I

PHYSICAL AND BIOLOGICAL INTERACTIONS

Does Physical Force have a Biological Counterpart?

Some scientists believe that the *human intelligibility of nature is fundamentally mathematical*. In other words, the apparent harmony of our universe is subtended by mathematics, the understanding of which alone can lead to a synthesis of the diversity of the phenomena observed. In fact, the history of physics has been marked by powerful syntheses such as Maxwell's theory of electromagnetism, Einstein's theory of relativity and, more recently, Hawking's theory of superstrings, which unifies the preceding theories. All these ideas emerged after long periods of scientific gestation that produced the mathematical formulation required to explain the phenomena of light through electromagnetic principles (Maxwell, 1872), gravitation through geometrical concepts (Einstein, 1923), and the evolution of the universe through gravitational laws and the quantum theory (Hawking, 1988). Rather unfortunately for most of us, the further we advance in science, the more abstract our image of the physical universe becomes, i.e. the mental representation we construct of it, try as we might to keep it as simple as possible (Espagnat, 1980).

THE MATHEMATICAL NATURE OF THE LIVING WORLD - The Power of Integration

I

Can physical and biological interactions be compared?

A very commonly observed physical interaction in nature is that of the force exerted by one body upon another. What exactly does the abstract notion of force correspond to? Physics deals with relatively simple systems that can be broken down into smaller systems, allowing an easier approach to the problem. In comparison, biological systems, as we shall see, are far more complex. One of our main objectives here is to distinguish the living from the nonliving. An important difference between physics and biology lies in the number of levels used for the description of natural phenomena. In a physical system, we need consider no more than two or three levels of description. For example, the properties of an electric current flowing in a conductor are explained by the transport of free particles, i.e. the electrons, in the conductor, for instance a copper wire. If we admit that the organization of particles of matter constitutes a structure, then we see that this explanation is based on only two levels of description: the macroscopic level (the electrical activity in the copper wire) and the microscopic level (the flow of electrons). In other words: on one hand, we have the set of copper atoms and, on the other, copper atoms with their peripheral electrons.

In a given physical system, the interplay between the sets of influences acting on the particles regulates the equilibrium of the system. These influences, called *forces*, stabilize the physical structure. The science of physics deals essentially with the forces acting between the particles, and their consequences, in solids and fluids, or even *within* elementary particles. Today, it appears evident that when a force acts on a particle and causes it to move, the system is in dynamic equilibrium. However, this only came to be understood when Newton first expressed the idea in a general, inductive form in his *Philosophiae naturalis principia mathematica* (1687). In his formulation, based on geometry, he enunciated three laws. The first corresponds to *Galileo's principle of inertia*,[2] which states that a body in motion

---

[2] Galileo Galilei, *Discouri* (Dialogues and mathematical demonstrations concerning the two new sciences of mechanics and local motion), (Leyden 1638). An interesting review of Galileo's life and work will be found in Geymonat's *Galilée* (1992).

continues moving at a uniform speed in a straight line, unless acted upon by an external force. The second law determines the variation of the velocity of the body under the action of an external force, the force being defined as the product of the mass of the body, which characterizes its matter, and the variation of its velocity, or acceleration. The third law states that the force F acting between two bodies varies directly as the product of the masses of the bodies, m and m', and inversely as the square of the shortest distance r between them. In mathematical terms, we have the equation: $F = mm'/r^2$. With his third law, Newton created the theory of gravitation. This was certainly a stroke of genius since the third law not only lays down the rule under which one body influences another but also gives a qualitative description of the influence by introducing the abstract concept of the *force* involved in the physical interaction.

> Newton's third law of motion expresses an idea that is at once simple and extraordinary. The concept of the force is valid not only for bodies in motion on the Earth but also for heavenly bodies. Moreover, the idea of the gravitational force describes the fact that one mass can act instantaneously on another at a distance, apparently with no transfer of any kind. It is this mysterious, non-local character of the law of gravitation, almost bordering on the occult, that caused Leibnitz, the rationalist, to oppose Newton, who insisted that far from trying to answer the *why* he was only addressing the *how* of the phenomena of gravitation. According to Newton's third law, the force of gravitational origin acting between two masses, called the gravitational interaction, satisfies the principle by which the action exercised by one particle A on another particle B is equal to that exercised by B on A. As readily seen in the equation above: the body of mass m being at x and the body of mass m' being at x', changing $m(x)$ to $m'(x')$ does not alter the intensity of the force since it is the square of the distance $r = |x - x'|$ that is involved. This is the *principle of action and reaction* contained in the law of gravitation. Another, more "operational" way of expressing this principle is to say that the momentum of a body, equal to the sum of the momentums of all its constitutive bodies, is conserved during its motion. The momentum is equal to the product of the mass of the body and its velocity. This principle of conservation explains the mechanism of propulsion of a rocket,

for example, in which case the mass m_A of the combustion gases ejected at a velocity v_A is equal to the mass m_B of the rocket propelled at a velocity v_B in the opposite direction, since $m_A v_A + m_B v_B = 0$. Since the mass of the gases is much smaller than that of the rocket, the velocity of the gas ejected will have to be proportionally high to propel the rocket at the required velocity.

The mathematical expression of Newton's laws of gravitation is rather simple, although the application of the laws to a system comprising several interacting bodies can be a formidable task. However, the fact that these laws can be formulated mathematically gives them an enormous advantage over other laws that cannot — at least for the time being — be so expressed. Thus, in human physiology, most of our knowledge can only be stated in qualitative terms, such as: "low temperatures lead to an increase in diuresis", or "emotional shocks trigger the secretion of adrenalin into the blood". The mathematical formulation of a law, in addition to providing quantitative prediction, allows the practical deduction of all the consequences. In the case of Newton's laws, this ensemble of consequences constitutes the classic science of mechanics. For example, the movement of a satellite around the Earth brings into play at least two bodies, the Earth and the satellite; but we also have to consider the action of the other planets of the solar system, together with their own satellites, as well as that of all the other celestial bodies in the universe. Thus, the application of even relatively simple physical laws to a system comprising more than two bodies soon becomes inextricable.

In his celebrated essay on the three-body problem, Poincaré (1890) proved the impossibility of resolving the equations of movement for three bodies, for example, the Sun, Jupiter and Saturn, while the problem could be readily solved for two bodies. Finally, the Kolmogoroff-Arnold-Moser theorem, developed between 1954 and 1967, showed that the stability of a three-body system could be determined only by imposing perturbations on the system. Thus, when a Hamiltonian system, i.e. a frictionless system, is subjected to a slight perturbation, although several initial conditions lead to a quasi-periodic, stable trajectory, a few can lead to a chaotic trajectory. This notwithstanding, one of the foundations of the organization of the physical structure of matter, composed of subatomic particles, atoms and molecules, is built on the abstract concept of *force*, and on the general, almost tautological principles of the conservation of the momentum and the electric charge of bodies.

Do biological systems involve elementary interactions playing a functional role analogous to that of a force in a physical structure? Among the many questions raised by the growth and development of a living organism, we may wonder how a cell "recognizes" the state of its development at a given moment, and how it "decides" to react to a particular situation. It is likely that, at the gene level, sequential and parallel sets of instructions, similar to those of computer programs, determine the precise instant at which a specific protein is synthesized. In turn, this protein will trigger a set of enzymatic reactions leading to other sets of biochemical reactions. Such sets, incorporating mechanisms of synthesis and their corresponding regulation, may be very complex indeed. Whether we consider the processes of embryonic development, or those involved in memory and learning, or in the regulation of physiological functions in higher organisms, the common point is the existence of molecules or signals emitted by one cell, called the *source*, acting at a distance on another cell, called the *sink*. This is illustrated, for instance, by the propagation of an electric potential along a nerve, the action of a hormone operating at a distance from its site of synthesis after being carried in the bloodstream, and the change produced in the shape of a molecule after it binds to another molecule. Table I.1 gives some examples of such interactions observed in human physiology.

