Chapter 2
Why Laws of Classical Physics Have Their Form

This chapter is devoted to the detailed analysis of the relationship between the principle of microscopic level reducibility and the formalism of classical physics, which has been used in Chap. 1. Initially I had no intention of writing this chapter and included its materials in Sect. 1.6. There, however, they turned out to be in conflict in style not only with this section but also with Chap. 1 as a whole. The matter is that Chap. 1 is devoted to conceptual discussion of the analyzed problems, whereas the materials to be presented below involve a large amount of mathematical details and manipulations. Therefore, I have decided to place them into an individual chapter and to keep in Sect. 1.6 only the main conclusions drawn based on these mathematical manipulations. Actually the main purpose of the present chapter is to justify these statements.

I have claimed that the formalism of Newtonian mechanics and statistical physics stems directly from the general principle, called the principle of microscopic level reducibility, which physical systems obey in the realm of classical physics. In this chapter I substantiate this statement and demonstrate directly how the formalism of differential equations, the notion of forces in Newtonian mechanics, the concept of phase space and initial conditions, the axioms of probability theory, etc. results from this principle.

2.1 Principle of Microscopic Level Reducibility

Dealing with objects of the inanimate world in the frameworks of classical physics, we admit the existence of the microscopic (elementary) level of their description. It means that in modeling such physical systems one can make use of the

I. Lubashevsky, *Physics of the Human Mind*, Understanding Complex Systems,
DOI 10.1007/978-3-319-51706-3_2

following premises to be referred further to as the *principle of microscopic level reducibility*.[1]

1. For any physical system there can be found a level of its *microscopic* description at which the system at hand is composed of individual *structureless* entities. The term 'structureless' is used to emphasize the fact that either these entities are really structureless or their internal structure does not change in time during analyzed phenomena and so can be treated as a fixed characteristics of the entities.
2. All the properties exhibited by the given system can be explained based on or derived from
   (a) individual properties of these entities *existing independently* of the presence of the other entities,
   (b) the properties of *pairwise* interaction between these entities meeting the *superposition principle*.

Further these *structureless* entities will be called particles for short.

Premise 2b may be replaced by another one using the concept of fields. Namely, instead of a long-distance interaction of particles a certain field, for example, electromagnetic field is introduced. This field is locally generated by particles, propagates in space, and, in turn, affects them. In these terms Premise 2b is read as

2. All the properties exhibited by the given system can be explained based on or derived from individual properties of its constituent structureless particles (item 2a) and
   b′) the own properties of some fields *freely* propagating through space as well as the properties of the *local* particle-field interaction obeying the *superposition principle* and being responsible for the field generation by the particles and in turn the effects produced by these fields on the particles.

It should be noted that the concept of particle interaction based on Premise 2b′ is much reacher in properties and potentiality in describing complex systems in comparison with that based on Premise 2b; it is clearly demonstrated in Chap. 4 and also noted in Sect. 2.6. However, if the dynamics of a certain system is characterized

[1]The two premises may be regarded as a particular version of the reductionism, a philosophical concept about the relationship between complex systems as whole entities and their constituent parts. In actual fact the concept of reductionism is more complicated and contradictory, for example, there are various versions of reductionism deserving an individual consideration. A detailed discussion of relationship between this principle and philosophical aspects of reductionism as well as holism of complex systems and emergent phenomena is postponed to Chap. 4. Nevertheless, it should be noted that the principle of microscopic level reducibility may be treated as one of the cornerstones in the research paradigm of physics, namely, it is a particular implementation of the general scientific methods based on decomposition analysis and synthesis (see, e.g., Beaney 2015).

by time scales much longer then the mean time during which the corresponding fields propagate in space over distances about the system size, Premise 2b′ is approximately reduced to Premise 2b. It will be used, for the sake of simplicity, in the following two sections, although the results to be obtained can be generalized to theories appealing directly to Premise 2b′. Besides, strictly speaking, the use of the fields leads to the necessity of modifying Premise 1 too because in this case a given physical system is decomposed not only into structureless particles but also fields existing on their own and which have to be treated as its constituent entities. However various aspects of these fields regarded as individual objects on their own, i.e., beyond the scope of the interaction between the particles that is implemented via these fields, do not belong to the subject-matter of this chapter.

The following two comments are also worthy of noting before passing directly to various consequences of the principle of microscopic level reducibility. First, Premise 2a concerns the properties that are ascribed to particles individually, i.e. independently of the presence or absence of other particles. In this sentence by the term "properties" I actually mean a certain collection of *types* of properties ascribed to the particles individually. For example, "being located at a spatial point" is a property type of point-like particles in classical physics, it characterizes a generic feature of all these objects. "Being able to restore its previous form once the forces are no longer applied" exemplifies another generic property which as a type is ascribed to all elastic springs. I noted this fact here to emphasize that particular *instantiations* of these properties, their *tokens*, can depend on the presence of other particles. For example, if a point-like particle A occupies a spatial point $\mathbf{r}$ then another similar particle B cannot be located at this point. Ferroelectricity also exemplifies this feature; in some crystals particular elastic deformation of the crystalline lattice inside a small region can be sustained by similar deformations in other regions via the formation of macroscopic electric field, giving rise to some specific deformation of the crystal as a whole. The difference and relationship between types and their tokens[2] is essential for elucidating the basic features of the general scientific methods based on decomposition analysis and synthesis (see, e.g., Beaney 2015). At the first step, the individual generic properties of physical particles can be studied dealing with one of them taken separately from the others. After that, at the second step, the complex behavior of ensembles of these particles can be reconstructed based on the found properties and the interaction between the particles which is specified by their spatial arrangement and the particular instantiations of their individual properties.

The second comment concerns the implementation of this step in reconstructing the behavior of many particle ensembles. In order to do this we need to know how to specify the interaction between the particles. Following the decomposition strategy it could be reasonable to analyze this interaction for a pair of these particles or, at least, a system consisting of a few particles taken separately. In this way,

[2] A detailed discussion about the distinction between a *type* and its *token* in various aspects can be found, e.g., in the article by Wetzel (2014).

however, we face up to a challenging problem of how the results to be bound can be generalized to the original many particle ensemble. It is solved within the framework of Premise 2b appealing to the superposition principle. This principle postulates that the interaction of an arbitrary chosen particle and all the other particles forming a certain ensemble, for example, the cumulative force with which the other particles act on the given one is just the algebraic sum of all the partial forces that can be found in the following way. We should consider a pair of the chosen particle and any one of the other particles assuming the remaining particles of this ensemble to be absent. Then the corresponding partial force is just the force with the second particle of the given pair would act on the first one in this case. In particular, the superposition principle allows us to reduce the interaction energy of a many particle ensemble to the sum of the energies of pair-wise interaction between individual pairs of its particles running over all the possible pairs in this ensemble. The superposition principle including its formulation relevant to Premise 2b′ is discussed in more detail in Sect. 2.6. Finally, I want to note that Premise 2b can be easily generalized to including also plausible three-body forces.

2.2 Thick Presentism and Formalism of Differential Equations in Classical Physics

In this section and the next ones I present some arguments about why Newtonian mechanics is based on the mathematical formalism of second order differential equations. At the first step appealing to the principle of microscopic level reducibility let us try to elucidate what general mathematical form the laws governing the dynamics of physical systems should have within the framework of classical physics.

2.2.1 Presentism and the Time Flow

The possibility of reducing a description of a physical system to *structureless* particles and interaction between them has an important consequence. These particles cannot remember their history or foresee their future; they just have no means to do this, so only the present matters to them. Therefore all the plausible quantities $\{Q\}_\alpha$ that can be used to describe the laws governing the motion of a given particle α have to be taken at the current moment of time. Naturally there should be other characteristics of the particles such as mass, charge, spin magnitude, etc. which, however, are treated as their internal properties not changing in time. Let us regard the dynamics of these particles as their motion in a certain N-dimensional space $\mathbb{R}^N$; for our world treated in the realm of classical physics $N = 3$. So the spatial position (spatial coordinates) x_α of the particle α has to enter the collection

$\{Q\}_\alpha$. The motion of this particle is represented by the time dependence $x_\alpha(t)$ of its position showing the points occupied previously and the points to be got in future according to prediction of its dynamics. However, for such particles

- the past no longer exists,
- the future does not exist yet,
- only the present matters to them and determines everything.

Thereby solely instantaneous characteristics of the particle motion trajectory $\{x_\alpha(t)\}$ may also enter the collection $\{Q\}_\alpha$. They are time derivatives of $x_\alpha(t)$ taken at the current moment of time t. In particular, it is the particle velocity $v_\alpha(t) = dx_\alpha(t)/dt$, its acceleration $a_\alpha(t) = d^2x_\alpha(t)/dt^2$, the time derivative of third order called usually the jerk or jolt $j_\alpha(t) = d^3x_\alpha(t)/dt^3$, and so on. However, in order to construct a time derivative we have to consider not only the current position $x_\alpha(t)$ of a particle but also its position $x_\alpha(t - \Delta)$ in the *immediate* past separated from the present by an infinitely short time interval $\Delta \to +0$. Indeed, for example, the particle velocity is defined as $v_\alpha(t) = \lim_{\Delta\to+0}[x_\alpha(t) - x_\alpha(t - \Delta)]/\Delta$. At this place an attentive reader may find some contradiction, in speaking about the present we actually deal with a certain kind of instants including not only the point-like current moment of time but also other time moments belonging to some neighborhood of the current time whose size may be an infinitely small value. It causes us to speak about the *thick* present.

The concept of *thick* present is worthy of special attention because it leads directly to the formalism of differential equations and the principle of least action playing a crucial role in modern physics. Therefore let us focus out attention on the philosophical doctrine usually referred to as *presentism* which can be employed to penetrate deeper into the concept of *thick* present.

Broadly speaking, *presentism* is the thesis that only the present exists. In the given form it is a rather contradictory and ambiguous proposition being one of the subjects of ongoing debates about the nature of time tracing their roots in ancient Greece. In particular, the problems of presentism are met in the famous paradoxes of motion (see, e.g., Huggett 2010) devised by the Greek philosopher Zeno of Elea (circ. 490–430 BC). Unfortunately, none of Zeno's works has survived and what we know about his paradoxes comes to us indirectly, through paraphrases of them and comments on them, primarily by Aristotle (384–322 BC), but also by Plato (428/427–348/347 BC), Proclus (circ. 410–485 AD), and Simplicius (circ. 490–560 AD). The names of the paradoxes were created by commentators, not by Zeno (Dowden 2016).

We confine ourselves to the *arrow* paradox primarily mentioned in the context of the time problem. This paradox is designed to prove formally that the flying arrow cannot move, it has to be at the rest and, so, the motion is merely an illusion. Citing Aristotle's *Physics* VI,

> [t]he third is ... that the flying arrow is at rest, which result follows from the assumption that time is composed of moments .... He says that if everything when it occupies an equal space is at rest, and if that which is in locomotion

is always occupying such a space at any moment, the flying arrow is therefore motionless.

Focusing our attention on the issue in question I want to interpret the arrow paradox as three logical steps:

- time is composed of instants—point-like moments of time—and the present is the current moment;
- only the present matters, i.e., all the properties of the flying arrow including its motion at a certain velocity are determined completely by its current state, i.e., the spatial point where it is currently located;
- whence it follows that the arrow motion is impossible because the state of any arrow, flying from the left to the right, in the opposite direction, or just being at the rest is the same if at the current moment of time it is located at the same spatial point; the arrow just does not "know" in which direction it has to move.

Aristotle was the first who proposed, in his book *Physics* VI (Chap. 5, 239b5–32), a certain solution to the arrow paradox. Since that time this paradox having been attacked from various points of view (see, e.g., reviews by Lepoidevin 2002; Huggett 2010; Dowden 2016), a detailed analysis of Aristotle's solution and its modern interpretation can be found in works by Vlastos (1966), Lear (1981), and Magidor (2008).

A naïve solution to the arrow paradox could be the proposal to include the instantaneous velocity in the list of basic properties characterizing the current arrow state. Unfortunately the instantaneous velocity, as well as the rate of time changes in any quantity, cannot be attributed to an instant—a point-like moment of time. The velocity is a characteristic of a certain, maybe, infinitesimal neighborhood of this time moment (Arntzenius 2000). For this reason Russell (1903/1937) rejects the instantaneous velocity at a given moment to be the body's intrinsic property having some causal power. Arguments for and against this view have been analyzed, e.g., by Arntzenius (2000) and Lange (2005).

As a plausible way to overcoming this causation problem of instantaneous velocity, a special version of presentism admits the present to have some duration (e.g., Craig 2000; Dainton 2010; McKinnon 2003). Following Hestevold (2008) it is called *thick* presentism. On the contrast, *thin* presentism takes the present to be durationless, which, however, immediately gives rise to logical puzzles like Zeno's arrow.

