

ABSTRACT

THE THERMODYNAMICS

Miroslav  
Holeček



# ABSTRACT THERMODYNAMICS

Miroslav Holeček

PLZEŇ - NYMBURK 2006

This publication is supported by the Czech Ministry of Education, Youth and Sports, the project MSM 4977751303, and by the Grant Agency of the Czech Republic, the projects 401/02/0579 and 106/05/0219.

© Miroslav Holeček, 2006  
ISBN 80-239-6075-X

TO MY PARENTS

# Preface

I still remember my first reading of a thermodynamic textbook (written by my teacher, J. Kvasnica, [16]). I cannot forget first impressions, especially a big surprise that thermodynamics is not only about heat engines. An amazing generality of the subject did fascinate me. On the other hand, the formulation of thermodynamics liked rather clumsily. Certain definitions did not seem to be able to avoid some circularity and the main law of the whole subject, the second law, was introduced by formulations resembling more some engineering rules than a fundamental law of nature.

These feelings acted to me as a challenge to search for better formulation of thermodynamics. Of course, many people had similar feelings and some of them took up this challenge in full seriousness. It is especially the school of the so-called rational thermodynamics, whose main idea is to lay a rational axiomatic foundation of thermodynamics that allows studying the subject by a rigorous mathematical language. However, the choice of its basic axiomatic schemas and the logic of whole construction does not seem to me to catch precisely the core of thermodynamics. It looks like that something, what is very important for classical thermodynamics, is missing in this approach.

I realized it very clearly when reading an interesting work by E. H. Lieb and J. Yngvason [21]. It is an axiomatic foundation of classical equilibrium thermodynamics but the *logic* of its construction is, in my opinion, essentially different from other attempts. It has influenced considerably my thinking about foundations of thermodynamics. After several years, I started to be more and more convinced that thermodynamics cannot be fully formulated in the spirit of Newtonian physics. Especially, the idea of an inert spatial background, existing by itself, seems to contradict the logic of thermodynamics. The Lieb's and Yngvason's approach outlines the way of escape from the Newton's heritage.

This book is an attempt to show it. The problem is, however, that I am still very far to have a comprehensive physical theory formulated fully in the thermodynamic spirit. Therefore, this book only outlines important ideas and explains their meaning in the context of such a possible theory. Since motivations come especially from thermodynamics, the sketched scheme of a physical theory not using the idea of spatial background is called the *abstract thermodynamics*. The abstract thermodynamics, however, is not intended to be a concrete physical theory. It is rather a *scheme* of a (general) physical description that is not formed in the Newtonian spirit.

The book could be hardly written without the support and encouragement of my friends and colleagues, Josef Rosenberg, Jaroslav Šesták, Josef Voldřich, and without many people who helped me to enter the field where physics and mathematics uncover their philosophical foundations, especially Jiří Fiala, Ivan M. Havel, Zdeněk Neubauer and Petr Vopěnka. I am also indebted to Jana Čepicková and Michal V. Hanzelín for their help and advice in preparation of the book.

# Contents

<b>Preface</b>	<b>i</b>
<b>Introduction</b>	<b>1</b>
<b>List of symbols</b>	<b>15</b>
<b>1 Physical systems</b>	<b>21</b>
1.1 Two levels of description . . . . .	22
1.2 Physical systems and subsystems . . . . .	25
<b>2 States and manipulations</b>	<b>29</b>
2.1 States of physical systems . . . . .	31
2.2 Manipulations with physical systems . . . . .	34
2.3 Processes . . . . .	40
2.4 The surrounding . . . . .	43
2.5 Global variables . . . . .	46
2.6 Temperature and equilibrium . . . . .	49
<b>3 Systems as sets of parts</b>	<b>53</b>
3.1 Structure . . . . .	56
3.2 Parts as subsystems and other parts . . . . .	62

3.3	Local variables . . . . .	67
3.4	Process variables . . . . .	72
3.5	State space . . . . .	73
3.6	The continuity rule . . . . .	76
3.7	Averaging and scale . . . . .	81
<b>4</b>	<b>Special structures</b>	<b>87</b>
4.1	Simple systems . . . . .	88
4.2	Thermomechanical systems . . . . .	93
4.3	Continuum . . . . .	100
4.4	State space of a continuum model . . . . .	106
<b>5</b>	<b>Space</b>	<b>113</b>
5.1	Space as a surrounding . . . . .	116
5.2	Symmetries of space . . . . .	123
5.3	Special symmetries . . . . .	125
5.4	Newton space . . . . .	128
5.5	Complete decomposition of $\epsilon$ . . . . .	133
5.6	Mach space . . . . .	135
5.7	Adiabatic space . . . . .	139
<b>6</b>	<b>Accessibility</b>	<b>141</b>
6.1	Relative motion . . . . .	145
6.2	Definition of ‘accessibility’ . . . . .	148
6.3	Time arrow . . . . .	151
6.4	Adiabatic accessibility . . . . .	153
6.5	Thermodynamics and entropy . . . . .	155
	<b>Bibliography</b>	<b>163</b>
	<b>Index</b>	<b>168</b>

# Introduction

Isaac Newton laid foundations of modern physics. The main reason of his triumph is the choice of a surprisingly fruitful *way of description* of matter and its interactions. An essential point of the Newton's scheme is a strict difference between objects and *empty space* between them. In Newton's viewpoint, space is something external what remains always similar and immovable [29]. It is identified with a 'container' holding material objects and having an independent, absolute existence. It should be emphasized that the Newton's approach represents the crucial change of the viewpoint comparing to the conception of his predecessor, the famous philosopher and mathematician R. Descartes [6]. For him, matter was not *located in* something extending but formed this extension by itself. This omnipresent extension is expressed by the concept of "res extensa". From the Cartesian point of view, space is thus merely a conceptual abstraction losing any meaning without bodies.

The Newton's concept of space as some independent entity of its own is a little metaphysical because there is no observable evidence preferring the idea of an absolute space *filled* by matter more than the idea of the extension *formed* by matter

itself<sup>1</sup>. We, however, cannot ignore the extreme usefulness of Newton's space in development of physics and remarkable prosperity of the Newton's programme. The reason consists also in the fact that the hypothesis about existence of an indifferent, unchangeable background allows formulating physical problems with extremely *simplified interactions*.

Namely, when studying an arbitrary physical system (a molecule, a pendulum on the Earth or a galaxy somewhere in the Universe) we need to define clearly the influence of surrounding matter. There is a huge number of extremely complicated objects of various natures in the Universe and thus the possibility of studying some of them needs a very simple assumption about influence of others. If we had no such an assumption, the physics would not be possible, because we would never know *which* objects should be taken into account.

The idea of empty space enables us to simplify the situation. Namely, we can imagine studied objects as if they were only in empty space without influence of anything else. Due to the assumption of existence of such an indifferent 'container' this consideration has a good operational meaning: we may ever imagine that the other matter is far enough not to influence the studied system. The system is thus *only* in space. Let us realize that such a consideration would not be possible when accepting the scheme in which there is only extending matter – then the system cannot be separated from surrounding matter because there is nothing else than the “surrounding matter”.

---

<sup>1</sup>The origin of the concept of absolute space lies more in theology than in physical considerations. Newton was influenced by the doctrine of Henry Moore (H. More, *Enchiridium metaphysicum*, in *Opera omnia*, London (1679), Vol II.) who proposed the conception of space as an actual manifestation of God. Since God has to exist everywhere, space must exist even where there are no bodies [18], p. 205.

Modern physics, however, brings many objections against the conception of absolute space as an inert arena on which other physical phenomena take place. In the general theory of relativity, for example, there is no background space (and time) – matter fields are not *in* space-time there are rather on ‘each others’<sup>2</sup>. The absence of a really empty space (vacuum) in the quantum-field physics is also interesting from this point of view. Such examples seem to show us that we are going back to a modification of the Cartesian description of the world.

Nevertheless, physics is rather Newtonian than Cartesian. Though there are many suspicions that the idea of an inert spatial background is problematic or even mistaken, we still believe that it is a good auxiliary hypothesis that works perfectly in an essential majority of physical situations. Physical problems in which the Newtonian conception collapses are indeed very interesting and appealing for theoretical physics but rather exotic and marginal for physics of everyday use. The physics of macroscopic objects (especially at human spatial and temporal scales) is formulated in the Newtonian spirit without any doubts. There are no indicia of disputableness of the Newtonian approach in this sort of problems. Really?

Namely, the only reason we fully believe to an inert spatial background in everyday physics is that we are not in a position of having a principally competitive approach. Any concepts or thought schemata of physical investigations of macroscopic objects are thus “grafted” into the Newtonian scheme and hence we are more and more insisted on its eminent usefulness. Possible indicia of dubiousness of the schema of description is thus

---

<sup>2</sup>Such considerations appear when thinking about a theory connecting the general relativity with quantum physics. Namely, such a theory - called the quantum gravity - could hardly have a fixed spatial background [35].

difficult to be noticed. We have to attend to of rather conceptual discrepancies than technical problems to find an incorrectness of the Newtonian way of descriptions. A possible case of such a discrepancy (of the Newtonian conceptual scheme with a suitable way of description of macroscopic objects) is presented in *thermodynamics*.

Thermodynamics is a peculiar physical discipline<sup>3</sup>. It came into being from highly practical problems concerning the efficiency and optimal designing of heat machines used in industry. The crucial work by Sadi Carnot, *Réflexions sur la Puissance Motrice de Feu, et sur les Machines Propres a Développer cette Puissance*<sup>4</sup> [3], published in 1824 (at Carnot's own expenses), however revealed some universal features touching an arbitrarily constructed machine. Carnot's considerations indicated clearly to a universal law of nature that is now known as the *second law of thermodynamics*. The way of Carnot's argumentation, however, is very unusual in comparison with typical schemas of the Newtonian physics. Although his considerations were essentially elucidated in subsequent works (done by E. Clapeyron, W. Thomson alias Lord Kelvin, R. Clausius and others – the story is briefly resumed in [15]), an element of something strange has remained in thermodynamics till now.

First of all, what is the essence of the second law? What does it exactly claim? Its verbal formulations resemble more engineering rules about impossibility of constructing machines realizing certain processes than statements of modern physics. The way of translation of these statements into precise physical formulations consists in existence of a full-valued physical quan-

---

<sup>3</sup>Compare language and approach to the subject in [2, 9, 23, 39, 40].

<sup>4</sup>*Reflections on the Motive Force of Fire and on the Machines Fitted to Develop that Power* [24], see also comments at [15], pp. 67-72.

tity, the entropy. It was introduced by R. Clausius (in 1865) who noticed that Carnot’s argumentation may be expressed mathematically as

$$\oint \frac{\delta Q}{T} = 0, \quad (1)$$

where  $\delta Q$  is the exchanged infinitesimal heat,  $T$  the absolute temperature and the integration is done over one cycle of a reversible process. It implies that the quantity  $S$ , defined by

$$dS \equiv \frac{\delta Q}{T}, \quad (2)$$

does not change after a reversible return to initial state. Clausius proposed to call  $S$  the *entropy* (from the Greek word  $\tau\rho\omicron\pi\eta$ , *change, transformation*). It means that the entropy is a *state variable*. This quantity thus represents similar characterization of physical systems as, say, momentum or energy, and gives a linkage to the way of thinking in the Newtonian spirit.

The problem, however, is that the definition of entropy is related to the concept of heat via (2). But what is ‘heat’ from the viewpoint of Newtonian physics? It cannot be identified with a kind of molecular motion. It is rather an “energy transfer to the hidden molecular modes” as observed by H. B. Callen ([2], p. 8). The essence of ‘heat’ is thus an energetic “communication” between macroscopic and microscopic scales. However, the concept of ‘macroscopic scale’ is difficult to be defined. Nevertheless, we can avoid the problem by defining the heat via the familiar thermodynamic relation<sup>5</sup>,

$$dE = -\delta W + \delta Q, \quad (3)$$

---

<sup>5</sup>See e.g. [23], p. 26.

where  $dE$  is an infinitesimal change of the energy of the system and  $\delta W$  the performed work. These quantities have an evident relation to Newtonian physics – the energy as well as the work are defined in mechanics. In thermodynamics, however, the concept of ‘work’ is not so simple because its mechanical definition as a ‘force times distance’ is insufficient. The question is *which* forces and *which* distances. Namely, if we calculated the sum of ‘forces times distances’ for all molecules of the system the work would always equal the change of the whole energy of the system and no ‘heat’ would appear in (3), see Fig. 1.

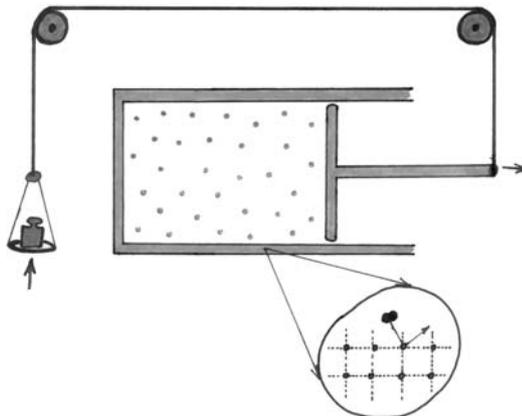


Figure 1: *What is work? If we calculated a sum of all ‘forces times distances’ including interatomic forces we would obtain a full energy exchange of the system with its surrounding. The work, however, is the ‘force times distance’ concerning only the movement of the piston.*

Let us illustrate this on the familiar thermodynamic system consisting of a gas in a box that includes a movable piston.

The piston is connected through a rope with a weight in gravitational field as illustrated in Fig. 1. The work is defined by the displacement of the piston. However, if this physical system is understood as a conglomerate of molecules of the gas and molecules of the box and piston, then the ‘displacement of the piston’ is only a special molecular movement (among many many others). Why *this* movement has a fundamental meaning to define the decomposition of the energy into work and heat? And how it could be generally defined? It is difficult to answer such questions from the point of view of the *system itself* in a different way than speaking about some “macroscopic modes” of the movement (see e.g. [2], Section 1.2) and their eminent relevance for thermodynamics. Thereby we again face the problem with an explanation of the word ‘macroscopic’.

Let us try to look at the system from the point of view of the other physical system in its surrounding – the weight that interacts with the system through a connecting rope (see Fig. 1). When displacing the weight, the piston begins to move. The weight thus *defines* what is meant by that special “mode” of the motion. It is enabled by a special arrangement that allows *manipulating* with the system by another system. The thermodynamic work then may be defined as a form of energy communication between the system and its surrounding which arises in special *manipulations* with the system<sup>6</sup>. The problem, however, is how these special manipulations should be defined.

In the pioneering work by R. Giles [9], that is attempting to

---

<sup>6</sup>S. Kauffman [14], a biologist, (when realizing in the discussion with a physician P. W. Anderson that the concept of work is far to be clear, p. 96) specifies work as “constrained release of energy” whereas by ‘constrains’ means external organization of energy transfer by various pistons, wheels, paddles and so on. These constrains thus realize various manipulations with the system.

formulate an axiomatic foundation of thermodynamics, the author introduces a class of “mechanical systems” that serve as suitable surroundings allowing to define the *adiabatic process* ([9], Chapter 5), i.e., the process when the only energy transfer is realized by the work. Recently, the Giles’ work has been successfully crowned by proposing a perfect axiomatic system of equilibrium thermodynamics by E. H. Lieb and J. Yngvason [21]. Though the authors are getting near the Giles’ work at many points, they avoid any attempts to define adiabatic processes (and work) in this way. Instead, they introduce the concept of adiabatic accessibility (in the spirit of Planck’s formulation of the second law) by using only a very special “mechanical system” in the surrounding of studied thermodynamic system – a weight in gravitational field<sup>7</sup> ([21], p. 17). We will present it in Chapter 6.

Thermodynamics thus seems as not being able to avoid the presence of something (a mechanical device, the weight in a gravitational field, etc.) in the *surrounding* of the studied system. If the system were only alone, the concept of heat, work or adiabatic accessibility would become superfluous or even nonsensical. In the Newtonian physics, ‘alone’ means to be only in space (so that influence of other bodies may be neglected). The (hypothetical) possibility for any system to be studied as ‘being alone’ is hence the leading idea of the Newton’s program: it allows simplifying considerable external interactions and studying inner dynamics of the system<sup>8</sup>.

---

<sup>7</sup>The other place where the concept of work is used, is the passage when the authors define simple systems. The ‘work’ is represented there only by a sort of special coordinates without a detail explanation ([21], p. 38).

<sup>8</sup>For example, when deriving the Kepler’s laws of motion of planets around the Sun, we imagine *one* planet (e.g. the Earth) and the Sun as a physical system which is alone in empty space.

Thus thermodynamics presents a small but remarkable conceptual controversy with the Newtonian scheme because its basic concepts cannot be *principally* defined when using only systems in empty space. It may be a source of many problems with understanding the entropy. Namely, we are tempted to translate any concept into the Newtonian scheme of thinking<sup>9</sup>. Hence the entropy is expected to be assigned to the system ‘alone in empty space’ (as any Newtonian quantity) in spite of its immediate relation to the concept of heat or adiabatic accessibility (which, however, *cannot* be meaningfully explained for systems being alone in empty space).

When studying thermodynamics, one has to notice its striking difference from other physical disciplines. Its methods, concepts and ways of defining physical problems are strange in comparison with other branches of classical physics, such as mechanics or electromagnetic theory. For example, a typical object of the thermodynamic study is a gas in a box equipped with a piston (as seen in Fig. 1). But *what* do we actually study? Are these the properties of the gas itself or those of the box, piston and all accompanying machinery (as the rope and weight in Fig. 1)? Many attempts have been made to “purify” thermodynamics from the baggage of various pistons, barriers, weights, etc., and formulate it similarly as, say, the theory of electromagnetic field<sup>10</sup>. Roughly speaking, statements of such formulations of thermodynamics are only about *gas itself*. Surprisingly, despite an enormous effort of its representatives to

---

<sup>9</sup>Recall problems with understanding quantum mechanics, where such a translation is impossible or highly problematic.

<sup>10</sup>Such a “pure” thermodynamics is sometime called the *rational thermodynamics*. Its representatives are B. D. Coleman, W. A. Day, M. E. Gurtin, W. Noll, D. R. Owen, J. Serrin, M. Šilhavý, C. A. Truesdell, and many others, see e.g. [5, 10, 31, 32, 40, 41, 42, 43].

convince others about the rightness of their formulation of thermodynamics (see e.g. [42]), this approach has never forced out the classical conception (compare the comment in [21], p. 10). In my opinion, the fact that thermodynamics defies (whatever it means) to be completely formulated in the Newtonian spirit is worth noticing. As if its strange formulation and certain inaptitude (in comparison with elegant formulations of other disciplines of classical physics) indicated the presence of some barriers. The barriers that close physics in a hardly visible prison. The prison that looks like empty space.

