. The dynamics of the quantum world was founded by Heisenberg’s and Schrödinger’s quantum mechanics, which became the fundamental theory of matter in 20th century physics.

The main concepts of quantum mechanics can be introduced heuristically by analogy with corresponding concepts of Hamiltonian mechanics if some necessary modifications depending on Planck’s constant are taken into account. This procedure is called Bohr’s correspondence principle (Fig. 2.18). So, in quantum mechanics, classical vectors like position or momentum must be replaced by some operators satisfying a non-commutative (non-classical) relation depending on Planck’s constant. If \( h \) disappears (\( h \rightarrow 0 \)), then we get the well known classical commutative relations of, e.g., position and momentum which allow us to measure both vectors together with arbitrary accuracy. An immediate consequence of non-commutative relations in quantum mechanics is Heisenberg’s uncertainty principle \( \Delta p \Delta q \geq \hbar /2 \). If one measures the position \( q \) with precision \( \Delta q \), then one disturbs the momentum \( p \) by \( \Delta p \). Thus, it is obvious that there are no trajectories or orbits in the quantum world which demand precise values of both the position and the momentum of a particle. Bohr’s popular electronic orbits are only very rough geometric visualizations [2.29].

According to Bohr’s correspondence principle, classical systems described by Hamiltonian functions must be replaced by quantum systems (e.g., electrons or pho-
tons) described by a Hamiltonian operator depending on operators (for position and momentum) instead of vectors. In classical physics, the states of Hamiltonian systems are determined by the points of a phase space. In quantum mechanics, the appropriate analogous concept is that of a Hilbert space. States of a quantum system are described by vectors of a Hilbert space spanned by the eigenvectors of its Hamiltonian operator.

In order to illustrate this mathematical remark a little bit more, let us imagine a single quantum particle. Classically, a particle is determined by its position in space and by its momentum. In quantum mechanics, every single position which the particle might have is a single alternative combined in a collection of all possible positions with complex-number weightings. Thus, we get a complex function of position, the so-called wave function $\psi(x)$. For each position $x$, the value of $\psi(x)$ denotes the amplitude for the particle to be at $x$. The probability of finding the particle in some small fixed-sized interval about this position is obtained by forming the squared modulus of the amplitude $|\psi(x)|^2$. The various amplitudes for the different possible momenta are also determined by the wave function. Thus, the Hilbert space is a complex state space of a quantum system.

The causal dynamics of quantum states is determined by a partial differential equation called the Schrödinger equation. While classical observables are commutative with always definite values, non-classical observables of quantum systems are non-commutative with generally no common eigenvector and consequently no definite eigenvalues. For observables in a quantum state only statistical expectation values can be calculated.

An essential property of Schrödinger’s quantum formalism is the superposition principle demonstrating its linearity. For example, consider two quantum systems which once interacted (e.g., a pair of photons leaving a common source in opposite directions). Even when their physical interaction ceases at a large distance, they remain in a common superposition of states which cannot be separated or located. In such an entangled (pure) quantum state of superposition an observable of the two quantum systems can only have indefinite eigenvalues. The superposition or linearity principle of quantum mechanics delivers correlated (entangled) states of combined systems which are highly confirmed by the EPR experiments. Philosophically, the (quantum) whole is more than the sum of its parts. Non-locality is a fundamental property of the quantum world which differs from classical Hamiltonian systems [2.30]. We shall return to this question in discussing the emergence of mind–brain and artificial intelligence (Chaps. 4–6).

Bohr’s correspondence principle lets the question arise of whether the existence of chaotic motion in classical Hamiltonian systems leads to irregularities in the corresponding quantum systems [2.31]. Our summary of basic quantum mechanical concepts gives some hints of changes which must be expected in passing from a classically chaotic system to its corresponding quantum mechanical version. In contrast to classical mechanics, quantum mechanics only allows statistical predictions. Although the Schrödinger equation is linear in the sense of the superposition principle and can be solved exactly, e.g., for a harmonic oscillator, and although the wave function is strictly determined by the Schrödinger equation, this does not
mean that the properties of a quantum state can be calculated exactly. We can only calculate the density of probability to find a photon or electron at a space-time point.

Because of Heisenberg’s uncertainty principle, there are no trajectories in the quantum world. Therefore, the determination of chaos with the exponentially fast separation of close trajectories is not possible for quantum systems. Another aspect of the uncertainty principle concerning chaos is noteworthy: remember a classical phase space with chaotical regions like in Fig. 2.16. The uncertainty principle implies that points in a $2n$-dimensional phase space within a volume $\hbar^n$ cannot be distinguished. The reason is that chaotic behavior smaller than $\hbar^n$ cannot be represented in quantum mechanics. Only the regular behavior outside these chaotic regions could be expected. In this sense, the tiny, but finite value of Planck’s constant could suppress chaos.

In quantum mechanics, one distinguishes between time-independent stationary and time-dependent Hamiltonian systems. For systems with stationary Hamiltonians the Schrödinger equation can be reduced to a so-called linear eigenvalue problem which allows one to calculate the energy levels of, e.g., a hydrogen atom. As long as the levels are discrete, the wave function behaves regularly, and there is no chaos. The question arises of whether there are differences between the energy spectra of a quantum system with a regular classical limit and a quantum system whose classical version displays chaos. Time-dependent Hamiltonians are used to described the time-evolution of, e.g., elementary particles and molecules.

According to Bohr’s correspondence principle, quantum chaos can be detected by starting with the investigation of some classical Hamiltonian systems. They may be integrable, almost integrable, or chaotic. Thus, the trajectories in the hyperplane of energy may be regular, almost regular, or almost chaotic. Quantizing the Hamiltonian function by replacing the vectors of position and momentum with the corresponding operators, we get the Hamiltonian operator of the corresponding quantum system. In the next step, the Schrödinger equation and eigenvalue equation can be derived. Now, we may ask if the properties of the classical system with its integrable, almost integrable, or chaotic behavior can be transferred to the corresponding quantum system. What about the spectrum, eigenfunctions, etc.? These questions are summarized under the title “quantum chaos”. For instance, there are calculations which show that the energy spectrum of a free quantum particle in a stadium, for which the classical motion is chaotic, differs drastically from that of a free quantum particle in a circle, for which the classical motion is regular.

In Fig. 2.19, the distribution of distances between neighboring levels is illustrated with two examples [2.32]. In Fig. 2.19a,b, a system consisting of two coupled oscillators is shown for two different values of the coupling coefficient. While the corresponding classical dynamics of Fig. 2.19a is regular, the classical dynamics of Fig. 2.19b is almost chaotic.

Figure 2.19c,d shows the example of a hydrogen atom in a uniform magnetic field. While the corresponding classical dynamics of Fig. 2.19c is regular, the classical dynamics of Fig. 2.19d is almost chaotic. The regular and chaotic cases can be distinguished by different distributions of energy levels (Poisson and Wigner distributions) which are calculated by solving the corresponding Schrödinger equation.
They are confirmed by several numerical models as well as by measurements in laboratories with laser spectroscopy. In this sense, quantum chaos is no illusion, but a complex structural property of the quantum world. Hamiltonian systems are a key to discovering chaos in the macro- and microcosmos. But, of course, we must not confuse the complex mathematical structure of deterministic chaos with the popular idea of disorder.

2.4 Conservative and Dissipative Systems and the Emergence of Order

Since Poincaré’s celestial mechanics (1892), it was mathematically known that some mechanical systems whose time evolution is governed by nonlinear Hamiltonian equations could display chaotic motion. But as long as scientists did not have suitable tools to deal with nonintegrable systems, deterministic chaos was considered as a mere curiosity. During the first decades of the 20th century, many numerical procedures were developed to deal with the mathematical complexity of nonlinear differential equations at least approximately. The calculating power of modern high speed computers and refined experimental techniques have supported the recent successes of the nonlinear complex system approach in natural and social sciences. The visualizations of nonlinear models by computer-assisted techniques promote interdisciplinary applications with far-reaching consequences in many branches of science. In this scientific scenario (1963), the meteorologist Edward Lorenz, a student of the famous mathematician Birkhoff [2.33], observed that a dynamical system with three coupled first-order nonlinear differential equations can lead to completely chaotic trajectories. Mathematically, nonlinearity is a necessary, but not sufficient condition of chaos. It is necessary condition, because linear differential equations can be
2.4 Conservative and Dissipative Systems and the Emergence of Order

solved by well-known mathematical procedures (Fourier transformations) and do not lead to chaos. The system Lorenz used to model the dynamics of weather differs from Hamiltonian systems à la Poincaré mainly by its dissipativity. Roughly speaking, a dissipative system is not conservative but “open”, with an external control parameter that can be tuned to critical values causing the transitions to chaos.

More precisely, conservative as well as dissipative systems are characterized by nonlinear differential equations \( \dot{x} = F(x, \lambda) \) with a nonlinear function \( F \) of the vector \( x = (x_1, \ldots, x_d) \) depending on an external control parameter \( \lambda \). While for conservative systems, according to Liouville’s theorem, the volume elements in the corresponding phase space change their shape but retain their volume in the course of time, the volume elements of dissipative systems shrink as time increases (compare Figs. 2.13, 2.14) [2.34].

Lorenz’s discovery of a deterministic model of turbulence occurred during simulation of global weather patterns. The earth, warmed by the sun, heats the atmosphere from below. Outer space, which is always cold, absorbs heat from the outer shell of the atmosphere. The lower layer of air tries to rise, while the upper layer tries to drop. This traffic of layers was modeled in several experiments by Bénard. The air currents in the atmosphere can be visualized as cross-sections of the layers. The traffic of the competing warm and cold air masses is represented by circulation vortices, called Bénard cells. In three dimensions, a vortex may have warm air rising in a ring, and cold air descending in the center. Thus, the atmosphere consists of a sea of three-dimensional Bénard-cells, closely packed as a hexagonal lattice. A footprint of such a sea of atmospheric vortices can be observed in the regular patterns of hills and valleys in deserts, snowfields, or icebergs.