In the framework of thick presentism there has been put forward a rather promising solution to the arrow paradox turning to the formalism of nonstandard analysis; for an introduction to this discipline a reader may be referred to Goldblatt (1998). Following (White 1982; McLaughlin and Miller 1992; McLaughlin 1994; Arntzenius 2000; Easwaran 2014; Reeder 2015) let us equip each point-like time moment t with some neighborhood of infinitesimal thickness 2ϵ, i.e., $t \to \mathbf{t} = (t - \epsilon, t + \epsilon)$ and understand time events as some objects distributed inside $\mathbf{t}$. Here ϵ is an infinitesimal—infinitely small hyperreal number of nonstandard analysis. Below I will use the term *bold* instants in order to address to such objects and not to

mix them with times intervals of finite thickness also conceded in some particular versions of thick presentism. It is worthy of noting that there is no contradiction between the notion of bold instants and the intuitive separability of time moments because for any two moments t_1 and t_2 separated by arbitrary small but finite interval the infinitesimal regions $\mathbf{t}_1$ and $\mathbf{t}_2$ do not overlap.

The notion of bold instants $\mathbf{t}$ opens, in particular, a gate to endowing the instant velocity with causal power just attributing the instant velocity to the left part $(t-\epsilon, t)$ of $\mathbf{t}$ and assuming that its effect arises in the right part $(t, t+\epsilon)$ (Easwaran 2014, a similar view was also defended by Lange (2002)). In this case, as it must, a cause and its effect are ordered in time; a cause precedes its effect.

Introducing the concept of bold instants we have to accept a special topological connectedness of time which is non-local on infinitesimal scales. Namely, for a time moment t at least all the previous time moments in the infinitesimal interval $(t-\epsilon, t)$ are to coexist, otherwise they cannot have causal power on it. Exactly this connectedness paves the way for properties that can be attributed only to time intervals including infinitesimals to have causal power (Lange 2002, 2005; Harrington 2011). Allowing the given multitude of time moments to exist we actually accept a special version of thick presentism called the *degree presentism* proposed by Smith (2002). His account assumes that all events have past and future parts whose existence degree (degree of reality) decreases to zero as their time moments go away from the present. Baron (2015b) has developed a related account of time called *priority presentism* according to which only the present entities exist fundamentally, whereas the past and future entities also existing are grounded in the present.

Any version of presentism has to explain how the flow of time is implemented in dynamical phenomena. In the framework of thick presentism Baron (2012) puts forward the *step-wise* model for the flow of time consisting in temporally extended (bold) instants. Each of these instants comes into and going our of existence in such a manner that successive thick instants partially overlap.

At the next step in describing dynamical processes in terms of thick presentism we face up to a problem of giving the meaning to *time* changes in the properties of some object for which its present partially contains its past and future parts. As a natural way to overcoming this problem, Smart (1949) introduces a complex structure of time containing in addition to the *physical* time a certain *meta-time*. Meta-time is necessary to deal with temporal properties of events embedded into the "river of time" when these properties themselves change in time and a meta-time is a place where these changes can occur. It should be emphasized that the introduction of two-dimensional time for thick presentism with bold instants does not lead to paradoxes arising in the time travel problem and used often as arguments against the possibility of two-dimensional time structure. A review of these arguments is given, e.g., by Richmond (2000), Oppy (2004), and Baron (2015a). The matter is that the difference between the physical time and meta-time becomes essential only within bold instants–the infinitesimal intervals—wherein time travels with non-zero length quantified by standard numbers are impossible.

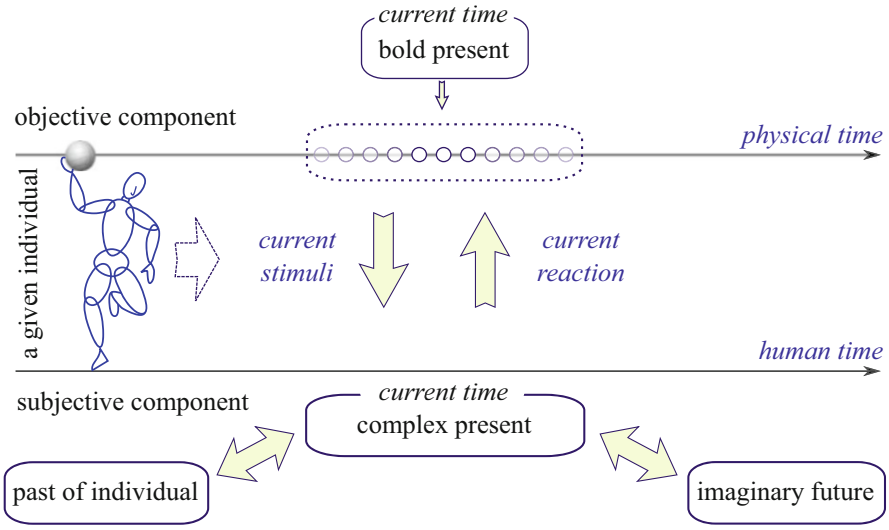

Fig. 2.1 Illustration of a plausible mechanism synchronizing the time flows in the subjective and objective components of human nature

Below I will outline my account of dynamical processes consisting in bold instants which is developed for explaining the use of differential equations for modeling dynamical phenomena in classical mechanics and the variational technique as a fundamental law governing dynamics of physical object. Before this, however, let me elucidate the further constructions using the relationship of human and physical time as a characteristic example.

In Sect. 1.7 two components of human nature, objective and subjective ones (Fig. 2.1), were considered. The objective component represents the world external for a given individual and embedded in the flow of the *physical* time. The subjective component representing the internal world of this individual is equipped with a more complex structure of time to be referred as to *human* time. It consists of the past retained in the memory, the imaginary future, and the complex (specious) present comprising all the moments of the physical time perceived by the individual as simultaneous. Because the past and future in human mind can affect our current actions we have to regard them as real objects existing in the subjective component. A detail discussion of these temporal components requires immersion in the modern theory of time which is beyond the capacity of the given section, so for a review a reader may be referred, e.g., to articles by Markosian (2014) and Hawley (2010) and paper-collections edited by Callender (2011) and Ciuni et al. (2013). Here we touch only the problem how these time components are related to each other, moreover, confine ourselves to the nearest past, the complex present, and the forthcoming future.

Generally speaking we may say that the complex present in the human time is synchronized with the present in the physical time via the direct interaction of the

given individual with the reality. However this synchronization is not an one-to-one map. In fact, the complex present may be conceived of as a certain interval $\mathfrak{T}_t$ with fuzzy boundaries containing the current moment t of the physical time. Its any point t' is perceived by the individual as the present in the human time with some degree $\Phi(t-t')$ decreasing to zero as the time gap $|t-t'|$ increases and exceeds the characteristic duration Δ of the complex present. It is necessary to emphasize that on scales about Δ the order of time moments in the physical time is not recognized by the given individual and so does not exist in the subjective component.

As far as the nearest past in the subjective component is concerned, it may be regarded as fixed. On the contrast, the forthcoming imaginary future permanently changes as its time moments t' come closer to the present, $t' \to \mathfrak{T}_t$, and becomes the fixed reality when the point t' goes into $\mathfrak{T}_t$. It is a result of permanent correction of the imaginary future based on the interaction between the individual and the reality.

The given example prompts me to put forward the following model of the time flow consisting in bold instants applied to describing dynamics of a certain physical system.

2.2.2 *Thick Presentism with Moving Window of Existence*

The non-stand analyses enables us to operate with infinitely small and infinitely large numbers in addition to standard ones. The set of these numbers forms a field, i.e., all the arithmetics operations (addition, subtraction, multiplication, and division), relations, and, thus, many mathematical functions are defined in it (for an introduction see, e.g., Goldblatt 1998). It allows us to deal with infinitesimal bold instants as ordinary intervals.

Using infinitesimals we can introduce a certain function $\Phi(|t'-T|/\epsilon)$ giving us the degree of existence for the events coming into being at a time moment t'. Here ϵ is the infinitesimal thickness of the bold instant $\mathbf{t}$ centered at T and $\Phi(|t'-T|/\epsilon) \to 0$ as the ratio $|t'-T|/\epsilon \to \infty$. The function $\Phi(|t'-T|/\epsilon)$ admits the interpretation as a certain characteristic function of the *window of existence* with fuzzy boundaries which is attributed to the bold instant $\mathbf{t}$. This window moves along the axis of the *physical* time due to the *flow of time*. Actually these constructions introduce a two dimensional (2D) time structure containing the meta-time governing the realization of physical systems on the corresponding time-space manifold (Fig. 2.2).

Having introduced the bold instants—time *intervals* even if they are of infinitesimal thickness—as the basic elements of time flow we have to modify the standard way of describing the dynamics of a certain system in the space $\mathbb{R}^N$. Within the standard description the system is specified by the point-like position $x(t)$ it occupies at the current moment of time t. As time grows the generated trajectory represents the system motion. In the realm of thick presentism we should ascribe a certain degree of existence not only to the point-like object $x(t)$ but also to the trajectory fragments $\{x(t)\}_{\mathbf{t}}$, where $t \in \mathbf{t}$. It means that the very basic level of the system description must consist in the trajectories, at least, their parts rather than point-like

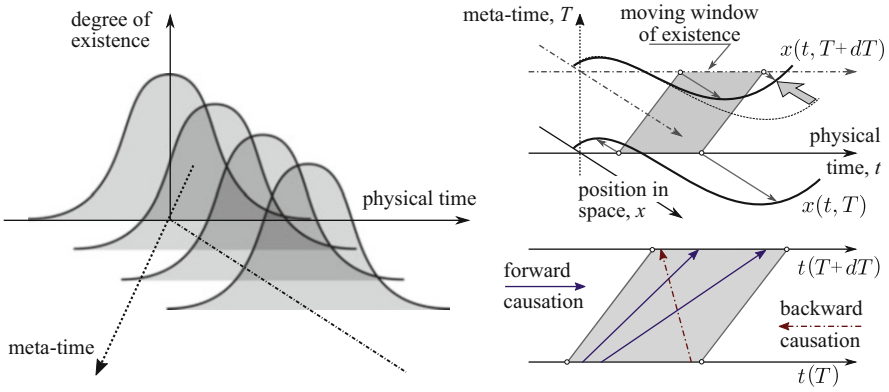

Fig. 2.2 Illustration of thick presentism with moving window of existence: (*left*) two-dimensional time structure and (*right*) the realization dynamics of some physical system

objects and causal power may be attributed only to these basic elements. We should to do this at each moment T of meta-time, otherwise, evolution and emergence as dynamical phenomena are merely a mirage—everything is fixed beforehand. In other words, the basic element of the system description in the 2D-time structure is given by the trajectory $\{x(t,T)\}$ or, speaking more strictly, its partition specified by bold instants $\mathbf{t}$. I have used the term *trajectory* to emphasize that these basic elements are certain functions of the argument t running from $-\infty$ to $+\infty$ rather than points of the space $\mathbb{R}^N$; here the meta-time T plays the role of a parameter.

In these terms the system dynamics at any moment T of meta-time is characterized by the following components:

- the past of the system: $\{x(t,T)\}$ matching $t < T$ and $t \notin \mathbf{t}_T$,
- the thick present of the system: $\{x(t,T)\}$ where $t \in \mathbf{t}_T$,
- the future of the system: $\{x(t,T)\}$ matching $t > T$ and $t \notin \mathbf{t}_T$

depending on T. It is worthy of noting that involving the past and further into consideration of causal processes affecting the system dynamics does not contradict the previous statement about their absence for structureless particles of classical physics. Such particles have no means to remember *individually* their past or to predict their future. However in the case under consideration the causal power of the past and future is due to the physical properties of the time flow itself rather than that of the particles and spends over temporal intervals of infinitesimal thickness only.

In the framework of thick presentism all the properties of the given system at the current moment T of meta-time must be determined completely by the trajectory $\{x(t,T)\}$, whereas the presence of its points in the reality is determined by the current position of the window of existence. It concerns also the property I call the *sensitivity* of the given system to the flow of meta-time or simply *meta-time sensitivity*. It quantifies the variation of the trajectory $\{x(t,T)\}$ *caused* by the meta-

time flow provided the corresponding part of the trajectory is present in the reality. The partial existence of a trajectory fragment in the reality decreases its variation so the governing equation for these trajectory variations can be written as

$$\frac{\partial x(t,T)}{\partial T} = P\left(\frac{|t-T|}{\epsilon}\right)\widehat{\Omega}\left[\{x(t,T)\}\right], \tag{2.1}$$

where the operator $\widehat{\Omega}\left[\{x(t,T)\}\right]$ specifies the meta-time sensitivity of the given system with the trajectory $\{x(t,T)\}$. Figure 2.2 (right fragment) illustrates the variations of the system trajectory as the meta-time grows. It should be noted that within the bold instant $\mathbf{t}_T$ the time moments may not be ordered in their effects, i.e., the variation of the system trajectory at moment $t \in \mathbf{t}_T$ can be partially caused by time moments preceding as well as succeeding it. In the latter case we can speak about backward causation (for a general discussion on the backward causation problem a reader may be referred to Faye 2010).