These examples enable us to grasp a concept that is fundamental to the understanding of biological phenomena: the concept of the *functional interaction*. In simple terms, this implies that *a product emanating from one entity acts on another at a distance*. Thus, a signal emitted by a given cell acts on another cell, which in turn emits a new signal after transforming the signal received. This defines the functional interaction between two cells since the action operates from one cell to the other.

Table I.1: Some examples of functional interactions in human physiology.

Physiological function: propagated signal	*Source*	*Sink*
Nervous activity: *potential*	A neuron	Another neuron
Thyroid metabolism: *thyroxin*	A thyroid cell	All cells
Allosteric activity: *change of shape*	A molecular subunit	Another molecular subunit
Gene regulation: *repressor*	A regulator gene	A structural gene
Pulmonary ventilation: *partial pressure of* O_2	A pulmonary region	Another pulmonary region

It is called "functional" because the function of the first cell is to *act* on the second by means of the signal it emits, causing the second cell to *produce* a new signal at a distance.

The functional interaction greatly resembles a mathematical function in that it consists of the application of one set on another, transforming an element of one set into an element of the other set. For this reason, we shall use the term *functional interaction* indifferently to describe the *action* of one biological structure on another as well as the *product* of this action, just as in mathematics we identify the function f: $x \rightarrow f(x)$ with its value $f(x)$. In fact, the functional interaction is the most elementary physiological function, a "building-block" with which the entire functional organization may be constructed. In this sense, and only by analogy with the construction of physical systems, we may say that the role of the functional interaction in biology is similar to that played by force in physics.

However, it is important to note the nature of the difference between the biological functional interaction, on one hand, and physical interactions, gravitational or electrical, on the other, as described below.

The gravitational interaction is *symmetrical*, in the sense of action and reaction, since it is expressed by the force: $F = \kappa mm'/r^2$, which, by linking two objects of masses m and

m', creates a physical structure, in other words, a dynamic assembly of matter. Similarly, the electrical interaction is also *symmetrical*, since it is expressed by the same type of equation: $F = \varepsilon qq'/r^2$, where the electric charges on the bodies replace the masses in the gravitational equation. Without going further into the fascinating question of the role played by symmetry in physics, a subject that has already received considerable attention (Wigner, 1970; Coleman, 1985), we would like to stress the profound influence of simple principles on the formulation of a problem. The property of symmetry must necessarily appear somewhere in the formulation even though, in the case of a highly complex system, it may not be as evident as in the formula: $F = mm'/r^2$.

Consider, for example, the energy of vibration and rotation of a molecular system due to the permanent vibration and rotation of the atoms of a molecule about their equilibrium position. This energy must remain invariant when the system undergoes symmetrical operations. Thus, in the case of a molecule composed of two vibrating atoms, such that the geometry remains unchanged when the atoms are interchanged symmetrically, say about a median plane, the energy must remain constant since the molecule remains identical to itself. The formulation of the property of the invariance of energy must therefore include the notion of symmetry. The property of invariance is reflected by the appearance of energy levels[3] that correspond exactly to the symmetry of the molecule. In mathematical terms, the symmetry groups[4] to which the molecule belongs define its energy levels. In practice, infrared or Raman spectroscopy is used to measure

---

[3] These energy levels can be calculated by quantum mechanics. They are given by the eigenvalues of the Hamiltonian operator representing the energy of the vibration and rotation of the molecular system. The operator, expressed in the form of a matrix, is factorized according to the representations of the group of symmetry. On simplification, zero-values appear in the matrix at specific points corresponding to the symmetry of the molecule.

[4] In mathematics, a *group* consists of a set associated with an operation, satisfying certain elementary operations. This mathematical structure, due to Evariste Galois (1811–1832), has proved very useful for the formalization of problems not only in physics but also in other fields, such as that of linguistics.

the energy levels of a molecule and establish its group of symmetry. This actually constitutes a kind of molecular fingerprint whereby the molecule is identified (Wilson, Decius & Cross, 1955). This method has been successfully applied even to large biological molecules, such as adenosine triphosphate (Chauvet, 1974). Methods based on symmetry are commonly used in physics since the property of symmetry is closely linked to the structural organization of matter, which is by essence of a geometrical nature. Thus, the laws of conservation of matter are associated with the properties of symmetry of the system. A law of conservation of matter is considered to represent a fundamental property of the system to such an extent that, in physical theories, the knowledge of the laws of conservation is considered sufficient to explain the system.

In contrast to physical interactions, the functional biological interaction is *non-symmetrical.* This property of non-symmetry should therefore be incorporated in the formulation of biological interactions. The biological functional interaction operates from one structural unit on another situated at a distance, i.e. from a source to a sink, although not *directly* from source to sink. This interaction is unidirectional since the molecule or signal, emitted at a given level of organization, will have no direct retroaction from sink to source. The consequences of the non-symmetry of the functional interaction on the development of biological systems and on the dynamics of biological processes in general are discussed in Chapters V and VI. As we shall see, these dynamics are governed by equations that include the property of non-symmetry.

The other important property of the functional biological interaction is that of *non-locality*, since the product emitted is transported at finite speed from source to sink. The property of non-locality is as fundamental to living organisms as that of locality associated with physical systems.[5] We shall see in Chapters V

---

[5] Putting it simply, we may say that a local phenomenon observed here and now may be deduced from what happened here just a little while ago, whereas a non-local phenomenon observed here and now has to be deduced from what happened far away from here and perhaps quite some time ago.

and VI, how the notion of non-locality, first introduced in physics by Brillouin (1952), is related to the representation of biological phenomena and, more precisely, how it results from the continuity underlying hierarchical systems (Chauvet, 1993a).

The concept of functional biological interactions raises a number of interesting questions. For instance, how do two functional interactions differ? The answer is that the difference is defined by the “nature” of the products of the interactions, i.e. the ions, molecules, hormones or potentials emitted, just as the difference between physical interactions is defined by the nature of the forces in play, be they gravitational, electrical or nuclear. However, in the case of two biological interactions, the properties of non-symmetry and non-locality have similar consequences and, in particular, lead to the structuring of the biological system in the form of hierarchical levels of organization of the physiological functions. This is analogous to the way the property of symmetry leads to the structuring of a molecular system with its different energy levels, though without any hierarchical organization. Thus, the properties of non-symmetry and non-locality lead to a novel situation in theoretical biology, obliging us to formulate physiological functions incorporating the idea of a hierarchical organization that must be reflected by the mathematical solution of the dynamic biological system.

Another question that may be asked is what happens when a given biological unit acts as a source for certain products and as a sink for others? This problem can be tackled mathematically and, as we shall see, the solution leads to other fundamental questions. We may also ask what happens when a source turns into a sink, or vice versa. In fact, this leads to the appearance of new properties related to the functional reorganization of the organism. Here also the dynamics of processes associated with this particular topology lead to specific properties. We shall address these questions in the following chapters and try to determine the elements that are common to physiological phenomena. We may then be able to deduce a fundamental criterion of biological evolution. To sum up, the functional biological interaction, an abstract mathematical entity representing, for instance, an electrical potential, a molecular concentration, a saturation level or some other biological parameter, involves three elements: the *source*, the *sink* and the *transformation* within the sink, and possesses at least the two properties of *non-symmetry* and *non-locality*.

