Chapter 1
What Is NONLINEAR?

Abstract The word nonlinear indicates something that conflicts with linear. It is not positively defined, but rather the “antithesis” of linear. Despite a lack of concrete contents, nonlinear is a powerful and productive key word indicating the direction of contemporary sciences. This is because growing criticism of linear theory is pushing various fields toward the study of nonlinearity. In this introductory chapter, we shall give an overview of the meaning of nonlinear as a mathematical structure, as well as of its impact on sciences.

1.1 Nature and Science

1.1.1 Natura Vexata

Nature is, originally, what is great, profound and capricious for human beings. Ancient people developed various idols to represent diverse aspects of nature—gods or monsters who tossed people back and forth. What was frightening was their tremendous energy and unpredictable behavior.

In modern ages, however, our view of nature has been fundamentally altered. Scientists have probed the anatomy of nature to reveal the “harmony” in the essence—one may express the “logic” of nature in the words of mathematics; one may harness various parts of nature by making machines. We began to dissect predictable (or reproducible) parts of nature and to cultivate them.

Predictability is not only a forecast in time future—its general meaning is the knowledge of “causal relations.” For instance, pharmacology starts from firsthand knowledge (experimental wisdom such that this medical herb is effective to that case), proceeds with componential analyses on the medical herb to find the effective chemicals, and pins down the mechanism of the medical effect at the molecular level. In this way, scientific research is directed toward universalities or principles.

However, can the sciences really tame nature within a predictable area?

A thing in the real world is a system that is a complex composition of plenty of elements. It is often difficult to study it directly. So we divide the system under examination into elements as small as we deem necessary and start by studying the simplest problem. In Discourse on the Method [6], René Descartes (1596–1650)
R. Descartes proposed a method of science that can develop a clear and distinct theory. His reductionism, described in his famous book *Discourse on the Method*, gave a guiding principle of modern science. Portrait by Frans Hals (unsigned) (Louvre Museum, Paris)

proposed a method of giving a solid platform from which a *clear* and *distinct* theory could be developed; he formulated “four precepts” describing the idea of dividing a system into elements (Fig. 1.1). By these statements, the *reductionism* in sciences is often connected to Descartes (though its origin actually dates back to ancient Greek philosophy). *Reduction* is a basic notion of philosophy, which generally means the replacement of a thing by something that we can directly manipulate. Reductionism is a strategy of science, which teaches us to start by separating elements from the complex real world and to manipulate each of them by experimental or theoretical methods.

The operation of separating an element from the real world means, in case of an experiment, the construction of an “experimental device,” and, in case of theory, the formulation of a “model.” The first work of an experimentalist is to cut out an element (object) from nature and to place it in the isolated space of his device.\(^1\) The device has some active manipulators by which the experimentalist can control the element and observe its responses. The element, effectively disconnected from nature, is under the perfect control of the device. The researcher tries to establish the

---

\(^1\) Francis Bacon (1561–1626) proposed a method of developing philosophy, the inductive reasoning from fact to law, which can free one’s mind from “idols” (*idola*). He pointed out that the description of nature in its free condition (*natura libera*) is not sufficient and claims the necessity of experimental method to observe afflicted nature (*natura vexata*) [1].
possibility of experimental observations by achieving perfect control of the object, i.e., by eliminating the uncontrollable element of freedom from the object.

A theorist also starts from the epistemological operation of separating an object from nature. For example, when a physicist discusses the motion of a “particle,” such an object is an abstract model that has eliminated all complicated relations with other elements as well as internal properties. A particle, called “point mass,” is represented only by its mass, position of the center of inertia, and time. The famous theory of Galileo Galilei (1564–1642) asserts: “A heavy object and a light object fall at the same speed, in the ideal condition”. Here, the ideal condition means that the interaction of the objects and air can be ignored. Experimentally, this condition can be satisfied when we put the objects in a vacuum vessel. Objects in the real world, however, fall in totally different ways; a bird’s wing, for instance, moves on an extremely complicated trajectory. It is still an impossible problem to calculate such an orbit that is affected by the interactions with the air, even with the help of a top-level computer. The notion of ideal condition eliminates such tremendous complexity and allows a theorist to assume an abstract model of an object.

We perceive here a sense of apprehension about scientific statements. There is a large gap between the problems in the real world and the understanding of the decomposed elements, the anatomized nature. We can control only slight movement of small elements; we are describing a small part of the universe—a falling apple, one planet, one group of animals, weather in a narrow region, etc.—on which we arbitrarily focus, as with the lens of a camera. The object of interest is, then, isolated from the rest of the universe. Science has described only such fragments of nature. And our naive images of the unpredictable and diverse aspects of nature have been hidden (or moved to the periphery) by careful and argumentative logic in the process of dissecting (rupturing) it. To restore our original view of nature, it is now necessary to polish our sense of science concerning composition/synthesis, as opposed to decomposition/reduction.

Linear theory explores the “composition” and “decomposition” of objects in their ultimate simplicity. The notion of linear is the generalization of the proportionality relation to higher dimensions (a larger number of variables)—the term “linear” is used because the graph of a proportionality relation is represented by a “line.” The proportionality relation is the simplest law that one may assume between two variables (see Sect. 1.2). For example, if the price of an apple is 50 cents, five apples cost $2.50. If an electric voltage of 1 V can drive a current of 0.1 A through a resistor, one may guess that the current will become 1.0 A when the voltage is increased to 10 V. Like these examples, it is quite conceivable to consider a proportionality relation between a pair of variables. The fundamental mathematical structure of the proportionality relation—that is, as we shall see in Sect. 1.3, the axioms of composition/decomposition—can be generalized for relations among

---

2 Galileo, of course, did not do the experiment in a vacuum vessel. Recognizing that the difference of the falling velocity originates from the friction of the air, he tried to reduce the difference by making samples of the same size and shape from wood and metal.
many variables. A variable consisting of multiple components is called a *vector*. Students may encounter the word vector, first, when they learn mechanics: A “force” is a vector that is represented by an “arrow,” and it may be composed/decomposed by geometrical manipulation using a *parallelogram*. Here, what we mean by drawing arrows and parallelograms to decompose a vector is the *proportional distribution* into the directions of the two sides of the parallelogram; composition is the inverse operation. Repeating this method, we can compose/decompose a vector in spaces of three or more dimensions

The eyes of scientists try to “reduce” an object into a *vector* by *parameterization* (i.e., measurement of some parameters that may represent the characteristics of the object). The real object is converted into a geometrical object and is projected onto the *vector space* (which is also called a *linear space*). For this “space,” where the object = vector is to be placed, the “linear structure,” i.e., an axiomatic system that enables the composition/decomposition of a vector (see Sect. 1.3.2), is prepared. Hence the notion of *linear law* = *generalized proportionality relation* coincides with the “frame of the space”; the straight graph of a linear law is, in itself, a linear (sub-) space.\(^3\)

However, there is a fundamental gap between the “theoretical space” (embodying “freely composable/decomposable object” or “space structured by a linear law”) and our naïve recognition of the “real world” (involving unpredictable phenomena and infinite diversity that are impossible to compose/decompose). Where does this unbridgeable gap start?