Equation (2.1) can relate to one another only the trajectory fragments corresponding to bold instants $\mathbf{t}$ that either contain the time moment T or are distant from it over scales about ϵ. So terms similar to

$$\int_{-\infty}^{+\infty} dt'\, K_i\left(\frac{t-t'}{\epsilon}\right) x(t',T) \tag{2.2}$$

should mainly contribute to the variation of the trajectory $\{x(t,T)\}$ at the point $t \in \mathbf{t}_T$ and the typical forms of the kernels $K_i(\ldots)$ are exemplified in Fig. 2.3. Such nonlocal effects can connect only time moments separated by infinitely small time lags whereas the motion trajectory of systems at hand are to be smooth curves. In this case the nonlocal operator $\widehat{\Omega}\left[\{x(t,T)\}\right]$ should reduce to a certain local function

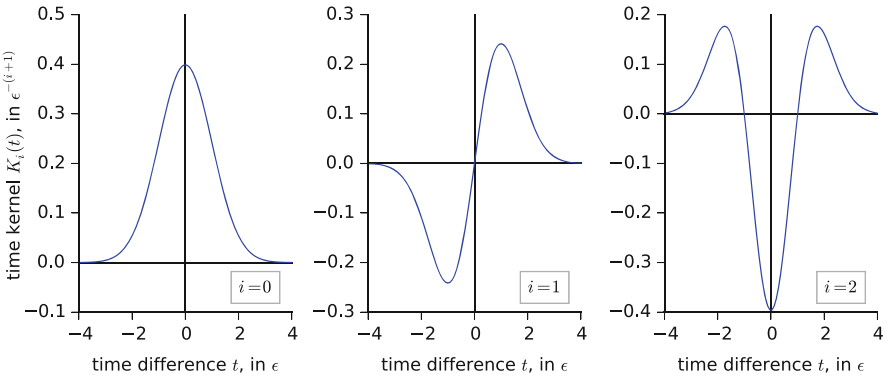

Fig. 2.3 Typical forms of kernels determining nonlocal contribution of different time moments to the system meta-time sensitivity

ω whose arguments are the current system position x and its various derivatives taken at the current moment t

$$\widehat{\Omega}\left[\{x(t,T)\}\right] \Longrightarrow \omega\left[x(t,T), \frac{\partial x(t,T)}{\partial t}, \frac{\partial^2 x(t,T)}{\partial t^2}, \frac{\partial^3 x(t,T)}{\partial t^3}, \ldots\right]. \tag{2.3}$$

The possible forms of this function and the corresponding consequences will be discussed in the following two sections. In the remaining part of this section I will explain the mechanism via which *steady-state* laws governing system dynamics can emerge in the realm of thick presentism.

The window of existence moves from past to future along both the time-axes and never returns to instants already passed. So in the realm of thick presentism the past of a given system cannot change as the meta-time grows, which stems directly from equation (2.1). However it does not mean that the future has no influence on the past. In the general case the past is formed during the trajectory transformation at time moments when the window of reality passes through them and the result depends on the "initial" details of the system trajectory $\{x(t,T)\}$ in the region $t > T$. In this case it is not possible to speak about universal laws governing the system dynamics in the standard interpretation. Nevertheless there a special case when it becomes possible.

According to equation (2.1) the change of the system trajectory $\{x(t,T)\}$ is characterized by two temporal scales. The first one is the thickness of bold instants, ϵ, specifying the duration of the time interval within which a given point of the system trajectory is in present. The second one is the time scale τ characterizing the rate of the conversion of forthcoming future into the nearest past within the current bold instant $\mathbf{t}_T$, i.e., the strength of the operator $\widehat{\Omega}\left[\{x(t,T)\}\right]$. When the ration

$$\frac{\tau}{\epsilon} \ll 1 \quad \text{or, moreover, is itself infinitesimal,} \tag{2.4}$$

the system trajectory gets equilibrium configuration (if it is stable) actually within the bold instant $\mathbf{t}_T$ which is described by the condition

$$\widehat{\Omega}\left[\{x(t)\}\right] = 0\,, \tag{2.5}$$

where the *steady-state* trajectory $\{x(t)\}$ does not depend on meta-time T.

In this case the system past is mainly determined by equality (2.5) and "forgets" completely the "initial" future of the system. As a result the laws describing the newly emerged past as the present may be of a universal form reflecting only the physics of a given system. Figure 2.4 illustrates this situation.

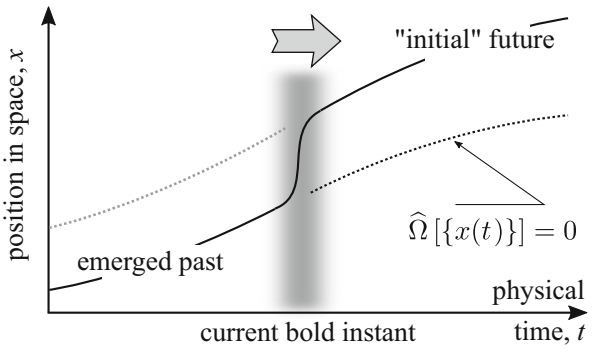

Fig. 2.4 Emergence of the system past matching the system dynamics governed by stable laws which are determined completely by the system physics

2.2.3 *Steady-State Laws of System Dynamics*

I will call condition (2.4) the limit of *steady-state laws* and will assume it to hold in our reality. In this case the system dynamics described in terms of the position $x(t)$ in the space $\mathbb{R}^N$ occupied by the system in the *immediate* past obeys the equation

$$\omega\left[x(t), \frac{dx(t)}{dt}, \frac{d^2x(t)}{dt^2}, \frac{d^3x(t)}{dt^3}, \ldots\right] = 0 \tag{2.6}$$

by virtue of (2.3).

Expression (2.6) is the main results of the present section. In particular, it explains why the laws governing the dynamics of systems in the framework of classical physics admits a representation in the form of some formulas joining together the time derivatives of the motion trajectory taken at the current moment of time. Thereby the formalism of differential equations is actually the native language of physics or, speaking more strictly, Newtonian mechanics. Naturally the question on whether differential equations are the very basic formalism of physics has been in the focus of long-term debates and attacked from various points of view, for a short review see, e.g., Stöltzner (2006) and references therein.

Causal relations can be also attributed to law (2.6), at least, when the list of arguments of the function $\omega(\ldots)$ is finite. In this case resolving equation (2.6) with respect to the highest order m time derivative of $x(t)$ we obtain the expression

$$\frac{d^m x}{dt^m} = \Phi\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}}\right) \tag{2.7}$$

which admits interpretation as a causal relationship between the lower order time derivatives

$$x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}} \tag{2.8}$$

playing the role of causes and the highest time derivative $d^m x/dt^m$ being their effect. Indeed, when the system trajectory undergoes sharp variations inside the bold instant $\mathbf{t}_T$ the highest derivative demonstrates changes most drastically and it is possible to say that the collection of quantities (2.8) finally cause the highest derivative to take value (2.7).

The accepted hypothesis on the finite number of arguments in the function $\omega(\ldots)$ can be directly justified if the field of hyperreal numbers is extended to a ring including nilpotent infinitesimals. Nilpotents are nonzero infinitely small numbers that yield zero when being multiplied by themselves for a certain number of times. So if the kernels $K_i(\ldots)$ contain nilpotent cofactors, the meta-time sensitivity operator $\widehat{\Omega}[\{x(t,T)\}]$ can comprise only finite order power terms with respect to quantities similar to (2.2). In a similar way Reeder (2015) uses nilpotents for constructing a novel solution to Zeno's arrow.

2.2.4 *Variational Formulation of Steady-State Dynamics*

There is a special case worthy of individual attention that admits the introduction of a certain functional

$$\mathcal{L}[\{x(t,T)\}]$$

to be call *action* following the traditions accepted in physics. This functional specifies the operator of meta-time sensitivity as its functional derivative

$$\widehat{\Omega}[\{x(t,T)\}] = -\frac{\delta S[\{x(t,T)\}]}{\delta x(t,T)}. \tag{2.9}$$

Because bold instants can couple only infinitely close time moments the action functional in the general form can be written as

$$\mathcal{L}[\{x(t,T)\}] = \int_{-\infty}^{+\infty} dt\, L\left[x(t,T), \frac{\partial x(t,T)}{\partial t}, \frac{\partial^2 x(t,T)}{\partial t^2}, \frac{\partial^3 x(t,T)}{\partial t^3}, \ldots\right], \tag{2.10}$$

where function

$$L\left[x(t,T), \frac{\partial x(t,T)}{\partial t}, \frac{\partial^2 x(t,T)}{\partial t^2}, \frac{\partial^3 x(t,T)}{\partial t^3}, \ldots\right] \tag{2.11}$$

is called the Lagrangian of a given system.

In the limit of stead-state laws the trajectory $\{x(t)\}$ meeting condition (2.5) should be stable with respect to small (infinitesimal) variations

$$x(t,T) = x(t) + \delta x(t,T)\,.$$

It means that the variations $\delta x(t,T)$ have to fade as meta-time T grows. This stability condition directly gives rise to the following requirement which has to be imposed on the corresponding form of the action functional and its Lagrangian.

Principle of Least Actions: Let a physical system admit the introduction of the action functional (2.10) describing its dynamics in meta-time. Then its steady-state trajectory $\{x(t)\}$ describing the system motion in the past including the immediate past matches the minimal value of the action functional among all the other possible trajectories

$$x(t) \Longrightarrow \min \int_{-\infty}^{+\infty} dt\, L\left[x(t), \frac{dx(t)}{dt}, \frac{d^2x(t)}{dt^2}, \frac{d^3x(t)}{dt^3}, \ldots\right]. \tag{2.12}$$

Actually this principle is in one-to-one correspondence with the principle of least actions well-known in physics provided the Lagrangian $L(x, dx/dt)$ depends only on the system position x and the velocity dx/dt.

2.3 Notion of Phase Space

In the previous section we have considered the general description of system dynamics in the framework of thick presentism and the limit of steady-state laws has been assumed to hold in our world. In this case the dynamics of a physical system conceived of as the motion of a point x in a certain space is governed by equation (2.6) joining together all the time derivatives of the trajectory $x(t)$ taken at the current moment of time t.

In what follows, first, we will confine ourselves to the case where the number of the time derivatives entering the right-hand side of (2.6) is finite for any physical object. Second, we will consider an ensemble of structureless particles whose individual motion can be represented as the motion of a point x_α in the space $\mathbb{R}^N$; in our world $N = 3$. This ensemble may be described as a point $x = \{x_\alpha\}$ of the

space $\mathbb{R}^{NM}$, where M is the number of particles in the given ensemble. Besides, for the sake of simplicity we will assume that for all the particles only the first $(m-1)$ derivatives of their coordinates x_α enter equation (2.6).[3]

Under these conditions equation (2.6) treated as some equality can be reserved with respect to the highest derivative $d^m x/dt^m$ which gives us expression (2.7). This expression may be interpreted as a causal type relationship between the time derivatives of order less than m (including the zero-th order derivative just being the particle positions) and the derivative $d^m x/dt^m$. For individual particles formula (2.7) takes the form

$$\frac{d^m x_\alpha}{dt^m} = \Phi_\alpha \left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}} \right). \tag{2.13}$$

where the particle index α is omitted at the list of arguments in the right-hand side of (2.13), which denotes that all the particles of a given ensemble should be counted here because of the particle interaction.

The fact that the right-hand side of equation (2.13) contains only the time derivatives of order less than m does not mean the mutual independence of these quantities. There could be conceived of some additional constrains imposed on this system such that one of these derivatives, mainly, $d^{m-1}x/dt^{m-1}$ is completely determined by the others. It actually reduces the number of arguments in (2.13). Therefore below we may assume the collection of quantities

$$\{Q\}_\alpha = \left\{ x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}} \right\}_\alpha, \tag{2.14}$$

to be *mutually independent* for all the particles $\{\alpha\}$. It means that for arbitrary chosen values there can be found an instantiation of this system such that during its motion these time derivatives take the given values at a given moment of time t.

Now we can introduce the notion of the *phase space*

$$\mathbb{P} = \left\{ x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}} \right\} \tag{2.15}$$

for the system at hand regarded as a whole. If we know the position of the system in the space $\mathbb{P}$ treated as a point θ with the coordinates

$$\theta = \left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}} \right), \tag{2.16}$$

[3]The further constructions can be easily generalized to the case when the state of different particles is characterized by different parameters m_α, which, however, over-complicates the mathematical expressions without any reason required for understanding the subject.