That is why we try to concentrate especially on those features of thermodynamics that are unusual and strange. This book is not *about* thermodynamics, it is only motivated by its special way of description. Typical features of this description (especially those that do not appear in other physical theories) are separated out and used in formation of a very abstract formulation of a general physical description of arbitrary physical systems. We call it the *abstract thermodynamics* to emphasize its similarity with the description introduced by classical thermodynamics. The main property of this abstract scheme of description is the absence of *any spatial background* as well as a *presupposed temporal scheme*. Instead of ‘space’ we have only another physical system, called the *surrounding* of the system, and the temporal scheme of standard physics is replaced by an abstract concept of *change*. The change is just a couple of states, whereas the concept of *state* has similar meaning as that in standard physics<sup>11</sup>. The crucial (and essentially ther-

---

<sup>11</sup>Notice that in thermodynamics (as well as in quantum mechanics), the concept of state plays much more important role than in newtonian mechanics. Surprisingly, the concept of state seems to be rather Cartesian than Newtonian as noticed, e.g., by L. Kvasz [17].

dynamic) point is that any change is understood as a result of certain *manipulation* with the system. Any manipulation is *caused* by an action in the surrounding of the system<sup>12</sup>, which represents the only *causality* in our approach. The surrounding, however, is only a special kind of the system, hence this “action” is nothing but a change defined on the surrounding. All these concepts are introduced and studied in Chapter 2.

In the Newtonian scheme, the idea of physical systems in empty space allows not only simplifying external interactions, but also offers a general way of *characterization* of states of the system. States may be characterized by current distributions of physical quantities in space. Since Newton’s space is a universal fixed background, such a description of the state is uniquely defined. In our approach, however, no such a universal background exists. That is why we introduce, in Chapter 3, a fixed set of some *parts* of the system. Individual parts may not be subsystems of the system – their role is only to characterize states of the system by a suitable mathematical description. The set of parts is thus an auxiliary tool of description of the system and may be chosen arbitrarily. Any choice of this set is called the *structure* of the system. In Chapter 4, important structures of equilibrium thermodynamics (simple systems) and continuum thermomechanics (discrete systems) are introduced. The concept of the *continuum* is then studied as a special structure of description of various macroscopic systems.

What about the Newton’s absolute space? In our approach, it becomes only a special kind of surrounding, called simply the *Newton space*, and therefore its properties may be studied without a baggage of unclear metaphysical assumptions. It is done in Chapter 5. Here we obtain an interesting result.

---

<sup>12</sup>Whereas a ‘manipulation’ may also mean to “do nothing”.

Namely, special properties of this surrounding (especially its symmetries) imply that it causes a strange manipulation with a physical system that is surrounded by it. There is only *one* manipulation caused by the Newton space that, however, gives an *infinite* number of changes. In other words, any state of the system may change into an infinite number of different states. It resembles the time evolution of the system when letting alone in empty space. We can conclude that the Newtonian ‘time’ is a *result* of symmetries of the Newton space.

However, the detail analysis of the problem shows that this “time interpretation” is only one possibility from many others. The way in which the Newton space “yields” the (Newtonian) time, indicates the presence of a more fundamental structure that may occur in various “modes” of realization. One of them is very similar to the Newtonian space and time but the ‘time’ has no meaning of an absolute entity flowing independently of anything else. It is rather defined by a set of clocks that are inherent part of the structure. Another realization is space with a weight in a gravitational field that does not include any ‘clocks’ and the concept of time is missing. Maybe, there are other interesting realizations but we do not search for them. The general structure yielding all those realizations is called the *Mach space* in honor of the famous physicist and philosopher Ernst Mach. Namely, his objections against the Newton’s absolute space seem to outline a channel towards this structure (Section 5.6).

In Chapter 6, we outline some features of a physical description that is not defined on a given spatial background. We go back to the Cartesian conception for which space is only a relation between bodies. This approach, however, brings serious problems because of its utter relativity. In Newtonian physics,

space includes fixed *places* (spatial points) allowing to define motion as going from one point of space to another point of space. In Cartesian physics, on the other hand, everything is in a relative motion towards anything else and no fixed "platform" exists. How can 'motion' be defined then? The problem with the definition of motion was one of reasons of the failure of the Cartesian physics.

That is why we start by analyzing the problem of definition of motion. Being motivated by the thermodynamic concept of 'adiabatic accessibility', we show that the essence of this concept offers an escape from the devastating relativity. Namely, motion may be defined even if there is no spatial background, but not at *any* situation. The idea consists in restricting *only* to situations when surrounding bodies behave *like* spatial places, i.e., they are fixed and remain in the same state. It is the core of the concept of *accessibility*. The accessibility replaces, in a way, the concept of 'dynamics' if no assumption about an (inert) spatial background is done.

This accessibility is motivated by the immobility and inertness of the Newton space. When replacing the Newton space by the Mach one, we obtain a more general accessibility called the *weak accessibility*. The weak accessibility has an understandable meaning in physics. If the realization of the Mach space with clocks is used, for instance, this kind of accessibility defines the *time arrow* (only transitions, when the datum on clocks increases, are accessible). The realization with a weight in a gravitational field then defines the adiabatic accessibility that has a direct relation to the entropy. A relation between 'entropy' and 'time arrow' then appears in a surprising context. At this point, we join the many times mentioned work by Lieb and Yngvason [21], and the book declines to its end (Fig. 2).

The book does not develop the abstract thermodynamics as a comprehensive axiomatic structure. The text describes by no means a complete physical theory. It rather tries to express some ideas of a possible physical description that essentially differs from the standard one. It could be globally characterized as a formulation of physics without a spatiotemporal “background”. Contrary to the general relativity or quantum gravity, the motivation for such a description and main ideas of its construction come from thermodynamics.

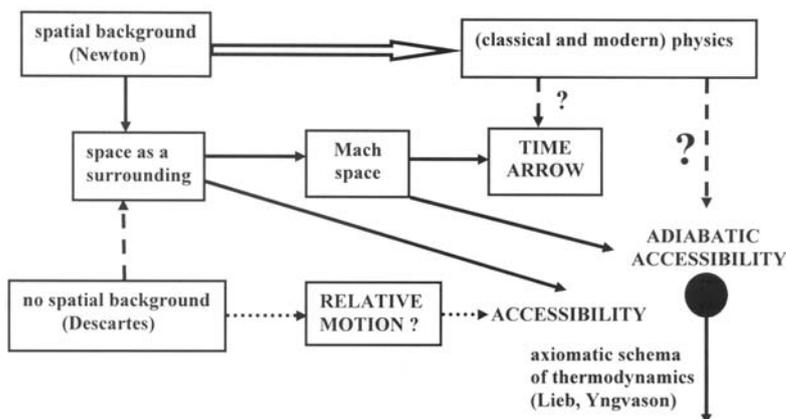


Figure 2: The diagram outlining main connections of our approach.

# List of symbols

## *Some mathematical symbols*

$[a, b]$	an ordered couple of objects $a$ and $b$
$\{x, P(x)\}$	means the set of objects having property $P$
$a \in A$	means ‘ $a$ is an element of the set $A$ ’
$a \notin A$	means ‘ $a$ is not an element of the set $A$ ’
$A \subset B$	means ‘the set $A$ is the subset of the set $B$ ’
$A \cap B$	the intersection of sets $A$ and $B$
$A \cup B$	the union of sets $A$ and $B$
$A \setminus B$	the complement of the set $B$ in the set $A$
$\text{Pow}(X)$	the set of all subsets of the set $X$
$f : A \rightarrow B$	the mapping from the set $A$ to the set $B$
$f^{-1} : B \rightarrow A$	the inverse mapping of the mapping $f$
$f(A) \subset B$	the range of values of the mapping $f$
$\mathbb{N}$	set of natural numbers $1, 2, 3, \dots$
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of positive real numbers,
$\mathbb{R}_{0,+}$	set of nonnegative real numbers,
$\langle r_1, r_2 \rangle$	the closed interval $r_1 \leq x \leq r_2$ ( $r_1, r_2, x \in \mathbb{R}$ )
$(r_1, r_2)$	the open interval $r_1 < x < r_2$ ( $r_1, r_2, x \in \mathbb{R}$ )

$\mathbb{E}_d$	$d$ -dimensional Euclidean space
$\mathbf{x}, \mathbf{x}' (\in \mathbb{E}_d)$	points of Euclidean space
$\mathbf{B}, \mathbf{B}_G, \mathbf{B}_q$	Banach spaces (complete normed vector spaces)
$\ \mathbf{b}\ $	the norm of a vector from a Banach space
$\mathbf{Q}, \mathbf{Q}^{-1}, \mathbf{Q}^\top$	a matrix, its inversion and transposition
$\mathfrak{A}$	$\sigma$ -algebra
$\mathfrak{A}^{\mathcal{X}}$	$\sigma$ -algebra that includes all sets from $\mathcal{X}$
$\mu, \mu\nu$	various measures
$\int_X f d\mu$	the integral of the function $f$ over the set $X$

*Symbols of concepts that have no mathematical description*

$\mathfrak{S}, \mathfrak{S}', \mathfrak{S}_i$	(physical) systems
$\mathfrak{S}' \stackrel{sub}{\subset} \mathfrak{S}$	means that $\mathfrak{S}'$ is a subsystem of $\mathfrak{S}$
$(\mathfrak{S}_1, \mathfrak{S}_2)$	the composition of systems $\mathfrak{S}_1$ and $\mathfrak{S}_2$
$\mathfrak{S}_{sur}$	the surrounding as a physical system

*Sets assigned to physical systems and derived concepts*

$\Xi(\mathfrak{S}), \Xi$	the set of representative subsystems of $\mathfrak{S}$
$\Sigma(\mathfrak{S}), \Sigma$	the state set of the system $\mathfrak{S}$
$f_{sub}$	the mapping between $\Sigma(\mathfrak{S})$ and $\Sigma(\mathfrak{S}')$
$s (\in \Sigma)$	a state (element of the state set)
$s_{sur} (\in \Sigma_{sur})$	a state of the surrounding
$\mathcal{M}(\mathfrak{S}), \mathcal{M}$	the set of manipulations with $\mathfrak{S}$
$\alpha, \beta, \gamma (\in \mathcal{M})$	manipulations
$id$	the identity manipulation
$(\alpha, \beta)$	the parallel composition of manipulations

$\pi$	the process (a subset of $\Sigma$ )
$\pi_{\tau_1, \tau_2}$	the parametrization of the process $\pi$
$\Pi(\pi, s)$	the tangent of the process $\pi$ at the state $s$
$\mathfrak{F}$	the mapping defining the surrounding
$G$	the global variable
$\mathcal{D}_G$	the definition domain of the global variable
$Q_g$	the set of global variables
$T$	the global variable ‘generalized temperature’
$f_{eq}$	the mapping defining the equilibrium

*Symbols describing parts and structures*

$\mathcal{P}(\mathfrak{S}), \mathcal{P}$	the set of parts
$p, p' (\in \mathcal{P})$	parts
$\mathcal{D}_{comp}$	the set of couples of disjunct parts
Comp	the mapping defining the composition
$(\mathcal{P}, \text{Comp})$	the structure
$(p_1, p_2)$	the composition of parts $p_1$ and $p_2$
$p_0$	the total part
$p' \overset{sp}{\subset} p$	means that $p'$ is a subpart of $p$
$\mathcal{P}_{atom}$	the set of all atomic parts from $\mathcal{P}$
$\mathcal{P}_{sub}$	parts corresponding to subsystems
$F_s$	the mapping defining the set $\mathcal{P}_{sub}$
$\mathcal{P}_{ch}$	externally defined parts
$\hat{q}, \hat{V}, \hat{m}, \hat{b}$	various (local) variables
$q_s$	the mapping corresponding to the variable $\hat{q}$
$\mathcal{D}_{q,s}$	the definition domain of $q_s$
$Q$	the set of local variables

$\hat{q} \ll \hat{q}'$	absolute continuity of variables
$\hat{q}^P$	the process variable
$q_{s,\Pi}$	the mapping corresponding to $\hat{q}^P$
$\mathcal{D}_{q,s,\Pi}$	the definition domain of $q_{s,\Pi}$
$\Gamma(p)$	the state space of the part $p$
$\Gamma$	the state space of atomic structures
$\Gamma_{full}(p_1, \dots, p_n)$	the full state space of $n$ parts
$\mathcal{A}$	the averaging mapping
$\text{Sc}$	the mapping defining the scale
$l$	the scale (a length)
$\mathcal{P}(l) (= \text{Sc}(l))$	l-space

*Symbols concerning special structures*

$v(p)$	the scaling parameter of extensive quantities
$\Sigma_{eq}$	the state space of equilibrium composition
$\Gamma^{(r)}$	the state space of a scaled copy
$rX$ ( $\in \Gamma^{(r)}$ )	a point in $\Gamma^{(r)}$
$\sigma, \partial\sigma$	a material sample and its boundary
$\hat{E}$	energy as the (local) variable
$\hat{V}$	volume as the (local) variable
$\hat{b}^q$	the flux of $\hat{q}$
$\hat{\delta}^q$	the supply of $\hat{q}$
$\mathcal{U}$	the set of elements modeling parts
$u$ ( $\in \mathcal{U}$ )	an element of this set
$\mathfrak{M}_s$	the mapping modeling parts as sets
$X_s$	the model of a part at state $s$
$X_s^0$	the model of the total part at state $s$

$\Upsilon$	the index set
$\xi (\in \Upsilon)$	an element of index set (a “label”)
$\eta_s$	the configuration at state $s$
$\mathfrak{C}$	the set of configurations
$\mu_{\mathcal{V}}$	a measure modeling the variable $\hat{\mathcal{V}}$
$\lambda_s^q$	a measure modeling the variable $\hat{q}$

*Symbols concerning space*

$\mathcal{E}$	the physical system called ‘space’
$\mathcal{C}_d$	the set of ( $d$ -dimensional) coordinate systems
$a, b (\in \mathcal{C}_d)$	coordinate systems
$f_{\mathcal{E},a}$	the mapping defining spatial points
$q_s^a$	a spatial field
$\psi, \tilde{\psi}$	fields as points of the state space
$\mathcal{G}$	a symmetry operation
$T_y$	translation
$R_Q$	rotation
$\mathcal{O}_r$	scaling
$\mathcal{E}_{Newt}$	the Newton space
$s_0$	the state of Newton space
$\epsilon$	the Newton manipulation
$\mathfrak{R}$	the set of generalized lengths
$\ell (\in \mathfrak{R})$	a generalized length
$\epsilon_{\ell}$	mappings decomposing $\epsilon$
$f_L$	the mapping assigned $\ell$ to couples of states
$\mathcal{E}_M$	the Mach space
$\epsilon_{\ell}^M$	mappings caused by the Mach space
$\hat{c}$	the clock variable

*Symbols concerning accessibility and entropy*

$N$	
$\succ$	the relation of accessibility
$\succsim$	the relation of weak accessibility
$E$	energy as a global variable
$t$	time as a global variable
$S$	entropy as a global variable
$\mathbf{w}$	the “vector” of work variables
$\sim$	the relation of thermal equilibrium

# Chapter 1

## Physical systems

If we study any problem we use a language. The more precisely and unambiguously the problem should be defined the more accurate language has to be used. The physics is an exact science and its ambition is to be precise as much as possible. Therefore the definite language of mathematics is the basement of physical thinking. Physics, however, is not a part of mathematics. Any physical concepts have a relation to reality around us and cannot be defined only within a framework of a mathematical theory. For example, the electromagnetic field cannot be fully identified with a set of vector functions on a suitably chosen geometric space. The field is a physical object that has to be also defined by a measuring procedure, what reflects its relation to other entities of reality<sup>1</sup>. The fact that studied physical object cannot be fully identified with its description within a mathematical structure is expressed here by introducing the concept of ‘physical system’. It means a part of reality that is studied from *physical* point of view in a *systematic* approach.

---

<sup>1</sup>In quantum mechanics, these questions play the crucial role.

This concept has no mathematical definition. To say what we mean by a concrete physical system is possible only in common language of everyday experience.

We try to differ strictly between the concept of physical system, that means an entity of physical reality, and its *description* by a suitable mathematical language. Namely, if we identify this physical concept with its description in mathematics we may accustom to its mathematical definition so firmly that we forget the way by which it has been transformed from the reality to mathematics. Then there is a danger of various misunderstandings and paradoxes. That is why we use two levels of the description: a nonformalized language of everyday experience and a strict mathematical language (in which physical theories are formulated).

## 1.1 Two levels of description

There are many various objects and material structures around us. Some of them are simply visible by naked eyes, others need more or less sophisticated device to be detectable in a way. They may be very large (such as the Earth or the whole galaxy) or extremely tiny (as atoms or elementary particles). In physics, we look at these objects and structures from special point of view whose purpose is to understand their physical properties, find objective laws of their behavior and possibly use them in some technological applications. This special viewpoint is called the *physical description*. In this description, structures and objects become *physical systems*.

The main task of any physical description is to find useable mathematical tools and structures enabling us to speak about physical systems in a correct mathematical language. There are

many steps by which we shift from an intuitively defined material structure to its mathematical description. Some of them are evident and clearly explainable but there are also steps that are very far to be obvious. These steps are very important since they often hide crucial events in the history of physical thinking. The assumption of ‘empty space’ in which physical objects ‘are’ is a very good example (whereas its role in mathematical description of arbitrary systems is colossal). Therefore, we explain now in which way the mathematical description is understood in our approach.

We differ between two levels of description. The first level includes only some primitive concepts as ‘physical system’ or ‘subsystem’ and their simple relations (e.g. ‘composition’ of physical systems). The physical system, however, cannot be characterized as a mathematical structure. It may be described, if necessary, only in common language. Macroscopic systems of our everyday experience may be described in great details at this level (a ‘car’ consists of ...), other systems may be described only very roughly (an ‘atom’ in a glass of water), or not at all (a ‘quark’). Sometime, a subsystem of the system is clearly describable (a ‘wheel’ is a subsystem of ‘car’), sometime it is quite problematic. Nevertheless, the description at this level is only intuitive and cannot form an essence of a mathematical description.