In a typical Bénard experiment, a fluid layer is heated from below in a gravitational field (Fig. 2.20a). The heated fluid at the bottom tries to rise, while the cold liquid at the top tries to fall. These motions are opposed by viscous forces. For small temperature differences \( \Delta T \), viscosity wins, the liquid remains at rest, and heat is transported by uniform heat conduction. The external control parameter of the system is the so-called Rayleigh number \( Ra \) of velocity, which is proportional to \( \Delta T \). At a critical value of \( Ra \), the state of the fluid becomes unstable, and a pattern of stationary convection rolls develops (Fig. 2.20b) [2.35].

Beyond a greater critical value of \( Ra \), a transition to chaotic motion is observed. The complicated differential equations describing the Bénard experiment were simplified by Lorenz to obtain the three nonlinear differential equations of his famous model. Each differential equation describes the rate of change for a vari-

![Fig. 2.20a,b. Bénard experiment: a heated fluid layer](image-url)
able $X$ proportional to the circulatory fluid flow velocity, a variable $Y$ characterizing the temperature difference between ascending and descending fluid elements, and a variable $Z$ proportional to the deviation of the vertical temperature profile from its equilibrium value. From these equations, it can be derived that an arbitrary volume element of some surface in the corresponding phase space contracts exponentially in time. Thus, the Lorenz model is dissipative.

This can be visualized by computer-assisted calculations of the trajectories generated by the three equations of the Lorenz model. Under certain conditions, a particular region in the three-dimensional phase space is attracted by the trajectories, making one loop to the right, then a few loops to the left, then to the right again, etc. (Fig. 2.21) [2.36].

![Fig. 2.21. Lorenz attractor](image)

The paths of these trajectories depend very sensitively on the initial conditions. Tiny deviations of their values may lead to paths which soon deviate from the old one with different numbers of loops. Because of its strange image, which looks like the two eyes of an owl, the attracting region of the Lorenz phase was called a “strange attractor”. Obviously, the strange attractor is chaotic. But which topological structure do the trajectories achieve by winding more and more densely without intersecting each other? An example illustrates the definition of so-called fractal dimensions [2.37]:

Let $M$ be the subset of the attractor in the $n$-dimensional phase space. Now, the phase space is covered by cubes with edge length $\varepsilon$. Let $N(\varepsilon)$ be the number of cubes which contain a piece of the attractor $M$. If $\varepsilon$ contracts to zero ($\varepsilon \to 0$), then the negative limit of the ratio of the logarithm of $N(\varepsilon)$ and the logarithm of $\varepsilon$, i.e., $D = -\lim \ln N(\varepsilon)/\ln \varepsilon$, is called the fractal dimension.
If the attractor is a point (Fig. 2.14a), the fractal dimension is zero. For a stable limit circle (Fig. 2.9) the fractal dimension is one. But for chaotic systems the fractal dimension is not an integer. In general, the fractal dimension can be calculated only numerically. For the Lorenz model, the strange attractor has the fractal dimension \( D \approx 2.06 \pm 0.01 \).

Another dissipative system in which chaotic motion has been studied experimentally is the Belousov–Zhabotinsky reaction. In this chemical process an organic molecule is oxidized by bromate ions, the oxidation being catalyzed by a redox system. The rates of change for the concentrations of the reactants in a system of chemical reactions are again described by a system of nonlinear differential equations with a nonlinear function. The variable which signals chaotic behavior in the Belousov–Zhabotinsky reaction is the concentration of the ions in the redox system. Experimentally, irregular oscillations of these concentrations are observed with a suitable combination of the reactants. The oscillations are indicated by separated colored rings. This separation is a fine visualization of nonlinearity. Linear evolutions would satisfy the superposition principle. In this case the oscillating rings would penetrate each other in superposition.

The corresponding differential equations are autonomous, i.e., they do not depend on time explicitly. For computer-assisted visualization it is often convenient to study the flow in a dynamical system described by differential equations of motion via discrete equations which construct the intersecting points of the trajectories with the \( (d - 1) \)-dimensional Poincaré map in the corresponding \( d \)-dimensional phase space (compare Fig. 2.16). The constructed points are denoted by \( x(1), x(2), \ldots, x(n), x(n + 1), \ldots \) with increasing time points \( n \). The corresponding equation has the form \( x(n + 1) = G(x(n), \lambda) \) for the successor point \( x(n + 1) \) of \( x(n) = (x_1(n), \ldots, x_{d-1}(n)) \). The classification of conservative and dissipative systems can be generalized from flows to Poincaré maps. A discrete map equation is called dissipative if it leads to a contraction of volume in phase space.

A famous example of a discrete map is the so-called logistic map with many applications in the natural sciences as well as the social sciences. The basic concepts of complex dynamical systems from nonlinearity to chaos can be illustrated by this map with rather simple computer-assisted methods. Thus, let us have a short glance at this example. Mathematically, the logistic map is defined by a quadratic (nonlinear) recursive map \( x_{n+1} = \alpha x_n(1 - x_n) \) of the interval \( 0 \leq x \leq 1 \) onto itself with control parameter \( \alpha \) varying between \( 0 \leq \alpha \leq 4 \). The function values of the sequence \( x_1, x_2, x_3, \ldots \) can be calculated by a simple pocket computer. For \( \alpha < 3 \) the sequence converges towards a fixed point (Fig. 2.22a). If \( \alpha \) is increased beyond a critical value \( \alpha_1 \), then the values of the sequence jump periodically between two values after a certain time of transition (Fig. 2.22b). If \( \alpha \) is increased further beyond a critical value \( \alpha_2 \), the period length doubles. If \( \alpha \) is increased further and further, then the period doubles each time with a sequence of critical values \( \alpha_1, \alpha_2, \ldots \). But beyond a critical value \( \alpha_c \), the development becomes more and more irregular and chaotic (Fig. 2.22c) [2.38].

The sequence of period doubling bifurcations which is illustrated in Fig. 2.23a is governed by a law of constancy which was found by Grossmann and Thomae for...
Fig. 2.22a–c. Logistic curve as nonlinear recursive map with control parameter $\alpha = 4\lambda = 2$ (a), $\alpha = 4\lambda = 3$, 2 (b), and $\alpha = 4\lambda = 4$ (c)

the logistic map and recognized by Feigenbaum as a universal property for a whole class of functions (the Feigenbaum-constant) [2.39]. The chaotic regime beyond $\alpha_c$ is shown in Fig. 2.23b.
2.4 Conservative and Dissipative Systems and the Emergence of Order

Fig. 2.23a,b. Sequence of period doubling bifurcations (a) and chaotic regime of the logistic map beyond $\alpha_c = 4\lambda_c$ (b)

In Fig. 2.24a–c the mappings of $x_n$ onto $x_{n+1}$ are illustrated for different control parameters, in order to construct the corresponding attractors of a fixed point, periodic oscillation between two points, and complete irregularity without any point attractor or periodicity.
It is rather astonishing that a simple mathematical law like the logistic map produces a complexity of bifurcations and chaos for possible developments as shown in Fig. 2.23a,b. A necessary, but not sufficient reason is the nonlinearity of the equation. In this context, the degrees of increasing complexity are defined by the increasing bifurcations which lead to chaos as the most complex and fractal scenario. Each bifurcation illustrates a possible branch of solution for the nonlinear equation. Physically, they denote phase transitions from a state of equilibrium to new possible states of equilibria. If equilibrium is understood as a state of symmetry, then phase transition means symmetry breaking which is caused by fluctuational forces.

Mathematically, symmetry is defined by the invariance of certain laws with respect to several transformations between the corresponding reference systems of
2.4 Conservative and Dissipative Systems and the Emergence of Order

an observer. In this sense the symmetry of Kepler’s laws is defined by its Galilean transformations (compare Fig. 2.6a). The hydrodynamical laws describing a fluid layer heated from below (Fig. 2.20a) are invariant with respect to all horizontal translations. The equations of chemical reactions (in an infinitely extended medium) are invariant with respect to all translations, rotations, and reflections of a reference system used by an observer [2.40].

Nevertheless, these highly symmetric laws allow phase transitions to states with less symmetry. For example, in the case of a Bénard experiment, the heated fluid layer becomes unstable, and the state of stationary convection rolls develops (Fig. 2.20b). This phase transition means symmetry breaking, because tiny fluctuations cause the rolls to prefer one of two possible directions. Our examples show that phase transition and symmetry breaking is caused by a change of external parameters and leads eventually to a new macroscopic spatio-temporal pattern of the system and emergence of order.

Obviously, thermal fluctuations bear in themselves an uncertainty, or more precisely speaking, probabilities. A particle which is randomly pushed back or forth (Brownian motion) can be described by a stochastic equation governing the change of the probability distribution as a function of time. One of the most important means to determine the probability distribution of a process is the so-called master equation. To visualize the process we may think of a particle moving in three dimensions on a lattice.

The probability of finding the system at point $x$ at time $t$ increases due to transitions from other points $x'$ to the point under consideration (“rate in”). It decreases due to transitions leaving this point (“rate out”). As the “rate in” consists of all transitions from initial points $x'$ to $x$, it is composed of the sum over these initial points. Each term of the sum is given by the probability of finding the particle at point $x'$, multiplied by the transition probability (per unit time) for passing from $x'$ to $x$. In an analogous way the “rate out” can be found for the outgoing transitions. Thus, the rate of change for the probability distribution of a process is determined by a stochastic differential equation which is defined by the difference between “rate in” and “rate out”.