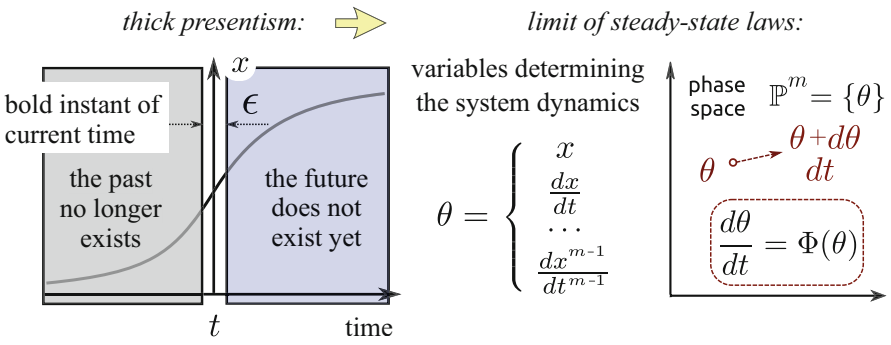

Fig. 2.5 Illustration of the phase space introduction starting from the principle of microscopic level reducibility in the framework of thick presentism

then the rate of the system motion in the phase space $\mathbb{P}$ is completely determined via relationship (2.13). Solving this equation we can construct the trajectory of the system motion. The aforesaid is illustrated in Fig. 2.5.

The phase space is one of the basic elements in describing such objects. In particular, specifying the system position in the phase space $\mathbb{P} = \{\theta\}$ we actually can calculate the velocity of the system motion in it. Indeed, the time derivatives entering the complete collection of mutually independent components for all the particles

$$\left\{ \left\{ x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}} \right\}_\alpha \right\}$$

may be treated as independent phase variables

$$\varphi_{\alpha,p} \stackrel{\text{def}}{=} \frac{dx_\alpha^{p-1}}{dt^{p-1}}, \qquad \text{for } p = 1, 2, \ldots, m,$$

in particular, the variable $\varphi_{\alpha,1} = x_\alpha$ just represents the spacial coordinates of the particle α, $\varphi_{\alpha,2} = v_\alpha$ is the velocity of its motion in the physical space, and $\varphi_{\alpha,3} = a_\alpha$ is its acceleration in it. The mutual independence of these variables is understood in the sense explained above; it is the principle possibility of finding a real instantiation of the system in issue at the state such that at a given moment of time all these derivatives take the corresponding values chosen arbitrary. Then the system dynamics governed by equation (2.13) may be represented as

$$\begin{aligned} \frac{d\varphi_{\alpha,p}}{dt} &= \varphi_{\alpha,p+1} \qquad \text{for } p = 1, 2, \ldots, m-1 \text{ and} \\ \frac{d\varphi_{\alpha,m}}{dt} &= \Phi_\alpha\left(\varphi_1, \varphi_2, \ldots, \varphi_m\right). \end{aligned} \tag{2.17a}$$

Here, as previously, omitting the index α at the arguments of the function $\Phi_\alpha\{\ldots\}$ denotes that its list of arguments should contain the phase variables $\{\varphi_{\alpha,p}\}$ of all the particles. These differential equations which symbolically may be written as

$$\frac{d\theta}{dt} = \Phi(\theta) \tag{2.17b}$$

determine all the laws of the system dynamics.

The existence of equations (2.17) endows the inanimate world in the realm of classic physic with a fundamental property described by two notions reflecting its different aspects. One of them is the notion of *initial conditions*.[4] Namely, if we know the system position θ_0 in the phase space $\mathbb{P}$ at an *arbitrary chosen* moment of time t_0 then, generally speaking, equations (2.17) possess the unique solution for $t > t_0$

$$\theta = \theta(t, \theta_0) \quad \text{such that at } t = t_0 \quad \theta(t_0, \theta_0) = \theta_0\,. \tag{2.18}$$

In other words, if the "forces" $\Phi(\theta)$ and the initial system position θ_0 are known, then the system dynamics can be calculated, at least, in principle. It means that inanimate systems have no memory; if we know what is going on with such a system at a given moment of time, then its "history" does not matter, which was claimed previously appealing to Premise 1.

The other one is the notion of the *determinism* of physical systems; if we repeat the system motion under the same conditions with respect to the initial position θ_0 and the "forces" $\Phi(\theta)$ acting on the system, then the same trajectory of system motion will be reproduced. Drawing this conclusion we actually have assumed implicitly that the "forces" $\Phi(\theta)$ do not depend on the time t. If it is not so, then we can expand the system to include external objects causing the time dependence of these "forces." The feasibility of such an extension is justified by the principle of microscopic level reducibility. In fact it claims that at the microscopic level describing completely a given system there are only structureless constituent particles and the interaction between them. So there no factors that can cause the time dependence of the "forces" $\Phi(\theta)$ and, in particular, endow them with random properties.

Brief digression: It should be noted that this determinism does not exclude highly complex dynamics of nonlinear physical systems manifesting itself in phenomena usually referred to as dynamical chaos. Dynamical chaos can be observed when the motion of a system in its phase space is confined to a certain bounded domain and the motion trajectories are unstable with respect to small perturbations. This instability means that two trajectories of such a system initially going in close proximity to each other diverge substantially as time goes on, and, finally, the initial proximity of the two trajectories becomes unrecognizable. These effects make the dynamics of such systems *practically* unpredictable. For example, in

[4] Actually the range of applicability of notion of initial conditions is much wider than Newtonian mechanics, which however is beyond the scope of our discussion.

numerical solution of equations (2.17) the discretization of continuous functions and round-off errors play the role of disturbing factors responsible for a significant dependence of the found solutions on the selected time step in discretization and particular details of arithmetic operations at a used computer. In studying systems with dynamical chaos in laboratory experiments the presence of weak uncontrollable factors is also inevitable. Moreover, there is a reason arguing for the fact that the notion of dynamical chaos is a fundamental problem rather than a particular question about practical implementations of system dynamics. The determinism of physical systems implies the reproducibility of their motion trajectories provided the *same* initial conditions are reproduced each time. However, in trying to control extremely small variations in the system phase variables we can face up to effects lying beyond the range of applicability of classical physic. So in studding various instantiations of one system it can be necessary to assume that each time the initial conditions are not set equal but distributed randomly inside a certain, maybe, very small domain. So determinism and dynamical chaos are not contradictory but complementary concepts reflecting different aspects of the dynamics of physical systems in the realm of classical physics.

2.4 Energy Conservation and Newton's Second Law

Appealing to the concepts of thick presentism it is not possible to find out the order $m-1$ of the derivative $d^{m-1}x/dt^{m-1}$ that determines how many components collection (2.14) contains, i.e. to specify the structure of the phase space $\mathbb{P}$ (2.15). From physics we know that $m=2$, i.e., for any ensemble of classical particles the phase space consists of the spatial coordinates and velocities of the particles making up it. Let us try to elucidate whether this type phase space endows the corresponding systems with unique properties via which such systems stand out against the other objects.

In the simplest case, i.e., when the value $m=1$, the phase space contains only the spatial positions of particles $\mathbb{P}^1=\{x\}$. In this instance the respective systems tend to go directly to spacial "stationary" points x_{eq} such that

$$\frac{dx_\alpha}{dt}=\Phi_{\alpha,1}\left(x_{\mathrm{eq}}\right)=0 \quad \text{for all } \alpha,$$

if, naturally, they are stable. This class of models, broadly speaking, is the heart of Aristotelian physics assuming, in particular, that for a body to move some force should act on it. There are many examples of real physical objects exhibiting complex behaviour that are *effectively* described using the notions inherited from Aristotelian physic. The complexity of their dynamics is due to the fact that all their stationary points turn out to be unstable and, instead, some complex attractors, i.e., multitudes toward which systems tend to evolve, arise in the phase space $\mathbb{P}^1$. Nevertheless, if our inanimate world were governed solely by Aristotelian physics it would be rather poor in properties. For example, if the motion of planets of a solar system obeyed such laws then they would drop to its sun and the galaxies could not form.

The next case with respect to the simplicity of phase spaces matches $m = 2$. It is our world; the phase space of physical particles, at least, within Newtonian mechanics consists of their spatial coordinates and velocities,

$$\mathbb{P}^2 = \left\{ x, v = \frac{dx}{dt} \right\}, \tag{2.19}$$

which together determine the next order time derivative, the particle acceleration,

$$a_\alpha = \frac{d^2 x_\alpha}{dt^2} = \Phi_{\alpha,2}\,(x, v)\,. \tag{2.20}$$

In other words, in Newtonian physics for a body to *accelerate* some forces should act on it, whereas in Aristotelian physics for a body to *move* some forces should act on it. The systems whose dynamics is described by the phase space $\mathbb{P}^2$ possess two distinctive features.

One of them, usually called the dynamics reversibility, is exhibited by systems where the regular "force" $\Phi_{\alpha,2}\,(x)$ depends only on the particle positions $\{x\}$. In this case the governing equation (2.20) is symmetrical with respect to changing the time flow direction, i.e. the replacement $t \to -t$. This symmetry is responsible for the fact that if at the end of motion the velocities of all the particles are inverted, $v_\alpha \to -v_\alpha$, then they should move back along the same trajectories.

The other one is the possibility of introducing the notion of energy for the real physical systems. At the microscopic level the energy, comprising the components of the kinetic and potential energy, is a certain function

$$\mathcal{H}(x, v) \tag{2.21}$$

whose value does not change during the system motion. Namely, if $x(t)$ is a trajectory of system motion then the formal function on t

$$H(t) \stackrel{\text{def}}{=} \mathcal{H}\left[x(t), v(t) = \frac{dx(t)}{dt} \right] \tag{2.22}$$

in fact does not depend on the time t. Such systems are called conservative. The existence of the energy $\mathcal{H}(x, v)$ does not necessary stem from the governing equation (2.20) but is actually an additional assumption about the basic properties of physical systems at the microscopic level. Naturally, it imposes some conditions on the possible forms of the function $\Phi_{\alpha,2}(x, v)$.

The two features endow physical systems with rich properties and complex behavior. For example, although in a solar system the planets are attracted by the sun, they do not drop on it because when a planet comes closer to the sun its kinetic energy grows, preventing the direct fall on the sun. Naturally, this planet should not move initially along a straight line passing exactly through the sun. The reversibility is responsible for this planet to tend to return to the initial state or its analogy

after passing the point at the planet trajectory located at the shortest distance to the sun. Broadly speaking, the existence of energy endows physical objects with a certain analogy of memory. Certainly, if the initial conditions for a given system are known, its further dynamics is determined completely, at least, in principle, so the previous system history does not matter. Nevertheless, the conservative systems "do not forget" their initial states in the meaning that the motion trajectories matching different values of the energy cannot be mixed.[5]

Summarizing this discussion about the systems with the phase space $\mathbb{P}^2$ we may claim that it is the *simplest situation* when the corresponding physical world is reach in properties.

As far as systems with a phase space containing time derivatives of higher orders are concerned, they seem not to admit the introduction of the energy at all in a self-consistent way within the standard interpretation. In order to explain this fact we reproduce the construction of the governing equations for such systems of particles using Lagrangian formulation of Newtonian mechanics based on the principle of least actions. It is worthy of noting that in some sense Lagrangian formulation of mechanics is more general than its formulation directly appealing to Newton's laws. Indeed in the latter case the existence of energy is an additional assumption imposing certain conditions on the forces with which physical particles interact with one another. In Lagrangian formulation the existence of some function, the Lagrangian L, reduced then to the system energy is the pivot point and the derived equation governing the system dynamics originally contain the forces meeting the required conditions.

As the general case, let us consider a system with the phase space

$$\mathbb{P}^m = \{\theta_m\} = \left\{x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}}\right\},$$

where m is a certain number not necessary equal to 2. The pivot point of Lagrangian formalism is the introduction of a certain functional $\mathcal{L}\{x(t)\}$ determined for any arbitrary trajectory $\{x(t)\}_{t=t_s}^{t=t_e}$ starting and ending at some time moments $t = t_s$ and $t = t_e$, respectively. The notion of functional means that for any given trajectory $\{x(t)\}$ we can calculate a certain number $\mathcal{L}\{x(t)\}$ which is treated as a measure of its "quality" in the realm of the Lagrangian mechanics. Since the systems at hand do not possess memory and cannot predict their future, all their significant

[5]First, it should be noted that a many-particle ensemble can exhibit so complex dynamics that it could be impossible to track its motion from a given initial state within physically achievable accuracy. In this case it possible to speak about the effective forgetting of the initial conditions. The latter also concerns extremely weak perturbations. Second, there are systems with highly complex dynamics whose description does not admit any energy conservation and their motion is irreversible; the term dynamical chaos noted before is usually used to refer to these phenomena. Nevertheless it does not contradict to the present argumentation because the corresponding irreversible description is obtained via the reduction of equation (2.20) and assuming the presence of a certain external environment weakly interacting with a system at hand.

characteristics including the "quality" of motion have to be determined by the local properties of the trajectory $\{x(t)\}$. In the given case, it is the collection θ_n of the phase variables x, dx/dt, $\ldots$, $d^{m-1}x/dt^{m-1}$. Therefore the functional $\mathcal{L}\{x(t)\}$ has to be of the integral form

$$\mathcal{L}\{x(t)\} = \int_{t_s}^{t_e} L\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{m-1}x}{dt^{m-1}}\right) dt\,, \tag{2.23}$$

where $L(\ldots)$ is some function of these n phase variables. The principle of least actions implies that "the Nature chooses the best trajectories to implement" the dynamics of mechanical systems. In other words, a real trajectory of system motion matches the minimum of functional (2.23) (or its maximum within the replacement $\mathcal{L} \to -\mathcal{L}$) with respect to all possible variations near this real trajectory (Fig. 2.6).