### 1.1.2 Syndrome

*Syndrome* is originally a technical term of medical science. The human body is a complex system where many organs (each of which has its own particular function) are cooperating. Slight sickness, caused by a little slump of an organ, may be cured by correcting the function of the organ or by a surgical operation removing a small part. When a serious problem occurs in some part, however, two or more organs are involved in a chain reaction and recovery becomes more difficult. Such a phenomenon is called a syndrome.

This medical word is often used as a rhetoric when a system loses control and starts unpredictable behavior. Our modern world is flooded with complex artifacts that human beings have added to nature. For example, a nuclear plant is a huge, complex system that is composed of a control system for nuclear reactions, a circulation system of core-cooling water, steam generators, turbine power generators, power transmission lines, the electric power network system, as well as the system operators. A “perturbation” in such a system can be amplified by operators’ errors,

---

\(^3\) The expression of the linear law can be simplified—decomposed into independent proportionality relations—by choosing a frame (set of the basis vectors) of the linear space to be amenable to the linear law. The best choice is found by solving the so-called eigenvalue problem; this central strategy of linear theory—structuring the space by the law itself—will be explained in Sect. 2.3.
and, in the most serious case, the trouble may reach core meltdown. Such a nuclear reactor accident is called a “China Syndrome.” This word expresses an absurd image that a nuclear accident in an American power plant will eject reacting hot nuclear fuel that will melt the rock and continue sinking to reach China, on the other side of the earth. Unfortunately, a severe core-melting accident, which gave (partial) reality to this rhetoric, occurred in Chernobyl nuclear plant in the old Soviet Union and caused an unprecedented disaster in 1986. The behavior of a system with some complexity, even if it was designed and made by human beings, can be uncontrollable—Chernobyl’s tragic episode tells the fundamental difficulty of predictions about a complex system.

“Environmental problems” are causing fears of syndromes that may destroy our modern civilized society. Does our consumption do great damage to the earth? Will the influences of our activities cause an unforeseen phenomenon? Will nature avenge our abuse with its massive power? Nature is still a threat to us due to its tremendous energy and unpredictable behavior. Here, the scales are the central problem; our concerns are the “scale” of the environmental impacts that can bring about a syndrome and the “scale” of the resultant environmental perturbations. The scientific content expressed by the word “environment” is the dynamics of a huge, complex system that combines many kinds of elements (atmosphere, ocean, and the ecosystem consisting of many kinds of species). The connections of these elements change variously, depending on the scales (magnitudes) of the perturbations, and make prediction of a syndrome very difficult.

In a system such as the human body, a complex “plant”, or the global environment, the connections relating many elements are modulated depending on the magnitude of the movement of each element. A system works as a united body, and its dynamics cannot be understood if it is decomposed into separate elements. Then, how can we analyze the composition of elements? How can we have a perspective concerning the scale?

Carelessness about decompositions (or ruptures) and unconcern about scales have been fostered by linear theory. Thus, nonlinear is crucial for contemporary science to regain an understanding of the scale of events that make the connections of elements irreducible; it is a challenge to the fictitious vision of abstracted events. Therefore we should not restrict our discussions to a matter of “dichotomy” of mathematical structures. The relation between linear and nonlinear should be analyzed as a problem retroactive to the genesis of science.

### 1.1.3 Déconstruction of Linear Theory

The scope of nonlinear science (the science devoted to nonlinearity) is not merely a “complement” of linearity. Motivated by criticisms of the narrowness and persistence of linear science, it has developed into a wider realm subsuming the linearity.

Borrowing the term coined by Jacques Derrida (1930–2004), we can say that the aim of nonlinear science is the déconstruction of linear science. Here, déconstruction means a strategy of philosophy that critically analyzes a “central vs. peripheral” relation (created in an implicit manner) of dichotomy and aims at reversing the
It does not intend to bring about “destruction” (collapse) of the system that has been structured by the “center”—the previous center will maintain its effectiveness, though its validity will be rather limited. Using this sophisticated strategy, we are going to include into the nonlinear science the solid mathematical structure of linear science.

The framework of linear theory is embodied by some mathematical laws that generalize the proportionality relation. There is no room for doubting the importance of linear theory as the starting point of all rational considerations. However, one should not expect that any law or principle holds unrestrictedly. A proportionality relation, indeed, distorts when the magnitudes of variables become to some extent large. Here, note that the concept of scale (or magnitude) intervenes in our discussion. Linear theory excludes the notion of scale from the arguments; this exclusion poses a fundamental, but hidden, limitation to its legitimacy. The déconstruction of linear theory (or “linearized” theory) starts by rehabilitating the notion of scale—focusing attention on the world of large scales where the proportionality relation distorts and linear theory ceases to apply.

The déconstruction of linear theory expands to the déconstruction of “Science.” Here, “complexity” is the key word for the déconstruction of the “central science” based on Cartesian reductionism. The reductionism that was the methodology to search for the “order” = “simplicity” of nature was, simultaneously, a strategy to bundle off the “disorder” = “complexity” to the periphery of science. Now, the exiled “complexity” tries to reverse its position by shaking the structure of the traditional science that holds “order” in the center—radical problems of contemporary sciences are throwing fundamental doubts on the validity of the science that shuts itself up in the area of order.

The dichotomy between the cosmos and the chaos has a parallel relation with the dichotomy between the linearity and the nonlinearity. Linear theory is, in essence, an exploration of “order”; it has constructed a microcosm of logic that is not violated by complexity. The déconstruction of linear theory—the most clear expression of order—will bring about the déconstruction of science, allowing the invasion of complexity.