Fluctuations are caused by a huge number of randomly moving particles. An example is a fluid with its molecules. So a bifurcation of a stochastic process can only be determined by the change of probabilistic distribution. In Fig. 2.25 the probabilistic function changes from a sharp centration at a single attractor (Fig. 2.25a) to

![Fig. 2.25a–c. Probabilistic function with single attractor (a), flat distribution (b), and two attractors as stochastic symmetry breaking (c)](image)
a flat distribution (Fig. 2.25b) and finally to a distribution with two maxima at two attractors (Fig. 2.25c), when the control parameter increases beyond corresponding critical values. Figure 2.25c illustrates stochastic symmetry breaking [2.41].

In this context, complexity means that a system has a huge number of degrees of freedom. When we manipulate a system from the outside we can change its degrees of freedom. For example, at elevated temperature the molecules of water vapor move freely without mutual correlation. When the temperature is lowered, a liquid drop is formed. This macroscopic phenomenon is formed when the molecules keep a mean distance between each other with correlated motion. At the freezing point water is transformed into ice crystals with a fixed molecular order. Since the early days of mankind people have been familiar with these phase transitions. The different aggregate states may have been a reason for philosophical ideas that water is a basic element of matter (compare Sect. 2.1).

Another example is taken from the material sciences. When a ferromagnet is heated, it loses its magnetization beyond a critical value. But the magnet regains its magnetization when the temperature is lowered. Magnetization is a macroscopic feature which can be explained by changing the degrees of freedom at the microscopic level. The ferromagnet consists of many atomic magnets. At elevated temperature, the elementary magnets point in random directions. If the corresponding magnetic moments are added up, they cancel each other. Then, on the macroscopic level, no magnetization can be observed. Below a critical temperature, the atomic magnets are lined up in a macroscopic order, giving rise to the macroscopic feature of magnetization (Fig. 4.9a). In both examples, the emergence of macroscopic order was caused by lowering the temperature. The structure is formed without loss of energy at low temperature. Thus, it is a kind of conservative (reversible) self-organization. Physically, it can be explained by Boltzmann’s law of distribution demanding that structures with less energy are mainly realized at low temperatures.

On the other hand, there are systems whose order and functioning are not achieved by lowering temperature, but by maintaining a flux of energy and matter through them. Familiar examples are living systems like plants and animals which are fed by biochemical energy. The processing of this energy may result in the formation of macroscopic patterns like the growth of plants, locomotion of animals, and so on. But this emergence of order is by no means reserved to living systems (compare Chap. 3). It is a kind of dissipative (irreversible) self-organization far from thermal equilibrium which can be found in physics and chemistry as well as in biology.

As is well-known from the second law of thermodynamics, closed systems without any exchange of energy and matter with their environment develop to disordered states near thermal equilibrium. The degree of disorder is measured by a quantity called “entropy”. The second law says that in closed systems the entropy always increases to its maximal value. For instance, when a cold body is brought into contact with a hot body, then heat is exchanged so that both bodies acquire the same temperature, i.e., a disordered and homogeneous order of molecules. When a drop of milk is put into coffee, the milk spreads out to a finally disordered and homo-
geneous mixture of milky coffee. The reverse processes are never observed. In this sense, processes according to the second law of thermodynamics are irreversible with a unique direction [2.42].

An example from hydrodynamics is the Bénard instability, which was already described in the beginning of Sect. 2.4. When the heated fluid layer (Fig. 2.20a) reaches a critical value, it starts a macroscopic motion (Fig. 2.20b). Thus a dynamic well-ordered spatial pattern emerges out of a disordered and homogeneous state as long as a certain flux of energy is maintained through the system.

Another example from fluid dynamics is the flow of fluid round a cylinder. The external control parameter is the Reynolds number $Re$ of fluid velocity. At low speed the flow happens in a homogeneous manner (Fig. 2.26a). At higher speeds, a new macroscopic pattern with two vortices appears (Fig. 2.26b). With yet higher speeds
the vortices start to oscillate (Fig. 2.26c–d). At a certain critical value, the irregular and chaotic pattern of a turbulent flow arises behind the cylinder (Fig. 2.26e). Figure 2.26a–e presents a survey of possible attractors with one and more fixed points, bifurcations, oscillating, and quasi-oscillating attractors, and finally fractal chaos [2.43].

A famous example from modern physics and technology is the laser. A solid state laser consists of a rod of material in which specific atoms are embedded. Each atom may be excited by energy from outside leading it to the emission of light pulses. Mirrors at the end faces of the rod serve to select these pulses. If the pulses run in the axial direction, then they are reflected several times and stay longer in the laser, while pulses in different directions leave it. At small pump power the laser operates like a lamp, because the atoms emit independently of each other light pulses (Fig. 2.27a). At a certain pump power, the atoms oscillate in phase, and a single ordered pulse of gigantic length emerges (Fig. 2.27b) [2.44].

The laser beam is an example of macroscopic order emerging by a dissipative (irreversible) self-organization far from thermal equilibrium. With its exchange and processing of energy, the laser is obviously a dissipative system far from thermal equilibrium.

In former days of history, scientists would have postulated certain demons or mystic forces leading the elements of these systems to new patterns of order. But, as in the case of conservative self-organization, we can explain dissipative self-organization by a general scheme which is made precise by well-known mathematical procedures. We start with an old structure, for instance a homogeneous fluid or randomly emitting laser. The instability of the old structure is caused by a change of external parameters, leading eventually to a new macroscopic spatio-temporal structure. Close to the instability point we may distinguish between stable and unstable collective motions or waves (modes). The unstable modes start to influence and determine the stable modes which therefore can be eliminated. Hermann Haken calls this process very suggestively a “slaving principle”. Actually, the stable modes are “enslaved” by the unstable modes at a certain threshold.

**Fig. 2.27a,b.** Wave patterns emitted from a lamp (a) and from a laser (b)
Mathematically, this procedure is well known as the so-called “adiabatic elimination” of fast relaxing variables, for instance, from the master equation describing the change of probabilistic distribution in the corresponding system. Obviously, this elimination procedure enables an enormous reduction of the degrees of freedom. The emergence of a new structure results from the fact that the remaining unstable modes serve as order parameters determining the macroscopic behavior of the system. The evolution of the macroscopic parameters is described by differential equations. In contrast to properties of the elements of a system at the microscopic level (for instance, atoms, molecules, etc.), the order parameters denote macroscopic features of the whole system. In the case of the laser, some slowly varying (“undamped”) amplitudes of modes may serve as the order parameters, because they start to enslave the atomic system. In the language of biology, the order parameter equations describe a process of “competition” and “selection” between modes. But, of course, these are only metaphoric formulations which can be made precise by the mathematical procedure mentioned above [2.45].

In general, to summarize, a dissipative structure may become unstable at a certain threshold and break down, enabling the emergence of a new structure. As the introduction of corresponding order parameters results from the elimination of a huge number of degrees of freedom, the emergence of dissipative order is combined with a drastic reduction of complexity. Dissipative structures are a fundamental concept of complex systems which are used in this book to model processes in natural and social sciences. The irreversibility of dissipative structures may remind us of Heraclitus’ famous quotation that nobody can enter a stream in the same state. Obviously, irreversibility violates the time-invariance symmetry which characterizes the classical (Hamiltonian) world of Newton and Einstein. But the classical view will turn out to be a special case in a steadily changing world. On the other hand, Heraclitus believed in an ordering law harmonizing irregular interactions and creating new order states of matter. We have to see whether the mathematical scheme of a dissipative system will satisfy the universal features of such a law.

A general framework for the evolution of matter would be based on a unified theory of all physical forces (Fig. 2.28). The standard models of cosmic evolution which are derived from Einstein’s general theory of relativity must be explained by the principles of quantum theory. Until today there are only several more or less satisfying mathematical models of cosmic evolution which can only partially be tested and confirmed by experiments. Nevertheless, it is the general idea of these models that the emergence of structures with increasing complexity (elementary particles, atoms, molecules, planets, stars, galaxies, etc.) can be explained by cosmic phase transitions or symmetry breaking [2.46].

In cosmic evolution an initial state is assumed to be nearly homogeneous and symmetric in the sense that in general no elementary particles can be distinguished, but they can be transformed into one another. During cosmic evolution, critical values have been realized step by step at which symmetries break down by deviations and fluctuations and new particles and forces emerge: “C’est la dissymétrie, qui crée le phénomène,” said Pierre Curie [2.47]. But we must be aware that the cosmic pro-
cesses of symmetry breaking and phase transitions are mathematical extrapolations from experiments and theories in high energy physics.

Nowadays, physics distinguishes four fundamental forces, the electromagnetic, strong, weak, and gravitational forces which are mathematically described by so-called gauge fields. Elementary particle physics aims to unify the four physical forces in one fundamental force corresponding to the initial state of the universe. Electromagnetic and weak forces have already been unified at very high energies in an accelerator ring at CERN (Fig. 2.28). Unification means that at a state of very high energy the particles that "feel" the weak force (electrons, neutrinos, etc.) and those that "feel" the electromagnetic force cannot be distinguished. They can be described by the same symmetry group \((U(1) \times SU(2))\), i.e., they are invariant with respect to transformations of this group. At a particular critical value of lower energy the symmetry breaks down into partial symmetries \((U(1)\) and \(SU(2)\)) corresponding to the electromagnetic and weak forces.