It is worthy of noting that in spite of its long-term history the fundamentality of the principle of least actions is up to now a challenging problem and there are a number of arguments for and against it from various points of view. Their brief review can be found, e.g., in Stöltzner (2006) as well as a detailed analysis of its ontological roots has be given by Stöltzner (2003, 2009), Katzav (2004), Smart and Thébault (2015), and Terekhovich (2015). Nevertheless its high efficient in many different branches of physics strongly argues for its real fundamentality. In Sect. 2.2.4 I have demonstrated that this principle can be derived based on the concepts of thick presentism for systems whose phase space contains hight order time derivatives as individual phase variables.

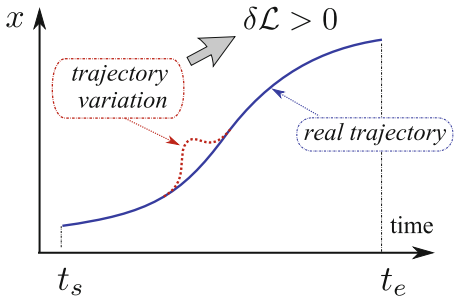

Fig. 2.6 Illustration of the principle of least actions describing the minimality of the functional $\mathcal{L}\{x(t)\}$ (2.23) taken at the real trajectory $\{x(t)\}$ with respect to its variations

In the case of Newtonian mechanics, i.e., for systems with the phase space $\mathbb{P}^2$ the principle of least actions leads to the governing equations of type (2.17) admitting the introduction of the energy given by the expression

$$\mathcal{H}(x, v) = v\frac{\partial L(x, v)}{\partial v} - L(x, v)\,. \tag{2.24}$$

In the general case (for $m > 2$) it is also possible to construct some function having properties similar to that of energy,[6] however we meet the following challenging problem.

Let $\{x(t)\}$ be a motion trajectory at which functional (2.23) attains its minimum (or maximum). Then for all small perturbations $x(t) + \delta x(t)$ of this trajectory similar to one shown in Fig. 2.6 the variation $\delta\mathcal{L} = \mathcal{L}\{x(t) + \delta x(t)\} - \mathcal{L}\{x(t)\}$ of this functional has to be equal to zero in the linear approximation in $\delta x(t)$. It reads

$$\delta\mathcal{L} = \int_{t_s}^{t_e} \left\{ \frac{\partial L}{\partial x}\cdot\delta x(t) + \frac{\partial L}{\partial x^{(1)}}\cdot\delta\left[\frac{dx(t)}{dt}\right] + \frac{\partial L}{\partial x^{(2)}}\cdot\delta\left[\frac{dx^2(t)}{dt^2}\right] + \ldots \right.$$
$$\left. + \frac{\partial L}{\partial x^{(m-1)}}\cdot\delta\left[\frac{dx^{m-1}(t)}{dt^{m-1}}\right] \right\} dt = 0\,, \tag{2.25}$$

where the symbol $x^{(p)}$ (with $p = 1, 2, \ldots m-1$) denotes the corresponding time derivative, $x^{(p)} = d^p x(t)/dt^p$, treated as the argument of the function $L(x, x^{(1)}, x^{(2)}, \ldots, x^{(m-1)})$. Using the identities

$$\delta\left[\frac{dx^p(t)}{dt^p}\right] = \frac{d^p[\delta x(t)]}{dt^p} = \frac{d}{dt}\left[\frac{d^{p-1}[\delta x(t)]}{dt^{p-1}}\right] \quad \text{(in the latter case } p > 1),$$

the rule of integration by parts

$$\int_{t_s}^{t_e} U(t)\frac{dV(t)}{dt}\,dt = \Big[U(t)V(t)\Big]_{t=t_s}^{t=t_e} - \int_{t_s}^{t_e} V(t)\frac{dU(t)}{dt}\,dt\,,$$

and choosing the trajectory perturbations $\delta x(t)$ such that (it is our right)

$$\delta x(t)|_{t=t_s,t_p} = 0\,, \qquad \left.\frac{d^p[\delta x(t)]}{dt^p}\right|_{t=t_s,t_p} = 0 \quad (\text{for } p = 1, 2, \ldots, m-1)$$

[6] An example of how to construct an "energy," i.e., Hamiltonian for systems with the phase space $\mathbb{P}^3 = \{x, v, a\}$ has been demonstrated, e.g., by Lubashevsky et al. (2003b).

equality (2.25) is reduced to

$$\delta\mathcal{L} = \int_{t_s}^{t_e} \left\{ \frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial x^{(1)}} + \left(\frac{d}{dt}\right)^2 \frac{\partial L}{\partial x^{(2)}} - \ldots \right.$$
$$\left. + (-1)^{m-1} \left(\frac{d}{dt}\right)^{m-1} \frac{\partial L}{\partial x^{(m-1)}} \right\} \delta x(t)\, dt = 0\,. \qquad (2.26)$$

Because equality (2.26) must hold for any particular perturbation of the trajectory $\{x(t)\}$ this trajectory has to obey the equation

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial x^{(1)}} + \left(\frac{d}{dt}\right)^2 \frac{\partial L}{\partial x^{(2)}} - \ldots + (-1)^{m-1} \left(\frac{d}{dt}\right)^{m-1} \frac{\partial L}{\partial x^{(m-1)}} = 0\,. \qquad (2.27)$$

The time derivative of the highest order contained in equation (2.27) is $d^{2(m-1)}x/dt^{2(m-1)}$; it enters this equation via the last term as the item

$$(-1)^{m-1} \frac{\partial^2 L}{\partial [x^{(m-1)}]^2} \cdot \frac{d^{2(m-1)}x}{dt^{2(m-1)}}\,.$$

When the derivative $\partial^2 L/\partial[x^{(m-1)}]^2$ is not equal to zero,[7] i.e., the Lagrangian $L(\ldots)$ is not a linear function with respect to its argument $d^{m-1}x/dt^{m-1}$, equation (2.27) can be directly resolved with respect to the derivative

$$\frac{d^{2(m-1)}x}{dt^{2(m-1)}}$$

and rewritten as

$$\frac{d^{2(m-1)}x}{dt^{2(m-1)}} = \Phi^*\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^{2m-3}x}{dt^{2m-3}}\right), \qquad (2.28)$$

where $\Phi^*(\ldots)$ is a certain function. Equation (2.28) governs the dynamics of the system in issue and can be regarded as its basic law written in the form of differential equation.[8]

[7] In the case of Newtonian mechanics with the phase space $\mathbb{P}^2$ the corresponding term is just the mass m of a given particle, $\partial^2 L/\partial[x^{(2)}]^2 = m$.

[8] Lagrangian mechanics with higher order time derivatives is well known and was developed during the middle of the nineteenth century by Ostrogradski (1850). So here I have presented the results in a rather symbolic form emphasizing the features essential for our consideration. Mechanics dealing with equation (2.28) with $m > 2$ as one describing some initial value problem faces up to the Ostrogradski's instability (see, e.g., Woodard 2007; Stephen 2008; Smilga 2009), which can be used for explaining why no differential equations of higher order than two appear

When the number of the phase variables forming the phase space $\mathbb{P}^m$ is larger than two, $m > 2$, the obtained governing equation (2.28) comes in conflict with the initial assumption about the properties of the given physical system. The matter is that the number of the arguments of the function $\Phi^*(\ldots)$ *exceeds* the dimension of the phase space $\mathbb{P}^m$ because $2(m-1) > m$ for $m > 2$. Thereby we cannot treat equation (2.28) as a law governing the deterministic motion of a certain dynamical system in the phase space $\mathbb{P}^m$. Indeed, to do this we need that the point $\theta_m = \{x, dx/dt, \ldots, d^{m-1}x/dt^{m-1}\}$ of the phase space $\mathbb{P}^m$ determine completely the rate of the system motion in it, in other words, the corresponding governing equation should be of form $d\theta_m/dt = \Phi(\theta_m)$ (see equation (2.17)). However, the obtained equation (2.28) stemming from the principle of least actions contradicts this requirement because its right hand side contains the time derivatives higher than $d^{(m-1)}x/dt^{(m-1)}$ for $m > 2$. Therefore in order to construct a solution of equation (2.28) dealing with only the phase space $\mathbb{P}^m$ we need some additional information about, for example, the terminal point of the analyzed trajectory. The latter feature, however, contradicts the principle of microscopic level reducibility because according this principle the current state of such a system should determine its further motion completely.

Summarizing this discussion we see that only in the case of $m = 2$ equation (2.28) following from the principle of least actions for trajectories in the phase space $\mathbb{P}^m$ admits the interpretation in terms of a certain dynamical system whose motion is *completely* specified within this phase space. If $m = 1$ the minimality of functional (2.23) does not describe any dynamics. Therefore, it is likely that the notion of the energy $\mathcal{H}(x, v)$ can be introduced in a self-consistent way only for systems with the phase space $\mathbb{P}^2 = \{x, v\}$ where the governing laws can be written as differential equations of the second order.

2.5 Probability Theory and Newtonian Mechanics

When the dynamics of a certain system becomes hardly predictable for whatever reasons the concept of probability is widely used to cope with this situation. Although the probability theory and the theory of dynamical systems dealing with equations of type (2.17) may be treated as individual and partly independent languages, the probability formalism inherits from Newtonian mechanics its fundamentals, including the notion of phase space. In the vocabulary of probability theory it is the sample space, $\mathfrak{S} = \{e\}$, i.e., the set of points $\{e\}$ specifying *all* the possible states of a given system. The sample space is just another notation for the set of points of the phase space at the level of microscopic description. In particular,

to describe physical phenomena (Motohashi and Suyama 2015). There are also arguments for the latter conclusion appealing to metaphysical aspects of time changes in physical quantities (Easwaran 2014).

leaving aside aspects of mathematical rigor, we may say that at any moment of time t the corresponding system can be found in one and only one state e.

The original element of this theory is the notion of the probability $p_e(t)$ of finding a given system at a state e at the current moment of time t. Formally it is a certain nonnegative value $p_e(t)$ ascribed to each elementary event e such that the equality

$$\sum_{e\in\mathfrak{S}} p_e(t) = 1$$

holds. The value $p_e(t)$ admits the following interpretation. Let us consider an ensemble of N identical copies of the given system which do not interact with one another. If the number of these copies is sufficiently large, $N \gg 1$, then the number of systems that currently occupy a state e is $N_e(t) = p_e(t)N$. In these terms the statement that the system at hand is known to occupy currently a certain state e means $p_e(t) = 1$ and $p_{e'}(t) = 0$ for $e' \neq e$.

Compound event is another fundamental notion of the probability theory. It corresponds to the concept of mesoscopic states to be discussed in Sect. 2.7 in detail. As the scale of description increases there could occur a situation when some elementary states making up a family $\mathfrak{E} = \{e\}$ become indistinguishable at a certain level of consideration. It allows us to treat such families as the corresponding mesoscopic entities and ascribe to them the probability

$$P(\mathfrak{E}, t) = \sum_{e\in\mathfrak{E}} p_e(t)\,. \tag{2.29}$$

Probabilistic properties of these families that stem from the set algebra may be rather intricate because the families $\{\mathfrak{E}\}$ can overlap with one another. Nevertheless, dealing with these objects we can return to the microscopic level and construct required relationships. In any case a possible formal construction of the probabilistic properties for a given system starting directly from its mesoscopic level must not contradict to the general relationships originating from the microscopic level.

A trajectory of motion of a probabilistic system in the space $\mathfrak{S} = \{e\}$ is a sequence of its "jumps" $\mathbb{P}_e(t, t_0) = \{e(t')\}_{t'=t_0}^{t'=t}$ between different points in the space $\mathfrak{S}$ during the time interval $t' \in [t_0, t]$. To say something meaningful about such trajectories it is assumed, maybe tacitly, that for some reasons the history of this system up to the moment t_0 does not matter and, in addition, its state $e_0|_{t'=t_0}$ is given. To accept this assumption for systems with memory the initial time moment t_0 should be placed in the distant past, which is equivalent to employing the formal limit $t_0 \to -\infty$. Since in the realm of probability any trajectory is possible, at least in principle, to quantify their "realizability" the probability of a given trajectory as

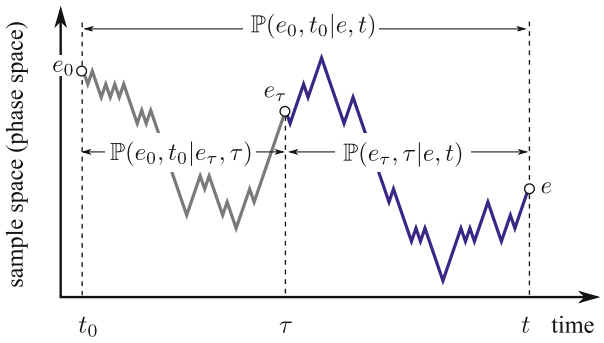

Fig. 2.7 Partition of a random trajectory $\mathbb{P}(e_0, t_0|e, t)$ in the sample space (phase space) into two parts $\mathbb{P}(e_0, t_0|e_\tau, \tau)$ and $\mathbb{P}(e_\tau, \tau|e, t)$

a whole is introduced[9]

$$\mathcal{P}\left[\mathbb{P}_e(t_0, t)\right]$$

and the trajectories $\{\mathbb{P}_e(t_0, t)\}$ are regarded as the basic elementary entities that can be used in constructing all the other probabilistic characteristics.