1.2 The Scale of Phenomenon / Theory with Scale

1.2.1 The Role of Scale in Scientific Revolutions

The great joy of a scientist is to find truth that no one noticed before. New knowledge of science emerges to overturn old common sense (idols). A revolution of science is

---

4 The word déconstruction is a coinage of Derrida [4, 5]. He perceives a fundamental inequality in various dichotomies (such as actuality/potentiality or parol/écriture) and criticizes the politics that utilizes the implicitly structured conflicts. He criticizes the Hegelian dialectic process that intends to reconcile conflicts by sublation (aufheben) as abstraction of actual problems.
the discovery of an unknown world of new *scale*—or, it may be more appropriate to say that a new *scale* is discovered to indicate a new world.

We, educated by modern science, know that the ground surface is not a plane, terra is not firm—the earth is a spherical object, inside of which is a hot fluid that convects actively in large space–time scales, and the continents are just thin and fragile husks. The earth is a tiny planet circulating around the sun. Our sun is an average star in an ordinary galaxy in the universe. Our universe emerged 10–20 billion years ago and is still expanding. There may exist other universes.

The notion of our “ancestors” can no longer assume the invariance of our species—human (Homo sapiens) evolved between 0.4 and 0.25 million years ago (the Homo genus diverged from the Australopithecus about 2 million years ago). About 100 million years ago, huge reptiles were ruling the earth. When fossils of huge bones were found from an old stratum, the forgotten history was revived with amazement. How long can the human genus survive? According to paleontology, the average longevity of a genus is several million years.

These are the statements of science narrating large space–time scales. Modern science has also widened our view to small-scale worlds.

On the micro-scale, matter is reduced to “particles.” The diversity of materials is explained by the variety of compounds or array structures of particles. From the viewpoint of materialism, biodiversity is explained as the degree of freedom of the particle array permitted to the polymer that is called DNA. In the micro-scale realm of quantum theory, a “particle” exhibits characteristics of a wave (it causes diffraction and interference), and the “existence” of a particle can be interpreted only in a “probabilistic” meaning.

The revolution of science may completely deny preceding theories, as Darwin’s theory of evolution did. However, in many cases (especially in physics), a new theory expands the horizon of knowledge by subsuming an old theory; the old theory keeps a narrow position as an “approximate theory” that holds in a special limit of *scale*. For example, Newton’s classical mechanics is still valid as the macroscopic (large energy) limit of quantum mechanics, as well as the microscopic (small energy) limit of relativity. The “turf” that Newton’s classical mechanics still reserves is the world of human’s first-hand scale; from there, physics began to development. Even after our scope of interest has expanded far beyond this turf, the old theory does not turn inside out, but instead maintains a limited value as a local truth.

In order to subsume an old theory, a new theory should involve a *scale*; an old theory is, then, localized in a limited domain of this scale. Because the old theory did not recognize such a *scale*, it was unaware of its “limitation.”

---

5 When the world of an old theory is reviewed from the world of a quite different scale (that was the “periphery” for the old authority), the old theory maintains a limited truth, moving backward to the seat of “approximate theory.” In the previous section, this strategy for reversing the central/peripheral perspective structured by classical theories was called *déconstruction*, borrowing the term from Derrida.
We can say that the aim of nonlinear science is to recover the scale of which linear theory lost sight. A more careful explanation of this point will become an introduction to nonlinear science.

### 1.2.2 The Mathematical Recognition of Scale

The literal meaning of scale is the unit by which we can quantify an observed quantity. The unit is what we only have to choose arbitrarily. However, we are going to reveal that the choice of scale is not freed from the “object”—the notion of scale includes a particularity of a phenomenon (or the theory describing a phenomenon). What does this mean?

Let us consider a pair of variables (parameterization of an object) $x$ and $y$, and denote by $\delta x$ and $\delta y$ their respective variations. When we find a certain relation between $\delta x$ and $\delta y$, we may claim that there is a law governing $x$ and $y$. If $|\delta x|$ and $|\delta y|$ are “small,” this relation is normally a proportionality relation, i.e.,

$$\delta y = a \delta x \quad (a = \text{constant}).$$  

(1.1)

Here, “smallness” is a relative notion—we can say “small” (or “large”) when we compare a parameter with a certain reference (or scale). However, we do not know a priori the actual reference that determines the “smallness” of the variation. Instead, we may say the variations $|\delta x|$ and $|\delta y|$ are small as far as relation (1.1) holds, i.e., we infer the scale from the limitation of the proportionality relation.

There are many examples of proportionality relations: elastic law of a spring, Ohm’s law of electric resistance, etc.—many elementary physics laws taught to children are proportionality relations. Newton’s equation of motion is also a proportionality relation between force and acceleration. The proportionality coefficients (the elastic constant, resistance, mass, etc.) may be assumed to be constant numbers for some magnitudes of variations of the parameters. However, when the magnitudes of the extension, the electric current, or the velocity become large, the coefficients are no longer constant. The scale that determines “smallness” is discovered when the proportionality relation is destroyed.\(^6\)

Complexity in the real world may develop in the regime where the proportionality relation—the law of simplicity—is violated. In various problems of contemporary sciences, such as climate change, change in ecosystems, and economic fluctuations, finding the scale (the basis to know if a variation is either small or large) is the key issue. Linear theory keeps silent about the scale—it is “unaware” of the scale. By

\(^6\) The scale is not an a priori number prescribed in a law; it is discovered when a new theory emerges to subsume the old theory, and, then, it measures the range where the old theory applies as the approximation of the new theory. As we discussed in Sect. 1.1.3, \textit{déconstruction} of an old framework switches the viewpoint from the center to the periphery; the periphery determines the scale of the landscape. The new theory, when it is constructed, pretends to have known the scale, however.
finding a deviation from a linear theory (proportionality relation), we can estimate the \textit{scale} that “localizes” the linear theory.

Let us formalize the foregoing arguments, invoking the theory of \textit{Taylor expansion}. We assume that a function \( f(x) \) is smooth (analytic) in a certain neighborhood of a point \( x = x_0 \). Then we can write \( f(x) \) as

\[
    f(x) = a_0 + a_1 \delta x + \cdots + a_n \delta x^n + \cdots, \tag{1.2}
\]

where \( \delta x = x - x_0 \), \( a_0 = f(x_0) \), \( a_n = f^{(n)}(x_0)/n! \) and \( f^{(n)} = d^n f/dx^n \).