However, why should this particular representation of biological phenomena in terms of functional interactions be considered more appropriate than some other? The simple answer is that the smallest number of concepts common to the largest number of biological phenomena may be expected to lead us to the largest number of properties interpretable by means of these concepts. In other words, the smallest number of very general principles should allow us to interpret the largest number of phenomena observed and, in particular, those that are still unexplained. This, of course, is the objective of the theory we are trying to construct. The goal is far from being attained, and we do not even know if it will ever be reached. It is true there are other possible representations of biological phenomena, such as those based on the concept of compartments. Considering the importance of this concept, we shall discuss it below in connection with that of the functional interaction proposed here.

Some consequences of the concept of functional interactions

The somewhat abstract considerations above allow us to deduce the structure of a general biological law involving functional interactions, at least in the case of *formal* biological systems defined on the basis of a small number of properties. Such formal systems correspond to an idealization of real biological systems. Physical laws governing gravitational interactions reveal the property of symmetry through the mathematical expression of the product of the masses and the relationship with the square of the distance between them. The symmetry of electrical interactions is expressed in a similar manner. Again, we find symmetry in the expression for entropy given by statistical mechanics in the description of molecular collisions. In contrast to physical laws, biological laws will have to explicitly include the property of non-symmetry between the sources and sinks involved in biological functional interactions. Then, if we admit that functional interactions constitute the basis of biological phenomena, we may go on to describe some of the consequences.

The first consequence concerns the mathematical properties of functional interactions. Sources, sinks and their interactions can be conveniently represented by a *mathematical graph*[6] (Figure I.1). Two points on a sheet of paper correspond to the source and the sink of a functional interaction, and the interaction itself is indicated by an oriented arc joining the two points, with an arrow on the arc to distinguish the source from the sink. If we generalize this procedure to any number of sources and sinks, we obtain a set of summits connected by oriented arcs, possessing mathematical properties. The summits are called "structural units" and the number of summits corresponds to the "degree of organization" of the structure. As we shall see, this kind of graph can be used to represent the functional organization of a biological organism. In fact, it constitutes the *topology* of the biological system. In other words, it corresponds to the *functional organization of a formal biological system* (O-FBS).

Let us now briefly recall some notions concerning *topology*. Founded by Listing in 1847, this branch of mathematics was

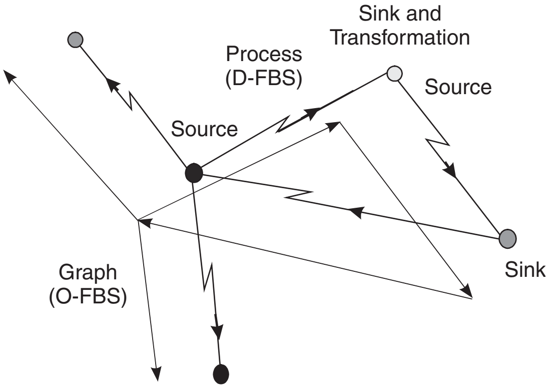

Figure I.1: A superposition of the graphs representing the topology of the organization of a formal biological system (O-FBS) and its dynamics (D-FBS).

---

[6] In the sense of the mathematical graph theory.

greatly developed by Poincaré (1890).[7] Topology studies the properties of space in terms of the notion of *neighborhood*, and consequently that of *continuity*. For example, objects that differ from the point of view of Euclidean geometry, such as the triangle, the circle or the square, are considered identical from the topological point of view. In contrast, a disk with a hole in it and a disk without a hole will be considered topologically distinct since the one cannot be transformed into the other by means of any continuous deformation. When we consider the position of points relative to each other without taking into account the distances between them, we obtain a topological graph. In other words, topology is what is left of geometry when the notion of the metric is removed. The connectivity between sources and sinks in a functional system is thus of a topological nature, but the interactions between them depend not only on the topology but also on the geometry of the system.

Mathematical graphs have often been used in fields as varied as economics, physics and biology to describe the complex relationships between the elements. Although these graphs, consisting of points and arcs, are apparently identical, they differ because of the properties imposed by the problem to be solved. Let us consider some examples.

The well-known puzzle of the Koenigsberg bridges proved to be one of the starting points of the mathematical theory of graphs (Bergé, 1983). The problem was to find a path in the town crossing each of its seven bridges only once. This corresponds to class of practical situations, e.g. the traveling salesman's problem, in which the traveler's route has to be optimized. Such problems are readily solved by combinatorial analysis and graph theory.

In a very different context, complex chemical systems may be analyzed by using *network thermodynamics* (Oster, Perelson and Katchalsky, 1973), the theory of which is derived from the generalization of the thermodynamics of irreversible processes

---

[7] See the Collected Works of Poincaré (1980) for a full account of his contribution to topology.

> (Glansdorff and Prigogine, 1970). The method introduces the use of *binding graphs* to calculate energy exchanges between the different elements of a given chemical system. In fact, the underlying principle is the unification of various scientific disciplines from a thermodynamic point of view. Thus, it has been applied to the transport of mass within biological tissues, as well as in the modeling of highly complex physicochemical phenomena in biology. In the case of binding graphs, the arcs represent the paths along which the energy circulates.
>
> Another example is that of the *interaction graphs*, in which the arcs represent the circulation of matter. The formalism applied to these graphs is the same as that of the transformation systems (Delattre, 1971a), which allowed the extension of compartmental analysis to the phenomena of radiation. As we shall see further on, compartmental analysis is also widely used in biology (Jacquez, 1985). Mathematical graphs have many applications, for instance, in electrical networks for the transport of energy, in Kirchoff's laws for the conservation of current at the junctions in an electric circuit, in telecommunications networks for the transport of information, in rail and road networks for the transport of goods, and so on. In short, mathematical graphs are very convenient for the representation of the displacement of matter or energy between the different reservoirs of a system, as well as for the analysis of the resulting global dynamics (Roy, 1970).

The second consequence of the existence of functional interactions concerns the formal expression of the dynamics of the process corresponding to the spatiotemporal fate of the product that, after being emitted by a source, will act at a distance on a sink. For example, a hormone is transported at a rate depending on the blood flow, a neurohormone migrates along the axon of a hypothalamic neuron for several days before reaching its target, an action potential is propagated at a velocity of eighteen meters per second along the larger nerve fibers, and so on. Since the delay with which the signal from the source reaches the sink depends on the localization of the source and the sink, the geometry of the biological system influences the dynamics of the process and is therefore implicitly included in this representation. The functional interaction being determined by its three elements, i.e. the source, the sink and the transformation of a product in the sink, the propagation of the signal and its action at a distance represent the dynamics of the process, describing *how* the interaction takes place (Figure I.1). Thus, the determination of the *dynamics of the formal biological system* (D-FBS), i.e. the representation of the action, in space and in time, of one structural unit of a system on another is clearly complementary to that of the *functional organization of the formal biological system* (O-FBS), i.e. its topology.

The *field theory* used in physics offers a convenient method for describing the propagation of the signal and its action at a distance. According to this theory, each point in physical space can be associated with a certain numerical value that is characteristic of the process analyzed. Consider, for example, the value at each point of an electric potential being propagated point by point over the membrane of a neuron. Such values may also be time-dependent. Thus, the action at a distance may be described by the spatiotemporal variation of a field, its value at a given point-sink being the result of the action of a *field operator* at the point-source at which the signal was emitted. Such values are represented by a *field variable* corresponding to the functional interaction, or more precisely, to the mathematical value of the functional interaction. This formulation in terms of the field theory may seem rather abstract but, as we shall see in Chapter VI, it offers considerable advantages in the study of biological systems.