In the general case the introduction of the trajectories as elementary entities brings the probabilistic theory out the scope of the Newtonian mechanics paradigm; it is illustrated in Fig. 2.7. Indeed, let us divide a certain trajectory $\mathbb{P}(e_0, t_0|e, t)$ into two parts $\mathbb{P}(e_0, t_0|e_\tau \tau)$ and $\mathbb{P}(e_\tau, \tau|e, t)$ fixing some time moment τ between the start time t_0 and the end time t, i.e. $t_0 < \tau < t$. For the first part $\mathbb{P}(e_0, t_0|e_\tau \tau)$ the second part $\mathbb{P}(e_\tau, \tau|e, t)$ is the future system motion, so its probability weight

$$\mathcal{P}_1\left[\mathbb{P}(e_0, t_0|e_\tau, \tau)\right]$$

cannot depend on the fragment $\mathbb{P}(e_\tau, \tau|e, t)$. For the second fragment the situation is just opposite, the fragment $\mathbb{P}(e_\tau, \tau|e, t)$ is its history and if the system is able to change the environment the details of the system motion during the time interval (t_0, τ) become the parameters of this environment. So in the general case the probability

$$\mathcal{P}_2\Big[\mathbb{P}(e_\tau, t_\tau|e, t)\Big/\mathbb{P}(e_0, t_0|e_\tau, \tau)\Big]$$

of the second fragment as an individual trajectory has to depend on the first one. As a result just counting the total number of trajectory realizations we may write the

[9] As well known, in the classical probability theory it is not possible to ascribe any probabilistic measure to a continuous random trajectory. However the formalism of nonstandard analysis enables to do this, Albeverio et al. (1986).

probability of the whole trajectory $\mathbb{P}(e_0, t_0|e, t)$ as the product of the probabilities of its parts

$$\mathcal{P}\left[\mathbb{P}(e_0, t_0|e, t)\right] = \mathcal{P}_1\left[\mathbb{P}(e_0, t_0|e_\tau, \tau)\right] \times \mathcal{P}_2\Big[\mathbb{P}(e_\tau, t_\tau|e, t)\Big/\mathbb{P}(e_0, t_0|e_\tau, \tau)\Big]. \tag{2.30}$$

Therefore we cannot choose any arbitrary time moment τ to specify the initial conditions for such a stochastic system because the properties of its motion after the fixed time moment τ should depend on the details of its previous motion. However, the proposition that a physical system does not change the environment during its motion, at least, within the framework of the microscopic level of description is one of the fundamental hypotheses maybe implicitly adopted in physics. If not so a component of the environment affected by the given system should be included into it. This assumption is taken into account within the concept of Markov stochastic processes which is no more than making this proposition explicit. It confines the analyzed objects to stochastic systems having no memory at the microscopic level. In the given example of trajectory division into two fragments the Markov concept supposes the probabilistic properties of the system motion represented by the second fragment to be independent of its history,[10] i.e.,

$$\mathcal{P}_2\Big[\mathbb{P}(e_\tau, t_\tau|e, t)\Big/\mathbb{P}(e_0, t_0|e_\tau, \tau)\Big] = \mathcal{P}_2\Big[\mathbb{P}(e_\tau, t_\tau|e, t)\Big]$$

The adoption of the Markov approximation returns the stochastic description to the paradigm of Newtonian mechanics. It becomes more evident when the stochasticity is introduced via a certain random "force" $\Phi_L(\theta, t)$ acting on the system such that equation (2.17) reads

$$\frac{d\theta}{dt} = \Phi_r(\theta) + \Phi_L(\theta, t)\,. \tag{2.31}$$

Here $\Phi_r(\theta)$ is a regular "force," meaning that the value $\Phi_r(\theta)$ is determined at any point θ of the phase space, whereas only the probabilistic characteristics of the random "force" $\Phi_L(\theta, t)$ called also the Langevin source are regarded as given. Usually the Langevin source is represented as the product of some function $G(\theta)$ (or vector function) and stochastic process $\xi(t)$ with known characteristics independent of the dynamics of the system at hand,

$$\Phi_L(\theta, t) = G(\theta) \cdot \xi(t)$$

[10] It should be noted that subdiffusive processes attracted much attention in physics during the last decades exemplify stochastic systems with essential memory. However this example does not contradict our statement about the absence of memory of physical objects at the *microscopic* level. The memory of subdiffusive process appears only at a mesoscopic level of description after reducing the complete phase space.

and the probability of realization is ascribed directly to the time pattern $\{\xi(t)\}$. In this way solving equation (2.31) for a given time pattern $\{\xi(t)\}$ we are able to construct the probability of the resulting trajectory $\{\theta(t)\}$ of the system motion

$$\mathcal{P}\{\xi(t)\} \Rightarrow \mathcal{P}\{\theta(t)\}\,.$$

Summarizing this section it is worthwhile to note once more that the existence of microscopic description for stochastic systems is a fundamental axiom of the probability theory. It is responsible for the deep relationship between the basic concepts of the probability theory and the paradigm of Newtonian mechanics. The possibility of starting consideration of a system at hand from the corresponding space $\mathfrak{S}$ of elementary states, on one side, enables us to analyze rather efficiently complex behavior of stochastic systems. On the other side, it impose certain frames on their possible properties. For example, the probability of a compound event has to be the sum of the probabilities of the elementary events forming this compound event (see Exp. (2.29)).

2.6 Superposition Principle of Particle Interaction

In the previous sections we considered inanimate systems as a whole not dividing them into parts. Generally speaking, dealing with many element ensembles we face up to a challenging problem. What should we do if a new element is added to a given ensemble; is it necessary to modify the previous model essentially or is there a general description where the number of element is arbitrary? In physics there is a simple recipe. One should describe the individual interaction of a chosen pair of particles ignoring all the other particles and sum up the results for all the possible pairs. This recipe is called the principle of superposition of particle interaction. It poses the question of whether the principle of interaction superposition is applicable to social systems. However before trying to answer this question it could be useful to understand the fundamental mechanisms by which the interaction superposition arises in the inanimate wold. Therefore in this section we return to the fundamentals of Newtonian mechanics (see Sect. 2.4) and discuss the principle of microscopic level reducibility with respect to its part concerning the particle interaction.

At first, we consider two particles α_1 and α_2 interacting only with each other. In this case, according to (2.20) the equation governing, for example, the dynamics of the first particle can be written in the general form as

$$m_{\alpha_1}\frac{d^2x_{\alpha_1}}{dt^2} = \Phi_{\alpha_1\alpha_2}\left(x_{\alpha_1}, v_{\alpha_1}|x_{\alpha_2}, v_{\alpha_2}\right), \tag{2.32}$$

where $v_\alpha = dx_\alpha/dt$ is the velocity of the particle α and the function $\Phi_{\alpha_1\alpha_2}(\ldots)$ is the force with which the second particle acts on the first one. The multiplier m_α called the mass of the particle α has been introduced to enable the aggregation

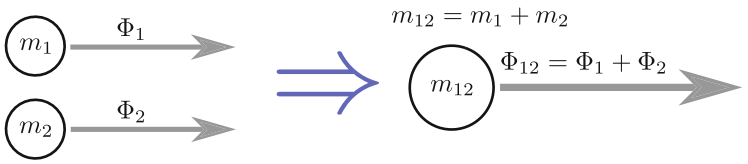

Fig. 2.8 Illustration of the aggregation of two particle moving within the same acceleration into one composed particle with an increased mass

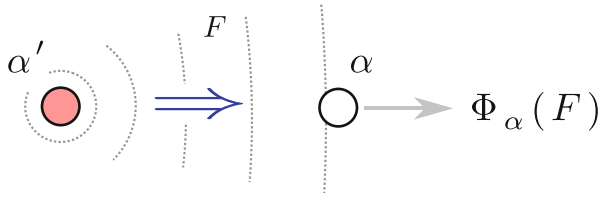

Fig. 2.9 Illustration of the interaction of particle via a field F generated by them

of two particles with the same dynamics, i.e., moving with the same acceleration $a = d^2x/dt^2$, into one composed particle, which is illustrated in Fig. 2.8 and must be feasible due to the superposition of interaction accepted by the principle of microscopic level reducibility.

In the case of an ensemble at hand consisting of more than two particle the principle of microscopic level reducibility allows us to write the cumulative force acting on a given particle α as the sum of all the forces with which each particle *individually* acts on the given one, i.e., if all the other particles except for the chosen pair would be absent,

$$\Phi_\alpha = \sum_{\alpha':\alpha' \neq \alpha} \Phi_{\alpha\alpha'}\left(x_\alpha, v_\alpha | x_{\alpha'}, v_{\alpha'}\right). \tag{2.33}$$

It is a very important assumption property, otherwise, the great progress in physics during the last three centuries would be impossible. Exactly the principle of superposition has allowed to study the dynamics of complex systems using the knowledge gained in the analysis of individual interaction of their constituents on their own.

Let us try to understand why it is so. To do this we need to make use another paradigm of particle interaction. The notion of forces belongs to the concept of nonlocal interaction of particles separated in space. In fact, the introduction of the force $\Phi_{\alpha\alpha'}\left(x_\alpha, v_\alpha | x_{\alpha'}, v_{\alpha'}\right)$ tacitly implies that the particle α' acts on the particle α immediately at the same moment of time when it appears at the point $x_{\alpha'}$.

An alternative to the nonlocal interaction paradigm is the concept of field; the electro-magnetic field exemplifies this concept. A particle *locally* generates a field F, this field propagates in the space, and getting another particle affects it (Fig. 2.9).

In a rather general form the dynamics of a field $F(x,t)$, i.e., its generation and propagation is governed by an equation of the following form

$$\underbrace{\widehat{\mathcal{D}}\{F\}}_{\text{field dynamics}} \quad = \quad \underbrace{\sum_{\alpha} \mathfrak{q}\big[x_\alpha, v_\alpha, F(x_\alpha,t)\big]}_{\text{field generation}}. \tag{2.34}$$

Here the left-hand side, some generally nonlinear operator $\widehat{\mathcal{D}}$ acting on the field F, describes the dynamics of the field F on its own or, in other words, local variations a given point as well as propagation in space caused by its own properties. The right hand side is the field generation by the ensemble of particles $\{\alpha\}$; naturally in the general case the generation rate $\mathfrak{q}\big[x_\alpha, v_\alpha, F(x_\alpha,t)\big]$ or, in other words, the source of the field F can depend on the local value $F(x_\alpha,t)$ of the field F. The description of the field dynamics in the form of this equation has a certain methodological advantage. This approach makes it clear that the field generation and field propagation are different phenomena deserving individual analysis.

Brief digression: The combination of the premise that only the present matters to physics and the concept of local interaction between particles and fields whose propagation through space is limited by some maximally possible speed c enables us also to select some plausible class of models for specifying the term $\widehat{\mathcal{D}}\{F\}$ in equation (2.34). This term describes the dynamics of the field F on its own, i.e., in the case when no particles are present. The accepted locality of the field dynamics requires that the law governing the field generation and propagation admit a representation in the form of equation (2.34), where both the sides are some functions determined at each spacial point x individually. As far as the left-hand side is concerned, it means that the term $\widehat{\mathcal{D}}\{F\}(x,t)$ maps in a certain way the *spatial distribution* of the field $\{F(x,t)\}$ as whole onto some value determined at each point x explicitly.