Physically, this kind of symmetry breaking means a phase transition which is connected with the emergence of two new physical forces and their elementary particles. The process of spontaneous symmetry breaking is well known. For instance, our breakfast egg is not stable in its symmetric position on its top. Any tiny fluctuation causes it to fall spontaneously down to an asymmetric, but energetically stable position. The phase transition of a ferromagnet from a non-magnetic to a magnetic state is caused by cooling down the temperature to a critical point. The elementary dipoles spontaneously take one of the two possible magnetic orientations, break the spin-rotation symmetry, and cause the emergence of a new macroscopic property (magnetization).
The complex variety of baryons (protons, neutrons, etc.) and mesons interacting via the strong force are constructed from the so-called quarks with three degrees of freedom, i.e., the so-called “colors” red, green, and blue. A baryon, for instance, is built up from three quarks which are distinguishable by three different colors. These three colors are complementary in the sense that a hadron is neutral (without color) to its environment. The mathematical symmetry group (SU(3)) characterizing the color transformation of quarks is well known.

After the successful unification of the electromagnetic and weak interactions physicists try to realize the “grand unification” of electroweak and strong forces, and in a last step the “superunification” of all four forces (Fig. 2.28). There are several research programs for superunification, such as supergravity and superstring theory. Mathematically, they are described by extensions to more general structures of symmetry (“gauge groups”) including the partial symmetries of the four basic forces. Technically, the unification steps should be realized with growing values of very high energy. But the “grand unification” demands states of energy which cannot be realized in laboratories. Thus, the high energy physics of grand unification could only be confirmed by certain consequences which could be tested in a laboratory or observed in the universe (e.g., the decay of protons). The superunification of all forces would demand infinitely increasing states of energy whose physical principles are still unknown.

The theory of the so-called “inflationary universe” assumes an early state of the universe with small size, but very high energy (“quantum vacuum”) which expands very rapidly to macroscopic dimensions driven by a repulsive force of the quantum vacuum state (“anti-gravity”). This cosmic phase transition allows one to explain some well-known properties of the observed universe such as the relatively homogeneous distribution of stars and matter. During the inflationary period, some tiny deviations from symmetry and uniformity would have been amplified until they were big enough to account for the observed structures of the universe. In the expanding universe the density of matter varied slightly from place to place. Thus, gravity would have caused the denser regions to slow down their expansion and start contracting. These local events led to the formation of stars and galaxies [2.48].

In general, the emergence of the structural variety in the universe from the elementary particles to stars and living organisms is explained by phase transitions, corresponding to symmetry breaking of equilibrium states (Figs. 2.29, 2.30). In this sense the cosmic evolution of matter is understood as a self-organizing process with the emergence of conservative and dissipative structures. But we must be aware that cosmic self-organization is today only a “regulative idea of research”, as Kant had said: we have more or less plausible dynamical models which are more or less empirically confirmed. The very beginning of cosmic evolution is still unknown.

If we only assume the classical principles of Einstein’s general relativity, then, as Roger Penrose and Stephen Hawking have mathematically proved, the standard models of cosmic evolution have an initial singularity which may be interpreted as the Big Bang, i.e., the emergence of the universe from a mathematical point. But if we assume a unification of the general theory of relativity (i.e., Einstein’s relativistic theory of gravitation) and quantum mechanics with imaginary (instead of real)
Fig. 2.29. The emergence of structural variety in the universe from elementary particles to galaxies [2.48]

Fig. 2.30. The evolution of matter with increasing and decreasing complexity [2.51]
time, then, as Hawking has mathematically proved, a “smooth” cosmic model is possible without any beginning, which simply exists, according to the mathematical principles of a unified relativistic quantum physics [2.49].

The singularity theorems of Penrose and Hawking started with predictions of small regions of space where space-time is so warped that gravity becomes infinitely great. The existence of such singularities, in the form of black holes, for example, suffer from a methodological disadvantage: Classical and relativistic laws of physics are not applicable in regions with infinite curvature, so it is not possible to predict events in time. That consequence is, of course, quite more dramatic than the exponentially increasing difficulties in predicting the long-term future of chaotic systems. That is why James B. Hartle and Stephen Hawking have suggested a singularity-free model of the universe, in which quantum theory and general relativity theory are unified, and the real time axis is replaced by an imaginary one (in the sense of real and imaginary numbers) [2.50]. In Hawking’s model, in contrast to Einstein’s relativistic theory, the three spatial axes, together with a complex time axis, lead to a closed early quantum universe that lacks boundaries and edges. This space-time not only would always have existed, but every physical event could be explained according to its laws. In this model, the traditional concepts of everything having somehow “begun” or been “created” are methodologically inappropriate and are revealed to be human imaginings stemming from our having adapted to the limited space-time facets of our everyday experience.

Hawking’s theory is not only mathematically consistent, but is also, at least in principle, experimentally testable. It is, therefore, a scientific theory and not mere speculation. Among the testable consequences of this singularity-free model is the prediction of black holes in which not all world lines of photons (“light beams”) disappear entirely, but are reemitted as measurable amounts of radiation. As in the explanation of the initial singularity of the universe, the reason lies in the possibility of quantum fluctuations rooted in the uncertainty relation. But radiating black holes lose energy and mass. In time, they will disintegrate and, with them, the history of their stars will be lost. In their place, memory gaps will appear in the universe. With the collapse of its galactic structures, a featureless universe expanding into a void is heading for a “cosmic Alzheimer’s disease”.

Philosophically, Hawking’s early quantum universe without a beginning reminds us of Parmenides world of unchangeable being. But the uncertainty principle of quantum mechanics implies that the early universe cannot have been completely uniform because there must have been some uncertainties or fluctuations in the positions and velocities of the particles. Thus, the universe would have undergone a period of rapid expansion which is described by the inflationary model, leading to our complex universe in the long run. The equilibrium of the Parmenidean world broke down and changed to the evolutionary and complex world of Heraclitus, caused by a basic principle of quantum physics under the hypothesis of a “smooth” time without singularities.

A cosmological model of an “eternal” universe without beginning and without end was already introduced by Hermann Bondi, Thomas Gold, and Fred Hoyle in 1948. These authors did not only assume spatial homogeneity and isotropy of the
universe at every time ("Cosmological Principle" of the standard models with Big Bang), but also temporal homogeneity and isotropy: Their "perfect Cosmological Principle" suggests that the universe globally looks the same not only at all points and in all directions, but at all times, leading to a steady state model. According to Hubble, there is a correlation between the red shift and the increasing distances of expanding galaxies. So if the average number of galaxies per unit proper volume is to remain constant, new galaxies must appear to fill up the holes in the widening comoving coordinate mesh. An ad-hoc hypothesis of steady state cosmology was the necessity of continuous creation of matter.

In recent quasi-steady state cosmologies, the strange assumption of a contingent and nonlocal creation of matter is explained by the local birth of new galaxies everywhere and at every time in the universe. The conditions of local big bangs are assumed to be realizable in the supermassive centers of old galaxies. The red shifts seem also to indicate the age of a galaxy. The uniform evolution with the sequential emergence of elementary particles, atoms, molecules, galaxies, stars, etc. after the global Big Bang (Fig. 2.29) is replaced by an autocatalytically self-reproducing universe without global beginning and without end, but with local births, growths, and deaths of galaxies. In this case, old dying galaxies create the matter of new galaxies like plants and organisms bearing the seed of new life. The universal dynamics would be a gigantic never ending nonlinear recycling process of matter [2.49].

But, perhaps, the laws of quantum mechanics open loopholes ("wormholes") of escape from the fate of our universe. According to general relativity theory, time travel cannot be faster than the speed of light. As light is curved by gravitational fields, time travelers must pass curved paths in space-time with high speed, limited by the speed of light. Therefore, in order to overcome disruption of space-time by gravitational fields, space-time regions would have to be explored using vast curved detours. According to Heisenberg’s principle of uncertainty, quantum fluctuations could open short-lived wormholes in space-time. So, the laws of quantum mechanics make it at least conceivable that wormholes can be employed as fleeting shortcuts between folded regions. However, if our universe is not alone but instead intertwined with a fractal multiverse, along with many other bifurcating universes, as was suggested in Andrei Linde’s inflationary theory, then wormholes could also be used as escape routes for fleeing a universe that is aging with cosmic Alzheimer’s disease and growing hostile to life as it loses energy.

From a theological point of view, these models do not need any creator, because their worlds simply have been and will be self-contained and self-organizing without beginning and without end. From a mathematical point of view, these models may be very elegant. But from a methodological point of view, we must conclude that we do not yet have a complete and consistent theory combing quantum mechanics and relativistic gravity which could explain the evolution of matter with its increasing complexity. Thus we are only certain of some of the properties by gravitational fields such a unified theory could have. Today, different approaches of string theories exist for achieving this unification on a sublevel of elementary particles. If all kinds of elementary particles of gravitational, strong, weak, and electromagnetic interactions are generated by oscillating strings, then there is even a chance to avoid the ultimate
loss of information in the black holes of an aging universe: The information could be stored by the vibrating membranes of more-dimensional strings on the substructure of matter [2.52].

2.5 Complex Systems of the Nano World and Self-Constructing Materials

In the evolution of matter, self-organizing processes can be observed from the level of elementary particles to the cosmic structures of galaxies. They are not only interesting from an epistemic point of view, but for applications in materials and life science, too. At the boundary between materials science and life science, supramolecular systems play a tremendous role. In this case molecular self-organization means the spontaneous association of molecules under equilibrium conditions into stable and structurally well-defined aggregates with dimensions of 1–10² nanometers (1 nm = 10⁻⁹ m = 10 Å).