To sum up, *the functional interaction is characterized by two complementary aspects, one being the triplet composed of the source, the sink and the arc between them, and the other being the spatiotemporal field describing the action of the source on the sink. The combinatorial analysis of the triplets gives the topology of the biological system, i.e. its functional organization, whereas the variation of the field determines the manner in which the products are exchanged in time and in space, thus defining the dynamics of the biological process.*

The use of graph theory to describe the O-FBS, combined with the use of the field theory to describe the D-FBS, are the two conceptual frameworks within which the execution of a set of functional interactions may be described. The combination of the two dynamic systems,[8] one corresponding to the O-FBS and the other to the D-FBS, describes the spatiotemporal operation of the biological system. Here, we have an important difference between physical and biological systems. Unlike physical systems, *in biological systems the topology, represented by the functional connections between the structural units, and the geometry, represented by the density of the structural units at each point in physical space, may both vary with time.*

For example, during biological development, the cellular density at a given point in the organism varies according to the rate of cell division. At the same time, new tissues are formed and new functional connections are established. Consequently, the dynamics of the new products exchanged undergoes a modification, bringing about a functional reorganization involving the topology as well as the geometry of the biological system. Again, when a biological organism suffers aggression from an antigen, the antibody-producing cells of the immune system first multiply rapidly to face the danger, and then decrease in number when normal conditions are restored. It is true that in physical systems we may observe analogous phenomena, for example during *phase transitions*, but these concern only the structural organization of matter without affecting its functional organization.

The conceptual framework proposed above allows the formalization of a certain type of problem. For example, a perturbation of the topology of a biological system will affect its evolution by provoking a reorganization of the functional interactions between its structures (O-FBS) while reestablishing the dynamics of the processes involved (D-FBS). The dynamic stability, which expresses the return of the system to

---

[8] In mathematical terms, a dynamic system is a set of differential equations describing a process, or the operation of a phenomenon, between two successive instants: t and $t + \mathrm{d}t$. In this case, we refer to the local dynamics, since the state of the dynamic system depends only on its immediately preceding state at a neighboring point. Of course, the biological system should be distinguished from the corresponding dynamic system.

its starting point after having been perturbed, also corresponds to the stability of its constituent systems, i.e. the O-FBS and the D-FBS. These two aspects of a formal biological system, each acting on the other, are inextricably linked. At a higher conceptual level, even a phenomenon as complex as the evolution of a biological species may be considered as the resultant of a coupling between the topology, the geometry and the dynamics of the associated processes.

How do biological principles differ from physical principles?

Biology differs from physics not only because of the nature of the entities studied by these sciences, but also because of the nature of the principles underlying the explanation of the phenomena observed. Whereas only a few principles of quantum physics and relativity are required for the construction of mathematical theories governing a wide range of physical phenomena, there are no such unifying principles in biology. Of course, biology has its own laws, for example those of the evolution of the species by selection and mutation, the doctrines of molecular biology, the principles of physiology, and so on. However, for the most part these laws are not mathematical and even when they are, as in the case of certain optimum principles in cardiovascular or respiratory mechanics, they do not lead to any truly *integrated* understanding of the functioning of a biological organism or its evolution.

Considering the differences between physics and biology, few biologists would willingly rely on physical models alone to explain biological phenomena. However, putting aside the mystical ideas of vitalism and admitting that all forms of life — from viruses to mammals — are but matter, it is not unreasonable to hold that living organisms must obey general laws of organization and functioning that differ from physical laws. We believe that the difference between living organisms and non-living matter may be accounted for by a theory constructed from general principles in a manner that has always proved satisfactory in the domain of physics. Moreover, such a theory would allow the simulation of biological phenomena by means of elementary mechanisms, thereby providing better insight into the processes involved. In fact, living organisms, no matter how complex, display the property of *stability* during their functioning as well as their development. In other words, the physiological state of the organisms is maintained in the neighborhood of a stationary state — either that of equilibrium or of one far from equilibrium. This kind of stability strongly suggests the existence of a few general principles in functional biology.

What exactly is a principle? Several laws, such as those of gravity, electromagnetism, hydraulics, nuclear interactions, and so on, govern physical phenomena. These laws share common points that may be stated in the form of great general principles. The principles of conservation of matter, momentum, energy, electric charge, the principles of symmetry and relativity, and the principle of least action, are a few of the examples in physics. These principles are so general and valid in so many different situations that they may actually appear tautological. However, as Poincaré (1854–1912) indicated in his essay on *Science and Hypothesis* (Poincaré, 1968), this is far from being the case:

> "What I have just said also sheds light upon the role of general principles such as the principle of least action or that of the conservation of energy. These principles are of the greatest value since they have been determined by seeking all that was common in the enunciation of several physical laws. In a way, they represent the quintessence of innumerable observations. However, their very generality leads to the consequence that they can no longer be verified. Since we cannot give a general definition of energy, the principle of the conservation of energy simply means that something remains constant. Then, no matter what new notions future experiments might bring, we may already be certain that there will be something that will remain constant — something that we may call energy. Does this mean that the principle has no meaning and that it disappears into a tautology? Of course, it does not; it means that the different things we call energy are in fact linked by a true relationship. However, if this principle has a meaning, it may be false. Perhaps we do not have any right to extend the applications of

> the principle indefinitely even though we may be sure it will be verifiable in the strict sense of the term. But how shall we know when the principle has been extended to its legitimate limit? Quite simply, when it ceases to be useful, i.e. as soon as it fails to provide a correct prediction for some newly observed phenomena. We may then be sure that the relationship claimed is no longer real, since otherwise the principle would still be fecund. Such an observation, without directly contradicting a further extension of the principle, will have nevertheless pronounced its death sentence."

In *The Value of Science*, Poincaré went on to add:

> "These principles result from the generalization of experimental observations, but from this very generality they seem to borrow a remarkable degree of certitude. In fact, the more general the principles the more often they can be tested, and the multiplication of the results in all their various forms — even the most unexpected — finally leaves no room for doubt."

Thus, in physics we have the *principle of least action*, which appears to be extraordinary, probably because of its apparently anthropomorphic nature. The principle states that a particle will choose a path between two points such that the *action* has a minimum value with reference to all the other possible paths between these points. This would suggest that the particle actually makes a mathematical decision based on the comparison of numbers corresponding to the paths along which it has never really traveled. This principle, which is valid not only in mechanics but also in optics and electromagnetism, may thus appear highly mysterious. But the mystery is quickly solved if we admit that the path of the particle depends on the geometry of the space it moves in, and that the particle moves along a *geodesic line*, i.e. the shortest distance between two points according to the nature of the space. For example, this would evidently be a straight line in a plane, or an arc of the great circle on the surface of a sphere.

We have mentioned that the main advantage of a principle in physics is its generality. Thus, Newton's laws of motion can be replaced by the unique, strictly equivalent postulate of the *variational principle* that is generally applicable in physical mechanics. This new formalism is elegantly simple and yet powerful enough to be used for complex systems. Furthermore, it can be effectively extended to disciplines other than that of Newtonian mechanics, such as quantum mechanics and classic field theory. As we shall often refer to this concept in the chapters that follow, it may be useful to give a brief description here.