The whole mathematics can be formulated in the language of the set theory. That is why we define the second level of the description by introducing several *sets* for each physical system. It is the *set of states* (whereas the ‘state’ itself is an undefined, primitive concept<sup>2</sup>), the *set of manipulations* (allowing to de-

---

<sup>2</sup>In the set-theoretical language, the set of states thus corresponds to a set of urelements, i.e. elements that are not sets.

scribe any external influence) and the *set of parts*, which forms a basement of the description of any system by physical variables. The parts and manipulations (defining the concept of ‘change’) generalize the standard spatiotemporal description in which systems are understood as being in empty space with an (absolute) flowing time.

These two levels of description cannot be mixed together. At the second one, all concepts and relations are describable in the framework of set theory (allowing to define various functions, functional spaces, etc.). The first, however, is only intuitive. If we speak about a physical system  $\mathfrak{S}$  we cannot do much more than to denote it and speak, if necessary, about some of its subsystems. If we speak about a property  $P$  of a physical system  $\mathfrak{S}$  we denote it simply as  $P(\mathfrak{S})$ . It means that this property may be (at least partially and approximately) described in the common language of the first level. This property may be a mathematical object (set, function) what means that this object is assigned to *this* system (and roughly described in common language).

A typical example is the state set,  $\Sigma$ , assigned to a concrete system  $\mathfrak{S}$ , i.e.  $\Sigma(\mathfrak{S})$ . We can describe various states of the system in the common language to demarcate roughly its state set, but it cannot be *defined* within this language. On the other hand, the state set is a primitive concept of our approach and its (approximate) characterization can be done *only* in common language. In this language, the state is more or less identified with some *configuration*, i.e. with a concrete arrangement of the structure<sup>3</sup>. This description cannot be strict – we put it more precisely into a mathematical language (by defining vari-

---

<sup>3</sup>Possibly endowed with vectors determining current velocities of individual parts.

ous sets and their relations) but this step is impossible without a previous pre-understanding done in the common language. If we study systems being unattainable to human experience (atoms, elementary particles, black holes, the Universe, etc.) the concept of state becomes more abstract and its description in the common language is only very rough and approximate.

Nevertheless, any physical description, whatever abstract, carries a trace of understanding it in the language of everyday experience. We speak about an atom as consisting of electrons occupying various configurations around the nucleus though we know that it is a very inaccurate picture. The choice of physical variables is also motivated to be close to our understanding of the state as a mechanical configuration as much as possible. That is why we introduce the concept of ‘part’ and describe systems by specifying them as they are build up from individual parts. Parts, however, need not be some material “pieces” – they may be chosen very abstractedly and artificially (e.g. parts as geometrical points or infinitesimal pieces of matter play important role in classical physics though they can be hardly understood as something approachable by our experience).

## 1.2 Physical systems and subsystems

If we study an object or structure by means of a physical description we call it the **physical system** or simply the **system**. It means that the material object or structure is studied by a systematic physical approach. It may be a box with gas, a car, a culture of living cells, a galaxy or a hurricane but also a molecule or an emitted radiation interacting with surrounding matter. There is a vast number of examples. Physical systems will be denoted by  $\mathfrak{S}$  and derived symbols.

Some parts of the studied object may be studied as physical systems, too. We say that the physical system  $\mathfrak{S}'$  is a **subsystem** of  $\mathfrak{S}$  and write it by means of a relation  $\overset{sub}{\subset}$ , i.e.

$$\mathfrak{S}' \overset{sub}{\subset} \mathfrak{S},$$

that has following simple properties, namely

$$\mathfrak{S} \overset{sub}{\subset} \mathfrak{S} \quad (1.1)$$

and

$$\mathfrak{S} \overset{sub}{\subset} \mathfrak{S}' \wedge \mathfrak{S}' \overset{sub}{\subset} \mathfrak{S} \Rightarrow \mathfrak{S} = \mathfrak{S}'. \quad (1.2)$$

This relation, however, does not only mean that ‘something’ is included in ‘something larger’. As well as the concept of physical system means also a special viewpoint from which material structures are studied, the fact that one physical system is *understood* as a subsystem of the other means that they both may be studied from the same viewpoint. A specification of several subsystems of a system (in common language) thus determines much better the way *in which* we do understand the system.

Therefore, we introduce the concept of a **representative set of subsystems**  $\Xi(\mathfrak{S})$  that includes the system  $\mathfrak{S}$  and such its subsystems which may be studied from the same viewpoint as the system alone. This set is thus only fuzzily defined<sup>4</sup> because usually there is no complete list of all “representative” subsystems<sup>5</sup>. From this point of view, it is describable only in common language. Nevertheless, we will use this concept also

---

<sup>4</sup>Excepting some special cases, e.g. if the system is formed by only a few mechanical parts that form a clearly given set subsystems.

<sup>5</sup>The ‘representative set’ is thus rather a semiset in the spirit of the definition of P. Vopěnka [45], see also Section 3.6.

in the abstract level of description where it means just a suitable set of subsystems (whose elements more or less coincide with “representative” subsystems)<sup>6</sup>.

Another important concept comes from the possibility of understanding two systems as one. We introduce the **composition** of two systems: the composition of a couple of systems  $\mathfrak{S}_1$  and  $\mathfrak{S}_2$  is a system  $\mathfrak{S}$  uniquely defined by the systems  $\mathfrak{S}_1$  and  $\mathfrak{S}_2$ . It is denoted as

$$\mathfrak{S} = (\mathfrak{S}_1, \mathfrak{S}_2)$$

and we accept that

$$(\mathfrak{S}_1, \mathfrak{S}_2) = (\mathfrak{S}_2, \mathfrak{S}_1), \quad (\mathfrak{S}, \mathfrak{S}) = \mathfrak{S}. \quad (1.3)$$

The composition is related to the concept of subsystem by the following relations:

$$\mathfrak{S}_1 \overset{sub}{\subset} (\mathfrak{S}_1, \mathfrak{S}_2), \quad \mathfrak{S}_2 \overset{sub}{\subset} (\mathfrak{S}_1, \mathfrak{S}_2) \quad (1.4)$$

and

$$\mathfrak{S}' \overset{sub}{\subset} \mathfrak{S} \Rightarrow (\mathfrak{S}', \mathfrak{S}) = \mathfrak{S}. \quad (1.5)$$

The composition of two systems resembles the union of two sets and the subsystem looks like subset. Nevertheless, physical systems cannot be understood as sets (what would be their elements?). The conflict is evident when we try to define a counterpart to the set intersection ( $\cap$ ) or complement ( $\setminus$ ). Namely, we can have two systems that have a common subsystem. A counterpart of ‘intersection’ then would be the “largest” common subsystem. It, however, cannot be defined because we have no list of all subsystems.

---

<sup>6</sup>Subsystems that are not representative exist, to be sure, but they have no relation to a chosen *form* of description of the studied system.

The concept of physical system is the basement of any physical description. It names the thing which is intended to be described. This ‘thing’, however, may be undefined or only very fuzzily defined. What means, for example, a hurricane or a living cell as a physical system? Any definition of such systems means more or less artificial separation of some part of reality from everything else. In our approach, such questions play a very important role. We will study them at various levels of description. First, we define the ‘surrounding’ as another physical system that has a deep relation to the studied system. For example, the surrounding air is the surrounding of the hurricane, other cells and an extracellular space is the surrounding of the cell. Though there is no clearly defined “border” between the system and its surrounding an interplay between these two “parts of reality” plays the crucial role.

Second, we define the concept of ‘parts’ of the system. These parts are *defined* “pieces” that *model* the system in a way. For example, the living cell may be understood as consisting of several parts – the membrane, intercellular liquids, cytoskeleton networks, etc. The system usually cannot be identified with a set of all its parts (the cell is not *strictly* defined by its membrane and the above described components) but parts allow *describing* the system by suitable mathematical tools. Especially, by assigning various variables to individual parts.

## Chapter 2

# States and manipulations

Any physical description of a real material structure means a “drastic” simplification of many complications and complex features of the studied object. For example, a piece of metal with an extreme rough, grainy polycrystal structure is represented by a smoothly deformable “continuum”, the Earth is understood as a “mass point” in the celestial mechanics, and so on.

Physical simplifications, however, do not originate only in neglecting details at essentially smaller scales than that of the chosen description but are also connected with a highly simplified influence of *surroundings*. A molecule, for example, is studied as being “alone” in the universe. While the first kind of simplification is well understood and seriously studied by physicists who try to estimate errors or corrections caused by neglecting details, the second is understood differently. We always describe a pure case without an influence of surrounding

and then add (if necessary) some defined forces of interactions as external conditions. We do it although we know that no such “pure cases” exist. This thinking has roots in the Newton’s program of physical description of nature and it is embodied in such concepts as ‘empty space’ (that is what remains if all external fields and objects are removed).

Nevertheless, an exception exists. It is the physical discipline called ‘thermodynamics’. In the thermodynamic description, the surrounding is never neglected or identified only with some empty space<sup>1</sup>. For instance, the Kelvin’s formulation of the second law may be paraphrased as follows [21]:

*No process is possible, the sole result of which is that a body is cooled and work is done.*

The meaning of the “sole result” or that “work is done” can be explained only when there is a communication between the system and its surrounding. Especially the concept of ‘work’ indicates an evident interconnection of a thermodynamic system with its surrounding. The work is important when defining the heat and, substantially, the concept of crucial importance, the entropy. It seems that thermodynamics cannot completely avoid some concepts describing a ‘communication’ of the system with its surrounding<sup>2</sup>.

In this book, the thermodynamic viewpoint is studied with a full respect. We supply to the concept of ‘state of the sys-

---

<sup>1</sup>Surprisingly, similar features are also presented in the Copenhagen formulation of quantum mechanics where the measuring device plays also the role of an external ‘thing’ that is nothing but a certain kind of a surrounding.

<sup>2</sup>In modern formulations of thermodynamics (e.g. [21]), the ‘work’ is replaced by the concept of ‘adiabatic accessibility’ that also cannot be defined without supposing existence of some objects outside the system.

tem' (that is crucial in any physical description) the concept of 'manipulation with the system'. The 'manipulation' is any influence coming from the surrounding of the system. The questions how it can be simplified or whether it may be fully neglected are not self-evident assumptions. They must be seriously studied without a baggage of metaphysical ideas about a fixed spatiotemporal background. An interplay between 'states' and 'manipulations' gives a rich structured description of physical systems that allows us to abandon several prejudices of newtonian physics.

## 2.1 States of physical systems

The 'state' is probably the most important concept in any physical description. Starting with description of any physical system we speak about its "states", changes of "states", suitable "description of states" and possible experimental "determination of states". On the other hand, to say what we exactly mean by the "state of the system" can bring us to terrible problems.

Remark: The problem of explanation of the concept of state is that usually the *state alone* is confused with the *description of the state*. However, the first is more philosophical concept which may be understood as a fact (attitude, viewpoint) that at any time moment the system has its own identity connected to a set of properties, qualities, specific interactions with surrounding, etc. This identity is called the state. At this level, the question if two systems can be at the same state or if one system can reach the same state at a latter moment is problematic and belongs to philosophical considerations (remember Heraclitus: "You cannot step twice into the same river"). There is no physical definition of "identity" or "state alone". For example, we could say that two systems are in the same state if *any* mea-

surement done on both systems gives the same results. But what means “any measurement”? What is, for example, a complete set of measurements saying us whether two glasses of water are in the same state? In description of simple, microscopic systems such a set of all measurements could be determined but, surprisingly, we obtain that measurements performed, say, on two same atoms (in the “same state”) may give different results as implied by quantum mechanics. On the other hand, the *description* of state is clearly defined physical question. We simply look for a suitable set of quantities that characterizes nothing else than the state of the system. But it does not explain the concept of state more than an identity of the system connected to a “set of properties, qualities, specific interactions with surrounding, etc.” So we must be careful. We should be aware that the concept of state is rooted in those philosophical considerations and is based on acceptance of some “identity” or “interchangeability” of the system with another one but our effort should be concentrated into the problem of searching for the best physical description of this (undefined but accepted) identity.

To avoid many conceptual problems with explaining the meaning of statements like a ‘system is in a state’, we will understand the ‘state’ as a primitive, undefined concept. Let  $\mathfrak{S}$  be a physical system. We assume that there is a set  $\Sigma$ , called the **state set**, that is assigned to this physical system, i.e.  $\Sigma(\mathfrak{S})$ . Any element  $s \in \Sigma$  is called the **state** (of the system  $\mathfrak{S}$ ).

At the intuitive level of the description, the state set is explained by a potentiality of the studied object to occur at various configurations (whereas individual configurations are related mutually by the “identity” of studied object). Recall that ‘physical system’ is not a material structure but rather a special viewpoint from which we do study such a structure. The state set is thus not absolutely given but it depends on that

what we *want to study*. If, for example, we study only various movements of a piece of chock the state set is considerable smaller then that arising when we also admit that the chock may be broken or heated, etc. At the level of the description in common language, the state set is thus a property of the systematic physical study what is the meaning of the expression “ $\Sigma(\mathfrak{S})$ ” at that level.

Nevertheless, can we specify at the abstract level of description whether, for example, the state of a glass of water means a ‘macroscopic arrangement’ or a ‘conglomerate of coordinates’ of all molecules? It depends on that *how* we understand the system, i.e. from *which* point of view we look at the studied material object (e.g. if it is understood as a conglomerate of molecules or as a macroscopic body). It has been said that a specification of some representative subsystems helps us to determine better the way *in which* we understand the system. An individual molecule, for instance, is not a representative subsystem of a macroscopically studied object.

On the other hand, states of the state set should express possibilities that are important from the chosen viewpoint. For example, if the body is studied as a macroscopic object then states differing only by coordinates of a molecule are not important possibilities and should not be understood as two states. That is why we relate the state set with the representative set of subsystems<sup>3</sup>,  $\Xi(\mathfrak{S})$ , by the following rule:

*For every  $\mathfrak{S}' \in \Xi(\mathfrak{S})$ , there is a surjective mapping  $f_{sub}$ ,*

$$f_{sub} : \Sigma(\mathfrak{S}) \rightarrow \Sigma(\mathfrak{S}'), \quad (2.1)$$

*uniquely determined by the set  $\Xi(\mathfrak{S})$  and the subsystem  $\mathfrak{S}'$ .*

---

<sup>3</sup>Here, the representative set of subsystems is understood as a well defined set because we use it in the fully formalized level of description.

The meaning of this rule is obvious: for *any* couple of different states of  $\mathfrak{S}'$ , say  $s'_1$  and  $s'_2$ , the states  $s_1$  and  $s_2$  defined by  $s'_1 = f_{sub}(s_1)$  and  $s'_2 = f_{sub}(s_2)$  are *different*, i.e. the state of  $\mathfrak{S}$  is “sensitive” on changes in  $\mathfrak{S}'$ . Hence if we understand that individual molecules are subsystems of the water then the state of water means a special configuration of all molecules. The state of *any* representative subsystem is thus determined by the state of the system.

Though a more detail structure of states from the state set is not introduced a state set may be constructed from other state sets so that its elements are derived from elements of them. An example is the state set of a subsystem whose elements may be understood as being constructed from elements of the system by the mapping  $f_{sub}$ . Another case is given by the following definition:

*We say that systems  $\mathfrak{S}_1$  and  $\mathfrak{S}_2$  are **mutually isolated** if the state set of its composition,  $\Sigma((\mathfrak{S}_1, \mathfrak{S}_2))$ , is a set of all couples of states from  $\Sigma(\mathfrak{S}_1)$  and  $\Sigma(\mathfrak{S}_2)$ , namely*

$$\Sigma((\mathfrak{S}_1, \mathfrak{S}_2)) = \Sigma(\mathfrak{S}_1) \times \Sigma(\mathfrak{S}_2). \quad (2.2)$$

Every state of the composition of mutually isolated systems is thus simply given by states of the both systems. On the other hand, any couple of states may realize a state of the composition. That means that mutually isolated systems do not influence each other, in the sense that the state of one of them is independent of the state of the other.

## 2.2 Manipulations with physical systems

The common physical study emphasizes the search for laws of inner dynamics that govern the system behavior when it is

let alone. In thermodynamics, however, studied systems are mostly not let “alone”. We rather study various manipulations with systems to find how effectively they may ‘produce’ work or how special changes of states can be realized. It is not only due to technological applications of thermodynamics, we cannot avoid some ‘manipulations with the system’ even in definitions of basic concepts, especially the ‘adiabatic accessibility’ that is deeply related to the concept of entropy (see Chapter 6). That is why we define the ‘manipulation’ as another primitive concept. Its meaning, however, is much more general than that used in the standard thermodynamic description.

We assign to the studied physical system  $\mathfrak{S}$  a set  $\mathcal{M}(\mathfrak{S})$  or simply  $\mathcal{M}$  (if there is no problem with misunderstanding). This set is called the **set of manipulations** and its elements are called **manipulations** (individual manipulations will be denoted by small Greek letters). They are defined as follows:

*Any manipulation  $\alpha \in \mathcal{M}(\Sigma)$  is a subset of  $\Sigma(\mathfrak{S}) \times \Sigma(\mathfrak{S})$ , i.e.*

$$\alpha \subset \Sigma(\mathfrak{S}) \times \Sigma(\mathfrak{S}),$$

*so that the set  $\mathcal{M}(\mathfrak{S})$  fulfills following conditions:*

- (i) The **identity manipulation**,  $id = \{[s, s], s \in \Sigma\}$ , belongs to  $\mathcal{M}(\mathfrak{S})$ .
- (ii)  $\emptyset \in \mathcal{M}(\mathfrak{S})$ .
- (iii) For every couple of manipulations  $\alpha, \beta \in \mathcal{M}(\mathfrak{S})$  there is a manipulation  $\gamma \in \mathcal{M}(\mathfrak{S})$  so that whenever  $[s_1, s_2] \in \alpha$  and  $[s_2, s_3] \in \beta$  then  $[s_1, s_3] \in \gamma$ .

The ordered couple of states  $[s, s'] \in \alpha$  is called the **change** under the manipulation  $\alpha$ . For simplicity, the change will be

sometime denoted as  $s_1 \rightarrow s_2$  (without a specification of the manipulation). The second condition describes the fact that if there is a change  $s_1 \rightarrow s_2$  and the change  $s_2 \rightarrow s_3$  there exists a manipulation  $\gamma$  that realizes the change  $s_1 \rightarrow s_3$ . Generally, there may exist many such manipulations. Each of them is called the **composition** of the manipulations  $\alpha$  and  $\beta$ .