To draw a conclusion about plausible forms of $\widehat{\mathcal{D}}\{F\}(x,t)$ let us appeal actually to the same argument that has been already used in Sect. 2.3 in reasoning why the laws of Newtonian mechanics should deal with time derivatives. Because only time moments belonging to an infinitesimally narrow *temporal* neighborhood of the present time t can contribute to the system properties, only the current values of the field intensity $F(t)$ and its time derivatives $\partial F/\partial t$, $\partial^2 F/\partial t^2$, ... may enter the term $\widehat{\mathcal{D}}\{F\}$. This argument may be also addressed to the spatial properties of the field F. Namely, the speed limitation in the field propagation through space leads to the requirement that only points belonging to some infinitesimally narrow *spatial* neighborhood of a given point x can contribute to the field properties at this point. During time interval $\Delta \to 0$ the field components at points distant from x by spatial scales $\ell \gg c\Delta$ just have no chance to get the point x. So only the quantities

$$F,\ \frac{\partial F}{\partial t},\ \frac{\partial^2 F}{\partial t^2},\ \dots$$

and their spatial derivatives taken at the given point x may enter the term $\widehat{\mathcal{D}}\{F\}(x,t)$. Generally the collection of these field quantities can be written as

$$\left\{\left(\frac{\partial}{\partial t}\right)^{n_t}\left(\frac{\partial}{\partial x}\right)^{n_x} F(x,t)\right\} \qquad \text{for } n_t,\ n_s = 0,1,2,\dots,$$

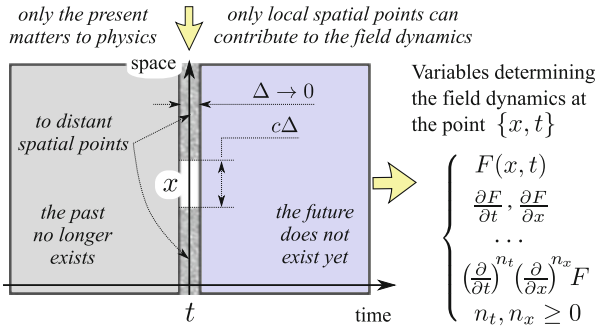

Fig. 2.10 Illustration of the basic factors determining the variables that are required for describing the field dynamics withing the concept of local field-particle interaction

where the zero-th derivative of a function denotes this function itself. So the left-hand side of equation (2.34) has to be some function

$$\widehat{\mathcal{D}}\{F\} = \Phi^F\left(F, \frac{\partial F}{\partial t}, \frac{\partial F}{\partial x}, \frac{\partial^2 F}{\partial t^2}, \frac{\partial^2 F}{\partial x^2}, \frac{\partial^2 F}{\partial t \partial x}, \ldots\right) \tag{2.35}$$

of the field F itself and its various spatio-temporal derivatives taken at a given spatial point x and time moment t. This feature is illustrated in Fig. 2.10.

To specify the form of the function $\Phi^F(\ldots)$ additional grounds are necessary. For example, keeping in mind the results of Sect. 2.4, I may hypothesize that for the notion of field energy can be introduced, the governing equation (2.34) has to be of the second order with respect to time derivative of the field F. It is the case when at leading order function (2.35) contains the second order time derivative of the field F. Besides, it seems that for the field F to be able to propagate freely through space, function (2.35) at leading order has to contain the second order spatial derivative of the field F, for example,

$$\widehat{\mathcal{D}}\{F\} = \frac{\partial^2 F}{\partial t^2} - c^2 \frac{\partial^2 F}{\partial x^2}\,.$$

As another alternative, the field F could have two components F_1, F_2 such that the governing equation (2.34) turns out to be of the first order with respect to their individual derivatives, but the interaction between these fields endows the system with properties required for the introduction of energy and allowing the field F_1, F_2 to propagate freely through space. The alternative is exemplified by the properties of electromagnetic field. However, a more detailed analysis of the plausible forms of the governing equation (2.34) is beyond the scope of the present section.

The development of the field concept had a strong impact on our insight into the inanimate world even in the framework of classical physics. From this point of view any physical system is an ensemble of particles, which *individually* do not interact with one another, and some fields generated by these particles and, in turn, affecting the particle motion. There is a fundamental distinction between the particles and field, at least, in classical physics. A particle has a "fixed" form so its dynamics can be regarded as spatial movement of a certain object without any change of its internal properties. In contrast, a field does not have any predetermined form and its

distribution in space, first, can vary in time and, second, reflects the dynamics of all the particles as well as the properties of the environment.

Let us return to equation (2.34) and consider it from the standpoint of classical physics. Classical physics deals with "weak" fields which cannot affect the properties of vacuum. It prompts us to use an approximation of equation (2.34) remaining only the terms belonging to the leading order in the field F. First, within this approach the left-hand side should be a linear operator acting on the field F in order for the field F be able to propagate in space. Second, on the right-hand side the field sources $\mathfrak{q}\big[x_\alpha, v_\alpha\big]$ may be considered not to dependent on the field itself. As a result the governing equation for the field dynamics becomes

$$\underbrace{\widehat{\mathcal{D}}F}_{\text{field dynamics}} = \underbrace{\sum_\alpha \mathfrak{q}\big[x_\alpha, v_\alpha\big]}_{\text{field generation}}. \tag{2.36}$$

This model leads us immediately to the superposition principle confining the microscopic description of physical systems to the pairwise interaction between the particles. In fact, if $F_\alpha(x,t)$ is the field generated individually by the particle α, i.e., obeying the equation

$$\widehat{\mathcal{D}}F_\alpha = \mathfrak{q}\big[x_\alpha, v_\alpha\big], \tag{2.37}$$

then the resulting field F generated by the cumulative action of all the particles is just a sum of all their individual components

$$F(x,t) = \sum_\alpha F_\alpha(x,t). \tag{2.38}$$

In other words, the superposition principle and the related concept of pairwise interaction accepted within the principle of microscopic level reducibility stem from the linearity of field properties assumed to hold in classical physics.[11] The field linearity implies that in the framework of the *microscopic description* individual particles generate fields independently of the current field intensity and the presence of the other particles as well as the field propagation in space is a linear process. If it were not so, the principle of microscopic level reducibility would not hold because the field dynamics would be irreducible to the individual properties of the constituent particle and the system dynamics would *directly* depend on the field intensity. In other words, in this case the fields must be treated as some additional constituent entities. As far as social systems are concerned, the applicability of

[11]The laser effect is one of the impressive phenomena demonstrating nonlinear properties of the interaction between electrons and electro-magnetic field. However it originates from quantum interaction of particles and fields and, so, a discussion of such phenomena is beyond our consideration.

the principle of interaction superposition to describing social phenomena will be considered in further chapters.

2.7 Mesoscopic Level of Description and Effective Theories

For complex physical systems comprising many particles similar in properties, usually called many particle ensembles or statistical systems, there are various levels of description. Broadly speaking, they are microscopic, mesoscopic, and macroscopic ones. The microscopic level represents the most detailed description of a given system within which each constituent particle is considered individually, the same concerns the individual interaction between these particles. The macroscopic level deals with this system as a whole or its parts comparable with the system in size. Naturally the necessity of introducing various levels of description as well as their specific features are determined by the key properties of given system a chosen approach has to cope with. For example, elementary particles and quantum fields are the basic entities of the microscopic description of an atom being a composed object, whereas for gases their atoms may be treated as the "elementary" entities of their microscopic description. When we analyze how the activity of our sun affects the climate on Earth the sun has to be treated as a highly complex object whereas in studying the dynamics of galaxies our sun as well as billions of other stars may be regarded as their "elementary" entities.

There is a fundamental problem in modeling properties of many particle ensembles. When the number of constituent particles is huge a detailed analysis of their individual dynamics becomes infeasible and, moreover, the detailed information about the state of one arbitrary chosen particle is highly redundant for us. So, on one hand, we are able to *construct* a formal mathematical model for a given system at its microscopic level employing our knowledge gained in studying individual properties of these particles. On the other hand, the ultimate goal is to understand the macroscopic properties of such ensembles. Thereby we have to find some mathematical formalism that enables us to convert the corresponding microscopic description into a model operating with quantities determined at the macroscopic level. Let us demonstrate it appealing, for example, to air contained in a room of volume $4 \times 4 \times 3\,\mathrm{m}^3$. Under the normal conditions this amount of air involves about $N_{\mathrm{mol}} \sim 10^{27}$ molecules of various gases. So to describe this system at the microscopic level we would need to specify the individual position and velocity of everyone of these N_{mol} molecules. Moreover these molecules move rather randomly and their velocities are distributed within a wide interval, so we cannot infer something interesting from the information about the position and velocity of *one* molecule. Actually that what we do need is the averaged properties of the gas molecules aggregated in quantities like the gas temperature and pressure.

For statistical physical systems with cooperative dynamics the problem of operating with such a tremendous amount of particles is overcame via the introduction of a mesoscopic level of their description. In such a system the main phenomena

and properties are governed by cumulative actions of many constituent particles. In other words, before getting remarkable the individual actions of many particles have to be accumulated, giving rise to a certain self-averaging on temporal and spatial scales exceeding microscopic ones. The scales on which this self-averaging becomes feasible will be referred to as the mesoscopic level of description. It should be noted that these mesoscopic scales could be rather small. Appealing again to the same example of air in a room we can estimate the lower boundary of the mesoscopic spatial scales as the mean free path of molecules, i.e., the average distance traveled by a moving molecule between successive collisions with other molecules. Under the normal conditions it is about 10^{-4} mm; the corresponding time scale is about 10^{-9} s.

At the mesoscopic level the detailed information about a many particle ensemble presented, for example, as its precise position in the complete phase space $\mathbb{P}$ describing the system at the microscopic level is redundant. To get rid of unnecessary fragments of this information there has been developed an approach usually called the quasi-adiabatic approximation having also many applications in other fields of physics. Within this approach the constituent particles are combined into relatively large groups treated as new "mesoscopic" entities. It is illustrated in Fig. 2.11, where the mesoscopic entities are shown as groups of $m \gg 1$ particles of spatial size λ much larger than the mean distance a between the constituent particles, $\lambda \gg a$. The value λ may be regarded as the characteristic spatial scale of the mesoscopic level of description. Then, appealing to some physical reasons or based on rigorous constructions, each of these mesoscopic entities is described as a whole by a few quantities often called order parameters. The order parameters characterize the averaged properties of the states and the relative arrangement of the corresponding particles inside one mesoscopic entity. Figure 2.11 just illustrates this averaging for some property Θ ascribed to the particles individually. Actually the quasi-adiabatic

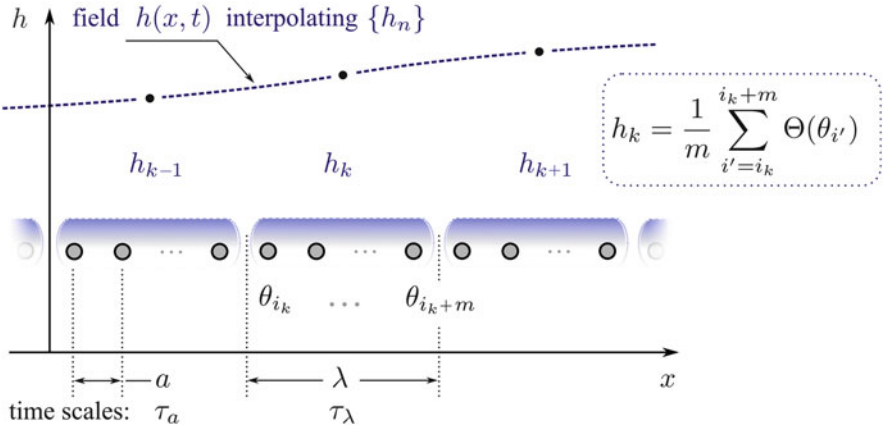

Fig. 2.11 Schematic illustration of constructing a mesoscopic level of description via aggregating the constituent particles of into mesoscopic entities

approximation and the notion of order parameters are based on the following three assumptions.

First, the constituent particles behave in a similar way and their arrangement is locally homogeneous on the mesoscopic scales λ, at least approximately. In other words, any fragment of such a system whose size $\delta x \gtrsim \lambda$ can be considered to be homogeneous and any its particle represents the characteristic properties of the behavior of all the other particles in the given fragment.

Second, a similar situation is the case with respect to time scales. Namely, the dynamics of relatively large mesoscopic entities as whole units may be characterized by time scales τ_λ much larger than the time scales τ_a of the microscopic level, $\tau_\lambda \gg \tau_a$. When the latter inequality holds, which is the second assumption, the order parameters vary in time slowly on the microscopic scales τ_a such that the constituent particles have enough time to attain a local "equilibrium" specified by the current values of the order parameters.

The existence of this "equilibrium" is actually the *third assumption* playing the crucial role in the quasi-adiabatic approximation. This "equilibrium" can be implemented in two possible scenarios. According to the first scenario, the individual dynamics of all the particles forming one mesoscopic entity turns out to be practically synchronized within the given group on scales τ_λ. So, as previously, any particle in this group represents the other particles in their dynamics. By the second scenario, in individual dynamics these particles are mutually independent of one another and their motion may be treated as random on scales about τ_λ. In this case the condition of the particle similarity is reduced to the requirement that the mean characteristics of this "random" motion is the same for all these particles. Then, due to the mesoscopic entities being formed of many particles, we can take into account only these mean characteristics in dealing with individual dynamics of the mesoscopic entities. In the latter scenario any particle belonging to a given mesoscopic entity again represents the dynamics of all its other particles but now after averaging over time scales about τ_λ. In both these cases the microscopic phase variables adiabatically "follow" the slow variations in the mesoscopic order parameters, directly or on the average, which is reflected in the name of this approximation.