If the Taylor expansion (1.2) has a non-zero radius of convergence,\(^7\) we can find a certain finite number \( r \) such that (see Problem 1.1)

\[
    \sup |a_n r^n| < 1. 
\]

By this \( r \), we can characterize the \textit{scale} of the function \( f(x) \)—using \( r \) as the \textit{unit}, we rescale \( x \) (simultaneously, we shift the origin of \( x \) to \( x_0 \)):

\[
    \tilde{x} = \frac{x - x_0}{r}. \tag{1.3}
\]

We call this process \textit{normalization}. Using the normalized variable \( \tilde{x} \), we can rewrite (1.2) as

\[
    f(\tilde{x}) = \tilde{a}_0 + \tilde{a}_1 \tilde{x} + \cdots + \tilde{a}_n \tilde{x}^n + \cdots, \tag{1.4}
\]

where \( \tilde{a}_n = a_n r^n \) (because \( |\tilde{a}_n| < 1 \), the radius of convergence of the power series (1.4) is greater than or equal to 1; see Problem 1.1). In the normalized Taylor series (1.4), we observe that \( |\tilde{x}^n| \ll |\tilde{x}| \) for \( |\tilde{x}| < 1 \) and \( n > 1 \) (thus, \( |\tilde{a}_n \tilde{x}^n| \ll 1 \)). Hence, within the range of \( |\tilde{x}| < 1 \) (i.e. \( |\delta x| < r \)), we may neglect the higher-order terms and approximate \( f(x) \) by a linear function \( f(\tilde{x}) \approx \tilde{a}_0 + \tilde{a}_1 \tilde{x} \). Note that \( r \) is the scale that determines the range of \( \delta x \) where \( f(x) \) can be approximated by a linear function. We can say that \( \delta x \) is “small” if \( |\delta x| < r \) (i.e., \( |\tilde{x}| < 1 \)); the “smallness” is judged by the proximity of \( f(x) \) to a linear function.

As seen in this formal argument, the mathematical model of a phenomenon, if it is not linear, has a \textit{scale} that characterizes the deviation from a proportionality relation (distortion from a linear graph). The measure of the \textit{scale} is often a number of absolute importance. In the law of motion, for example, the mass (the coefficient relating force and acceleration) may be assumed to be a constant number (i.e., Newton’s law of motion holds) within the range of velocity that is much smaller than the speed of light \( c \)—the \( c \) is the absolute measure of the \textit{scale} in the theory of motion. Einstein’s relativity theory, correcting Newton’s law, claims that the mass of a particle moving with a velocity \( v \) must be

\[^7\text{Let us consider a power series } \sum_n a_n x^n. \text{ If } R^{-1} = \limsup_{n \to \infty} |a_n|^{1/n} < \infty, \text{ the power series converges for } |x| < R. \text{ We call } R \text{ the radius of convergence. If the Taylor expansion (1.2) has a non-zero radius of convergence, } f(x) \text{ is said to be analytic in the neighborhood of } x_0.\]
\[ m = \frac{m_0}{\sqrt{1 - (v/c)^2}}, \]  

(1.5)

where \( m_0 \) is the mass of the particle at rest. Defining \( \tilde{v} = v/c \), and Taylor-expanding \( \gamma(\tilde{v}) = m(\tilde{v})/m_0 \) in the neighborhood of \( \tilde{v} = 0 \), we observe

\[ \gamma = 1 + \frac{1}{2} \tilde{v}^2 + \frac{3}{8} \tilde{v}^4 + \cdots + \frac{\Gamma(v + 1/2)}{\sqrt{\pi} v!} \tilde{v}^{2v} + \cdots. \]  

(1.6)

The radius of convergence of (1.6) is shown to be 1 (see Problem 1.1). Hence, if \( \tilde{v} \) (the velocity normalized by \( c \)) is sufficiently small (now the “smallness” is quantitatively in the sense of \( |\tilde{v}| \ll 1 \)), we may approximate \( m \approx m_0 \) and use Newton’s law as an “approximate law.” Because the speed of light \( c \approx 3 \times 10^8 \text{ m/s} \) is a huge number in the scope of the physics of Newton’s age, the constancy of \( m \) did not come under question. The limitation of Newton’s theory—the nonlinearity in the relation between force and acceleration—became apparent after the scope of physics extended to the scale comparable to \( c \), and the theory found \( c \) as an absolute number to normalize (scale) the velocity.\(^8\)

### 1.3 The Territory of Linear Theory

#### 1.3.1 Linear Space — The Horizon of Mathematical Science

Mathematical science attends to the deep structure of various phenomena; events are projected onto the horizon of mathematics by abstraction into parameters, i.e., by measuring the object and expressing it by numbers = parameters; the “structure” is, then, the relations among the measured parameters. The measurement of an object and its subsequent representation by a set of parameters is called parameterization.

When we parameterize an object \( x \) with a number, we denote the number by \( \tilde{x} \) and distinguish it from the \( x \) itself. The evaluation of the number \( \tilde{x} \), which means the measurement of \( x \), is mathematically written as

\[ x = \tilde{x} \, e, \]  

(1.7)

where \( e \) is the unit of the number \( \tilde{x} \). To put it another way, the unit is defining the basis for the parameterization. We can freely choose \( e \), and, if we change \( e \), the value \( \tilde{x} \) changes.

Generally, we need multiple parameters to parameterize an object. For the time being, we call a set of multiple variables a “vector.” The number of the variables is

\(^8\) The \( c \) is the largest possible speed to relate two events in space–time. Light (electromagnetic wave propagating in vacuum) can propagate at this maximum speed, because it is a massless particle.
called the *dimension* or the *degree of freedom*. The *basis* of a vector is the set of *units* (mathematical definition of these notions will be refined in the next subsection).

For example, to analyze the motion of a “particle” (or the center of inertia of a body), we have to measure its *position* which we denote by $\mathbf{x}$. For the measurement, we first define coordinates. In our space, we need three independent coordinates. In the so-called Cartesian coordinates, we invoke three mutually orthogonal unit vectors (the unit length is, for example, 1 m). Using this set of unit vectors as the *basis*, the position is parameterized as

$$
\mathbf{x} = \hat{x} \mathbf{e}_x + \hat{y} \mathbf{e}_y + \hat{z} \mathbf{e}_z.
$$

(1.8)

Here, the basis is equivalent to the coordinate system.

On a more abstract level, any object is identified as a vector; parameterization (measurement) of the object = vector is done by specifying a set of units = basis and quantifying the set of parameters.