We introduce two useful auxiliary concepts. The set  $\alpha|_1$ ,

$$\alpha|_1 \equiv \{s \in \Sigma, \exists s' \in \Sigma : [s, s'] \in \alpha\},$$

and the set  $\alpha(s)$  for every  $s \in \alpha|_1$  as

$$\alpha(s) \equiv \{s' \in \Sigma : [s, s'] \in \alpha\}.$$

The meaning of manipulations is obvious: any ordered couple  $[s_1, s_2] \in \alpha$  denotes a possible change of the state under the manipulation  $\alpha$ . It is worth stressing that it may not be defined for all states, i.e.  $\alpha|_1$  may differ from  $\Sigma$ . Namely, a concrete influence on the system may be applied only in special states. For example, the manipulation defined by “putting the body between two pistons with a given distance  $\Delta x$  and pushing it by a prescribed force” is possible only at states in which the corresponding size of the body has the dimension  $\Delta x$ .

On the other hand, we can have more “results” of the same manipulation under the system at the same state, i.e.  $\alpha(s)$  may have more elements than the only one. The reasons may be various. For instance, the manipulation is “roughly” defined and there are various finer details that “decide” which state will be eventually realized. The other possibility is that there are some inner bifurcations in the system allowing the realization of more states (such as in the problem of buckling, see Fig. 2.1).

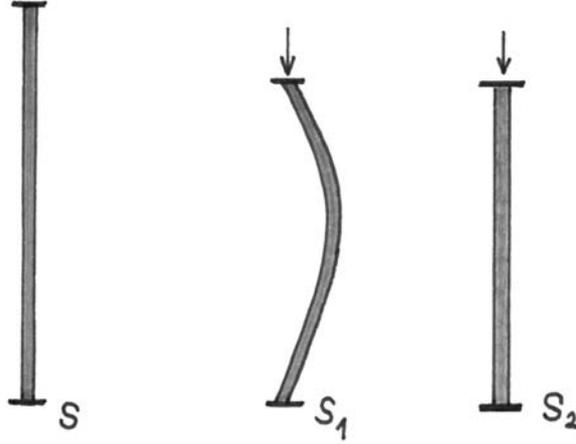


Figure 2.1: The manipulation defined by applying a given load on a bar may cause different deformations of the bar, e.g.  $s \rightarrow s_1$  or  $s \rightarrow s_2$ , (some of them may not be stable, however).

We say that the manipulation  $\alpha$  is **perfect** if  $\alpha$  is a mapping, i.e.

$$\alpha : \alpha|_1 \rightarrow \Sigma.$$

To describe the “roughness” of the manipulation (i.e. how much it differs from a mapping) we define the following concept:

We say that the manipulation  $\alpha \in \mathcal{M}$  is a **parallel composition** of manipulations  $\beta, \gamma \in \mathcal{M}$  if  $\alpha|_1 = \beta|_1 = \gamma|_1$  and  $[s, s_1] \in \beta$  and  $[s, s_2] \in \gamma$  imply that  $[s, s_1], [s, s_2] \in \alpha$ . We denote

$$\alpha = (\beta, \gamma).$$

Obviously, the parallel composition is commutative, i.e.  $(\beta, \gamma) = (\gamma, \beta)$ . The definition of parallel composition of an arbitrary

amount of manipulations  $\beta, \gamma, \dots$  (finite, infinite or even uncountable) is obvious. We denote again  $\alpha = (\beta, \gamma, \dots)$ .

A manipulation  $\alpha$  which gives multiple valued results (i.e. changes one state into two or more different states) may be only roughly defined and we look for finer defined ones (see Fig. 2.2). That is, we search for perfect manipulations whose parallel composition gives the ambiguous manipulation  $\alpha$ . It is a very important activity in experimental physics because we ever try to obtain a unique result of an experiment (i.e. a special manipulation with the system). In theoretical physics, we also prefer a description of the problem (i.e. the choice of suitable variables) in which transformations of states into other states is unique as much as possible. That is why we define:

*The manipulation  $\alpha$  has a **complete decomposition** if there is a set of perfect manipulations  $\alpha(\iota)$  so that*

$$\alpha = (\dots, \alpha(\iota), \dots),$$

*where  $\iota$  is from a linearly ordered set of indices so that individual items in  $(\dots)$  are ordered according to these indices.*

To find a complete decomposition of a manipulation is a very important physical problem which will play the crucial role in this book. It concerns the problem of space that will be studied in Chapter 5.

Remark: The identity mapping does not mean that “nothing has happened”. The change  $[s, s] \in id$  means only that the ‘initial’ and ‘final’ states are the same. On the other hand, we cannot define more than this condition because there is no “finer structure” underlaid the change. We have no ‘time parameter’ that allows us to identify any change with a part of time evolution (i.e. an infinite succession of states parameterized by a time parameter).



Figure 2.2: The manipulation characterized as a “man’s hammer blow” is very roughly defined: there are thousands of ways by which a man can hit with a hammer. A complete decomposition of this manipulation would be, say, a set of precise device that can impact by a strictly defined way.

The concept of manipulation is quite general. It includes also the situation when the system is “let alone”. Namely, it means also a preparation of some special external conditions, namely those that minimize interactions with surroundings. In our approach, however, there is no implicit assumption about such a possibility. It is worth stressing that the study of a system ‘being alone without any external influence’ is a pure idealization that can never be fully guaranteed. It is deeply related with the Newtonian concept of ‘empty space’. We will study these interesting questions in Chapters 5,6.

Manipulations are various external influences with the system. If we manipulate with the system we manipulate with each subsystem, too. The manipulation  $\alpha$  with a system  $\mathfrak{S}$  is represented by a set of couples  $[s, s']$ , where  $s, s' \in \Sigma(\mathfrak{S})$ . If  $\mathfrak{S}'$  is a subsystem of  $\mathfrak{S}$  then the mapping  $f_{sub}$ , given by (2.1), defines the manipulation  $\alpha'$  with  $\mathfrak{S}'$  as the set of cou-

ples  $[f_{sub}(s), f_{sub}(s')]$ . Nevertheless, the set of manipulations with a subsystem may not be completely determined by the set of manipulations with the system. Let us imagine, for example, a molecule inside a macroscopic body. Understanding the body as a system  $\mathfrak{S}$  we can have a broad set of manipulations with this body,  $\mathcal{M}(\mathfrak{S})$ , that may describe various macroscopic influences on the body (defined by forces on its boundary, external fields, heating, etc.). These manipulations, however, may not describe manipulations with the molecule that is defined by completely different processes and actions. If the molecule is understood as a subsystem of  $\mathfrak{S}$  (i.e. the states of  $\mathfrak{S}$  are sensitive on states of the molecule) then there is an evident disproportion between the definition of the state set of  $\mathfrak{S}$  (it is defined very finely) and the definition of the set of manipulations with  $\mathfrak{S}$  which is defined very low. That is why we define:

*Let  $\Xi$  be the representative set of subsystems and  $\mathfrak{S}' \in \Xi$ . We say that  $\mathcal{M}(\mathfrak{S})$  is **consistent** with  $\Sigma(\mathfrak{S})$  if there are no other manipulations in  $\mathcal{M}(\mathfrak{S}')$  than those represented by couples  $[f_{sub}(s), f_{sub}(s')]$  where the set of couples  $[s, s']$  represents a manipulation from  $\mathcal{M}(\mathfrak{S})$ . The consistence of states and manipulations thus defines a surjective mapping  $g$ ,*

$$g: \mathcal{M}(\mathfrak{S}) \rightarrow \mathcal{M}(\mathfrak{S}'), \quad (2.3)$$

so that

$$g(\alpha) = \{[f_{sub}(s), f_{sub}(s')], [s, s'] \in \alpha\}. \quad (2.4)$$

### 2.3 Processes

There is no idea of flowing time in which changes happen. We have no temporal determination underlying the concept of ma-

nipulations. If  $[s_1, s_2] \in \alpha$  then we may say that the state  $s_1$  is “earlier” than the state  $s_2$  but it has no deeper meaning than a formal description of the ordering of two states. It has by no means an absolute validity. The physical system is not “immersed” in an independent temporal background. We even do not suppose an external flow of time. If something like the “time flow” is described it has to be derived from the structure of manipulations and not conversely (see Chapter 5).

This comment is very important in understanding the concept of ‘process’. In thermodynamics, we speak often about various processes because many concepts (like heat) may be defined only in dependence on some processes. Also in our abstract thermodynamic schema, processes play an important role. We also accept the existence of various processes as continual successions of states corresponding to a continual change of states of the system. We, however, do not explain these successions as states in a continually flowing time. Our definition sounds as follows:

*The **process**  $\pi$  is a subset of states,  $\pi \subset \Sigma$ , so that there exists a one-to-one mapping*

$$\pi_{\tau_1, \tau_2} : \langle \tau_1, \tau_2 \rangle \rightarrow \Sigma \quad (2.5)$$

where  $\langle \tau_1, \tau_2 \rangle$  is a closed real interval with  $\tau_1 < \tau_2$ ,

$$\pi_{\tau_1, \tau_2}(\langle \tau_1, \tau_2 \rangle) = \pi,$$

and for any couple  $\tau, \tau' \in \langle \tau_1, \tau_2 \rangle$ ,  $\tau < \tau'$ , there is a manipulation  $\alpha \in \mathcal{M}$  so that  $[\pi_{\tau_1, \tau_2}(\tau), \pi_{\tau_1, \tau_2}(\tau')] \in \alpha$ .

The mapping  $\pi_{\tau_1, \tau_2}$  is called the **parameterization** of the process. The set of possible parameterizations is restricted as follows. If  $\pi_{\tau_1, \tau_2}$  and  $\pi_{\tau'_1, \tau'_2}$  are two parameterizations then the

mapping  $\pi_{\tau_1, \tau_2}^{-1} \circ \pi_{\tau'_1, \tau'_2}$  is a monotone, continuous function from  $\langle \tau'_1, \tau'_2 \rangle$  onto  $\langle \tau_1, \tau_2 \rangle$ . It implies that if  $\pi_{\tau_1, \tau_2}^{-1}(s)$  is an inner point of the interval then it is an inner point in any parameterization.

*The state  $s \in \pi$  is said to be **inner state of the process** if  $\pi_{\tau_1, \tau_2}^{-1}(s)$  is an inner point of the interval  $\langle \tau_1, \tau_2 \rangle$ .*

The parameter  $\tau$  is not a time. It is an arbitrary real parameter that *models* the continual ordering of states. Any process may be always approximated by an arbitrary large set of manipulations and corresponding (finite) succession of changes. Namely, if  $\pi_{\tau_1, \tau_2}$  is a parameterization of a process then for any division of the interval  $\langle \tau_1, \tau_2 \rangle$  onto intervals  $\langle \tau^i, \tau^{i+1} \rangle$ ,  $i = 0, \dots, N$  and  $\tau^0 = \tau_1$  and  $\tau^N = \tau_2$ , we have a set of manipulations  $\alpha^1, \dots, \alpha^N$  so that their composition forms the change from  $\pi_{\tau_1, \tau_2}(\tau_1)$  to  $\pi_{\tau_1, \tau_2}(\tau_2)$ .

Nevertheless, we cannot say that the process is a “continuous change” because no such concept has been introduced. It corresponds to everyday observations: we perceive only changes of a state to another state – the feeling of a continuity arises when we notice that any such change can be decomposed onto finer ones.

The process is always defined for a finite real interval. We would like to characterize processes that are defined on arbitrarily small intervals (that allows to define some concepts assigned to processes more locally). We do it as follows:

*Let  $s$  be an inner state of the process  $\pi$ . The class of processes  $\Pi(\pi, s)$  is called the **tangent of the process  $\pi$  at  $s$**  if for any  $\pi' \in \Pi(\pi, s)$  the state  $s$  is an inner point of the process  $\pi'$  and the intersection  $\pi \cap \pi'$  is a process, too.*

The tangent is thus a class of processes that are identical at  $s$  and its arbitrarily small vicinity.

## 2.4 The surrounding

The origin of any manipulation with the system lies in a surrounding of the system. However, how the ‘surrounding’ can be defined physically? The most natural idea is to identify the surrounding with another physical system, denoted by  $\mathfrak{S}_{sur}$ . The question is how such a system should be specified. If we identify it with something like a ‘rest of the Universe’ we get a highly complicated system without a clear definition and specification. We want rather to work with a surrounding whose structure is as simple as possible.

That is why we define the surrounding as a sufficient “source” of all manipulations with the system. The word ‘source’, however, has to be defined with the help of concepts characterizing any system (i.e. the system  $\mathfrak{S}_{sur}$ , too). How is a manipulation with the system realized? First, we prepare specific external conditions (e.g. prepare a measuring device interacting with the system into a special state or protect the system against an external influence by a barrier, etc.). These specific external conditions may be characterized as a special state  $s_{sur}$  of the system  $\mathfrak{S}_{sur}$  (i.e.  $s_{sur} \in \Sigma(\mathfrak{S}_{sur})$ ). Afterwards we manipulate with the device (e.g. only by keeping the protection) that causes a manipulation with the system. The surroundings of  $\mathfrak{S}$  thus may be characterized as the system  $\mathfrak{S}_{sur}$  so that *any* manipulation with  $\mathfrak{S}$  corresponds to an (initial) state of  $\mathfrak{S}_{sur}$  and a subsequent manipulation from  $\mathcal{M}(\mathfrak{S}_{sur})$ . The complete definition sounds as follows:

*The **surrounding** of a system  $\mathfrak{S}$  is another system  $\mathfrak{S}_{sur}$  which is not a subsystem of  $\mathfrak{S}$  as well as  $\mathfrak{S}$  is not a subsystem of  $\mathfrak{S}_{sur}$ . The set of states and manipulations of  $\mathfrak{S}_{sur}$  is related to those of  $\mathfrak{S}$  so that for every couple  $[s_{sur}, \alpha_{sur}]$ , where  $s_{sur} \in \alpha_{sur}|_1$ ,*

$\alpha_{sur} \in \mathcal{M}(\mathfrak{S}_{sur})$ , there exists a uniquely given manipulation from  $\mathcal{M}(\mathfrak{S})$ . Moreover the mapping  $\mathfrak{F}$ ,

$$\mathfrak{F}: \Sigma(\mathfrak{S}_{sur}) \times \mathcal{M}(\mathfrak{S}_{sur}) \rightarrow \mathcal{M}(\mathfrak{S}), \quad (2.6)$$

defined by this assigning is surjective, i.e. for every manipulation  $\alpha \in \mathcal{M}(\mathfrak{S})$  there exists at least one couple  $[s_{sur}, \alpha_{sur}]$  so that  $\alpha = \mathfrak{F}([s_{sur}, \alpha_{sur}])$ .

That means that for every manipulation from  $\mathcal{M}(\mathfrak{S})$  (i.e. a manipulation with the system  $\mathfrak{S}$ ) there exists a state  $s_{sur} \in \Sigma(\mathfrak{S}_{sur})$  and a manipulation from  $\mathcal{M}(\mathfrak{S}_{sur})$  (i.e. a manipulation with its surrounding) that realizes the manipulation with the system. We say that the manipulation  $\alpha \in \mathcal{M}(\mathfrak{S})$  is **caused** by  $\mathfrak{S}_{sur}$ . This word reflects the (only) **causality** accepted in our approach.

Any couple of different manipulations with  $\mathfrak{S}$  corresponds to different elements of the set  $\Sigma(\mathfrak{S}_{sur}) \times \mathcal{M}(\mathfrak{S}_{sur})$ . Manipulations with the surrounding are thus fair sensitive to cause any manipulation with the system. On the other hand, the couple  $[s_{sur}, \alpha_{sur}]$  may not be defined uniquely (i.e. there may be more “initial” states of the surroundings and more appropriate manipulations with the surroundings). It may mean that the surroundings of the system is needlessly “large”. Maybe, it includes less states and manipulations to realize any manipulation with the system. The following Lemma is useful in this context:

**Lemma:** Let  $\mathcal{M}(\mathfrak{S}_{sur})$  be consistent with  $\Sigma(\mathfrak{S}_{sur})$ . If a subsystem  $\mathfrak{S}'_{sur}$  of  $\mathfrak{S}_{sur}$  is a surrounding of  $\mathfrak{S}$  then  $\mathfrak{S}_{sur}$  is the surrounding of  $\mathfrak{S}$ , too.

**Proof:** The consistence means that for each subsystem  $\mathfrak{S}'_{sur} \stackrel{sub}{\subset} \mathfrak{S}_{sur}$  there is a surjective mapping  $g: \mathcal{M}(\mathfrak{S}_{sur}) \rightarrow \mathcal{M}(\mathfrak{S}'_{sur})$  defined by

(2.3). The assigning described by the function  $\mathfrak{F}$  is then given as

$$\mathfrak{F}([s_{sur}, \alpha_{sur}]) = \mathfrak{F}'([f_{sub}(s_{sur}), g(\alpha_{sur})]), \quad (2.7)$$

where  $[s_{sur}, \alpha_{sur}]$  is an arbitrary element of  $\Sigma(\mathfrak{S}_{sur}) \times \mathcal{M}(\mathfrak{S}_{sur})$ ,  $f_{sub}$  is the function defined by (2.1) and  $\mathfrak{F}'$  is the mapping (2.6) defined for  $\mathfrak{S}'_{sur}$ .  $\square$

The surrounding  $\mathfrak{S}'_{sur}$  introduced in the Lemma is said to be the **simpler surrounding** than  $\mathfrak{S}_{sur}$ .

The surrounding is a source of all manipulations with the system. This may be, however, a source of misunderstandings. When we completely change the surrounding (e.g. when putting a living cell from an organisms into a test tube) the new surrounding may not be a source of the same manipulations as before. However, we cannot forget that the concept of physical system itself is not defined absolutely – it depends also on surrounding (the cell in organisms is different physical system than the same cell in the test tube). Nevertheless, to explain physically what means a “change of the surrounding” while the system “remains the same”, we may assume, for example, that the state set of a “new” physical system, say  $\mathfrak{S}_{new}$ , (formed by a new surrounding) remains the same as that of the original physical system  $\mathfrak{S}$ , i.e.

$$\Sigma(\mathfrak{S}_{new}) = \Sigma(\mathfrak{S}). \quad (2.8)$$

The set of manipulations, however, may be different, i.e.

$$\mathcal{M}(\mathfrak{S}_{new}) \neq \mathcal{M}(\mathfrak{S}). \quad (2.9)$$

Remark: Notice that the state set is the concept that helps us to fix a “material identity” of the system. The change of surrounding, defined by (2.8), thus cannot “destroy” the system in the meaning that some

states “vanish” from the state set or some new ones “appear”. It is exactly that what is supposed in experimental physics – an experiment means a change of a “natural” surrounding of the studied system. Nevertheless, we believe that the system can be in same states as before. Notice also that realizations of experiments is usually intended to enlarge the set of manipulations, i.e.  $\mathcal{M}(\mathfrak{S}_{new}) \supset \mathcal{M}(\mathfrak{S})$  in (2.9).