Formalisms in analytical mechanics

In mechanics, a system that is *free*, i.e. one with no external force acting on it, and *conservative*, i.e. one in which the total energy is constant, obeys two fundamental laws. The first is the law of conservation of momentum since, on one hand, the action and the reaction being equal and opposite cancel the sum of the internal forces taken two by two, and on the other, the system being free, the resultant of the external forces is null. The second is the law of conservation of energy during the movement, since the work done by the forces is independent of the paths taken by the material points of the system, depending only on the starting and finishing points of the movement. In other words, in a conservative system the total mechanical energy E, i.e. the sum of the kinetic energy T and the potential energy V, is a constant: $E = T + V =$ constant. According to Leech (1961):

> "The practical importance of the concept of energy lies in the fact that all the mechanical properties of a complex system can be summed up by putting into analytical form a certain number of scalar functions designated by the term energy. Analytical mechanics is based on the generalization of this idea."

The principles of conservation express the fact that a certain quantity calculated according to a given rule will always give the same result whatever the natural conditions. In the physical universe, some quantities (such as energy, electrical charge and the number of particles) are conserved, whereas other quantities (such as frictional energy and heat) are not, but the final balance must satisfy a law of variation with time. *Conservative systems* are of great importance from a theoretical point of view since they are relatively simple to formulate in terms of general principles. Thus, it can be shown that the forces acting on a conservative system depend upon the potential function in a very special way. These forces may be considered as being derived from a potential,[9] such that $F = -\partial V/\partial x$. In contrast to conservative systems, almost all real physical systems are *dissipative systems*. In such systems, there is a spontaneous loss of energy, as in the case of an oscillating pendulum, which will eventually stop moving because of the frictional forces at work. This is why dissipative systems require a specific formulation.

Furthermore, mechanical systems may be subject to other complications such as the limitation of movement imposed by conditions of a geometrical nature, materialized by *liaisons*. In the case of the pendulum, the weight must move over a constant distance at each oscillation. Such liaisons introduce additional unknown forces into the equations of movement. Although the notion of a liaison is an abstract, mathematical concept, it facilitates the resolution of problems in mechanics. As Leech (1961) put it: the notion of a liaison is but the simplified image of a more complex, real phenomenon. To deal with these complications, a new postulate going beyond Newton's laws of motion was laid down: the *principle of virtual work*. This principle states that *the total virtual work of all the forces acting on a system of material points is null, whatever be the virtual, infinitesimal movements of these points, starting from an equilibrium state of the system*. In other words, the work done by the forces of liaison acting in a system is positive or null for all virtual displacements compatible with the liaisons. Here again we have a principle that is not easy to accept immediately. Although it deals with virtual displacements, i.e. displacements that are not real, the principle actually allows us to describe reality and has proved indispensable in the formulation of several problems in analytical mechanics.

---

[9] For example, water flowing downhill will follow the path corresponding to the "greatest slope", implying the operation of an optimum principle.

The two main principles of mechanics, due to the work of *d'Alembert* (1717–1783) and *Hamilton* (1805–1865), are applicable far beyond the field of Newtonian mechanics. According to d'Alembert's principle, *the total virtual work of the effective forces, i.e. the applied forces and the forces of inertia, acting on a system of material points is null, whatever be the infinitesimal, virtual movements, reversible and compatible with the liaisons between these points.* From d'Alembert's principle, which may be considered as the most condensed and general statement of the laws of mechanics, we can deduce not only Newton's laws of motion and Hamilton's principle but also the law of the conservation of energy when the virtual movement coincides with a real movement.

The formalism due to Lagrange (1736–1813) is strictly equivalent to that of Newton. The equations of motion are deduced from the function: $L = T - V$, called the Lagrange function or the Lagrangian of the system. This function depends on the parameters of position q, which in turn depend on the coordinates of positions x, y and z. A system composed of N particles will therefore be governed by $3N$ equations, called the Lagrangian equations. This formalism can also be applied to a non-conservative system by replacing V with a specific function M to take into account the dissipative nature of the system. An example of this is the analysis of the motion of a charged particle in an electromagnetic field.

The Hamiltonian formalism adds coordinates p, called *conjugated momentums*, to the parameters of position q. These supplementary coordinates, defined by the relationship: $p = \partial L / \partial q$, correspond to quantities of movement when the system is conservative. Whereas q and $\dot{q}$ are dependent, since $\dot{q}$ is the derivative of q with respect to time, the variables p and q can be considered independent. Thus, the $3N$ equations of the Lagrangian formalism are replaced by the $6N$ equations of the Hamiltonian formalism. The trajectory of a system can then be represented by that of a point having $6N$ coordinates in a space with $6N$ dimensions, called the *phase space*. The N trajectories in physical space are thereby replaced by a single trajectory in phase space. One of the major advantages of the Hamiltonian formulation is that it provides a perfect framework for quantum mechanics as well as statistical mechanics. As in the case of the Lagrangian formalism, the Hamiltonian formalism introduces the Hamilton function: $H=\sum_{pq}-L$, also called the *Hamiltonian*, which generally represents the total energy of the system. For a conservative system, it can be shown that: $H=T+V=E$, if the transformation of the Cartesian coordinates into the coordinates of position, i.e. $x\equiv x(q)$, does not depend explicitly on time. The equations of motion written using the Hamiltonian formalism, called the canonic equations[9], may also be used for a dissipative system. However, in this case the Hamiltonian H will no longer correspond to the total energy of the system, which is certainly one disadvantage of the formalism.

The mysterious principle of least action

The background sketched in above allows us to present the *variational principles* that we propose to apply to the important problem of the evolution of biological systems. Variational principles, commonly used in physics, are noteworthy for the elegance of their expression and the ease with which they can be used to solve physical problems. Let us consider the origin of variational principles. Figure I.2 shows that, given a function: $y=f(x)$, movement may take place from one point $A(x_1,y_1)$ to another $B(x_2,y_2)$ along different paths, such as (C) or (C′). Each of these paths may be assigned a certain value represented by

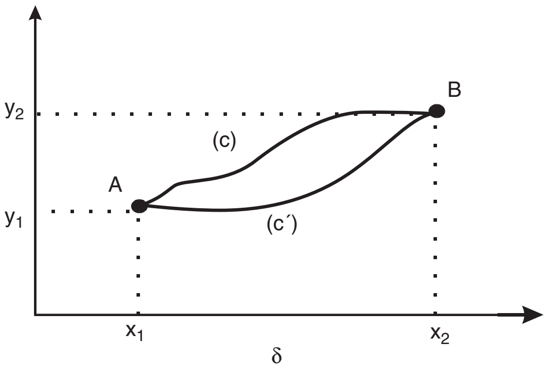

Figure I.2: The principle of least action for a null variation with time. The extremum curve passes through A and B.

the function I, and a path may be selected such that the function I be a stationary function, i.e. a function for which the variation, or the first-order derivative, is null. This particular path, called the *extremum curve*, corresponds to the *extremum function*. The variation is denoted by the symbol δ so that the inertia of the extremal function I is given by the expression: $\delta I = 0$. This function, which represents the global effect between two points A and B, in fact integrates a succession of local effects, say F, between two points: x et $x + \mathrm{d}x$, that are infinitesimally close together.

The extremum curve represents the shortest path under conditions of minimum stationariness. In most problems, the extremum is required to be a minimum, but in some cases the extremum desired may correspond to the maximum. In mechanics, a problem enunciated in these terms leads to an equation similar to that of Lagrange,[10] but in which the Lagrangian L is replaced by the function F described above. Then the function I, in this case called the *Hamiltonian action*, noted S, represents the integration of the quantities: $L = T - V$, between the two instants corresponding to the beginning and the end of the movement. *Hamilton's principle* establishes that the variation δ of the action is null, or $\delta S = 0$. Thus, Hamilton's principle, which is a consequence of Newton's laws of motion, may be considered as a fundamental principle from which the Lagrange equations and the whole field of mechanics may be derived by means of a concise and elegant formulation. Applied to Newtonian dynamics, Hamilton's principle corresponds to a principle of least action.