For example, a basket of fruit is, from a mathematician’s perspective, a vector—let us denote it by a symbol $\mathbf{x}$. Suppose that $\mathbf{x}$ contains one apple, two lemons, and three pears. Representing one apple by $A$, one lemon by $L$, and one pear by $P$, we can write

$$
\mathbf{x} = 1A + 2L + 3P.
$$

In business, one may be more concerned with more precise amounts. So, one can measure the weight of each component. Let us represent the apple, lemon, and pear of 1 g by $e_1$, $e_2$, and $e_3$, respectively. Using the set of these symbols as the basis, we can write

$$
\mathbf{x} = x_1e_1 + x_2e_2 + x_3e_3,
$$

where $x_1$, $x_2$, and $x_3$ are, respectively, the weight of each component. A dietitian may see the fruit basket from a quite different viewpoint; defining the basis by the vitamin A, vitamin C, fructose $\cdots$ of 1 g and denoting them by $g_1$, $g_2$, $g_3$, $\cdots$, the fruit basket $\mathbf{x}$ is parameterized as

$$
\mathbf{x} = \xi_1g_1 + \xi_2g_2 + \xi_3g_3 \cdots.
$$

We note that the parameterization of an object = vector is to decompose (resolve) it in terms of the *basis* = set of *units*; the choice of the basis is done in accordance with the observer’s “subject” to see the object—we emphasize this point here to strengthen the argument of Sect. 1.1.1.

We have to use the *proportionality relation* in order to compose or decompose fruit baskets. Calculation of the price, too, is done by the proportionality relation. These algebra are the so-called *law of vector composition*. Before giving a precise and general definition of these concepts, we first see the naïve connection between the notion of the vector space (linear space) and the proportionality relation. Let us
consider two fruit baskets \( x = x_1e_1 + x_2e_2 + x_3e_3 \) and \( y = y_1e_1 + y_2e_2 + y_3e_3 \). If \( \alpha \) of \( x \) and \( \beta \) of \( y \) are added, we obtain a combined fruit basket

\[
\alpha x + \beta y = (\alpha x_1 + \beta y_1)e_1 + (\alpha x_2 + \beta y_2)e_2 + (\alpha x_3 + \beta y_3)e_3.
\]  

(1.9)

Here, we have applied the proportionality relation to each component (element). The calculation of the price has to be based on the proportionality relation, too. Suppose that the unit price (for 1 g) of each fruit is \( p_1, p_2, \) and \( p_3, \) respectively. The price of the basket \( x \), then, is given by \( p_1x_1 + p_2x_2 + p_3x_3 \), which expresses the proportionality relation for each component. This calculation is the so-called inner product of the vector \( x \) (= fruit basket) and an adjoint vector \( p = p_1e_1 + p_2e_2 + p_3e_3 \) that is the “price list.” Here, each \( e_j (j = 1, 2, 3) \) is a unit vector representing the unit of the price of the corresponding fruit of 1 g. We have the orthogonality relation \( e_j \cdot e_k = \delta_{jk} \). The price is, using the conventional notation of inner product, \( p \cdot x \).

Generalizing the foregoing arguments, we can define what we call “measurement of an object” (or “description of an event”)—it is a sequence of processes (1) select \( n \) variables that can represent the object, (2) give a unit to each variable, and (3) evaluate the number of each variable based on the unit. An object in nature (or society) is, by measuring it in an appropriate coordinate system, projected into a vector space of dimension \( n \), and it becomes a mathematical object that can be studied with geometric methods. \(^\text{10}\)

\[\text{1.3.2 The Mathematical Definition of Vectors}\]

Up to now, a set of multiple variables has been called a “vector.” However, the concept of a vector should be defined independently of (prior to) its “measurement” (or description in terms of a set of parameters). The measurement (description) belongs to the side of “subjectivity”; one can select an arbitrary way of measurement. For example, (1.8) gives one possible description of the position vector \( x \)—we are not defining the left-hand side (the position vector) by the right-hand side (its description). The right-hand side changes when we change the basis (coordinate system). Therefore, prior to describing a vector by its components (elements), we need to define what a vector is. We shall define that a vector is a member of a linear space (or, synonymously, vector space) where the law of vector composition applies. This law is basically the generalization (to higher dimensions) of the proportionality law, by which we can manipulate vectors; the measurement of a vector will be done by such manipulations. By axiomatizing the algebra based on the proportionality

\(^9\) \( \delta_{jk} \) is the Kronecker delta that is defined as \( \delta_{jk} = 1 \) for \( j = k \) and \( \delta_{jk} = 0 \) for \( j \neq k \).

\(^{10}\) Galileo said that nature is a book written in the language of mathematics. He thought that phenomena in nature could be studied using geometric methods; prior to this, the object must be mapped to a geometric object—a vector—by the “measurement.” We note that the measurement is based on a choice of a certain scale (unit), and here, “subjectivity” concerning the scale influences the theory. Because of this problem, a deeper analysis is needed for the discussion about the scale, which will be discussed in Chap. 4.
(linear-graph) relation, the “linear structure” is marked to the “space” where we describe/analyze/manipulate the object = vector; the space where we develop theory is, thus, called a linear space.

We denote by \( \mathbb{R} \) the real number field (the totality of real numbers) and by \( \mathbb{C} \) the complex number field (the totality of complex numbers). A set \( X \) is called a linear space if it is endowed with the following law of vector composition: Let \( K = \mathbb{R} \) or \( \mathbb{C} \) (which we call the field of scalars). For arbitrary members \( x \) and \( y \) of \( X \), and an arbitrary number \( \alpha \in K \), we define the sum \( x + y \) and the scalar multiple \( \alpha x \) that satisfy

\[
\begin{align*}
(a) \text{ if } x, y \in X & \text{ and } \alpha \in K, \text{ then } x + y \in X \text{ and } \alpha x \in X, \\
(b) x + y = y + x & \quad (x, y \in X), \\
(c) (x + y) + z = x + (y + z) & \quad (x, y, z \in X), \\
(d) \text{ for every } x \text{ and } y, \text{ there is a unique } z \text{ such that } x + z = y, \\
(e) 1 x = x & \quad (x \in X), \\
(f) \alpha(\beta x) = (\alpha \beta)x & \quad (x \in X, \alpha, \beta \in K), \\
(g) (\alpha + \beta)x = \alpha x + \beta x & \quad (x \in X, \alpha, \beta \in K), \\
(h) \alpha(x + y) = \alpha x + \alpha y & \quad (x, y \in X, \alpha \in K).
\end{align*}
\]

A member of \( X \) is called a vector, and a number in \( K \) is called a scalar (or a coefficient). A linear space is said to be real or complex according as \( K = \mathbb{R} \) or \( \mathbb{C} \).