## 2.5 Global variables

The core of any physical description lies in assigning various quantities to the system. When varying the state of system (i.e.  $s \in \Sigma$  changes to  $s' \in \Sigma$ ) we obtain various values of these quantities and thus they are usually called ‘variables’. We define special variables that are assigned to the whole system.

*The **global variable**  $G$  is a mapping*

$$G : \mathcal{D}_G \rightarrow \mathbf{B}_G, \quad (2.10)$$

*where  $\mathbf{B}_G$  is a Banach space (of this variable) and  $\mathcal{D}_G \subset \Sigma$ .*

The Banach space is a suitable structure for defining values of variables because it is a complete normed vector space<sup>4</sup>. Usually,  $\mathbf{B} = \mathbb{R}^n$ . Notice the important feature of the mapping (2.10): there may be states for which the global variable is not assigned to the system (e.g. ‘equilibrium temperature’ is not defined at non-equilibrium states as illustrated in Fig. 2.3).

Let us assign to the system  $\mathfrak{S}$  a **set of global variables**  $Q_g(\mathfrak{S})$  (or  $Q_g$ ) so that for any global variable  $G \in Q_g(\mathfrak{S})$  the set of states  $\mathcal{D}_G$  is not empty (nevertheless the set  $Q_g(\mathfrak{S})$  itself may be empty). Global variables included in  $Q_g$  characterize

---

<sup>4</sup>That means that its elements have a size (a nonnegative real number) called the norm. The norm of an element  $\mathbf{b}$  is denoted by  $\|\mathbf{b}\|$ , [8].

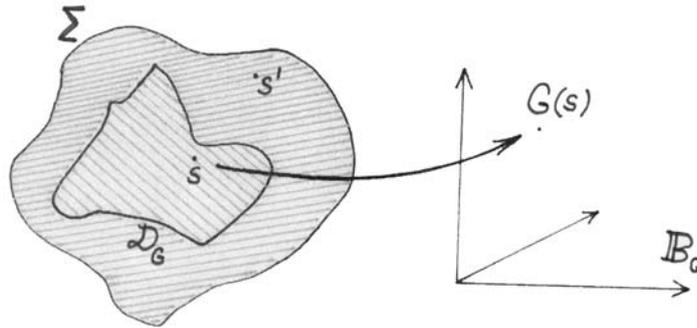


Figure 2.3: A global variable  $G$  is a mapping from an abstract state set  $\Sigma$  to a corresponding Banach space  $\mathbf{B}_G$ . An important feature of the presented approach is that a variable may not be assigned to all states – its definition domain may be only a “small” subset of  $\Sigma$ .

the system as a whole. Their values (or changes of these values when the state of the system is changed) should be determined somehow. This determination is done by various experiments that means a form of “communication” of the system with its surrounding. The value of a global variable (or its change) should be in principle detectable by the surrounding<sup>5</sup>. We introduce two important classes of variables whose values have a clear relation to a suitable class of surroundings.

The first class are global variables whose values only “flow over” the system and its surrounding and their total values remain the same. These variables correspond to some conservative properties, such as energy, momentum, angular momen-

<sup>5</sup>Recall that a change of the surrounding may change the physical system, too. We look for such a surrounding that determines the system for which a corresponding global variable is definable.

tum, mass, electric charge, etc. It motivates us to the following definition:

*We say that the surrounding of a system  $\mathfrak{S}$  has the **conservative property** with respect to the global variable  $G$  if this variable is defined both for the system and the surrounding, and for every  $s_{sur} \in \Sigma_{sur}$ ,  $\alpha_{sur} \in \mathcal{M}_{sur}$  and every couple  $[s, s'] \in \mathfrak{F}(s_{sur}, \alpha_{sur})$  the condition*

$$G(s) + G(s_{sur}) = G(s') + G(s'_{sur}) \quad (2.11)$$

*holds whenever  $[s_{sur}, s'_{sur}] \in \alpha_{sur}$ .*

Notice that the conservative property is not a property of the global variable  $G$  alone but rather its relation to a concrete surrounding. It may seem a little surprising because we understand, say, the momentum as a conservative variable in an absolute meaning. However, if a fixed massive wall is part of the surrounding then the momentum of the system is *not* a conservative variable<sup>6</sup>.

The second class of variables having a clear relation to special surroundings is called the ‘generalized temperature’. In special conditions, values of these variables only “copy” their values on the surrounding and thus may be simply determined. These “special conditions” are called the ‘equilibrium’ and play a very important role in physics. Let us study them in next section.

---

<sup>6</sup>If, for example, the system is a microscopic particle and the surrounding is formed by macroscopic objects (walls, etc.). We may assert that there is a change of the momentum of the wall after the impact of the particle though, but this change is not definable on the macroscopic surrounding. We concede that this kind of thinking is not familiar to standard physics, where such concepts as ‘macroscopic surrounding’ are not taken into account quite seriously.

## 2.6 Temperature and equilibrium

The temperature is a very important physical quantity that is deeply related to the second law of thermodynamics. Namely, its basic property is a “tendency” to be the same over the whole system. Be motivated by this property, we define the concept of generalized temperature as follows:

*Let  $\mathfrak{S}$  be a system and  $\Xi$  the representative set of its subsystems. The global variable  $T$  is called the **generalized temperature** if it belongs into the set of global variables of every subsystem from  $\Xi$ , and has the following property: For every  $s \in \Sigma$  and every subsystem  $\mathfrak{S}' \in \Xi$  holds that if  $s \in \mathcal{D}_T$  then  $f_{sub}(s) \in \mathcal{D}'_T$ , where  $\mathcal{D}'_T \subset \Sigma(\mathfrak{S}')$  is the definition domain of  $T$  on the subsystem, and  $f_{sub}$  is defined by (2.1). Then the relation*

$$T(s) = T(f_{sub}(s)) \quad (2.12)$$

*is always valid.*

The meaning of the definition is obvious. The generalized temperature of a system is defined only if it is defined for every subsystem from the class of subsystems  $\Xi$  of the system and its value equals to values of this global variable on all subsystems from  $\Xi$ .

The equilibrium temperature defined in thermodynamics is the best case of such a global variable. Its high importance is evident: even a special law – the zeroth law of thermodynamics – is formulated to define this variable (see Section 6.5).

Remark: The generalized temperature may not be the thermodynamic temperature. It may be any quantity whose values on the system always match with those on its (representative) subsystems. A nontrivial example is ‘time’ if its values (data measured by a clock) are understood as values of a global variable (why not, we will show

this possibility in Section 6.3). In this case, the definition domain covers all states of the system and the condition (2.12) is trivially valid.

The thermodynamic temperature is defined in equilibrium. Let us explain this concept and show its relation to the generalized temperature. If we want to specify some physical or chemical properties of the water in a vessel we simply take a small sample of water and do corresponding measuring that determine these properties. Afterwards, we have information about the water in the vessel. However, we have obtained properties of *this* sample only and not those concerning the water that *remains* in the vessel. We simply believe that the water in the sample is *the same* as the surrounding water. The similar situation happens when we measure the temperature by a mercury thermometer. The datum of the thermometer is nothing but the temperature of the mercury. We believe again that it corresponds to the temperature of the surrounding air.

Such belief relies on a very important physical assumption. Namely, that there is some *equilibrium* between the thing (physical system) whose properties we really measure (a sample of the water, the mercury within the thermometer, etc.) and its surrounding. The equilibrium then guarantees that *some* physical properties are the same, e.g. temperature, density, pressure, chemical composition, etc.

The concept of equilibrium is thus some relation between the system and its surrounding. It plays a very important role in thermodynamics. In our abstract approach, we will specify a basic feature of this relation, namely the fact that any state of the surrounding defines a concrete state of the system.

The surrounding  $\mathfrak{S}_{sur}$  is said to be in **equilibrium** with the system  $\mathfrak{S}$  if there is a one-to-one mapping,

$$f_{eq} : \Sigma_{sur} \rightarrow \Sigma,$$

so that the mapping (2.6) is then defined by the prescription

$$\mathfrak{F}([s_{sur}, \alpha_{sur}]) = f_{eq}^{\mathcal{M}}(\alpha_{sur}), \quad (2.13)$$

where the mapping  $f_{eq}^{\mathcal{M}} : \mathcal{M}_{sur} \rightarrow \mathcal{M}$  is defined by the condition

$$[s_{sur}, s'_{sur}] \in \alpha_{sur} \Leftrightarrow [f_{eq}(s_{sur}), f_{eq}(s'_{sur})] \in f_{eq}^{\mathcal{M}}(\alpha_{sur}). \quad (2.14)$$

Moreover, if there is a generalized temperature  $T$  defined for  $\mathfrak{S}_{sur}$  then the same global variable is defined for  $\mathfrak{S}$  and

$$T(s_{sur}) = T(f_{eq}(s_{sur})) \quad (2.15)$$

for every state  $s_{sur} \in \Sigma_{sur}$ .

The function  $f_{eq}$  represents the above-mentioned ‘‘coupling’’ of the surrounding with the system. This coupling forms an extremely simple structure of manipulations with the system – the system copies changes of the surrounding.

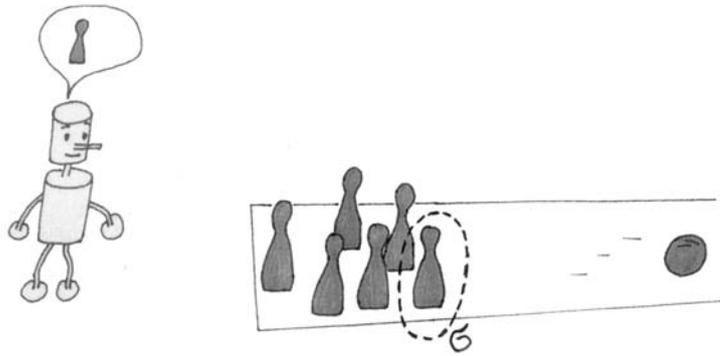


Figure 2.4: The picture may serve as a summary of main concepts introduced in this chapter. We choose a part of reality (e.g. a pin) and study it as a physical system,  $\mathfrak{S}$ . A “material identity” of the system is characterized by its state set,  $\Sigma(\mathfrak{S})$  (we understand what means the ‘falling pin’, ‘laying pin’, ‘pin at a different position’, etc.). However, individual states, *per se*, are not *defined* (they may be only *described* by various variables – see the next chapter). The dynamics of the studied system is defined by various changes – couples of states. As such, they are a property of the system. All possible changes form the set of manipulations,  $\mathcal{M}(\mathfrak{S})$ . The source of all manipulation are various bodies around the system (other pins, a rolling ball, the Earth, etc.). Namely, these bodies may cause various changes of states. All such bodies form another part of reality, a physical system called the surrounding,  $\mathfrak{S}_{sur}$ .

## Chapter 3

# Systems as sets of parts

If we decide to study a thing as a *physical system* we choose a special viewpoint from which the thing is thought of. The basement of our thinking is an idea of some *set of states* that means various “modes of being” of the studied thing. The state set is a field of potentialities of its occurrence: different states means different “configurations” of the *same* thing. But how may this “configuration” be precisely defined and described? That is, how should the concept of state be connected with its description?

In classical physics, the state is characterized by a current distribution of some properties in space. At a moment, the body has a concrete distribution of the mass density or electric charges, velocities of individual parts localized at different spatial points, special spatial configurations of various fields, and so on. A different state is then characterized by *different* distributions of all these properties. This way of understanding depends essentially on the idea of a fixed space. The distribution of properties means the distribution *in space*.

In our approach, however, we try to avoid the concept of space from many reasons. That is why we have to introduce a compensation of this fixed background by something what may also serve as a fixed reference but what is not absolute and unique and, in particular, is not so encumbered by many metaphysical assumptions. We introduce some ‘parts’ that are fixed building blocks of detail description of states of a system. For any physical system, a fixed set of such parts is introduced. This set, in a way, replaces space – states are not characterized by distributions of properties in space but by properties on individual parts. The property, as usual in physics, is described by a variable whose values are points of a mathematical structure (usually a set of real or complex numbers). These variables, also called ‘local variables’, differ from the global variables defined in Chapter 2. While global variables are supposed to represent some fundamental features of studied systems including some relations to their surroundings (conservative properties, equilibrium, etc.), the (local) variables *characterize* quantitatively individual states of the system with respect to a chosen description. They are defined in dependence on the way of description (a choice of the set of parts) and their main purpose is to characterize the state set as well as possible<sup>1</sup>.

What do we mean by ‘parts’? They may be formed by various pieces of matter and identified with a chosen subsystems of the system. For example, individual bodies of a many-body sys-

---

<sup>1</sup>In physics, the state of the system is often fully *identified* with a set of values of all variables that is represented by a point in the so-called *state space*. Here we try to avoid such an identification because then two different levels of the description are mixed together. Namely, the same system with the same state set may be described differently by choosing different sets of parts. It influences the choice of variables and we may obtain different state spaces.

tem may be understood as ‘parts’<sup>2</sup>. Alternatively, they may be artificially chosen as pure auxiliary concepts having no relation to natural “pieces” of the system. An extreme example of such parts are ‘spatial points’ – parts originated from the concept of space as will be explained in Chapter 5. Another example are various ‘boundaries’ used in continuum thermomechanics as it is described in Chapter 4.

Parts may be composed together to form a larger part. The idea of ‘composition’ comes from macroscopic systems that consist of a whole hierarchy of various material parts, that are composed from smaller parts, etc. Nevertheless, its meaning is essentially broader. It is a very important operation in equilibrium thermodynamics, continuum thermomechanics and in such fields of physics where the description of a system at various scales (‘scale’ means a minimal or typical size of the description) is systematically studied<sup>3</sup>. These questions are deeply related to a possibility of a natural definition of the concept of ‘averaging’ that plays the key role in physics of structures including a huge number of “microscopic” degrees of freedom.

The understanding of a physical systems as a set of parts is always a *way of description*. There is no formal reason to be restricted only on macroscopic systems which may be naturally

---

<sup>2</sup>It should be emphasized that if parts are chosen to be individual particles of a many-particle system we do not need the concept of space *in which* they ‘are’. Namely, their positions (as well as velocities and other spatiotemporal characteristics) may be understood as *variables* assigned to them in dependence on concrete states.

<sup>3</sup>The same system is described differently at various scales. The way of transition between these descriptions plays important role and is studied by various methods. The most famous is the so-called renormalization group approach that was systematically introduced by K.G. Wilson for applications in statistical physics and the quantum field theory [46, 47].

understood as actual sets of parts. If we find a sensible way in which, say, an atom may be understood as a set of parts in our way of description then there is no formal reason to refuse the developed conceptual framework. It opens an appealing question if some concepts introduced and used only for macroscopic objects may have a sensible meaning also for microscopic ones.

### 3.1 Structure

Physical systems introduced in Chapters 1,2 are determined only by their global properties. States from  $\Sigma$  are states of the whole system, global variables are assigned to the whole system and the manipulation causes the change with the whole system, too. Only the presence of subsystems indicates that there is some inner structure but there is no idea how subsystems should be put together to form the system. Moreover, there is no clearly defined set of *all* subsystems (of a system). In a standard physical approach a studied system is *localized* in space by occupying a region (whereas the region may consist only from several points as it is in newtonian mechanics of mass points). Hence a ‘substructure’ of the studied system is simply given by “data” obtained from various subparts of this region.

Since the idea of ‘space’ in which the system is located is not introduced in our approach, we choose a ‘structure of description’ of the system by introducing a formal set,  $\mathcal{P}$ , whose elements are various parts. Certain couples of parts (‘disjunct’ parts) may be combined together to form a ‘larger’ part. Thus larger and larger parts arise till a total (‘largest’) part  $p_0 \in \mathcal{P}$ , including any other parts, is made up. Such a part, however, may not exist. Parts and their compositions are understood in an abstract level – there is no spatiotemporal characterization

of the concept of part and composition. We emphasize again that the structure of parts means a *scheme of description* of a physical system which is based on some primitive concepts such as ‘states’, ‘parts’, ‘composition’, etc. Nevertheless, we can imagine many other ways of description underlying other aspects of the studied physical structure.

Let  $\mathfrak{S}$  be a physical system. We introduce an abstract set  $\mathcal{P}(\mathfrak{S})$  called the **set of parts** of this system (the specification of the system will be omitted if there is no problem with misunderstanding). The basic operation is the **composition** of two parts to form the one part. This operation may be defined as a mapping,

$$\text{Comp} : D_{\text{comp}} \rightarrow \mathcal{P}, \quad (3.1)$$

where  $D_{\text{comp}} \subset \mathcal{P} \times \mathcal{P}$ . Parts  $p_1, p_2 \in \mathcal{P}$  forming the couple belonging to the domain of Comp, i.e.  $[p_1, p_2] \in D_{\text{comp}}$ , are said to be **disjunct**<sup>4</sup>. The composition of parts  $p_1, p_2$  is a part that is denoted as  $(p_1, p_2)$ . The following rules describe the basic properties of the composition operation:

- (i)  $[p_1, p_2] \in D_{\text{comp}} \Rightarrow [p_2, p_1] \in D_{\text{comp}}$  and  $(p_1, p_2) = (p_2, p_1)$
- (ii)  $[p_1, p_2], [(p_1, p_2), p_3] \in D_{\text{comp}} \Rightarrow [p_2, p_3], [p_1, (p_2, p_3)] \in D_{\text{comp}}$  and  $((p_1, p_2), p_3) = (p_1, (p_2, p_3))$ .

The second property allows defining the composition of more mutually disjunct parts, namely

$$(p_1, p_2, p_3) \equiv ((p_1, p_2), p_3), \quad \text{etc.}$$

---

<sup>4</sup>There is an important difference between the composition of systems and the composition of parts. The first may be defined for an arbitrary couple of systems while the later only for disjunct parts.

The set of parts endowed with the mapping  $\text{Comp}$  is called the **structure** of the system  $\mathfrak{S}$ . The structure thus may be understood as a couple  $(\mathcal{P}, \text{Comp})$ .