Nanostructures may be considered as small, familiar, or large, depending on the viewpoint of the disciplines concerned. To chemists, nanostructures are molecular assemblies of atoms numbering from 10³ to 10⁹ and molecular weights of 10⁴ to 10¹⁰ daltons. Thus, they are chemically large supramolecules. To molecular biologists, nanostructures have the size of familiar objects from proteins to viruses and cellular organelles. But to materials scientists and electrical engineers, nanostructures are at the current limit of microfabrication and thus they are rather small [2.53].

In the beginning of nanoscience there was the vision of an ingenious physicist. In an article entitled “There’s Plenty of Room at the Bottom”, Richard Feynman declared:

The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It would be, in principle, possible … for a physicist to synthesize any chemical substance that the chemist writes down … How? Put the atoms down where the chemist says, and so you make the substance. The problems of chemistry and biology can be greatly helped if our ability to see what we are doing, and to do things on an atomic level, is ultimately developed – a development which I think cannot be avoided. [2.54]

Feynman proclaimed his physical ideas of the nanoworld in the late 1950s. The belief in a new world needs new instruments of observation and measurement for confirmation. Since the start of the 1980s, the nanoworld could actually be explored using the scanning tunnel microscope. At the end of the 1980s, Eric Drexler described a revolutionary vision of technological applications:

Nature shows that molecules can serve as machines because living things work by means of such machinery. Enzymes are molecular machines that make, break, and rearrange the bonds holding other molecules together. Muscles are driven by molecular machines that haul fibers past one another. DNA serves as a data-storage system, transmitting digital instructions to molecular machines, the ribosomes, that manufacture protein molecules. And these protein molecules, in turn, make up most of the molecular machinery just described. [2.55]
With nanotechnology, atoms will be specifically placed and connected in a fashion similar to processes found in living organisms. Complex organisms, such as plants and animals, make use of molecular machinery to manufacture and undertake repairs at the cellular and subcellular levels. A cell can be considered a factory of nanomachines consisting of molecular prototypes such as protein, nucleic acid, lipid, and polysaccharide. They are used for energy production, information processing, self-replication, self-repairing, and moving. A ribosome, for example, is a cellular nanomachine that reads information off a RNA strand in order to construct the amino acids of a protein. It reminds us of the assembly-like production of cars by robots in the motor industry. Biological micro-organisms have been understood as cellular systems driven and controlled by nanomachines. For example, bacteria such as Escherichia coli use whip-like tails for moving around in fluids. The tails like a propeller fueled by biochemical nanomachines. These nanomachines consist of proteins in membranes generating the rotation of whip-like tails. They use motor shafts and armatures like electric motors. But the similarity of nanomachines and electric motors is only illustrative. A biochemical nanomachine does not use electric current to generate a magnetic field; it changes the shape of molecules by biochemical procedures, such as decomposing ATP, in order to rotate the shaft [2.56].

Genetic engineering and computer programming have begun to inspire the development of new materials. Using special bacterium-sized assembler devices, nanotechnology should permit the exact control and fast manipulation of molecular structures. A fast enzyme can process almost a million molecules per second, even without conveyors and power-driven mechanisms to slap a new molecule into place as soon as an old one is released. Drexler assumed that an assembler arm would be about fifty million times shorter than a human arm and, accordingly, would be able to move back and forth about fifty million times more rapidly. According to Feynman’s vision, such machines would seize individual atoms using selectively sticky manipulator arms, then plug those atoms together like Lego blocks until chemical bonding took place. Following the line of computer programming, one would expect general-purpose chemical synthesizers acting like a general-purpose computer using nanotechnology. The desired molecules would be modeled on a computer screen and an appropriate assembler would allow the mass-production of the desired substances. Perhaps someday, specially designed nanodevices the size of bacteria will be programmed to destroy arterial plaque or cancer cells, or to repair cellular damage caused by aging. They could be injected into the body with an induction to self-destruct or integrate themselves into the body’s cells. Finally, it still seems to be science fiction that smart nanodevices distributed throughout the brain might permit the copying of thought patterns and mind uploading, so that a copy of a person’s personality and memories could be placed in storage, or even run as a form of naturally created artificial intelligence.

Nanostructures are complex systems which evidently lie at the interface between solid-state physics, supramolecular chemistry, and molecular biology. It follows that the exploration of nanostructures may deliver hints about the emergence of life and about the fabrication of new materials. But engineering of nanostructures cannot be mastered in the traditional way of mechanical construction. There are no
man-made tools or machines for putting together their building blocks like the elements of a clock, motor, or computer chip. We must thus understand the principles of self-organization which are used by nanostructures in nature. Then, we only need to arrange the appropriate constraints under which the atomic elements of nanostructures associate themselves in a spontaneous self-construction: The elements adjust their own positions to reach a thermodynamic minimum without any manipulation by a human engineer.

Historically, the idea of supramolecular interactions dates back to a famous metaphor of Emil Fischer (1894), who described a selective interaction of molecules as the lock and key principle. Today, supramolecular chemistry has far surpassed its original focus. Molecular self-assemblies combine several features of covalent and noncovalent synthesis to make large and structurally well-defined assemblies of atoms. The strengths of individual van der Waals interactions and hydrogen bonds are weak relative to typical covalent bonds and comparable to thermal energies. Therefore, many of these weak noncovalent interactions are necessary in order to achieve molecular stability in self-assembled aggregates. In biology, there are many complex systems of nanoscale structures such as proteins and viruses which are formed by self-assembly. Living systems sum up many weak interactions between chemical entities to make large ones. How can one make structures of the size and complexity of biological structures, but without using biological catalysts or the informational devices coded in genes?

Many nonbiological systems also display self-organizing behavior and furthermore provide examples of useful interactions. Molecular crystals are self-organizing structures. Liquid crystals are self-organized phases intermediate in order between crystals and lipids. Micelles, emulsions, and lipids display a broad variety of self-organizing behavior. An example is the generation of cascade polymers yielding molecular bifurcational superstructures of fractal order [2.57]. Their synthesis is based on the architectural design of trees. Thus, these supramolecules are called dendrimers (from the Greek word *dendron* for tree and polymer). The generation of dendrimers has followed two basic procedures for monomer addition. A divergent construction begins at the core and builds outward via an increasing number of repeating bifurcations. A convergent construction begins at the periphery and builds inward via a constant number of transformations. The divergent construction transforms the chemical reaction centers from the center into the periphery, generating a network of bifurcating branches around the center. The bifurcations increase exponentially up to a critical state of maximal size. They yield fractal structures such as molecular sponges which can contain smaller molecules, which can then be dispersed in a controlled way for medical applications.

Examples of cave-like supramolecules are the Buckminsterfullerenes, forming great balls of carbon [2.58]. The stability of these complex clusters is supported by their high geometric symmetry. The Buckminsterfullerenes are named after the geodesic networks of ball-like halls which were constructed by the American architect Richard Buckminster Fuller (1895–1983). The cluster C_{60} of 60 carbon atoms has a highly Platonic symmetry of atomic pentagons forming a completely closed spheroid.
Cave-like supramolecules can be arranged using chemical templates and matrices to produce complex molecular structures. Several giant clusters comparable in size to small proteins have been obtained by self-assembly. Figure 2.31 shows a ball-and-stick model of the largest discrete cluster (700 heavy atoms) ever characterized by X-ray structure analysis. This cluster containing 154 molybdenum, 532 oxygen, and 14 nitrogen atoms has a relative molecular mass of about 24 000. The highly symmetric “big wheel” was synthesized by Achim Müller and coworkers [2.59]. Giant clusters may have exceptional novel structural and electronic properties: There are planes of different magnetization which are typical for special solid-state structures and of great significance for materials science. A remarkable structural property is the presence of a nanometer-sized cavity inside the giant cluster. The use here of templates and the selection of appropriate molecular arrangements may well remind us of Fischer’s lock and key principle.

Molecular cavities can be used as containers for other chemicals or even for medicaments which need to be transported within the human organism. An iron-storage protein that occurs in many higher organisms is ferritin. It is an unusual host-guest system consisting of an organic host (an aprotein) and a variable inorganic guest (an iron core). Depending on the external demand, iron can either be removed from this system or incorporated into it. Complex chemical aggregates

Fig. 2.31. Giant supramolecular cluster (“big wheel”) in a ball-and-stick representation: An example of a complex near-equilibrium system [2.59]
2.5 Complex Systems of the Nano World and Self-Constructing Materials

like polyoxometalates are frequently discovered to be based upon regular convex polyhedra, such as Platonic solids. But their collective electronic and/or magnetic properties cannot be deduced from the known properties of these building blocks. According to the catchphrase “from molecules to materials” supramolecular chemistry applies the “blue-prints” of conservative self-organization to build up complex materials on the nanometer scale with novel catalytic, electronic, electrochemical, optical, magnetic, and photochemical properties. Multi-property materials are extremely interesting.

The exploration of the nanoworld and applications in nanotechnology depend on better instruments of observation and measurement. The scanning force microscope is a further development of the scanning tunnel microscope and can be used like a fountain pen to write down molecular structures of nano size. A thin film of thiolmolecules is used as “nano ink”. In a tiny drop of water the thiolmolecules organize themselves as mono layer. Nanocrystals of a few hundred atoms can organize themselves with cadmium ions, selen ions, and organic molecules into a ball-like structure (Fig. 2.32). In ultraviolet light they fluoresce with a certain color. Thus, they could be used as markers (“quantum dots”) of molecules, cells, and substances in medicine, for example. Complex systems of carbon molecules can organize themselves as tiny tubes of 1nm diameter according to certain catalysts and templates. Their symmetric order of bonding results in great hardness and toughness. Carbon nanotubes might be used as conductors for miniaturized chips beyond the limits of silicon technology.