Let us now consider another type of *principle of least action*, which has several interesting applications. This is obtained from a more general variation, noted Δ. The variation Δq is null at the limits q_1 and q_2, whereas the variation Δt is not null (Figure I.3).[11] Then, if the system is conservative, the principle of least action leads to a new function W, called the *Maupertuis function*, of

---

[10] The Lagrange equation is written: $\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial q}\right) = 0$, with $\delta \int_{t_1}^{t_2} (q, q, t) = 0$.

[11] The relationship between the two variations is given by: $\Delta = \delta + \Delta \mathrm{t}\ \frac{d}{dt}$

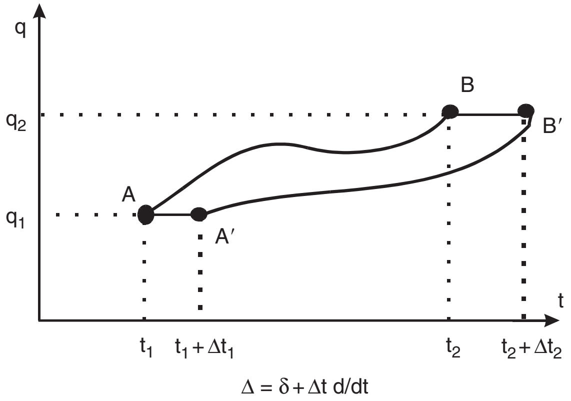

Figure I.3: Here there is a variation of time as well as a variation of position: the variation includes the displacements AA' and BB'.

which the variation Δ is null.[12] This is a more general form of the principle of least action since it takes into account the variation of time as well as the variation of the parameters of position *q*.

The principles indicated in Table I.2 are of great importance in dynamics since they express a very subtle, fundamental physical property. As Feynmann (1969) said in one of his informal lectures on physics:

> "Problem: Find the true path. Where is it? One way, of course, is to calculate the action of millions and millions of paths and look at which one is lowest. When you find the lowest one, that's the true path.
>
> That's a possible way. But we can do it better than that. When we have a quantity that has a minimum, for instance, in an ordinary function like the temperature — one of the properties of the minimum is that if we go away from the minimum in the *first* order, the deviation of the function from its minimum value is only *second* order. At any place else on the curve, if we move a small distance the value of the function changes also in the first order. But at a minimum, a tiny motion away makes, in the first approximation, no difference.
>
> That is what we are going to use to calculate the true path. If we have the true path, a curve that differs only a little bit

---

[12] This function may be written as: $W = \int \sum pq \, dt$

Table I.2: Principles of least action.

	Local effect	*Global effect Extremum function*	*Stationariness*	*Consequences*
Geometry	F Element of length	I Length	$\delta I = 0$ Shortest path	Great circle of the sphere
Mechanics	L Lagrangian	S Hamiltonian action	$\delta S = 0$ Hamilton's principle	Law of inertia Principle of relativity
Optics	n Refractive index	N Optical action	$\delta N = 0$ Fermat's principle	Descartes' Law of Refraction

> from it will, in the first approximation, make no difference in the action. Any difference will be in the second approximation, if we really have a minimum.

The principle of least action applied to the dynamics of a material point allows us to retrieve Galileo's principle of relativity. In effect, the Lagrange function for a free body, not subjected to an external force, will have the form $L = T$ since $V = 0$. We then deduce that the action S is given by the integral of v^2 between the initial and final instants. According to the principle of least action, the function S, which is at a minimum, corresponds to a constant velocity of the body between the initial and final instants. This corresponds to Newton's first law of dynamics, i.e. the law of inertia, which states that a free movement is rectilinear and uniform, describing a straight line in Euclidean space. Since the Lagrangian remains unchanged from one reference frame to another in which the same rectilinear, uniform movement occurs, the same laws of motion will apply in both reference frames. Therefore, the spatiotemporal properties and the laws of mechanics will also remain unchanged. This corresponds precisely to *Galileo's principle of relativity*.

An important principle that has completely changed our understanding of the universe is Einstein's principle of relativity (Einstein, Lorentz, Weyl and Minkowski, 1952). The consequences of this principle have led to the construction of what must be considered the most beautiful theory in physics today. Einstein's extension of the law of inertia established that "straight lines" belong to a universe possessing a Riemannian metric space.[13] In this space, the "straight lines" are in fact *geodesics*, or curves corresponding to the shortest paths between two points on any surface, taking into account the forces of gravity.

Another example of the principle of least action is to be found in the laws of optics enunciated by Fermat (1601–1665). More precisely, in this case we should refer to the principle of least time since among all the possible paths that light may

---

[13] The Riemannian metric space is given by: $ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2$, where x, y, z are the coordinates of a point in space, t the time, and c the velocity of light.

take to go from one point to another, the path chosen will be the one requiring the shortest time. A ray of light refracted through different media obviously does not take the shortest path. Fermat therefore suggested that the path taken might be that requiring the shortest time. A century later, the definition of the Lagrange function for this type of system led to the general formulation of the *optical path*, i.e. the path that light would take in a vacuum during the time actually taken. This would be cT, where T is the interval of time and c the velocity of light. Here, the action N is given by the integral of the index of refraction n, which is a continuous function of the position, corresponding to rate of propagation of the phase wave from one point to the other. Fermat's principle can thus be expressed by $\delta N = 0$. From this we can readily deduce Descartes' well-known law of optics: $n \sin i =$ constant, where i is the angle of incidence.

Many other applications of the principle of least action could be cited, for example in electrostatics, electromagnetism, quantum mechanics or relativity. But what is interesting is the power of this principle in situations that may be described as virtual. In the example above, Descartes' law of optics enables us to understand the behavior of light when, traveling in a given medium, it strikes another medium. It is easy to realize that something occurs at the interface, leading to the refraction of light, so that the phenomenon is perfectly understandable. However, with the principle of least time, the understanding becomes problematic. Here is what Feynmann (1969) had to say about the principle of least time:

> The principle of least time is a completely different philosophical principle about the way light works. Instead of saying that it is a causal thing, that when we do one thing, something else happens, and so on, it says this: we set up the situation, and *light* decides which is the shortest time, or the extreme one, and chooses that path. But *what* does it do, *how* does it find out? Does it *smell* the nearby paths, and check them against each other? The answer is, yes, it does, in a way. That is the feature which is, of course, not known in geometrical optics, and which is involved in the idea of *wavelength*; the wavelength tells us

approximately how far away the light must “smell” the path in order to check it.

Such problems can be solved by *variational calculus*. To take a simple example from geometry, a circle is described as the locus of all the points situated at a constant distance from a fixed point. This is the simplest definition, but the circle could be defined in another way. We could say it is a curve of a given length, surrounding the largest surface. In the examples considered above, we looked for the minimum value, whereas here we are interested in the maximum.

It is surely extraordinary that the fundamental laws of physics can be expressed in terms of a principle of least action. True, this principle applies only in the case of conservative systems for which all the forces can be calculated by means of a potential function. However, it should be noted that non-conservative systems exist only at the macroscopic level. In fact, at the microscopic level there are no dissipative systems. Frictional phenomena, for example, appear at the macroscopic level simply because we ignore the underlying microscopic activity of the vast number of particles involved.