We say that the part  $p'$  is a **subpart** of the part  $p$  if there is a part  $\tilde{p} \in \mathcal{P}$  so that  $p = (p', \tilde{p})$ . Then we write

$$p' \overset{sp}{\subset} p.$$

Notice that  $p' \overset{sp}{\subset} p$  means that there is another part, say  $p''$ ,



Figure 3.1: Parts resemble a suitable chosen “bricks” that may be put together to model the structure of the system as much as possible. A physicist describing a system as a set of parts thus resembles a child trying to find a suitable composition of puzzle blocks.

that, put together with  $p'$ , gives  $p$  (part is not a subpart of itself!). Parts thus resemble bricks that are put together to form the system. Let us define several useful concepts:

The part  $p_0$  is called the **total part** if for every  $p \in \mathcal{P}$ ,  $p \neq p_0$ ,

$$p \overset{sp}{\subset} p_0.$$

The part  $p$  is said to be **atomic** if there are no parts that are subparts of  $p$ . The set of all atomic parts from  $\mathcal{P}$  is denoted as  $\mathcal{P}_{atom}$ .

If the part includes a subpart then it is said to be **composed**.

The structure is said to be **atomic** if any  $p \in \mathcal{P}$  is a composition of parts from  $\mathcal{P}_{atom}$ .

The structure is said to be **finitely atomic** if it is atomic and  $\mathcal{P}_{atom} = \{p_1, \dots, p_N\}$ .

The structure that does not include any atomic parts, i.e.  $\mathcal{P}_{atom} = \emptyset$ , is called the **non-atomic structure**.

In non-atomic structures there are only composed parts. It means that any part  $p \in \mathcal{P}$  includes an infinite succession of subparts  $p_1, p_2, \dots$  so that

$$\dots \overset{sp}{\subset} p_{i+1} \overset{sp}{\subset} p_i \overset{sp}{\subset} \dots \overset{sp}{\subset} p. \quad (3.2)$$

In atomic structures, every part from  $\mathcal{P}$  is a composition of finite number of parts from  $\mathcal{P}_{atom}$  because the composition is defined only for a finite set of parts. The situation when parts are composed from an infinite number of atomic parts cannot be formalized in our approach<sup>5</sup>. Instead we define:

The structure is said to be **weakly atomic** if every composed part from  $\mathcal{P}$  includes a part from  $\mathcal{P}_{atom}$  as its subpart.

Though weakly atomic structures include an “underlay” set of atomic parts, they have some features of non-atomic ones as shown in the following Lemma.

**Lemma:** If the structure is weakly atomic, but not atomic, then there is a part  $p \in \mathcal{P}$  that includes an infinite succession of composed parts,  $p_1, p_2, p_3, \dots$ , fulfilling (3.2).

---

<sup>5</sup>We would like to avoid the use of a form of transfinite induction, etc.

**Proof:** Let us take an arbitrary composed part  $p \in \mathcal{P}$ . If structure is weakly atomic then there is  $p' \in \mathcal{P}_{atom}$  so that  $p = (p', p_1)$ . If the part  $p_1$  is composed then there is an atomic part  $p'_1$  so that  $p_1 = (p'_1, p_2)$ . If  $p_2$  is composed we go further and obtain a chain of parts  $p_i \overset{sp}{\subset} \dots \overset{sp}{\subset} p_1 \overset{sp}{\subset} p$  till the last part of the composition is atomic too. This situation, however, cannot happen for every composed part of  $\mathcal{P}$  because then every part would be a composition of atomic parts. Therefore there is a part for which the chain is an infinite succession fulfilling (3.2).  $\square$

Let us present several examples.

**Structure of mass points:** We have a set of  $N$  mass points (Newtonian particles). Individual particles form atomic parts  $p_1, \dots, p_N$  and the structure is finitely atomic. This structure is the most often used structure in newtonian mechanics, statistical physics and so on. Parts arising by compositions of atomic parts (i.e. consisting of more mass points) play important role when various interactions of particles are described. Special variables describing these interactions (e.g. an interaction energy) are assigned to couples (or triplets or larger collections) of particles.

**Structure of equilibrium cells:** The structure is formed by  $n$  atomic parts – simple thermodynamic systems in equilibrium (see Section 4.1). It is used in equilibrium thermodynamics. It is also finitely atomic. The fact that compositions of two atomic parts may form another part that is described by same means as atomic ones is crucial – see the beginning of the Callen's book [2] where the operation of removing of some inner constrains (e.g. a barrier) explains main features of thermodynamics. The removing of constrains is nothing but a formation an equilibrium cell as a composition of several atomic parts.

**Structure of spatial points:** Let  $\mathcal{P}$  be identical with  $d$ -dimensional Euclidean space,  $\mathbb{E}_d$ , i.e. parts are points of  $\mathbb{E}_d$ . Since points have no subparts the structure consists only of atomic parts and is atomic (but not finitely atomic). This structure plays crucial role when describing the system by spatial fields (see Chapter 5).

**Structure of sets of spatial point:** If the previous structure is enriched by various sets (finite or infinite) of spatial points it becomes weakly atomic. This structure is seldom used in physics.

**Structure of regions:** We understand the ‘region’ as an arbitrary measurable subset of points of  $\mathbb{E}_d$  so that its  $d$ -dimensional Lebesgue measure (volume) is nonzero. Let us have a region  $\Omega$  and form the set of all subsets of  $\Omega$  that are regions, too. A region  $\Omega_1$  is a subpart of a region  $\Omega_2$  if  $\Omega_1$  is a proper subset of  $\Omega_2$ . This structure has no atomic parts because any region includes a proper subset that is a region, too. This structure plays eminent role when modeling continuum systems (see Chapter 4).

The structure (parts and their compositions) may also *model* the fact that the one physical structure has usually an infinite number of levels of description; whereas the ‘level of description’ means a scale under which all minute details are neglected or, say more precisely, *averaged*. The composition operation thus represents the passage to less accurate descriptions of the system. But why do we try to model cascades of more and more inaccurate descriptions? The reasons are two, at least. First, each physical description of a concrete system begins with neglecting or smoothing out some details (inner structure, microscopic degrees of freedom, some processes) at lower scales. Even a description at a deep microscopic level has to neglect

something (quarks, fluctuations of a space-time foam or something else). Therefore we are ever somewhere “between” a more detail or more inaccurate description and the understanding of steps ‘up’ or ‘down’ is of the crucial importance. Second, we are not able to model complex systems at an arbitrarily detail level and the search for a “compromise” description, which models appropriately behavior of the systems, plays the key role in applied physics. The systematic way of such a search means the understanding of the passage to less accurate description at larger scales.

Atomic structures play important role in physical thinking. They model the situation in which there is a scale that has an eminent meaning. Processes *under this scale* have small or negligible influence on the global behavior of the studied system. The description *at this scale* thus seems to be the fundamental one and usually means a conscious or unconscious suppressing of other descriptions. The discovery of the atomic structure of matter is a good example – we believe that there is a possibility to explain all macroscopic behavior at this fundamental level. Nevertheless, we have no method how to do it actually and that is why we eventually construct more inaccurate (but realistic) descriptions at higher scales.

### 3.2 Parts as subsystems and other parts

Introducing the concept of parts and their compositions is motivated by the mechanical understanding of the system as a structure consisting of various parts (and the state as a configuration of these parts). Physics follows this thought line not only in mechanics and thus physical variables are assigned to individual parts of the system, whatever is their concrete meaning.

This procedure enables us to “catch” the system by sophisticated mathematical tools. A typical example is the approach used in continuum theories, where a decomposition of the system into a set of infinitesimal cells allows us to describe states of the system by a suitable chosen set of spatial functions. The set of parts *defines* the description of the system and the ‘system’ (described only in the common language) thus becomes a mathematically definable structure (see Fig. 3.2). Parts themselves,

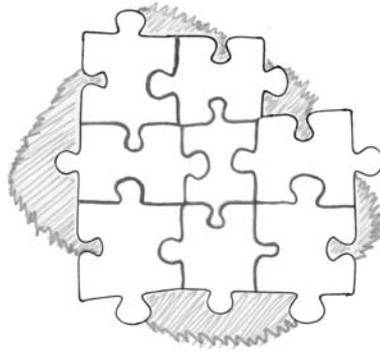


Figure 3.2: The system is illustrated by a fuzzy region to express its only “verbal” definition by a phrase like “let us study a hurricane” or “we describe a living cell”, etc. From the viewpoint of a precise mathematical language, the ‘system’ is not clearly described and defined. Therefore a structure of parts is introduced to “cover” them as well as possible to form a definable structure.

however, *may or may not* be identified with certain subsystems of the system. For example, a part in the mentioned continuum description (an infinitesimal cell) can be hardly identified with a macroscopic subsystem (in continuum description the system as well as any representative subsystem has to be macroscopic).

Any subsystem is a full-valued system which implies that it has its state set. On the other hand, the state set for a part is not generally defined (contrary to the so-called *state space* which is defined for individual parts, see Section 3.5). We only assign values of some variables to parts in dependence on the state of the *whole system*. It may be a source of various misunderstandings. Namely, we need not suppose that the part is in a (physically defined) state because the “state of part” is not introduced and have no sense for parts that are not subsystems.

We divide the set of parts into two subsets,

$$\mathcal{P} = \mathcal{P}_{sub} \cup \mathcal{P}_{ch}. \quad (3.3)$$

The first set,  $\mathcal{P}_{sub}$ , may be identified with a subset of the set  $\Xi$  (a set of representative subsystems of the studied physical system) and some supplementary parts. The second set,  $\mathcal{P}_{ch}$ , includes some other parts, whose definition will be concretized lately.

First, let us define the part  $\mathcal{P}_{sub}$ . The idea is as follows. There is a set  $\Xi'$ ,  $\Xi' \subset \Xi$ , and a one-to-one mapping  $F_s$  from  $\Xi'$  into a subset of  $\mathcal{P}$ ,

$$F_s : \Xi' \rightarrow \mathcal{P}, \quad (3.4)$$

so that for every  $\mathfrak{S}_1, \mathfrak{S}_2 \in \Xi'$  the following relations are valid. Namely, if  $\mathfrak{S}_1 \neq \mathfrak{S}_2$  then

$$\mathfrak{S}_1 \overset{sub}{\subset} \mathfrak{S}_2 \Leftrightarrow F_s(\mathfrak{S}_1) \overset{sp}{\subset} F_s(\mathfrak{S}_2), \quad (3.5)$$

and if  $p_1, p_2 \in F_s(\Xi')$  then

$$[p_1, p_2] \in D_{comp} \Rightarrow (p_1, p_2) = F_s(F_s^{-1}(p_1), F_s^{-1}(p_2)) \quad (3.6)$$

(notice that  $(F_s^{-1}(p_1), F_s^{-1}(p_2))$  is a composition of subsystems). The relation (3.5) implies that there must be a part  $p \in \mathcal{P}$  so that  $F_s(\mathfrak{S}_2) = (F_s(\mathfrak{S}_1), p)$  (see the definition of subparts). However, no subsystem corresponding to the part  $p$  may be included in  $\Xi'$ . That is why we complete the set  $F_s(\Xi')$  by adding all such parts. It forms the set  $\mathcal{P}_{sub}$ . That means that if  $p \in \mathcal{P}_{sub}$  then either  $\mathfrak{S}' \in \Xi'$  exists so that  $p = F_s(\mathfrak{S}')$  or there exist  $\mathfrak{S}_1, \mathfrak{S}_2 \in \Xi'$  so that  $F_s(\mathfrak{S}_2) = (F_s(\mathfrak{S}_1), p)$ . In the second case we will say that  $p$  is a **complement part**.

If  $\mathcal{P}_{ch} = \emptyset$ , parts from  $\mathcal{P}$  model individual subsystems or complement parts. Contrary to subsystems alone, the concept of part has features of spatial determination – if  $p_1$  and  $p_2$  are disjunct (i.e. the couple  $[p_1, p_2]$  belongs to  $D_{comp}$ ) then they may be understood to be spatially separated. The relation (3.6) then transfers this property onto subsystems (which are assigned to some parts).

Parts from  $\mathcal{P}_{sub}$  are defined by a set of subsystems. It means that they have a natural physical meaning (parts corresponding to subsystems have clear physical meaning, complementary parts are in fact defined by couples of subsystems). Parts from  $\mathcal{P}_{ch}$ , if any, do not have such a natural status. They may be said to be chosen *artificially* for purposes of the description. Let us explain it in the case when the total part,  $p_0$ , belongs to  $\mathcal{P}_{ch}$ , i.e. if it cannot be identified with the system alone by the relation  $F_s(\mathfrak{S}) = p_0$ . Imagine, for example, a closed, firmly chosen container including a piece of a special substance, say a liquid. The liquid can be in a vast number of states. For example, it can occupy various parts of the container, change its phase and become a gas expanding to the whole container, etc. Though the system is the liquid, we usually use the description whose variables are defined rather for the container alone than

for the liquid, e.g. by understanding *its* volume as a variable. In our language, the container defines the total part  $p_0$  which is not identical with the system alone.

Remark: There are some puzzling questions whose origin lies in identification of the system with a defined total part. For instance, the entropy of ideal gases depends on their volumes (which has an effect if two gases are mixed together). But if we identify the system - ideal gas - with a collection of its molecules, the dependence of the entropy on the volume is puzzling – is the entropy a property of empty space between molecules?<sup>6</sup> In thermodynamics, the difference between the system (as a material “holder” of physical properties) and externally defined parts is important and has to be clearly defined. Several problems connected with the definition of entropy lie in this difference.

We generalize this example by postulating that parts from  $\mathcal{P}_{ch}$  mean something *externally defined*, chosen for purposes of the most suitable description. The external definition may be realized by various pistons, walls, special boundaries or barriers controlling (and defining) some parts of the system exogenously. All such “machinery” namely should be understood more as belonging to the surrounding than to the system alone. That is why we accept that the set of parts  $\mathcal{P}_{ch}$  is a *property* of the surrounding, i.e.

$$\mathcal{P}_{ch}(\mathfrak{S}_{sur}),$$

what means that these parts are describable by concepts concerning the surrounding. It may seem strange if the artificially defined parts from  $\mathcal{P}_{ch}$  are only mentally delimited by some thought barriers somewhere inside the system. However, as we

---

<sup>6</sup>An evident manifestation of this puzzle is the so called Gibbs paradox.

will see later, the possibility of doing that is supported by the idea of (empty) space. And space alone may be understood as a very special kind of surrounding. It will be shown in Chapter 5.

### 3.3 Local variables

Global variables express some basic quantitative characterization of individual states of the system. They are introduced on a rather abstract level of the description because their definition relies only on the idea of state. When introducing the structure of the system – a set of parts that are chosen to *define* the description of the system – a new sort of variables may be introduced. We call them the ‘local variables’ or only ‘variables’.

The set of parts,  $\mathcal{P}$ , is *fixed*, i.e. it does not depend on the state of the system (it is an important property of the description using space as a *fixed* background – parts only generalize this conception). The dependence of the structure of system on its state is then expressed by various values of these variables that are assigned to individual parts.

Remark: Imagine a physical system formed by a water (in a vessel) and a piece of floating ice. The ice is a clearly definable subsystem of the system. The state set of the system includes also states when there is no ice (because it has melted). At these states, this subsystem becomes indefinable (we may identify it with a collection of molecules dispersed somewhere in the water but it may be hardly understood as a good definition of a subsystem). Nevertheless, the piece of ice may be defined as a part from  $\mathcal{P}$ . There are various variables assigned to this part (mass, momentum, energy, temperature, etc.). At states when there is no ice, this part exists, but values of these variables correspond to the way in which we understand this

part. E.g. they are zero or equal those of the water, and so on. This approach seems rather strange and artificial, but notice that we do the same when working with spatial background (space exists permanently only values of physical variables at its individual points may change and sometime have some special values, e.g. zero mass density, etc.).

We define the ordered set  $Q$  called the **set of variables**. An element  $\hat{q} \in Q$ , called the **variable** or **local variable**, defines for every state of the system  $s \in \Sigma$  the mapping  $q_s$ ,

$$q_s : \mathcal{D}_{q,s} \rightarrow \mathbf{B}_q, \quad (3.7)$$

where  $\mathbf{B}_q$  is a Banach space whose elements are values of the variable  $\hat{q}$  (e.g.  $\mathbf{B}_q = \mathbb{R}^n$ ). The domain of the mapping,  $\mathcal{D}_{q,s}$ , is a subset of  $\mathcal{P}$ .

The question how large is the domain  $\mathcal{D}_{q,s}$  plays the crucial role. Namely, the fact that the composition of parts forms another part is thus connected with the question whether a variable can or cannot be defined for this composition (i.e. if the composition of parts belongs into the domain, too). It is of the crucial importance. The repeated composition is a passage to larger scales. Variables which “survive” this passage are those which play crucial role in the macroscopic description of matter.

The question, if a (local) variable may be identified with a global one, is nontrivial. For example, the global variable of the crucial importance is the (global) energy. We believe that it may be assigned to any state of an arbitrary system (at least in classical physics) and we are able to measure its changes during various processes. On the other hand, we are able to define the (local) energy of various parts of the description (e.g. as the sum of kinetic and potential energy of a particle). Nevertheless,

the definition of a concrete form of the local energy may not be in agreement with results of those measurements. Then we have to search for a better form of local energy and (maybe) a better choice of a set of parts.

The variable  $\hat{q}$  may be formally enlarged to a global variable  $G \equiv G(\hat{q})$  if the structure includes the total part  $p_0$ . Namely, the total part is defined uniquely that allows defining uniquely the mapping (2.10) by choosing  $\mathcal{D}_G = \{s \in \Sigma, p_0 \in \mathcal{D}_{q,s}\}$  and defining

$$G(\hat{q})(s) = q_s(p_0). \quad (3.8)$$

The essential problem, however, is to find a local variable corresponding to a given global one. Namely, global variables express some fundamental characterization of systems and our choice of local variables should correspond with that. We could define formally the mapping (3.7) for  $\mathcal{D}_{q,s} = \{p_0\}$  (if there is a total part) by the prescription (3.8) but the variable has to be defined also for smaller parts (as guaranteed by the continuity rule formulated afterwards). The search for local variables corresponding to global ones is a very important problem of physics.