As has already been said, we have to define the law of vector composition without invoking the components of a vector, so the axiomatic definition (1.10) is rather complicated in comparison with the aforementioned naive rule (1.9) in which the sum and the scalar multiple are defined by the component-wise proportional calculations (evidently, (1.9) is consistent with the general definition (1.10)).

The law of vector composition enables us to compose/decompose vectors, and by these manipulations, we can parameterize a vector and represent it in terms of the components (elements)—in fact, the parameterization is just the decomposition of a vector into a set of elementary vectors, each of which is directed parallel to one of the basis vectors. Let us formulate the process of parameterization explicitly.

Let \( X \) be a linear space. We choose vectors \( e_1, \ldots, e_n \in X \) (\( n \) is a certain integer), and define a system \( B = \{e_1, \ldots, e_n\} \). This \( B \) is said to be the basis of \( X \) if every \( x \in X \) can be written as

\[
x = \sum_{j=1}^{n} x_j e_j \quad (x_1, \ldots, x_n \in K),
\]

11 However, the abstractness enables us to consider various kinds of vectors. A function may be regarded as a vector of a linear space (which is called a function space); for functions \( f(x) \) and \( g(x) \), we can define \( f(x) + g(x) \) and \( \alpha f(x) \) appropriately, and compose/decompose functions as vectors (see Note 1.1).
with a unique set of scalars (coefficients) \( x_1, \ldots, x_n \). In (1.11), the right-hand side is the *parameterization* of the left-hand side vector \( x \) (cf. (1.8)). Displaying only the coefficients of (1.11), we may write as\(^1\) \[ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = t(x_1, \ldots, x_n). \] (1.12)

The number \( n \) of the elements of the basis is the *dimension* (or the *degree of freedom*) of the linear space \( X \). A member of an \( n \)-dimensional linear space can be represented by \( n \) components (coefficients) \( x_j \in \mathbb{K} \) \((j = 1, \ldots, n)\). So we denote by \( \mathbb{K}^n \) the \( n \)-dimensional linear space over the field \( \mathbb{K} \).

The parameterization of a vector is simple if the components are mutually independent. Here “independence” means, geometrically, the orthogonality of the basis vectors; to axiomatize these concepts, we have to first define the *inner product*. Let \( X \) be a real (or complex) linear space. We define a map from an arbitrary pair \( x, y \in X \) to a real (or complex) number, which we denote by \((x, y)\). This \((x, y)\) is called the *inner product* if the following conditions are satisfied:

\[
\begin{align*}
(x, x) &\geq 0, \text{ and } (x, x) = 0 \text{ is equivalent to } x = 0. \\
(x, y) &= (y, x). \\
(a_1x_1 + a_2x_2, y) &= a_1(x_1, y) + a_2(x_2, y).
\end{align*}
\]

(1.13) (1.14) (1.15)

The inner product of real vectors \( x, y \in \mathbb{R}^n \) are often written as \( x \cdot y \).

Two vectors \( x \) and \( y \) are said to be *orthogonal* if \((x, y) = 0\). An *orthonormal basis* is a system of basis vectors \( e_1, \ldots, e_n \) such that \( e_i \cdot e_j = \delta_{ij} \). \(^1\)

When we choose an orthonormal basis, the components of a vector can be easily evaluated:

\[ x_j = (x, e_j) \quad (j = 1, \ldots, n). \] (1.16)

Recalling (1.8), this calculation of the components (parameterization) may be regarded as the “measurement” of an object = vector; the basis means the system of *units* to measure the object. \(^1\)

We may choose a general basis \( B = \{e_1, \ldots, e_n\} \) consisting of vectors which are not necessarily orthogonal to each other. Then we have to generalize the relation

\[ (x_1, \ldots, x_n) \]

\(^1\) In this book, we normally represent a vector as a *column vector* (displaying the coefficients in a vertical array). When we need to save space, we write it as the transpose of a *row vector*, i.e., \( t(x_1, \ldots, x_n) \).

\(^1\) The word “normal” means that the length of each vector (given by \( (e_j, e_j)^{1/2} \)) is 1, and the word “ortho” means the orthogonality \( (e_j, e_k) = 0 \) \((j \neq k)\).

\(^1\) Endowing \( \mathbb{R}^n \) with an orthonormal basis and the inner product \( x \cdot y = \sum_{j=1}^n x_jy_j \), we can identify it as an \( n \)-dimensional Euclidean space.
between the basis and the coefficients of a vector with introducing the notion of dual space. Given the $B$, its dual basis $B^* = \{e^1, \cdots, e^n\}$ is the system of vectors that satisfy the orthogonality condition:

$$(e_j, e^k) = \delta_{jk}.$$ 

The linear space span by $B^*$ is called the dual (or adjoint) space of $X$ and denoted by $X^*$. An arbitrary vector $x \in X$ can be decomposed as

$$x = \sum_{j=1}^{n} x^j e_j \quad [x^j = (x, e^j)].$$  \hfill (1.17)

On the other hand, $y \in X^*$ is decomposed as

$$y = \sum_{j=1}^{n} y_j e^j \quad [y_j = (y, e_j)].$$  \hfill (1.18)

Following Einstein’s rule of summing over one upper and one lower index, we abbreviate $\sum_j a_j b^j$ to $a_j b^j$. Then we may write $x = x^j e_j = x^j e^j$.

In this book, bases are chosen to be orthonormal unless otherwise specified. We use lower indexes both for the basis vectors and for the coefficients (to avoid confusion with powers).

### 1.3.3 Graphs—Geometric Representation of Laws

We are now going to compare nonlinear with linear in a linear space. As pointed out in the preceding subsection, the “linear structure” is marked to a linear space by axiom (1.10) of the law of vector composition. Therefore nonlinearity has already been discriminated against as “distortion” in the linear space.

How is “linearity” integrated with space? And what is “distortion” of nonlinearity? For the clarification, we study the structure of a graph immersed in space. Here, a graph is a geometric object representing a relation among parameters.

Let us consider $n$ variables $x_1, \cdots, x_n$ (assumed to be real-valued) which are related to each other by an equation:

$$F(x_1, \cdots, x_n) = 0.$$  \hfill (1.19)

The set of variables satisfying relation (1.19) appears as a graph that is a hypersurface in the linear space $\mathbb{R}^n$ (see Fig. 1.2). \hfill 15

---

15 A hypersurface is an $(n-1)$-dimensional manifold that is immersed in the $n$-dimensional Euclidean space. Here an $m$-dimensional manifold (topological manifold) is a geometric object that
Fig. 1.2 A mathematical law of science can be represented by a graph—a manifold immersed in a linear space. This figure shows a graph that has a “pleat.” When the graph is folded in some direction ($x_1$ in this figure), relation (1.19) is multi-valued in determining $x_1$ as the function of the other parameters.