The formalization of physical phenomena

There is evidently a close relationship between physics, the science of the material universe, and mathematics. It is certainly remarkable that the intelligibility of nature should be mathematical. How does the mathematical formulation actually work with a scientific construct? First, there is the experiment that allows us to discover the relationships between the results of the measurements. Then, these results are used to define the physical quantities corresponding to mathematical entities. For instance, with Newton’s laws of motion, the measurement of the time taken by a sphere to roll down an inclined plane, from one point to another, reveals its *acceleration*, expressed mathematically as a uniform increase in velocity. Here, we can already see how complicated and misleading it might be to express the idea of acceleration in ordinary words. What exactly is a uniform increase in velocity? We could, of course, say that the velocity increases by a certain amount during each interval of time. However, the velocity itself is the variation of the distance during the same interval of time. Again, we might say that the acceleration is the variation of the variation of the distance with time. In fact, we can define acceleration only because a quantified measurement can actually be made. Formalization can then be used to express the acceleration more conveniently. In the physical sense, such variations are mathematically described by differentials, i.e. infinitesimal elements of a variable that, in this case, may be time or space. In fact, it was the physical discovery of the phenomenon of acceleration that led to the development of the corresponding mathematical tools, i.e. differential and integral calculus. However, as we shall see, newly observed physical phenomena can often be introduced within the framework of an established mathematical theory for further investigation.

How exactly do we construct a mathematical formulation for natural phenomena? How does the mathematical process of integration then allow us to make predictions concerning the phenomena observed? Of course, the relationships obtained by the analysis of the phenomena must be reproducible, i.e. measurements made under identical conditions should always have the same values. Thus, in the case of the sphere rolling down an inclined plane, the same measuring instruments, for example, a stopwatch or a pair of photoelectric cells, should indicate the same time intervals. Only reproducible results such as these will allow us to conceive the scientific structure underlying the observation. The first consequence of the existence of a reproducible relationship is that it is valid in all points of time and space. Moreover, this relationship must remain valid even though certain factors, or parameters, may vary with space. In our example, the factor g, the gravitational constant that intervenes in the movement, depends on the place where the experiment is carried out. Between the force acting on a body, located at a given point in space, and the movement of the body from this point to a neighboring point in space, there exists a simple relationship in which the constant of proportionality is the mass of the body, as indicated by Newton's laws of motion. Here we have the secret of the process of integration. This local relationship, existing between two neighboring points, is repeated successively from point to point, and the sum of all the local variations then gives the integrated, global variation between the starting and finishing points of the movement. Metaphorically, we might say that mathematics tells us exactly what will happen a little further in space and an instant later in time since we already know what is happening here and now. There is clearly a close correspondence between the physical phenomenon, i.e. the variation observed in space between two successive instants, and the mathematical solution of the differential equation, since this correspondence exists between the different physical elements, e.g. distance, time, force and mass, and the mathematical symbols of the equation. This one-to-one correspondence may be called *the natural law*, since it allows us to make predictions concerning natural phenomena. According to Ullmo (1969):

> Prediction is possible as soon as there is a parallelism between the variations in the measurements of certain physical quantities and the variations of certain mathematical symbols. This is the principal form of the mathematical expression of natural phenomena. The parallelism, ensured by the constant testing of the reproducibility of results in the physical universe, is imposed by the mathematician.

Here we have the principle underlying the mathematical formulation of a physical phenomenon. We see that it is essentially based upon the superposition of a large number of elementary phenomena, all of which are identical, so that the physically observable global phenomenon corresponds to the mathematical solution of a differential equation.

Once we grasp the mechanism of the elementary phenomenon, the integration of all the elementary phenomena will give us an insight into the complex phenomenon observed. The mathematical formulation is useful because of the simplicity of the elementary phenomenon. For instance, Newton's law of gravitation is simple since it expresses the elementary phenomenon of attraction between two bodies. Again, Coulomb's law expresses the elementary phenomenon of attraction between two electrical charges with opposite signs, and that of repulsion when the signs are the same. When these laws are applied to sets of masses or charges, the mathematical combination of the elementary phenomena becomes extremely complex. However, mathematical reasoning enables us to move from the local to the global aspect of the phenomenon mainly because of certain "natural" properties of non-living matter, such as its *homogeneity* (the physical properties of matter being the same at all points), its *locality* (the effects produced at a distance being negligible), and the *simplicity* of the elementary mechanism. The integration of the elementary mechanisms leads to a higher level of the organization of matter. In this sense, the global phenomenon is explained by the *mathematical integration* of local phenomena.

Evidently, the discovery of new elementary mechanisms is likely to become increasingly difficult as we advance in our knowledge of the physical universe. The three characteristics of non-living matter, i.e. homogeneity, locality and simplicity, may not all be present at a certain level of structural organization. However, certain fundamental principles, such as the conservation of matter and energy, still allow us to describe the physical universe. In contrast, the absence of these characteristics in living organisms may suggest the impossibility of a mathematical representation of biological phenomena. As we shall see in the chapters that follow, biology cannot be reduced merely to physics. The structural organization of the non-living world depends on interactions called forces. In living organisms, such interactions are compounded by a complex set of functional biological interactions that can be physically described. Finally, the combination of these functional interactions produces the self-organization of a biological system.

The formalization of natural phenomena is therefore subject to important constraints. For instance, let us consider the differential equation representing Newton's law: ($d^2x/dt^2 = F/m$). This is a second-order differential equation with respect to time. However, why do the other derivatives not appear in this equation? Furthermore, why is there no variation with respect to space? It is not easy to answer these questions. In this particular case, experimental observation revealed the local physical variation that is expressed by the differential equation. However, it can be shown mathematically that a second-order differential equation is equivalent to a system of two first-order differential equations. In general terms, a differential equation of order *n* is equivalent to a system of *n* first-order differential equations. Such systems are called *dynamic systems*. To return to Newton's law, the mathematical formulation actually introduces a new physical concept — that of the quantity of movement, or *momentum*.

The first question concerns the existence of higher derivatives. In fact, the formulation of certain physical phenomena requires that such terms be introduced into an *n*-order differential equation corresponding to the dynamic system. The mathematical theory of dynamic systems can then be applied to obtain, if not the solutions themselves, then at least the conditions for the stability of these solutions in the neighborhood of the singular points, i.e. the points of the dynamic system at which the first derivative is null. In other words, in the equation: $\mathrm{d}X/\mathrm{d}t = f(X)$, we can determine the values of X such that $f(X) = 0$. The behavior of these non-linear differential equations is often rather strange, sometimes producing periodic solutions describing the variation of the physical system with time. Since this behavior depends on the values of the parameters associated with the equations, the variation of the system with time can be mathematically predicted in a way that would otherwise be impossible. Thus, in the case of dissipative chemical structures, subjected to an open thermodynamic regime, the periodic structuring of the medium with time was successfully predicted by solving the dynamic system of equations representing the phenomenon (see Chapter V).

The second question concerns the existence of variations with respect to space. Some equations contain terms, such as $\mathrm{d}X/\mathrm{d}x$, $\mathrm{d}^2X/\mathrm{d}x^2$, and so on. We then have partial differential equations. As in the case involving variations with respect to time, these equations represent reproducible local, spatiotemporal relationships. In general, partial derivatives of an order higher than the second order play no role in the mathematical representation of physical phenomena.