The way in which concrete values of a variable are assigned to individual parts is a complex physical problem which may be studied only in concrete cases. Generally, a different way of this assignment is expected if the part is from  $\mathcal{P}_{sub}$  or  $\mathcal{P}_{ch}$ . In the first case, the variable is assigned to a subsystem and thus it may (but may not) correspond to a global variable defined on the subsystem (e.g. the energy of subsystem in continuum thermomechanics). If the part is from  $\mathcal{P}_{ch}$  it is defined externally by means of a surrounding and thus values of a variable defined on this part should be also defined by means of this surrounding (we will discuss it in more detail in Chapter 4).

Remark: One concrete physical body (e.g. a piece of chalk) may be described by an infinite number of various choices of a set of parts and a set of global and local variables. We can choose an extremely rough description representing the system by one part (e.g. when studying it as a rigid body) or a very fine description in which the system is understood as an infinite set of parts (in a continuum description). The choice of a set of global and local variables depends on the set of states (e.g. when supposing only pure mechanical displacements in space or if the system may change temperature, may be broken to several pieces and so on), but neither  $Q_g$  nor  $Q$  strictly depend on  $\Sigma$ . Namely, the chosen set of variables expresses (as well as a splitting the system into parts) a degree of simplification of a real structure. We can accept a broad variety of states but describe them by a small set of variables. On the other hand, a rich group of variables may show itself as superfluous because there are no meaningful changes of some of them on the variety of studied states.

Some special kinds of variables are very important to be defined:

*The variable  $\hat{q}$  is said to have the **averaged property** if for any couple of disjunct parts  $p_1, p_2$  the following implication is true in any state  $s$ , namely*

$$p_1, p_2 \in \mathcal{D}_{q,s} \Rightarrow (p_1, p_2) \in \mathcal{D}_{q,s}. \quad (3.9)$$

Variables with averaging property thus “survive an arbitrary long trip” to higher scales and play an eminent role in description of macroscopic systems. A very important class of variables with averaging property is formed by the so-called additive variables. They are defined as follows:

*The **additive variable**  $\hat{q}$  is a variable with averaging property (i.e. it obeys (3.9)) that fulfills moreover the following relation*

for every state  $s \in \Sigma$ ,

$$q_s(p_1, p_2) = q_s(p_1) + q_s(p_2), \quad (3.10)$$

where we omit the double parenthesis and write simply  $(p_1, p_2)$  in place of  $((p_1, p_2))$ .

Now we define the relation of two special variables that is very important when parts are modeled by a set structure:

Let values of a variable  $\hat{q}'$  be only nonnegative real numbers (i.e. for every state and every part,  $q'_s(p) \geq 0$ ). We say that  $\hat{q}$  is **absolutely continuous** with respect to the variable  $\hat{q}'$  if for every state  $s \in \Sigma$  the following statement is valid: For every  $\epsilon > 0$  exists  $\delta > 0$  so that for every part  $p \in \mathcal{D}_{q,s} \cap \mathcal{D}_{q',s}$  for which  $q'_s(p) < \delta$  holds that  $\|q_s(p)\| < \epsilon$  where  $\|\cdot\|$  is the norm in  $\mathbf{B}_q$ . We denote

$$\hat{q} \ll \hat{q}'.$$

The meaning of absolute continuity is obvious<sup>7</sup>. If  $\hat{q} \ll \hat{q}'$  then the value of  $\hat{q}$  at an arbitrary part tends to zero whenever the value  $\hat{q}'$  tends to zero. Additive variables play important role for systems whose parts are modeled by some sets. Such modeling (studied in Chapter 4) allows then to define various measures on parts and additive variables are good candidates for such “measures”. The concept of absolute continuity is then very important when comparing various measures mutually.

---

<sup>7</sup>The concept of absolute continuity originates in the measure theory where it is defined as a comparison of different measures, e.g. [36], see Chapter 4.

### 3.4 Process variables

There is no problem of assigning a variable to a process. Let  $\pi$  be a process. For a concrete parameterization  $\pi_{\tau_1, \tau_2}$ , every variable  $\hat{q}$  and every part  $p \in \mathcal{D}_{q,s}$  where  $s \in \pi$ , we may define

$$q(\tau) \equiv q_{s(\tau)}, \quad (3.11)$$

where  $s(\tau) = \pi_{\tau_1, \tau_2}(\tau)$  and  $\tau_1 \leq \tau \leq \tau_2$ .  $q(\tau)$  is thus a mapping

$$q(\tau) : \langle \tau_1, \tau_2 \rangle \rightarrow \mathbf{B}_q. \quad (3.12)$$

It models an “evolution” of the local variable  $\hat{q}$  in the parameter  $\tau$  which means nothing but values of the variable  $\hat{q}$  on states from the set  $\pi \subset \Sigma$ .

On the other hand, there are also variables that cannot be assigned to individual states alone but only to certain processes. A typical example is the heat. The (exchanged) heat is defined only in a processes and not at individual states (we have no ‘content of heat’ in the body). Nevertheless, there is a possibility of defining slightly modified state variables that correspond in a way to such variables. The idea is as follows. For a state  $s$  and any process  $\pi$  that includes  $s$  as its inner state, there is the tangent  $\Pi(s, \pi)$ . It may be understood as the process  $\pi$  in an “infinite vicinity” of  $s$ . If we assign variable to the tangent we mean that it depends only on what “happens” (“during” the process  $\pi$ ) at  $s$  and no at other states.

*The variable  $\hat{q}^P$  is called the **process variable** if it defines for a state  $s \in \Sigma$  and a tangent  $\Pi$  at  $s$  the mapping*

$$q_{s, \Pi} : \mathcal{D}_{q, s, \Pi} \rightarrow \mathbf{B}_q. \quad (3.13)$$

The domain  $\mathcal{D}_{q, s, \Pi}$  is, similarly as for ordinary (local) variables, a part of  $\mathcal{P}$ . The concept of process variable is more general

then that of (local) variable. Namely, any (local) variable may be understood as a process variable whose domain  $\mathcal{D}_{q,s,\Pi}$ , as well as the mapping  $q_{s,\Pi}$ , do not depend on  $\Pi$ .

### 3.5 State space

The **state space of a part**  $p$ , denoted by  $\Gamma(p)$ , is defined as

$$\Gamma(p) = \{[q_s(p)]_Q, s \in \Sigma\}, \quad (3.14)$$

where  $[q_s(p)]_Q$  is an ordered set (e.g. a vector) of values  $q_s(p)$  of all variables<sup>8</sup> from  $Q$  so that  $p \in \mathcal{D}_{q,s}$ . The definition is obvious – we simply form a set of all values of all variables which may be assigned to a part of the system in all states.

Values of any variable on the part  $p$  then may be understood as values assigned to points of the state space  $\Gamma(p)$  by the following prescription,

$$q(X_p) = q_s(p), \quad (3.15)$$

where  $X_p = [q_s(p)]_Q$  and (3.15) thus means a choice of corresponding “coordinates” of the “vector”  $[q_s(p)]_Q$ .

Let the structure of the system be finitely atomic. Then the **state space of the system**  $\Gamma$  is defined as

$$\Gamma = \Gamma(p_1) \times \dots \times \Gamma(p_N), \quad (3.16)$$

where  $\{p_1, \dots, p_N\} = \mathcal{P}_{atom}$ . Every point  $X \in \Gamma$  thus represents values of all variables on all atomic parts of the system.

---

<sup>8</sup>Values of process variables may be included into the definition of state space, too. But only if there is a unique process – for example velocities defined by the time evolution.

If the structure is atomic but not finitely atomic, then there is a straightforward possibility of generalizing the concept of state space by replacing  $N$ -tuples of values from state spaces of atomic parts by a mapping assigning to individual atomic parts these values. Let us suppose that the state space of all atomic parts (3.14) is the same, i.e. there is a given set  $\Gamma_0$  so that for every  $p \in \mathcal{P}_{atom}$

$$\Gamma(p) = \Gamma_0. \quad (3.17)$$

Then the state space  $\Gamma$  may be introduced as a set of mappings,  $\Gamma$  so that any its element  $f$ ,

$$f : \mathcal{P}_{atom} \rightarrow \Gamma_0, \quad (3.18)$$

represents values of all variables on all atomic parts.

In a standard physical description, the state space is a basic concept that plays the role of our state set. Namely, if we assume that every couple of different states,  $s_1, s_2 \in \Sigma$ , is represented in  $\Gamma$  by a couple of *different* points,  $X_1, X_2 \in \Gamma$ , then the state space may play the role of the state set. This assumption, however, depends on the choice of a sufficiently rich set of variables (points of state space are defined by values of variables). It motivates us to the following definition:

*We say that the choice of the set of local variables,  $Q$ , on the system with atomic structure is **complete** with respect to a state set  $\Sigma$  if every couple of different states from  $\Sigma$  is represented by different points of the state space  $\Gamma$ .*

The completeness of the set of local variables is a very important assumption. It allows us to characterize an abstract state set by a mathematical structure (for example the set  $\mathbb{R}^{kN}$  where  $k$  is the number of components of the “vector”  $[q_s(p)]_Q$  or a special class of mappings). That is why we always want to have a

complete set of variables. Its determination, however, may be a very complicated problem. It means to find a sufficient set of variables that characterize completely all states of the system. The search for such a set is a full-valued physical problem.

Why is the state space defined only for atomic structures? Namely, these structures have a clearly defined set of some “basic” parts so that any part of the structure is a composition of these parts. Hence, if we have information about these parts (values of all variables) then we believe that we have a complete information about the whole system. This belief is substantial and have the essential influence on the whole physics. In our approach, however, there cannot be excluded cases when some important information about the state of the system “emerges” on some larger parts only. That is why we introduce an important rule that excludes such a situation. It will be studied in the next section.

But what happens if the structure is not atomic? Then there are infinite “chains” of subparts of subparts (3.2) and no clearly defined “basic” parts exist. The state space may be introduced only for concrete parts but not for a whole system. That is why we search for a more sophisticated description in which an atomic structure is introduced. Since there are no some “smallest” subsystems, we construct atomic parts artificially, i.e. atomic parts are included in  $\mathcal{P}_{ch}$ . This procedure is deeply related to the concept of space, where these additional atomic parts are spatial points. We will study it in detail in Chapter 5. Another possibility of constructing an artificial set of atomic parts makes use of a representation of parts of the structure on a system of sets. This procedure, however, may not work perfectly because a given set of variables may not be complete. It is studied in Chapter 4.

### 3.6 The continuity rule

For the time being, we have not embodied in our abstract schema the fact that the description of the system on a finer set of parts means a better information about the state of the system. The information about any state of the system is given only in a form of the set of values of all variables on individual parts. If we have two disjunct parts,  $p_1, p_2 \in \mathcal{P}$ , and its composition  $p_3 = (p_1, p_2)$ , we have at a state  $s$  values of a variable  $\hat{q}$  on those parts,  $q_s(p_1)$ ,  $q_s(p_2)$  and  $q_s(p_3)$  (if they are defined). It is obvious that information about the composition  $(p_1, p_2)$  should be better if we have information about its subparts,  $p_1$  and  $p_2$ , individually. It is expressed also by the fact that the set of numbers “ $\{q_s(p_1), q_s(p_2)\}$ ” is larger than “ $\{q_s(p_3)\}$ ”. It remains valid also if the value  $q_s(p_3)$  is not defined.

But the problem is if the value of  $\hat{q}$  is defined for the composition  $p_3$ , but it is not for its subparts  $p_1$  and  $p_2$ . Therefore we will always suppose that the system of variables and their assigning to individual parts is chosen to be rich enough so that the following statement is always guaranteed, namely

$$\forall \hat{q} \forall s : (p_1, p_2) \in \mathcal{D}_{q,s} \Rightarrow p_1 \in \mathcal{D}_{q,s} \vee p_2 \in \mathcal{D}_{q,s}. \quad (3.19)$$

It ensures that the composition of parts is described by the same or less number of values than there is the total number of values describing its individual parts. Moreover, it means that some ‘new’ variable cannot immediately arise if we form a composition of two parts and hence it guarantees a continuity of the description when forming compositions of parts. We call the statement (3.19) the **continuity rule**.

Remark: To fulfill this rule, we have to choose suitable variables. For example, let us have the structure of mass particles. If we introduce

the interaction energy between two particles as a variable, the rule is not fulfilled because this variable is not defined on single parts (individual particles). That is why we have to choose, instead, the energy (including both kinetic, potential and interaction energy) as the suitable variable.

This rule, however, seems to contradict the physical description of macroscopic systems. Imagine, for instance, an atomic structure whose atomic parts are individual molecules of a gas in a box. The temperature cannot be assigned to individual molecules, but it emerges as a suitable variable for some ‘macroscopic’ parts, i.e. compositions of atomic parts including, say,  $10^{10}$  or more molecules. Nevertheless, the rule remains valid because there is really no division of this part onto two subparts, so that the temperature is not defined for the both parts (at least one part has to be macroscopic!)<sup>9</sup>.

The fact that there is a variable defined for huge compositions of atomic parts but not for small ones, is not a contradiction with the rule alone, but with our thinking that uses a form of induction principle. Namely, we think as follows. The temperature is not defined for any atomic part (molecule). It implies (due to the continuity rule) that it cannot be defined for the composition of two atomic parts. Now, if the temperature is not defined for a composition of  $n$  atomic parts then (due to (3.19) again) it cannot be defined for its composition with an arbitrary atomic part (to form the composition of  $n + 1$  atomic

---

<sup>9</sup>If we have two parts with defined temperature then their composition may not have defined the temperature too (e.g. if the parts are not in mutual thermal equilibrium). But the situation, when the temperature of a part is defined while that of its macroscopic subpart not, contradicts the concept of temperature. See the definition of generalized temperature in Section 2.6.

parts). As a result, the temperature cannot be defined for *any* composition.

The problem with the induction consists in the fact that it is obvious for small collections of objects and we automatically uses it for arbitrarily large collections. This, however, is not so obvious and may lead to paradoxical results (as the problem with the interpretation of the continuity rule for large collections of objects).

Remark: The induction principle is formalized by the axioms of Peano arithmetic. It is embodied in the induction axiom,

$$[\phi(1) \wedge \forall n : \phi(n) \Rightarrow \phi(n')] \Rightarrow \forall n : \phi(n),$$

where  $\phi(n)$  is a property and  $n'$  is a (uniquely defined) successor of  $n$ . The quantifier  $\forall n$  means all states which can be reached by a stepwise procedure  $n \rightarrow n'$ . However, the pure Peano arithmetic (PA) cannot tell us *how far* can we get by this procedure! For example, by using the induction axiom only in the framework of PA, we cannot formally deduce that the stepwise procedure exceeds a very large number, say,  $10^{20}$ .

We need a *model* of PA. In several last decades, the results in logic and the set theory have shown that there is not only the standard model of PA, picturing the set of familiar natural numbers  $\mathbb{N}$ , but also some nontrivial nonstandard models. All nonstandard models have the structure schematically figured as the line fuzzily divided into the parts [44] (see Fig. 3.3) The whole line models the structure of PA, but it does also the shorter line which is formed by standard natural numbers  $\mathbb{N}$ . Thus this (essentially shorter!) part itself has to go to infinity. Consequently, there is something beyond the infinity! In the well-known Robinson's nonstandard arithmetic [34], the part  $\mathbb{I}\mathbb{N}$  consists of some *infinite natural numbers* obeying all arithmetic laws but being larger than any natural number.

There is a very nontrivial interpretation of the nonstandard model of PA that is relevant to our problem with the continuity rule. It has been found out by P. Vopěnka [45] who realized that our ability to see or grasp enormous collections of objects as a whole (for example

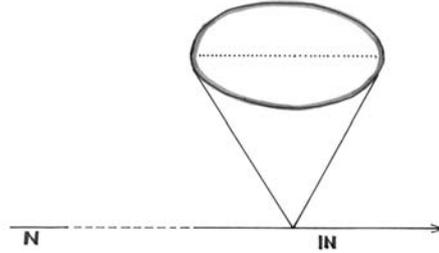


Figure 3.3: Structure of a nonstandard model – if we “look” at any piece of the part  $\mathbb{IN}$ , we “see” a linear succession of numbers without any beginning

the table consisting of a huge numbers of atoms) does not automatically mean that we are really able to go through these objects by a stepwise procedure. Otherwise speaking, he sensed that the shorter

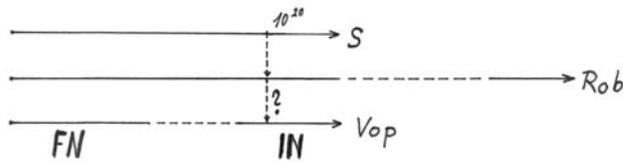


Figure 3.4: Standard model (S), Robinson’s non-standard model (Rob) and the Vopěnka’s model (Vop) whose “standard” part (FN) is essentially smaller than the standard natural numbers.

part of the nonstandard model, denoted him by  $\mathbb{FN}$ , may not reach to an absolute infinity but only to a *horizon* fuzzily limiting our ability to continue each stepwise procedure. Because the way to the horizon never finishes (we ever may do an other step) he called this phenomenon the *natural infinity*. Thus in the Vopěnka’s approach, large

numbers, such as  $10^{20}$ , may be understood to be infinite and not accessible by any step-by-step process and, of course, by an induction proof! The comparison among the standard model and Robinson's and Vopěnka's nonstandard ones are illustrated in Fig. 3.4.

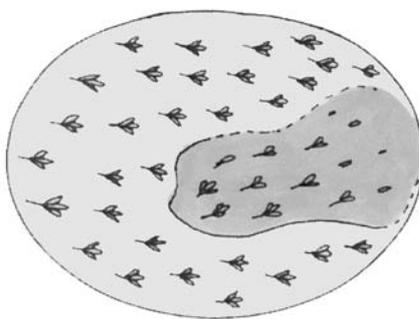


Figure 3.5: A semiset as a fuzzily defined part of a (large) set. It may be defined, for instance, as a group of leaves on a given tree that are not green. If the set represents a macroscopic collection of molecules then the semiset may be interpreted as the “largest” collection of molecules that does not form a macroscopic sample.

While Robinson's model is created by a very complicated procedure in the framework of the standard set theory, Vopěnka embodied the idea of natural infinity directly to a new set theory called the *alternative set theory* where the nonstandardness is implemented from the beginning by using the concept of *semisets* [37, 45]. The semisets are parts of sets which are not sets alone. They model the horizon included in large collection of objects (they correspond to the part  $\mathbb{FN}$ ). This approach thus allows us to formalize such concepts as a “macroscopic collection of molecules”. Namely, any macroscopic set of molecules includes a semiset. The horizon included within semisets

then prevents the use of the induction principle to “carry” a property (e.g. that the temperature cannot be defined) of small collections to the whole set including a semiset. See Fig. 3.5.