![Fig. 2.32. Self-organizing nanocrystals (“quantum dots”) [2.60]](image_url)

Supramolecular transistors are an example that may stimulate a revolutionary new step in the development of chemical computers. Actually, there is a strong trend towards nanostructures in electronic systems which may realize small, fast devices and high-density information storage. But one can also imagine nonelectronic applications of nanostructures. They could be used as components in microsensors or as catalysts and recognition elements in analogy to enzymes and receptors in living systems. In natural evolution very large complex molecular systems are also produced by stepwise gene-directed processes. The conservative self-organization processes of nanomolecular chemistry are non-gene-controlled reactions. Only a clever combination of conservative and non-conservative self-organization could have ini-
tiated prebiotic evolution before genes emerged. But even during the evolution of complex organisms, conservative self-organization must have occurred. Open (“dissipative”) physical and chemical systems lose their structure when the input of energy and matter is stopped or changed (e.g., laser, BZ-reaction). Organismic systems (like cells) are able to conserve much of their structure at least for a relatively long time. On the other hand, they need energy and matter within a certain interval of time to keep their structure more or less far from thermal equilibrium. In the technical evolution of mankind, the principles of conservative and dissipative self-organization have once more been discovered and open new avenues of technical applications.

The complex systems approach enables engineers to endow materials with more and more of the attributes of living organisms. Self-regulation and self-adaptation to a changing environment are well-known capabilities of living systems. They can be considered as specific forms of self-organizing open systems in a changing environment. Analogously, engineers aim to create complex materials systems that can sense their own state, the state of their environment and respond to it. Dramatic examples are materials for bridges that could detect and counter corrosion before a pylon gives way, buildings that could brace themselves against seismic waves, or skins of aeroplanes that could spontaneously react against dangerous material fatigue.

Actuators are materials which can change their features according to changing states of the system [2.61]. Examples are piezoelectric ceramics and polymers acting either as pressure sensors or as mechanical actuators. The electrical polarity of their crystal or molecular structures allows a transformation of mechanical forces exerted on them into electrical current or, conversely, a transformation of electrical stimuli into vibrations. Piezoelectric polymers could be embedded in the skin of a robotic hand in order to get a high degree of sensitivity (e.g., to decipher braille).

Other examples are alloys with a so-called shape memory that can be used as actuators. Below a certain control value of transition temperature, a shape-memory wire will take any shape it is bent into. When the wire is heated beyond the transition, it returns to its original shape. Engineers propose the incorporation of a shape-memory metal into a material system in its low-temperature shape. It exerts a force whenever it is heated. The force-generating transition takes place as the atoms in the alloy’s crystal grains toggle between different geometric arrangements. Damage-resistant bridges or airplane wings would be possible applications of these control structures.

There are even actuator materials that can reversibly transform their mechanical properties from a liquid to a solid state. They consist of fine polarizable particles of ceramic or polymer suspended in a liquid such as silicone oil. When subjected to strong electric fields, such fluids organize themselves into filaments and networks which stiffen the material into a gel-like solid. When the electric field is removed, the organization dissipates, and the material becomes fluid again. Other applications are optical fibers acting as sensor materials. The properties of these hair-thin fibers are affected by changes in temperature, pressure or other physical or chemical conditions within the materials. They can be considered as “glass nerves” providing optical signals of the material’s internal “health”.
Sometimes, modern materials scientists call their systems smart or even “intelligent” materials. The goal of their research is sometimes described as “animation of the inanimate world” [2.62]. From a philosophical point of view, this slogan seems to hark back to alchemistic traditions. Some philosophers of science may perhaps criticize the vocabulary of materials scientists as non-scientific animism. But, from the viewpoint of complex systems, there is a hard scientific core. Properties of self-organization are not necessarily combined with conscious behavior on the basis of nervous systems. They even do not necessarily depend on biological catalysts or the informational devices coded in genes. Thus, there is no break between the so-called inanimate and animate world. In the evolution of matter, we observe systems with more or less high degrees of organization. It is clear that we have only made the very first steps in understanding their full potential.

Concerning the future of technology, the question arises, how realistic is the vision of self-replicating nanorobots? They would be the equivalent of a new parasitic life form. Pathogenic bacteria and cancer cells are dangerous examples of self-replicating biological systems. Computer viruses with self-replicating strings of bits are the first, at least virtual, examples of artificial self-organizing systems. Bill Joy, the chief scientist of Sun Microsystems, has already raised concerns about the societal implications of proliferating nanobots [2.63]. In an artificial evolution, Joy says, hostile agents could evolve into populations of embodied biochemical agents of nano size. As autonomous, self-interested beings, they could attack the foundations of human life. Richard E. Smalley, who received the Nobel Prize in chemistry for the discovery of fullerenes, dismisses the notion of out-of-control nanorobots [2.64]. Following Feynman’s slogan, “There’s plenty of room at the bottom”, Smalley argues that not much room is needed to manipulate atoms one by one with nano-sized atomic instruments. He calls these constraints the fat and sticky fingers problem: The nanobot’s manipulating “fingers” are not only too large (“fat”) but also too sticky, because their atoms will adhere to the atom that is being moved. Smalley’s picture of fingers underlines the fact there are no counterparts of our today technology at nanometer sizes. In living systems, evolution has developed examples of biochemical nanomachines, and there is no reason to believe that there cannot be others on different material grounds. But the technological strategy should follow the natural idea of self-organization under appropriate constraints, not the old-fashioned mechanical idea of picking and placing atoms with nanoscale pincers. We should not look for assemblers, but self-assemblers. From the point of view of computer science, the idea of a universal fabricator of any kind of structure, including itself, is not strange. A universal Turing machine (compare Sect. 5.2) is already embodied by our general-purpose computers, which process all kinds of programs. Why not on the nanoscale?

### 2.6 Time Series Analysis, Fractals, and Multifractals

Understanding complex systems and nonlinear dynamics in nature seems to yield appropriate models for the evolution of matter. But how can we be sure that our
models are correct? The mathematical theory of nonlinear dynamics distinguishes different types of time-dependent equations, generating different types of behavior, such as fixed points, limit cycles, and chaos. For application, they are related to natural systems in the micro, nano, and macroworld. We use our understanding of the special mechanisms to write an appropriate dynamical equation. For example, Lorenz’s understanding of the dynamics of weather led to his famous nonlinear equations, which were also applied to biological and economic systems by people familiar with those fields.

From a methodological point of view, this is the top-down approach to model building: We start with an assumed mathematical model of a natural system and deduce its behavior by solving the corresponding dynamical equations under certain initial and secondary conditions. The solutions can be represented geometrically as trajectories in the phase space of the dynamical system and classified by different types of attractors. They forecast the types of behavior that we are likely to observe in a specialized field of research. Especially chaotic dynamics can be derived from the given equations if certain criteria are satisfied. But, in practice, we often must take the opposite, bottom-up approach. Physicists, chemists, biologists, or physicians start with data mining in an unknown field of research. They only get a finite series of measured data corresponding to time-dependent events of a dynamical system. From these data they must reconstruct the behavior of the system in order to guess its type of dynamical equation. Therefore, the bottom-up approach is called time series analysis [2.65]. In many cases, we have knowledge of the system from which the data came. Time series analysis then aims to construct a black box, which takes the measured data as input and provides as output a mathematical model describing the data. In practice, the realistic strategy of research is a combination of the top-down approach with model building and the bottom-up approach with time series analysis.

The bottom-up approach starts with data as results of measurements, not with the idealized variables of a model. The measurements approximate the variables of a dynamical model. Their difference is called measurement error, which can be caused by several factors of noise. Noise of measurements refers to fluctuations of data that differ from a well-defined average behavior and arise from chance. While measurement noise is caused by the intrinsic behavior of the real system, the outside influence of the system also affects a kind of noise. Many variables of outside influence must be excluded in order to reduce the complexity of model building. The outside influence on the actual behavior of a system is considered random noise affecting the measured variables of the model.

In classical measurement theory, measurement error is analyzed by statistical methods, such as correlation coefficient and autocorrelation function. But these standard procedures are not able to distinguish between data from linear and nonlinear models. In nonlinear data analysis, the measured data are used in a first step to reconstruct the dynamics of the system in a reconstructed phase space. A simple example is the finite difference equation of the logistic map, which we studied in Sect. 2.4: The nonlinear equation

$$x_{t+1} = f(x_t)$$

describes a relationship between $x_{t+1}$ and $x_t$. In Fig. 2.24, the coordinates are plotted $x_{t+1}$ versus $x_t$. If there is no measurement
noise, we can identify the measurement data \( D_t \) and the variable \( x_t \) at time \( t \). It is no surprise then, that a scatter plot of the measured data \( D_{t+1} \) versus \( D_t \) delivers the same relationship as the model.

If data are collected from a continuous-time dynamical system with differential equations, rather than finite-difference equations, the corresponding phase plane or phase space must be reconstructed from the measured data of the continuous system. The heuristic idea is that the measured data in the reconstructed phase space show the same dynamical behavior as the trajectories in the phase space of the dynamical model. Consider, for example, the data generated by a harmonic oscillator with the 2nd order differential equation \( \frac{d^2x}{dt^2} = -bx \). The corresponding phase plane is given by the variables \( x \) and \( y \), which are determined by the two 1st order differential equations \( \frac{dx}{dt} = y \) and \( \frac{dy}{dt} = -bx \). We suppose that a time series \( D(t) = x(t) \) is measured without measurement noise. In order to reconstruct the phase plane from the measured data, we remember that the state of the system at any instant \( t \) is represented by the position \( (x, y) \) on the phase plane. The time series of measurements yields us only one coordinate \( D = x \) at every instant. But we can calculate the other coordinate \( y = dD/dt \) from the 1st order differential equation of the phase space. A plot \( dD/dt \) versus \( D \) generates a continuous phase plane. In the reconstructed discrete phase plane of the measured data \( D_{t+1} \) versus \( D_t \) the trajectory shows the same cyclic behavior as in the continuous phase plane of the model.