An example of formalization: the phenomenon of diffusion and the propagation of heat

A classic example of an equation of the type considered above is the reaction-diffusion equation. This equation describes a physical phenomenon in which, during the time dt, the variation of $X(t)$ (for example, the concentration of a chemical solution) is the result of the combined action of two mechanisms: on one hand, the diffusion due to the Brownian motion, or molecular agitation, which is a purely thermodynamic effect,[14] and, on the other, the chemical reaction represented by the function $X(t)$. Since $X(t)$ is a vectorial function, we obtain a system of partial derivative equations.

Although it is rather complex because of the mechanisms involved, this example of the coupling between the reaction and the diffusion is a good illustration of what we have called *the natural law*, i.e. the correspondence between the physical elements and their mathematical symbols. The diffusion process is represented by: $D\nabla^2X$, where D is the diffusion coefficient for the chemical species, the concentration of which in the medium is X, the reaction being represented by $f(X)$. The equation of the variation of X with respect to time may then be written as: $\partial \mathrm{X}/\partial \mathrm{t} = D\nabla^2X + f(X)$. This equation also shows the locality of the physical variations. In general, this type of equation expresses a physical principle, for instance that of the *conservation of mass* in a small volume situated about a point x between two infinitely close instants: t and $t+\mathrm{d}t$. Thus, the formalization is underpinned either by experiments describing the phenomenon through a set of reproducible relationships, or by a few great principles that are sufficiently general to be applicable in a wide range of situations.

---

[14] Einstein (1905) showed that Brownian motion, corresponding to the random movement of small particles suspended in a fluid, could be explained by the effect of the collisions between the molecules of the fluid and the particles.

I

In classic dynamics, the principle of conservation of energy is expressed simply by saying that an isolated dynamical system conserves its energy during its variation with time. The principle may be formulated in terms of the two types of energy known, i.e. the potential energy that depends only on the position of the body, and the kinetic energy that depends only on its velocity. The physical significance of this is far-reaching indeed — the dynamical system varies with time such that the potential energy is transformed into kinetic energy, the sum of the two remaining constant. However, the problem becomes inextricable if we consider a system containing large numbers of particles, since the determination of the dynamics of the system would require the knowledge of the dynamics of each of the particles. For instance, even in a tiny drop of a reactive solution, this might mean analyzing the behavior of 10^{20} particles by solving three times this astronomical number of simultaneous equations — an almost impossible task for the time being.

If this is the case, how do we interpret a physical experiment as simple as that of recording the propagation of heat in a metal bar strongly heated at one end? What we observe in practice is the tendency towards a homogeneous distribution of temperature in the bar, i.e. a thermal equilibrium. According to Fourier (1768–1830), who established the theory of the propagation of heat in solids, the flux of heat between two bodies is proportional to the temperature gradient between them. Mathematically, this is described by an equation containing second-order partial derivatives with respect to space, and first-order partial derivatives with respect to time.

In the case of the metal bar, the solution to this equation corresponds to the distribution of heat according to the well-known statistical law due to Laplace (1749–1827) and Gauss (1777–1855). This remarkable law governs certain random phenomena under very general conditions, as often found *normally* in nature, which is why it is usually called the *normal law of distribution*. The calculation of probability shows that this law applies to a randomly repeated phenomenon, each occurrence of which is independent, and which is subject to the influence of a large number of independent, random factors. For example, if we make repeated measurements of some quantity, under the same conditions, then the values obtained will satisfy the normal law of distribution.

Let us now consider the problem of the random walk, which simulates Brownian motion. In this model, a molecule is supposed to move in a random direction over a unit distance in each unit time-interval. If we calculate the probability of finding the molecule in a given point at a given time, we shall have solved the diffusion equation that, we may recall, expresses a law of conservation of the number of particles. We obtain the same type of solution for the heat equation. The *irreversibility* of the phenomenon of diffusion is thus included in the solution of these models, since the solutions are found to be identical. Although the problem of the propagation of heat and the problem of the random walk involve apparently very different phenomena, the profound analogy between them is expressed by the identity of the solutions obtained. Here we may admire the power of mathematics that has led to the discovery of a fundamental unity that would otherwise have remained hidden.

Mathematics and physical reality

The description of the physical universe from the standpoint of general principles is of such elegant simplicity that it appears truly beautiful — at least from a theoretician's point of view. This is clearly the case for the theory of relativity that showed that the velocity of light is a physical constant, for quantum mechanics that revealed the wave-particle duality of radiation, for statistical mechanics that brought out the remarkable relationship between entropy and probability, to mention only a few of the most striking examples.

The esthetic aspect of a mathematical theory is no less important than that of any great work of art. In both types of endeavor, the subjective beauty we behold springs from our ability to encompass the extraordinary nature of the work in a single glance. Nevertheless, the quality of simplicity we admire is merely apparent. In fact, behind the simplicity of a fugue written by Bach or a Mass composed by Mozart, the shapes and colors of abstract art or the beautifully restrained examples of Romanesque architecture, there lies hidden an extraordinarily complex organization.

The mathematical formulation of a physical phenomenon gives us an immediate insight into its subtle complexity. As our understanding of the physical universe deepens, we keep on discovering new laws based on abstract concepts. Thus, the description of forces acting on elementary particles is far removed from Newton's law since, at this level of description, we have to use the principles of symmetry, the laws of conservation, the mathematical theory of groups and elaborate operators. Nevertheless, the essential simplicity will remain, although a high price may have to be paid through relentless efforts at abstraction. However, in the end we shall be richly rewarded since the mathematical principles involved will surely accentuate our belief in the fundamental harmony of the physical universe.

We may wonder to what extent the abstract world of mathematics might be identified with physical reality. Could the profundity of mathematical concepts actually correspond to some intrinsic truth in our universe? In effect, the mental constructions of mathematicians do appear to have a certain "reality", as if such constructions were more akin to discoveries than to inventions. Nevertheless, we may ask how a new mathematical structure might be discovered through the examination of the initial physical structure. In contrast, the mental construction of reality appears to be sufficiently "natural", leading us to believe in the reality of the mathematical world, i.e. the Platonic world, in the sense of the reality of the physical world, and finally, to believe in the essential harmony between the two worlds. In the chapters that follow, we hope to extend this Platonic view of reality to the biological world. Evidently, this approach is based on pure intuition, or speculation, as some might call it. However, the coherence of the results obtained in biology is such that it could hardly be due to mere chance. Thus, Penrose (1989) claimed that his little finger told him that the human mind was surely more than just a "collection of microscopic wires and circuits", and added:

> Because of the fact that mathematical truths are necessary truths, no real information, in the technical sense of the term, is conveyed to the discoverer. All the information was there all the time. One had simply to put things together to see the answer! This fully agrees with Plato's idea that a discovery (a mathematical one, for instance) is merely a form of *reminiscence*! I have often been struck by the resemblance between the simple fact of being unable to recall someone's name and the inability to find the right mathematical concept. In each of these cases, the concept we looked for is, in a sense, already present in the mind, although in the form of unusual words in the case of a mathematical idea that has not yet been discovered.
>
> If this point of view is to be of use in mathematical communication, we have to imagine that interesting and profound mathematical ideas somehow have a more forceful existence than those that are uninteresting or trivial.

Without going as far as Penrose with regard to the "spirit" of matter, we shall restrict ourselves to the more prosaic "functioning" of matter so as to seek some general principles underlying the "collection of microscopic wires and circuits" sometimes mentioned in connection with the working of the human mind. However, even considering only the concrete aspects of biology, putting aside all notions of the conscious mind, it is rather difficult to do otherwise than believe in the mysterious bond between the Platonic world of mathematics and that of the living world. After all, is there any reason why what is true for physical reality should not also be true for biological reality?