When the quasi-adiabatic approximation holds the order parameters $\{h\}$ rather adequately characterize the state of such a system and its dynamics. Therefore, the set $\{h\}$ may be regarded as a certain point of a reduced phase space $\mathbb{P}_h$ being the image of the complete phase space $\mathbb{P}$ after aggregating the phase variables of the microscopic level into the slow variables $\{h\}$. To complete the system description at the mesoscopic level governing equations dealing with the order parameters only, i.e., containing no additional microscopic phase variables have to be constructed. It can be done appealing to some physical reasons or strictly derived from the corresponding microscopic equations. Such models belonging to the mesoscopic level are often called phenomenological or effective theories to underline the fact that their construction is aimed at describing specific phenomena observed on scales exceeding substantially microscopic ones. In philosophy of science these models are usually regarded as epistemological and methodological issues telling us how

to describe a phenomenon at hand and what concepts and formalism can be used to capture its properties. On the contrary, the corresponding microscopic models may be categorized as ontological ones because they are about the nature of the analyzed phenomenon. In Chap. 4, nevertheless, I will try to convince you that phenomenological theories can be as fundamental as microscopic description.

Posing the question about mesoscopic governing equations we touch one of the fundamental problems in modeling complex systems, including physical and social ones, and human behavior as well. It is the question about a possible form of these equations, in particular, whether they can be written in the form of classic differential equations similar to (2.17) or other mathematical notions have to be used. The matter is that the mesoscopic entities are no longer structureless. So the previous arguments leading to equation (2.17) are not strictly applicable in the realm of the mesoscopic level description. Nevertheless, as far as physical objects are concerned with, the mesoscopic governing equations written in the form

$$\frac{dh}{dt} = \Psi(h)\,. \tag{2.39}$$

are widely met in many applications and constructed based on relatively rigorous procedures or, sometimes, just appealing to physical reasons. Exactly in this case the notion of the reduced phase space $\mathbb{P}_h$ is justified because it means that if we know the current position of the analyzed system in this phase space then the current rate of the system motion in $\mathbb{P}_h$ is completely specified.

The partition of the initial set of structureless particles into the mesoscopic entities is fuzzy in some sense. For example, dealing with two mesoscopic entities neighboring in space an original particle belonging to one of them can partly contribute to a property ascribed to the other one. It is possible to avoid such ambiguity by introducing a certain effective field $h(t, x)$ spatially interpolating the discrete collection $\{h\}$ (Fig. 2.11). However, it does not mean that now we deal with a phase space whose points are functions and, so, its dimension is infinite. The matter is that the field $h(t, x)$ introduced in this way cannot exhibit substantial variations on individual scales of the mesoscopic entities and such short scale variations must be ignorable within phenomenological models operating with the effective fields $h(t, x)$.

As a rule, a many particle ensemble does not exist on its own, i.e., is not a closed system but embedded into some environment. The interaction between this environment and the constituent particles usually is individual and cannot be governed directly by the mesoscopic characteristics of the given ensemble. So to take into account this interaction between the ensemble and environment at the mesoscopic level, its mesoscopic effects are described in terms of random Langevin forces $\Lambda\{h\} \circ \xi(t)$ whose intensity $\Lambda\{h\}$ depends generally on the order parameters. In other words the "force" $\Psi\{h\}$ in equation (2.39) is replaced by the sum of its regular component $\Psi_r\{h\}$ and random one

$$\Psi\{h\} \Rightarrow \Psi_r\{h\} + \Lambda\{h\} \circ \xi(t)\,. \tag{2.40}$$

Besides, if even a system at hand is closed, i.e., there are no external objects interacting with it, not all the detailed features of its dynamics at the microscopic level can be governed directly by the characteristics of the mesoscopic level. Mesoscopic models based on governing equations similar to equation (2.39) just ignore such features. Nevertheless, there are systems where extremely weak perturbations lying outside their mesoscopic description affect the system dynamics substantially endowing it with unpredictability at the mesoscopic level. The Langevin force model (2.40) also takes into account these effects, at least, qualitatively. In other words, the complex structure of mesoscopic entities can be responsible for random factors arising at the mesoscopic level even the given system is completely deterministic at its microscopic level.

The constructions discussed above are summarized in Fig. 2.12. At the microscopic level we have a collection of equations governing the individual dynamics of particles. Increasing the scales of consideration, spatial and temporal ones, we construct some mesoscopic entities made of many particles whose characteristic properties ascribed to them as whole units form the collection of the order parameters $\{h_k\}$. As a fundamental assumption widely accepted in such theories, the order parameters are supposed to "control" the dynamics of the microscopic phase variables $\{\theta_i\}$ such so the latter follow variations in the order parameters

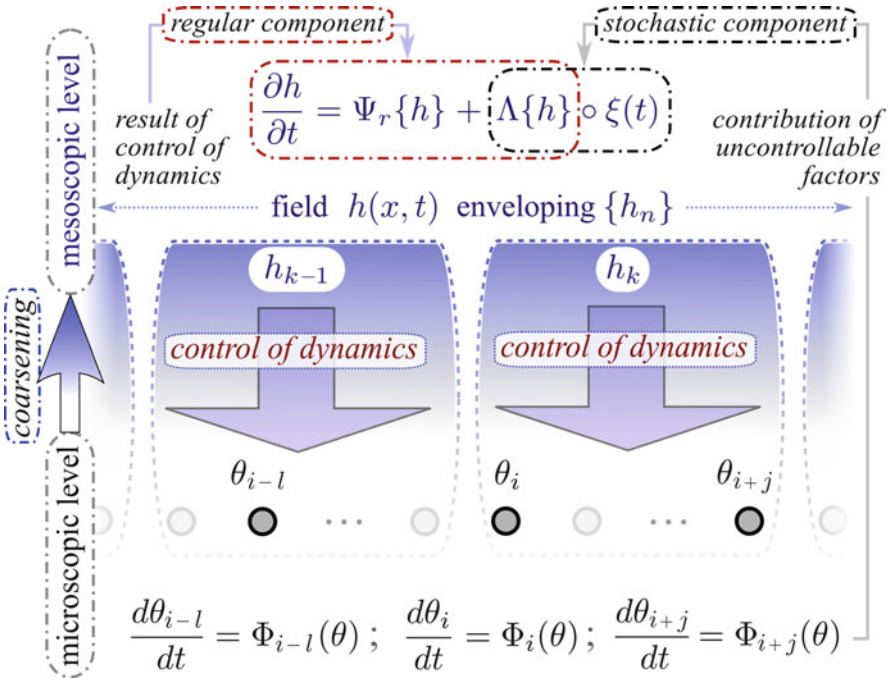

Fig. 2.12 Schematic illustration of constructing an effective model for a statistical physical system with cooperative dynamics

adiabatically. It endows the mesoscopic governing equations with the form matching the paradigm of Newtonian mechanics. The effects of microscopic dynamics laying beyond the adiabatic approximation are taken into account in terms of random Langevin sources because of their fast variations in space and time on the microscopic scales.

Concluding this part of our discussion about the mesoscopic level of description, I should underline the follow.

The reduced phase space $\mathbb{P}_h$ and the Newtonian type governing equations similar to (2.39) with the Langevin random forces (2.40)) **do not form** the general mathematical realm of the effective theories even for physical objects. They are just the approximation used in many applications. In particular, by now there have been found many complex systems in physic, chemistry, engineering, etc. whose mesoscopic description deals with such notions as long-term memory and, thus, does not match the paradigm of Newtonian mechanics. Hydrodynamic turbulence, martensites, media with subdiffusion are their examples.[12]

2.8 Conclusion

In Chap. 2 I have presented arguments for the relationship between the notions and formalism used in the basic laws of classical physics and the existence of the microscopic level of description of the corresponding physical systems which obeys the principle of microscopic level reducibility.

According to this principle, first, any system belonging to the realm of classical physics admits the representation as an ensemble of structureless particles with certain properties ascribed to them individually. Second, the interaction between these particles is supposed to be determined completely by their individual properties and to meet the superposition principle. In the given case the superposition principle is reduced either to (*i*) the model of long-distant pair-wise interaction between particles or (*ii*) the model of local interaction between particles and some fields with linear properties. Within the former model the dynamics of any system is determined completely by the current values taken by the individual properties of its particles. Within the latter model this statement also holds provided the particle properties instantiated withing some time interval are taken into account.

Whence I have drawn the following conclusions, where, for short, the particle long-distant interaction is implied to be the case if a specific model of particle interaction is not noted explicitly.

[12]There is a vast literature about physical systems with anomalous dynamics laying beyond the paradigm of Newtonian mechanics and recently summarized, e.g., in Uchaikin (2013).

- Laws governing the dynamics of such systems can be written within the formalism of ordinary differential equations dealing with time derivatives of the particle's individual properties. It has been justified appealing to the concepts of thick presentism regarding the flow of time as a sequence of bold instants. In the model of local particle-field interaction these equations should be completed with partial differential equations for the corresponding fields.
- Dynamics of such systems can be described as their motion in the corresponding phase space $\mathbb{P}^m$. A point of this phase space admits interpretation as a collection of all time derivatives of the particles' individual properties whose order is less than a certain integer m; naturally, this collection includes these properties too. The position of a given system in its phase space $\mathbb{P}^m$ determines the rate of system motion in this phase space, which enables us to introduce the notion of initial conditions and the concept of determinism of physical systems.
- The energy conservation should be a consequence of some general laws governing various systems of a given type rather than particular circumstances. Within this requirement, for a system to admit the introduction of energy, its phase space $\mathbb{P}^2$ must comprise only the individual properties of the constituent particles and the corresponding time derivatives of the fist order. The dynamics of such systems is described by differential equations of the second order with respect to time derivatives, which is exactly the case of Newtonian mechanics. In these sense the systems belonging to the realm of classical physics take the unique position among the other plausible models.
- The fundamentals of Newtonian mechanics are widely employed in the probability theory developed for coping with systems whose dynamics is not strictly predictable. In particular, it is the notion of the phase space, i.e., the complete set of the possible elementary states $\mathfrak{S} = \{e\}$ of a given system. All its probabilistic properties are supposed to be some functions of the probabilities $\{p_e(t)\}$ of finding the system at a state e at time t. Besides, stochastic processes are widely described within Markov approximation appealing directly to a probabilistic generalization of the principle of microscopic level reducibility. Namely, it may be formulated as the proposition that:
  - any probabilistic system can be represented as an ensemble of structureless particles (or particles whose structure is fixed) possessing individual properties on their own;
  - possible instantiations of these properties specify the set of the elementary states $\mathfrak{S} = \{e\}$ of the given system with the probabilities $\{p_e(t)\}$ *directly* ascribed to them;
  - the gist of Markov approximation: all the properties of a given system including its probabilistic dynamics on infinitesimal time scales are determined completely by its current state.
- The superposition principle provides a universal approach to describing the interaction of particles on scales of their ensemble as a whole. It reduces this interaction to a certain algebraic sum of the *individual* interactions between a

pair of particles. There are two implementations of the superposition principle.

- In the model of long-distance interaction, for example, the potential energy of an ensemble of particles is reduced to the sum of the individual potential energies between all their pairs. This interaction is regarded as immediate, there is no time delay in the response of one particle to variations in the state of another distant particle.
- In the model of local particle-field interaction a certain field being *linear* in properties is assumed to be *locally* generated by a particle, to propagate freely in space, and after getting another particle to affect it also *locally*. The linearity of such fields means that the field generation rate as well as the speed of field propagation are independent of their intensity and the force with which a filed acts on a particle is strictly proportional to its intensity. In this case the resulting interaction of particles is delayed by the time interval required for the field to pass the distance between them.

When the characteristic time scales of system dynamics are much larger then the time required for the corresponding fields to pass the distance about the system mean size, the long-distance interaction is a reasonable approximation of the local particle-field interaction.

- A wide variety of cooperative phenomena in complex physical systems admit effective mesolevel description in terms of quantities often called order parameters. They may be treated as continuous fields whose local values aggregate in some way the detailed properties of the corresponding particle clusters. Usually mesoscopic theories of such complex phenomena inherit the basic features of the corresponding microlevel models, however, the constrains imposed on plausible mesolevel models are weaker than that of microlevel. It is due to the fact that the internal structure of particle clusters, playing the role of basic mesolevel entities, may change in time.

  As a result, within a *mesolevel description* a complex physical system can acquire properties impossible at its microlevel, for example, long-term memory and indeterminism caused by intrinsic stochasticity. It is not a paradox because these properties aggregate in some way complex dynamics of many microlevel variables.

http://www.springer.com/978-3-319-51705-6

Physics of the Human Mind

Lubashevsky, I.

2017, XIV, 380 p. 83 illus., 41 illus. in color., Hardcover
ISBN: 978-3-319-51705-6