If we can solve (1.19) for some variable (for example $x_1$) to obtain a relation

$$x_1 = f(x_2, \cdots, x_n),$$

law (1.19) is translated into a map from $(x_2, \cdots, x_n) \in \mathbb{R}^{n-1}$ to $x_1 \in \mathbb{R}$. The function $f(x_2, \cdots, x_n)$ which is determined by relation (1.19) is called the implicit function. In general, however, the relation between $x_1$ and $x_2, \cdots, x_n$ may not be single-valued (see Fig. 1.2). Therefore the expression of law in the form of (1.19), which can describe multi-valued relations, is more general than the form of (1.20) which is restricted to single-valued relations.

We may consider a more general case where $n$ variables $x_1, \cdots, x_n$ are related by simultaneous $\nu$ ($< n$) relations:

$$F_k(x_1, \cdots, x_n) = 0 \quad (k = 1, \cdots, \nu).$$

If we can solve (1.21) for $x_1, \cdots, x_\nu$, we obtain a system of implicit functions:

$$x_k = f_k(x_{\nu+1}, \cdots, x_n) \quad (k = 1, \cdots, \nu),$$

is locally (i.e., in the neighborhood of an arbitrary point on the manifold) identical (homeomorphic) to a Euclidean space and is a Hausdorff space (i.e., the second axiom of separation holds so that distinct points have disjoint neighborhoods).

16 The implicit function theorem gives the condition under which the implicit function is uniquely determined; it basically says that a critical point becomes the obstacle in defining the implicit function (cf. Sect. 1.4.4).
1.3 The Territory of Linear Theory

The graph of a set of simultaneous relations is the intersection of hypersurfaces (each of them is the graph of one equation). If we give two relations for three variables \((x_1, x_2, x_3) \in \mathbb{R}^3\), the graph is a \((3-2=1)\)-dimensional manifold, that is a curve in \(\mathbb{R}^3\)

which determines a map from \((x_{\nu+1}, \cdots, x_n) \in \mathbb{R}^{n-\nu}\) to \((x_1, \cdots, x_\nu) \in \mathbb{R}^\nu\) (see Fig. 1.3).

When we represent a graph by a relation \(y = f(x)\), we call \(x\) the independent variable and \(y\) the dependent variable. When the linear space \(X\) of independent variables and the linear space \(Y\) of dependent variables are defined separately, the graph is a subset (manifold) in the product space \(X \times Y\). Here a product space is the linear space of the combined variables \(\{x, y\}\) \((x \in X, y \in Y)\), where the law of vector composition (defining the sum and the scalar multiple) are induced by those of \(X\) and \(Y\):

\[
\{x_1, y_1\} + \{x_2, y_2\} = \{x_1 + x_2, y_1 + y_2\}, \quad \alpha\{x, y\} = \{\alpha x, \alpha y\}. \quad (1.23)
\]

We may identify \(\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}\).

The map \(f(x)\) is not necessarily defined everywhere on \(X\). The subset of \(X\) where \(f\) is defined is called the domain of \(f\). The totality of the values of \(f(x)\) (i.e., the image of the domain of \(f\)) is called the range of \(f\), which is a subset of \(Y\).

Up to now, we have discussed the general relations between graphs and maps (functions). Hereafter, we shall study the special features of graphs in linear theory.

First of all, we need to define what linear law is. As mentioned in Sect. 1.1.1, linear law is a generalization (to higher dimensions) of proportionality relation; this generalization can be formally done as follows. Let \(f\) be a map from a domain \(U\) to a range \(V\). We assume \(U \subset X\) and \(V \subset Y\), where both \(X\) and \(Y\) are real (or complex) linear spaces. This \(f\) is said to be a linear map (or linear operator), if, for every \(x, x' \in U\) and arbitrary scalars \(a, b \in \mathbb{K} = \mathbb{R}\) (or \(\mathbb{C}\)), a relation
\[ f(ax + bx') = af(x) + bf(x') \]  

(1.24) holds. To demand this relation for all \( x, x', a \) and \( b \), both \( U \) and \( V \) must be linear spaces. In what follows, we assume \( U = X \) and \( V = Y \).

We denote by \( G \) the set of variables \( \{x, y\} \in Z = X \times Y \) satisfying the relation \( y = f(x) \) for a linear map \( f \). We can easily verify that \( G \) (the graph of \( f \)) satisfies axiom (1.10) of the linear space (see Problem 1.2). On the other hand, for the graph \( G \) to be a linear (sub-)space, the condition (a) of (1.10) demands that relation (1.24) holds. Therefore the fact that \( f \) is a linear map and the fact that its graph \( G \) is a linear space (represented by a “plane” immersed in the linear space \( Z = X \times Y \)) are equivalent.\(^{17}\) As will hereinafter be described, the geometric representation of a plane graph \( G \) gives the explicit form of a linear map \( f \).

We first define a basis of \( Z = X \times Y \) (here we assume that both \( X \) and \( Y \) are real, for simplicity). Let \( \{e_1, \cdots, e_n\} \) and \( \{e_1, \cdots, e_m\} \) be the orthonormal bases of \( X \) and \( Y \), respectively. We represent \( z \in Z \) in terms of the components: \( z = t(x_1, \cdots, x_n, y_1, \cdots, y_m) \). A plane graph \( G \) is either a hyperplane (when \( m = 1 \)) or the intersection of \( m (> 1) \) hyperplanes. A hyperplane that includes the origin of \( Z \) can be defined as the totality of the vectors that are perpendicular to a certain vector \( a \) (see Fig. 1.4). The condition for \( z \) and \( a = t(a_1, \cdots, a_n, b_1, \cdots, b_m) \) to be perpendicular is that

\[ a \cdot x_1 \quad a \cdot x_2 \quad a \cdot x_3 \]

Fig. 1.4 The graph of a linear law is represented by a linear subspace immersed in the space \( X \times Y \). The graph is given as a hyperplane or the intersection of hyperplanes; each hyperplane is characterized by its normal vector \( a \)