In general, the dynamics on a phase plane are given by a pair of coupled differential equations \( \frac{dx}{dt} = f(x, y) \) and \( \frac{dy}{dt} = g(x, y) \). Sometimes we can only measure \( x \). But then we can calculate \( \frac{dx}{dt} \) and get the value \( f(x, y) \), which also contains some information about \( y \). This information is often sufficient to reconstruct the dynamics of trajectories in the \((x, y)\) phase plane. The 1st derivative of \( x \) at time \( t \) is calculated using the well-known formula

\[
\frac{dx(t)}{dt} = \lim_{h \to 0} \frac{[x(t + h) - x(t)]}{h}.
\]

A time series of measurements \( D(t) = x(t) \) without noise consists of measurement data \( D_0, D_1, D_2, \ldots \) at discrete times \( t = 0, 1, 2, \ldots \). The derivative of \( x \) at time \( t \) can be approximated by differences of corresponding measurement data

\[
\frac{dD_t}{dt} = [D_{t+h} - D_t]/h \quad \text{with } h = 1, 2, \ldots
\]

The smallest useful value of \( h \) is 1. But sometimes it is appropriate to select a larger time-lag \( h \). By plotting \( D_{t+h} \) versus \( D_t \) the phase plane dynamics of a system can often be reconstructed from measurements \( D_t \) without the direct measurement of the variable \( y \) of the model. In this case, the dynamics in the reconstructed \((D_t, D_{t+h})\) phase plane are similar to the original \((x, y)\) phase plane of the dynamical system.

Nonlinear dynamical systems generating chaos must be determined by at least three equations. As an example, the Lorenz attractor (Fig. 2.21) is generated in a phase space with three coordinates \( x(t), y(t), \) and \( z(t) \), which are determined by three nonlinear differential equations. Figure 2.33a shows a time series of measured data \( D_t \) from the Lorenz system. If only one variable \( D(t) = x(t) \) can be measured, a Lorenz attractor in a \((D_t, D_{t-h}, D_{t-2h})\) phase space (Fig. 2.33c) can be reconstructed with great similarity to the original Lorenz attractor of the \((x, y, z)\) phase.
space (Fig. 2.33b). In general, a time series can be embedded in a $p$-dimensional space with $p$-coordinates $D_t = (D_t, D_{t-h}, D_{t-2h}, \ldots, D_{t-(p-1)h})$ and time-lag $h$. According to Takens’ embedding theorem [2.66], the reconstructed dynamics are geometrically similar to the original for both continuous-time and discrete-time systems. The sequence of points created by embedding a time series is called the trajectory of the time series.

In practice, decisions about chaotic dynamics are rather difficult. How can we decide that a time series of measured data is not generated by noisy irregularity but by highly structured chaotic attractors? A chaotic attractor is determined by a trajectory in a bounded region of a phase space with aperiodic behavior and sensitive dependence on initial conditions. These criteria – determinism, boundedness, aperiodicity, and sensitivity – can be checked by several techniques of time series analysis. A system is called deterministic when future events are causally set by past events. For example, a finite-difference equation like $x_{t+1} = f(x_t)$ is deterministic if $f(x_t)$ has only one value for each value of $x_t$ and the future value $x_{t+1}$ can be calculated from the past value $x_t$ by function $f$.

How can we decide that measured data of past events $D_t$ determine the future events $D_{t+1}$? We suppose that measurements are made up to time $T$, and that a prediction of the value at time $T + 1$ should be made. Again, we use the afore mentioned procedure to embed

![Fig. 2.33a. Measured time series of Lorenz system [2.67]](image)

![Fig. 2.33b,c. Trajectory in $(x, y, z)$ phase space (b) and reconstructed trajectory in $(D_t, D_{t-h}, D_{t-2h})$ phase space with time-lag $h$ of Lorenz attractor (c) [2.68]](image)
the time series in a $p$-dimensional space with time-lag $h$. The embedding point at time $T$, representing measurements of past events, is $D_T = (D_T, D_{T-h}, \ldots, D_{T-(p-1)h})$. We look through the finite rest of the embedded time series for the closest point to $D_T$, which is called $D_c$ at time $c$. $D_c$ represents the past events to the measurement $D_{c+1}$. As $D_T$ is close to $D_c$, the measured value $D_{c+1}$ is expected to be close to $D_{T+1}$ in deterministic dynamics. Thus, the prediction of $D_{T+1}$ is identified with the measured value $D_{c+1}$. The difference between the prediction and $D_{T+1}$ is the prediction error which indicates the quality of the prediction. A more meaningful indication of determinism uses the average of many prediction errors.

Dynamics are bounded if they stay in a finite range of the phase space and do not approach $+\infty$ and $-\infty$ when time increases. In the case of noise, the trajectories spread unbounded all over the phase space. A chaotic attractor is always bounded in a certain region of the phase space. But practically measured data are, of course, always in a finite range, because the physical universe is finite. Thus, boundedness of measured data is related to the concept of stationarity. A time series is stationary if the mean and standard deviation remain the same throughout the time series. Aperiodicity means that the states of a dynamical system never return to their previous values. But values of states may return more or less to previous values. Thus, aperiodicity is a question of degree. How can we determine the degree of aperiodicity in measured data?

Again, we embed the time series of measurements in a $p$-dimensional space with time-lag $h$. Each point $D_t = (D_t, D_{t-h}, \ldots, D_{t-(p-1)h})$ represents the state of the dynamical system at time $t$. The distance of two states is measured by the distance between two points at times $i$ and $j$ by $\delta_{ij} = |D_i - D_j|$ (Fig. 2.34). If the time series is periodic with time $T$, the values of states are repeated after $T$ values for several times. In this case, the distance $\delta_{ij}$ of the points representing times $t$ and $j$ is zero for $|i - j| = nT$ with $n = 0, 1, 2, \ldots$. The degrees of periodicity and aperiodicity can be studied in recurrence plots of points $(i, j)$ if the distance of $D_i$ and $D_j$ is smaller than a given distance $r$.

Such plots depict how the reconstructed trajectory recurs or repeats itself. The number of dots in a recurrence plot shows how many times the trajectory came

**Fig. 2.34a,b.** Recurrence plots with periodicity for quadratic map $x_{t+1} = 3.52x_t(1-x_t)$ ($p = 2, r = 0.001$) (a) and aperiodicity for chaotic map $x_{t+1} = 4x_t(1-x_t)$ ($p = 2, r = 0.001$) (b) [2.69]
within a distance $r$ of a previous value. The correlation integral $C(r)$ defines the density of points $(i,j)$ in a recurrence plot where the measured time series $D_i$ and $D_j$ are closer than $r$ for $i \neq j$ (Fig. 2.34). The correlation integral is an effective concept of chaotic time series analysis [2.70]. If the distance $r$ increases, more dots appear in the recurrence plots with an increasing density $C(r)$. The characteristic curves of $C(r)$ are flat for a periodic system, with a gentle slope for a chaotic system, and with a steeper slope for a random system.

There is an important relationship between the correlation integral and the concept of fractal dimension (compare Sect. 2.4). Consider the scattered points in an area within a distance $r$ to a reference point on a 2-dimensional surface (e.g., a circle with radius $r$ and area $\pi r^2$) or in a 3-dimensional space (e.g., a sphere with radius $r$ and volume $4/3 \pi r^3$). In general, for points scattered throughout an object in a $\nu$-dimensional space, the number of points closer than distance $r$ to a reference point is proportional to $r^\nu$. The correlation integral was introduced as a measure for the density of scattered points within a distance $r$ to a reference point of a recurrence plot. Thus, the correlation integral of a scattering of points throughout a $\nu$-dimensional object is proportional to $r^\nu$, i.e., $C(r) = qr^{\nu}$ with a constant $q$ of proportionality. The correlation dimension $\nu$ of the $\nu$-dimensional object can be calculated by the logarithm of this equation, i.e. $\log C(r) = \nu \log r + \log q$. In order to find the correlation dimension $\nu$, we can plot $\log C(r)$ versus $\log r$ and determine the slope of the resulting line. This procedure can also be used to find the fractal dimension of an object.

In time series analysis the correlation dimension is sometimes used to find attractors. It is well known that chaotic attractors are often self-similar with fractal dimension. If a time series is generated by a chaotic system, the trajectory of the time series, which is reconstructed from the measurement data by embedding, has the same topological properties as the original attractor of the system, as long as the embedding dimension is large enough. Takens proved a method for finding an appropriate embedding dimension for the reconstruction of an attractor: If the original attractor has the dimension $\nu$, then a dimension $p = 2\nu + 1$ is adequate for the embedding space of the reconstructed attractor. But this method yields no procedure for finding a chaotic attractor, because its existence has been assumed in order to determine its dimension from the measurement data.