\(^{17}\) The word “linear” originates form the fact that the graph of a proportionality relation is given by a “line.” In a higher-dimensional space, an arbitrary “cross-section” of the graph of a linear law is a straight line (see Fig. 1.4).
1.3 The Territory of Linear Theory

\[ \mathbf{a} \cdot \mathbf{z} = \sum_{j=1}^{n} a_j x_j + \sum_{k=1}^{m} b_k y_k = 0. \]  \hspace{1cm} (1.25)

Let us first assume \( m = 1 \). If the basis vector \( \varepsilon_1 \) is not perpendicular to \( \mathbf{a} \), i.e., if \( b_1 \neq 0 \), we can solve (1.25) for \( y_1 \) and define the implicit function \( f : (x_1, \cdots, x_n) \mapsto y_1 \). Introducing a \( 1 \times n \) matrix

\[ L = (\alpha_1, \cdots, \alpha_n) \quad (\alpha_j = -a_j/b_1; \ j = 1, \cdots, n), \]

we may write \( f \) as

\[ f(x) = Lx. \]  \hspace{1cm} (1.26)

When \( m \) vectors \( \mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(m)} \) are given, and they are mutually independent (i.e., none of them are parallel to each other), we can define \( m \) hyperplanes \( G_1, \cdots, G_m \) such that \( G_j \) is perpendicular to \( \mathbf{a}^{(j)} (j = 1, \cdots, m) \). The intersection of \( G_1, \cdots, G_m \) is the totality of the points satisfying

\[ \mathbf{a}^{(\ell)} \cdot \mathbf{z} = \sum_{j=1}^{n} a_j^{(\ell)} x_j + \sum_{k=1}^{m} b_k^{(\ell)} y_k = 0 \quad (\ell = 1, \cdots, m). \]  \hspace{1cm} (1.27)

Solving (1.27) for \( y_1, \cdots, y_m \), we obtain the representation of the implicit function \( f : (x_1, \cdots, x_n) \mapsto (y_1, \cdots, y_m) \). Introducing matrices

\[ A = \begin{pmatrix} a_1^{(1)} & \cdots & a_n^{(1)} \\ \vdots & \ddots & \vdots \\ a_1^{(m)} & \cdots & a_n^{(m)} \end{pmatrix}, \quad B = \begin{pmatrix} b_1^{(1)} & \cdots & b_m^{(1)} \\ \vdots & \ddots & \vdots \\ b_1^{(m)} & \cdots & b_m^{(m)} \end{pmatrix}, \]

and assuming that \( \det B \neq 0 \), we define an \( n \times m \) matrix \( L = -B^{-1} A \). With this \( L \), the implicit function \( f(x) \) is written in the form of (1.26). Thus we find that every linear map \( f(x) \) can be associated with a matrix \( L \) to be written as (1.26).

The linearity condition (1.24) demands that the graph \( G \) must include the origin, i.e., \( f(0) = 0 \). Parallel shift of the graph, allowing it to deviate from the origin, is a trivial transformation. So, a law that is represented by such a shifted graph is also called a linear law. When we have to distinguish such a generalized linear law from the previously defined one, the generalized one is called an inhomogeneous linear law. The map representing a generalized linear law is given by adding an inhomogeneous term (a constant vector \( \mathbf{c} \in \mathbf{Y} \)) to the right-hand side of (1.26).

We have clarified that a linear law is represented by a graph that, in itself, is a linear subspace; by decomposing a subspace from the space, we obtain a graph of

---

18 If the bases of \( X \) and \( Y \) are not orthonormal, we have to first apply a linear transformation to orthogonalize the bases. Then, the linear map can be cast into the form of (1.26).
In this sense, the graph of a linear law can be regarded as a part of the “structure” of the “space,” or, to put it in the opposite way, we can structure the space by the linear law—this will give us a strong strategy of linear theory in elucidating the order implied by the law (see Sect. 2.3).

In comparison with the “plane graph” of a linear law, the graph of a nonlinear law is distorted (the plane tangent to the curved graph is the “linear approximation” of the nonlinear law). The theme of nonlinear science is the phenomena that only distorted graphs can describe. In the next section, we shall study the basic “patterns” of nonlinearity. Before that, though, we have to learn the exponential law which, together with the proportionality law, constitutes the core of linear theory.

### 1.3.4 Exponential Law

The exponential law (in its higher-dimensional generalization) plays the central role in the linear theory of dynamics (study of temporal evolution).

First, let us explain the connection between the proportionality relation and the exponential law by recalling the calculation of “compound interest.” Suppose that, for a principal $x_0$, an interest $\alpha x_0$ is given every year ($\alpha$ is a constant). The point is that the interest (increment) is proportional to the principal at every moment when it is evaluated. Adding the interest to the principal, the balance becomes $(1 + \alpha)x_0$ after 1 year and becomes the principal for the next 1 year. After $n$ years, the balance will become $(1 + \alpha)^n x_0$. This geometric progression—also known as Malthusian law in the theory of population dynamics—is a “discrete” exponential law where the increment occurs stepwise.

In the limit of continuous time, i.e., if the interest is added to the principal at each infinitesimal passage of time, the evolution of the balance is described by the exponential function $e^{\alpha t} x_0$, where $t$ is the continuous time and $x_0$ is the initial ($t = 0$) value. The coefficient $\alpha$ (or, sometimes, its reciprocal $\alpha^{-1}$) is called the time constant, which determines the time scale of the exponential function.

We may generalize the exponential function allowing both $\alpha$ and $x_0$ to be complex numbers. When $\alpha$ is a pure imaginary number, we find that the exponential law describes “oscillation”; writing $\alpha = i\omega$ ($\omega \in \mathbb{R}$),

$$ e^{it} = \cos (\omega t) + i \sin (\omega t). $$

For a general complex $\alpha$, the real part of $\alpha$ (denoted by $\Re \alpha$) yields exponential growth (when $\Re \alpha > 0$) or damping (when $\Re \alpha < 0$). We say that $e^{\alpha t}$ is stable when $\Re \alpha \leq 0$ and unstable when $\Re \alpha > 0$. The imaginary part of $\alpha$ (denoted by $\Im \alpha$) represents the angular frequency of oscillation.

The exponential function is derived by a linear differential equation. The rate of change of a certain quantity $x(t)$ is given by the differential coefficient $dx(t)/dt$. If this rate of change is proportional to $x(t)$, the evolution of $x(t)$ is described by

$$ \frac{d}{dt} x = ax, $$

(1.28)
Nonlinear Science
The Challenge of Complex Systems
Yoshida, Z.
2010, XI, 211 p., Hardcover
ISBN: 978-3-642-03405-3