Another way to characterize chaotic dynamics is to measure the strength of their sensitive dependence on initial data. Consider two trajectories starting from nearly the same initial data. In chaotic dynamics only a tiny difference in the initial conditions can result in the two trajectories diverging exponentially quickly in the phase space after a short period of time (Fig. 2.35). In this case, it is difficult to calculate long-term forecasts, because the initial data can only be determined with a finite degree of precision. Tiny deviations in digits behind the decimal point of measurement data may lead to completely different forecasts. This is the reason why attempts to forecast weather fail in an unstable and chaotic situation. In principle, the wing of a butterfly may cause a global change of development. This “butterfly effect” can be measured by the so-called Lyapunov exponent. A trajectory $x(t)$ starts with an initial state $x(0)$. If it develops exponentially fast, then it is approximately
given by \( |x(t)| \sim |x(0)|e^{\Lambda t} \). The exponent \( \Lambda \) is smaller than zero if the trajectory is attracted by attractors, such as stable points or orbits. It is larger than zero if it is divergent and sensitive to very small perturbations of the initial state.

Let us consider a finite-difference equation \( x_{t+1} = f(x_t) \) with two nearby initial positions \( x_0 \) and \( y_0 \) in the phase space. By iterated application of the function \( f \) we get \( x_t = f(x_{t-1}) = f^t(x_0) \) and \( y_t = f(y_{t-1}) = f^t(y_0) \) with \( t = 0, 1, 2, \ldots \). If the positions \( x_t \) and \( y_t \) are separated exponentially fast by iterations, then their distances are \( |y_t - x_t| = |y_0 - x_0|e^{\Lambda t} \) with \( \Lambda > 0 \). For increasing \( t \to \infty \) it follows \( (1/t)|y_t - x_t|/|y_0 - x_0| \to \Lambda \). If the path of the trajectory is within a bounded region, the exponential separation only occurs when the initial positions are very close to each other. In this case, we decrease the difference \( |y_0 - x_0| \) before we determine the limit for increasing \( t \to \infty \). The Lyapunov exponent of the trajectory \( x_t = f^t(x_0) \) can then be defined by the constant

\[
\Lambda = \lim_{t \to \infty} \frac{1}{t} \lim_{|y_0 - x_0| \to 0} \ln \frac{|y_t - x_t|}{|y_0 - x_0|} = \lim_{t \to \infty} \frac{1}{t} \ln |f^t(x_0)|/|y_0 - x_0| = \lim_{t \to \infty} \frac{1}{t} \ln |df^t(x_0)/x_0| = \lim_{t \to \infty} \frac{1}{t} \sum_{i=0}^{t-1} \ln |df(x_i)/x_i|.
\]

For continuous dynamical systems with differential equations, the trajectory is a vector \( \mathbf{x}(t) \) with a Lyapunov exponent \( \Lambda = \lim \sup 1/t \ln |\mathbf{x}(t)| \). The Lyapunov exponent provides a measure for the mean convergence and divergence rate of neighboring trajectories of a dynamical system. For an \( n \)-dimensional system, the \( n \) Lyapunov exponents \( \Lambda_1 \geq \Lambda_2 \geq \ldots \geq \Lambda_n \) describe different types of attractors. For non-chaotic attractors we can distinguish asymptotically stable equilibrium with \( \Lambda_i < 0 \) \((i = 1, \ldots, n)\), asymptotically stable limit cycle with \( \Lambda_1 = 0 \) and \( \Lambda_i < 0 \) \((i = 2, \ldots, n)\), asymptotically stable two-torus with \( \Lambda_1 = \Lambda_2 = 0 \) and \( \Lambda_i < 0 \) \((i = 3, \ldots, n)\), and asymptotically stable \( m \)-torus with \( \Lambda_1 = \ldots = \Lambda_m = 0 \) and \( \Lambda_i < 0 \) \((i = m + 1, \ldots, n)\). A chaotic system must have at least one positive Lyapunov exponent. In the 3-dimensional case, the only possibility for chaos is \( \Lambda_1 > 0, \Lambda_2 = 0, \Lambda_3 < 0 \) with \( \Lambda_3 < -\Lambda_1 \).

Dynamical systems can be classified by attractors with increasing complexity from fixed points, periodic and quasi-periodic up to chaotic behavior. This classification of attractors can be characterized by different methods, such as typical patterns...

\[\text{Fig. 2.35. Exponential dependence on initial conditions measured by Lyapunov exponent } \Lambda [2.71]\]
of time series, their power spectrum, phase portraits in a phase space, Lyapunov exponents, (fractal) dimension, and a measure of their information flow (Kolmogorov-Sinai-Entropy), which will be discussed in Sect. 5.3 in more detail. Table 2.1 yields an overview of these degrees of dynamic complexity, which form the framework for the complex dynamical approach of this book.

One of the most significant concepts is that of the fractal dimension, a measure of the roughness of an object. Fractality seems to be a natural feature of reality. Rocky coastlines consist of cliffs and crannies. Rocks with rough surfaces erode. In organic growth, such as that for the airways of the lungs, a fractal process of iterative division is the natural outcome of the genetic rules for animal development. In Euclidean geometry, we are familiar with the single dimension of a straight line, or the two dimensions of a plane. An example of a fractal dimension is that of Koch’s curve (Fig. 2.36). In order to measure its length, one starts with a ruler that is one-third of the breadth of the object (the curve). This ruler corresponds to each line inside the curve in the top panel. The line fits inside the curve four times. The ruler is then shortened to a third of its original length, as shown in the bottom diagram. Because this shorter ruler can fit into more “crannies” of the curve, the length of curve obtained using this ruler is greater than that given by the original ruler, (by four-thirds). For each change in state, the length measured is multiplied by the same fraction, four-thirds. The fractal dimension is then defined as the ratio of the logarithm of 4 to the logarithm of 3, or 1.2618... The intuitive sense of this “fractal” number is obvious: the curve is crinkly, so it fills more space than a one-dimensional straight line does. However, it does not completely fill the two-dimensional plane.

<table>
<thead>
<tr>
<th>phase portrait</th>
<th>time-history response</th>
<th>power spectrum</th>
<th>auto-correlation</th>
<th>Lyapunov-exponents</th>
<th>(fractal) dimension</th>
<th>KS-entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(−, −, −)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0, −, −)</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0, 0, −)</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(+, 0, −)</td>
<td>2 &lt; h &lt; 3</td>
<td>eKS &gt; 0</td>
</tr>
</tbody>
</table>

Table 2.1. Dynamic complexity of attractors for 3-dimensional systems [2.72]
Using analog recursive procedures, we can construct Hilbert and Sierpinski curves that fill a plane with self-similar patterns, in iterated steps of increasing density. These curves seem to be more than lines, with one dimension, but less than planes, with two dimensions. Their dimensions are “fractions” between the integers one and two. The fractal dimension can be illustrated by the geometrical dimension $D$ of similarity. For a Euclidean object of dimension $D$, the length, area, or volume of an object with edge length $\varepsilon$ is proportional to $\varepsilon^D$. For example, a square with edge length $\varepsilon$ has an area of $\varepsilon^2$, while a cube has a volume of $\varepsilon^3$. For self-similar objects, one way to measure the length, area, or volume of an object is to count the number of self-similar copies. If there are $N$ copies each with an edge length $\varepsilon$, then the length, area, or volume of the object is related to its dimension: $N$ is proportional to $\varepsilon^D$. Thus, one obtains $D \approx \log N / \log \varepsilon$. For Koch’s curve, the number of self-similar copies is $N = 4$ and the edge length is $\varepsilon = 3$. If phase portraits of chaos attractors have a fractal dimension (Table 2.1), they are termed “strange.” Time series are sometimes characterized by statistical self-similarity on different scales (e.g., Fig. 8.16). Thus, a fractal dimension could hint at chaos, but its presence alone does not indicate chaos.

Self-similar mathematical objects consist exclusively of smaller self-similar copies of themselves. Our procedure for calculating the dimension of a fractal object is only useful if we know the number $N$ of self-similar copies and the size $\varepsilon$ of the original relative to each copy. For practical applications (e.g., a map or picture of a fractal object or real objects in the three-dimensional world), we need a better procedure for estimating the fractal dimension. The following procedure comes directly from the definition of the fractal dimension. In a first step, all points in the object are covered with $N(\varepsilon_0)$ or cubes of edge length $\varepsilon_0$. This step is repeated with squares or cubes of edge length $\varepsilon_1 = \varepsilon_0/2$, then with $\varepsilon_2 = \varepsilon_1/2$, and so on. By doing this, we obtain a function $N(\varepsilon)$ sampled at the values $\varepsilon = \varepsilon_0, \varepsilon_1, \ldots$ In theory, the
Fig. 2.37. Multifractal process of turbulence

dimension $D$ is defined by $\lim N(\varepsilon) = k \cdot \varepsilon^{-D}$, with a constant $k$. In practice, $D$ can be estimated as $D \approx \left( \frac{\log(N(\varepsilon_{i+1})/N(\varepsilon_i))}{\log(\varepsilon_{i+1}/\varepsilon_i)} \right)$. However, the squares or cubes should not be made smaller than the cells or particles that are considered to be the building blocks of the object [2.73].

Intuitively, a fractal is a pattern or object whose parts echo the whole, only scaled down. By contrast, a multifractal has more than one scaling ratio in the same object. Some parts of the object shrink quickly, others slowly. Multifractals resemble the way in which many aspects of nature really work more closely than fractals. Different clusters are formed on the surface of the Earth in multifractal processes; they are not always distributed and scaled in the same way. On a stormy day, the wind velocities form clusters of high gusts interspersed with gentler breezes. One can think of a multifractal as being composed of an infinite hierarchy of different fractal sets. An example is given in Fig. 2.37, which shows a hierarchy for the vertical cross-section of stratified turbulence. The generic multifractal process of turbulence is a cascade of cells with different distributions of whirls. Mathematically, multifractals are defined by two groups. One determines the statistics (more precisely, they vary as a function of scale), while the second defines the notion of scale itself [2.74].