### 3.7 Averaging and scale

Let us have a set of  $n$  mutually disjunct parts,  $p_1, \dots, p_n$ . Each part has its own state space,  $\Gamma(p_i)$ ,  $i = 1, \dots, n$ . Denoting their composition as  $p$ , i.e.

$$p = (p_1, \dots, p_n),$$

we can construct the state space of this part,  $\Gamma(p)$ , as given in (3.14). On the other hand, due to the continuity rule (3.19), the information about states of all parts  $p_i$  gives a better description of  $p$  than a set of values of variables assigned to  $p$ , i.e. a point in  $\Gamma(p)$ .

It motivates us to define the **full state space** of  $p$  with respect to  $p_1, \dots, p_n$  as done when defining the state space of an atomic structure (see (3.16)):

$$\Gamma_{full}(p_1, \dots, p_n) = \Gamma(p_1) \times \dots \times \Gamma(p_n). \quad (3.20)$$

The **averaging** then may be defined as a passage from the full state space to the state space  $\Gamma(p)$ , i.e. as a mapping

$$\mathcal{A} : \Gamma_{full}(p_1, \dots, p_n) \rightarrow \Gamma(p). \quad (3.21)$$

The averaging thus means a “loss of information” due to a change of description of individual parts to their composition. It is the crucial concept in any physical theory of large (e.g. macroscopic) systems. Let us employ the concept of ‘scale’

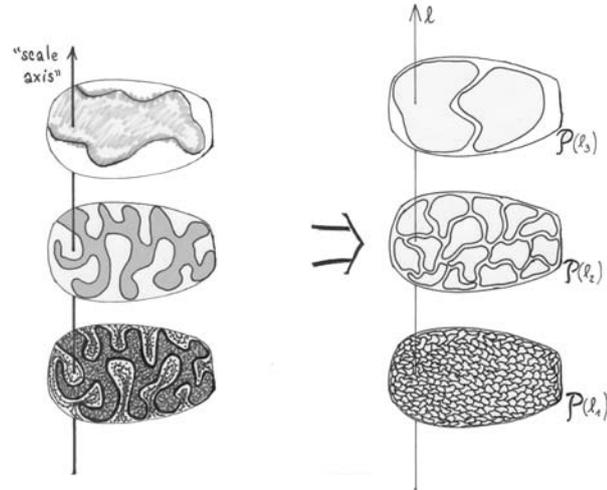


Figure 3.6: The system looks differently at various scales. When zooming out, various details vanish and some new features may emerge. It corresponds to the description of the structure by more roughly defined parts – compositions of parts defined at smaller scales. Individual  $l$ -spaces are formed by parts defined at the same scale.

in our abstract description<sup>10</sup>. To do that we introduce the

---

<sup>10</sup>The scale is understood as a characteristic (or smallest) length taken into account in a physical model. In the last century, this concept started to play a very important role due to various reasons: the presence of divergences in quantum field theories (the ‘cutoff’ removing these divergences corresponds to this length – see [19, 33]), the impossibility of managing microscopic models of systems with many degrees of freedom (e.g. large fluctuations in critical phenomena have been described by employing the concept of scale in the works of K.G. Wilson and others [47]), the so-called “size-scaled” effects in macroscopic physics (e.g. the failure of materials

mapping

$$\text{Sc} : \mathbb{R}_+ \rightarrow \text{Pow}(\mathcal{P}), \quad (3.22)$$

where  $\text{Pow}(\mathcal{P})$  is a set of all subsets of  $\mathcal{P}$ . The subset  $\mathcal{P}(l) \equiv \text{Sc}(l)$  is called the **l-space** if

- (i)  $l \neq l' \Leftrightarrow \mathcal{P}(l) \cap \mathcal{P}(l') = \emptyset$ ,
- (ii) if  $p_1, \dots, p_n \in \mathcal{P}(l)$  then there is a scale  $l' > l$  depending on  $n$  so that  $(p_1, \dots, p_n) \in \mathcal{P}(l')$

(see Fig. 3.6). The  $l$ -space may be empty because the empty set belongs to  $\text{Pow}(\mathcal{P})$ , too. Moreover, if the system has the total part  $p_0$  and it belongs into  $\mathcal{P}(l_0)$  then for any  $l > l_0$  the subset  $\mathcal{P}(l) = \emptyset$ . Let us present the defined concepts on two simple examples.

**Structure of mass points.** The set of parts consists of three parts,  $\mathcal{P} = \{p_1, p_2, p_3\}$ , where  $p_3 = (p_1, p_2)$  and  $\text{Comp} = \{[p_1, p_2], [p_2, p_1]\}$ . We suppose that there are two vector variables assigned to *every* part at any state, namely  $[\mathbf{q}_i, \mathbf{p}_i]$ ,  $i = 1, 2, 3$  where the first is a position and the second a momentum. This structure represents two mass points with constant masses,  $m_1$  and  $m_2$ . If such description of a real structure is realistic or not, depends on the set of studied states,  $\Sigma$ , which represents a physical “background” of the problem.

---

and structures [4]) and many others. Nevertheless, the scale is usually understood as only helpful, “artificial” concept enabling us to construct reasonable models describing nature at chosen levels and formulate relations between these descriptions. An exception is the approach of L. Nottale [30] who develops the concept of scale relativity being motivated by the Einstein’s theory of relativity. Another example of grasping the scale as a more fundamental concept is presented by I.M. Havel [11] who introduces the concept of ‘scale axis’ (see Fig. 3.6, [12]).

A full description of this system (the maximal information about the structure in *this* description) is given by the double vector  $[\mathbf{q}_1, \mathbf{p}_1, \mathbf{q}_2, \mathbf{p}_2]$ . The demand that composition of two parts (i.e. the third part) is described by the same set of variables,  $[\mathbf{q}_3, \mathbf{p}_3]$ , means an averaging defined on the structure. Namely, the step

$$[\mathbf{q}_1, \mathbf{p}_1, \mathbf{q}_2, \mathbf{p}_2] \rightarrow [\mathbf{q}_3, \mathbf{p}_3] \quad (3.23)$$

expresses the essence of averaging. The state set  $\Sigma$  is supposed to include all positions and momenta of the parts  $p_1$  and  $p_2$  so that the state space of each mass point is defined as

$$\Gamma_1 = \Gamma_2 = \{[\mathbf{q}_i, \mathbf{p}_i]\} \subset \mathbb{R}^6 \quad (3.24)$$

and

$$\Gamma_{full} = \Gamma_1 \times \Gamma_2. \quad (3.25)$$

The averaging (3.23) thus can be expressed as

$$\Gamma_{full} \rightarrow \Gamma_3, \quad (3.26)$$

where  $\Gamma_3$  is the state space of the part  $p_3$ . We may suppose that  $\Gamma_3 = \Gamma_1 = \Gamma_2$ , too. The state spaces  $\Gamma_{full}$  and  $\Gamma_3$  describe the *same* system but at different scales because we may define two l-spaces:  $\mathcal{P}(l_1) = \{p_1, p_2\}$  and  $\mathcal{P}(l_2) = \{p_3\}$  where  $l_1 < l_2$ .

The concrete form of the averaged value  $\mathbf{p}_3$  is simple to be determined, namely

$$\mathbf{p}_3 = \mathbf{p}_1 + \mathbf{p}_2. \quad (3.27)$$

The reason is that the momentum is an additive quantity that describes the global behavior of any mechanical system. The

averaged position may be, in principle, defined by various manners, e.g.  $\mathbf{q}_3 = \mathbf{q}_1 + \mathbf{q}_2$ , but it would be a rather awkward definition. The suitable one defines the averaged position as a position of the center of gravity of the system, i.e.

$$\mathbf{q}_3 = \frac{m_1 \mathbf{q}_1 + m_2 \mathbf{q}_2}{m_1 + m_2}. \quad (3.28)$$

**Atomic structure of geometric cells.** We choose an atomic structure with  $\mathcal{P} = \{\Omega_1, \Omega_2, \Omega\}$  where  $\Omega_1 \cap \Omega_2 = \emptyset$  and  $\Omega_1 \cup \Omega_2 = \Omega$ . We postulate that  $\Omega = (\Omega_1, \Omega_2)$ . The state set  $\Sigma$  is chosen in such way that each of the parts  $\Omega_1$  and  $\Omega_2$  alone is in thermal equilibrium at every state from  $\Sigma$ . We have the only variable - the thermodynamic temperature  $\hat{T}$  whose values are positive real numbers. The state space of the first and second part is thus

$$\Gamma_1 = \Gamma_2 = \Gamma = \mathbb{R}_+. \quad (3.29)$$

The temperature  $\hat{T}$  of the composed part  $\Omega$  is not defined if the parts are not in a mutual thermal equilibrium. The definition of local variables and their values may be done as follows:

1. We simply “ban” the definition of the temperature on the part  $\Omega = (\Omega_1, \Omega_2)$  whenever parts are not mutually in thermal equilibrium. The mapping  $T_s$  is then defined on the part  $\Omega$  only if  $s \in \Sigma_{eq}$  where  $\Sigma_{eq} \subset \Sigma$  are states in which the temperatures of the parts  $\Omega_1$  and  $\Omega_2$  are the same. Then

$$T_s(\Omega) = T_s(\Omega_1) = T_s(\Omega_2). \quad (3.30)$$

If  $s \notin \Sigma_{eq}$  then  $\mathcal{D}_{T,s} = \{\Omega_1, \Omega_2\}$ .

2. We change the set of variables to a more suitable one. For example, we use the fact that the inner energy of the system in thermal equilibrium,  $E_s$ , is the function of the temperature,  $E_s = E(T_s)$  (if no other variables are taken into account). The fact that the inner energy is an additive variable allows us to define its value on  $\Omega$  in any state. Namely,  $E_s(\Omega) = E_s(\Omega_1) + E_s(\Omega_2)$  (see Section 6.5).

Let us discuss another possibility consisting in defining the temperature of the part  $\Omega$  at *any* state from  $\Sigma$  as an averaged concept. That means, we suppose that  $T_s(\Omega)$  is defined by an averaging procedure,  $\mathcal{A} : \Gamma \times \Gamma \rightarrow \Gamma$ ,

$$T_s(\Omega) = \mathcal{A}(T_s(\Omega_1), T_s(\Omega_2)), \quad (3.31)$$

that projects values from the full state space onto the state space of the part  $\Omega$ . For example, by the prescription

$$T_s(\Omega) = \frac{V_1 T_s(\Omega_1) + V_2 T_s(\Omega_2)}{V_1 + V_2}, \quad (3.32)$$

where  $V_i$  are volumes of the parts  $\Omega_i$ .

The variable  $\hat{T}$ , however, cannot be understood as a global variable because it does not obey the definition of generalized temperature. Namely, if we define a global variable  $T$  by the relation (3.8), i.e.  $T(s) = T_s(\Omega)$ , then  $T$  cannot be the generalized temperature because its value generally differs from its values on subsystems (if  $\Omega_i$  represents subsystems of  $\Omega$ ). Nevertheless, when  $\hat{T}$  is understood as a local variable only then there is no problem with understanding it as an averaged concept. Any definition of such a ‘non-equilibrium temperature’, however, needs a careful physical explanation [25, 32].

## Chapter 4

# Special structures

In this chapter, several structures are introduced and studied in detail. Chosen examples illustrate two main schemas of description used in thermodynamics and continuum thermomechanics, but not solve concrete physical problems. Especially, we concentrate into choice of parts, introduction of variables and construction of the state space.

We start by the structure used in equilibrium thermodynamics. Its importance is crucial because it is the *only* schema of description for which the variable ‘entropy’ has been rigorously introduced. Second example is the structure used when describing various macroscopic processes (thermal, mechanical, hydrodynamical or electromagnetical) in macroscopic media. The discipline studying such processes and their combinations is called the continuum thermomechanics. We outline its methods of description and especially the way of formulating continuum description by the use of non-atomic structures.

## 4.1 Simple systems

Look back at the consideration presented in Section 2.6. We have noticed there an important assumption allowing to interpret results of many measuring procedures that are done on *other* systems than those whose properties should be determined. Let us remind ourselves the experiment when a sample of water is taken from a vessel and its physical and chemical properties are determined. The assumption of an equilibrium between the sample and remaining water enables us to assign the measured data to the water what remains in the vessel.

In Section 2.6, we have formally defined what means that the system and its surrounding are in equilibrium. But what is the system and what its surrounding in the experiment with a sample of water? The water in the vessel *before* the measurement is a physical system, say  $\mathfrak{S}_0$ . When taking a sample, we form two new physical systems,  $\mathfrak{S}_{sample}$  and  $\mathfrak{S}'_0$ , representing the sample and remaining water respectively. What is a relation of  $\mathfrak{S}_{sample}$  and  $\mathfrak{S}_0$ ? First of all,  $\mathfrak{S}'_0$  *is not* a surrounding of  $\mathfrak{S}_{sample}$  because the sample of water loses any contact with the remaining water (no manipulation with the water in the vessel can influence the sample afterwards it was displaced).

The water in the vessel is a surrounding of the sample of water *only* so long as the sample is within the water. But what does the ‘sample of water’ mean until it is defined by taking it into a test tube? To determine a moment when the sample is both *defined* and in *equilibrium contact* with surrounding water, we assume that a closed wall separating the sample from surrounding water suddenly appears somewhere within the water, see Fig. 4.1. At this moment, the water outside the wall plays the role of surrounding of the sample. Moreover, this sur-

rounding is *in equilibrium* with the sample in the spirit of our definition (the state of the surrounding water defines the state of the sample).

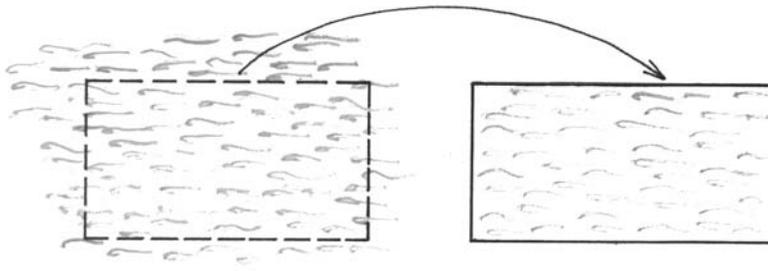


Figure 4.1: A sample of water is taken off to form a system defined by walls separating it from surrounding water.

The situation may be described in terms of ‘parts’. Namely, the existence of a subsystem  $\mathfrak{S}_{sample}$  of the water in the vessel that lately forms the sample of water may be supposed *before* than it is defined by the wall (e.g. as a conglomerate of certain molecules). But this subsystem has no clear definition (we cannot write a “list” of all molecules belonging to it). We may describe this situation as the fact that there is no part from  $\mathcal{P}_{sub}$  defined for this system. That is, though the system  $\mathfrak{S}_{sample}$  exists, there is no part corresponding to this system (any part has to be *defined*). Only the presence of the wall defines clearly a part – the water inside the space defined by the wall. This part is thus defined *externally* and has all features of a part from  $\mathcal{P}_{ch}$ . Writing more formally,  $\mathcal{P}(\mathfrak{S}_{sample}) = \mathcal{P}_{ch} = \{p\}$ , because there is only one part defined by the wall. It motivates us to the following definition:

The system  $\mathfrak{S}$  is called the **simple unit** if the following conditions are valid:

- (i) Its surrounding,  $\mathfrak{S}_{sur}$ , is in equilibrium with  $\mathfrak{S}$ ,
- (ii) the system of parts includes only one part,  $p$ , that belongs to  $\mathcal{P}_{ch}$ , i.e.

$$\mathcal{P}(\mathfrak{S}) = \mathcal{P}_{ch} = \{p\}, \quad (4.1)$$

- (iii) for each variable  $\hat{q} \in Q$ , there is a global variable  $G_q \in Q_g(\mathfrak{S}_{sur})$  so that for every part  $p \in \mathcal{P}$ ,

$$q_s(p) = (v(p))^{k_q} G_q(f_{eq}^{-1}(s)), \quad (4.2)$$

where  $v(p)$  is a positive real number determined by the part  $p$ ,  $k_q = 0$  or  $k_q = 1$ , and  $f_{eq}$  is the mapping defining the equilibrium situation (see Section 2.6).

The variables for which  $k_q = 0$  are said to be **intensive** and those for which  $k_q = 1$  are said to be **extensive**. The global variable  $G_q$  corresponding an extensive variable  $\hat{q}$  is usually called the **density** of variable  $\hat{q}$ .

If  $\mathfrak{S}'$ ,  $\mathfrak{S}$  are simple units defined by the same surrounding then  $\mathfrak{S}'$  is called the **scaled copy** of  $\mathfrak{S}$ .

The meaning of this concept is clear from the definition of simple unit. The both systems have the same set of variables defined by the common surrounding. The only difference between values of variables on  $\mathfrak{S}$  and  $\mathfrak{S}'$  is thus in the value of  $v$  as implied by (4.2). All intensive variables thus have the same values on the both systems. Let us denote

$$r \equiv \frac{v(p')}{v(p)}.$$

If  $k_q = 1$  then

$$q_s(p) = rq_s(p), \quad (4.3)$$

i.e. values of extensive variables are  $r$ -times scaled. If  $\Gamma(p)$  is the state space of a simple unit then we usually denote by  $\Gamma^{(r)}(p)$  the state space of its scaled copy. If  $X_p \in \Gamma$  we denote the corresponding state of the scaled copy as  $rX_p$  (we take over the nomenclature from [21]).

Let us form a composition of  $n$  simple units,  $\mathfrak{S}_1, \dots, \mathfrak{S}_n$ , that are pairwise mutually isolated. It means that the state set of this composition is

$$\Sigma \equiv \Sigma((\mathfrak{S}_1, \dots, \mathfrak{S}_n)) = \Sigma(\mathfrak{S}_1) \times \dots \times \Sigma(\mathfrak{S}_n). \quad (4.4)$$

The structure of this composition may be understood as being formed by  $n$  atomic parts, .

$$\mathcal{P}_{atom} = \{p_1, \dots, p_n\},$$

and their various compositions. We suppose that there is a generalized temperature  $T$  defined for every surrounding of individual simple units,  $\mathfrak{S}_1, \dots, \mathfrak{S}_n$ . This global variable defines an intensive local variable  $\hat{T}$  on the system  $\mathfrak{S} = (\mathfrak{S}_1, \dots, \mathfrak{S}_n)$ , called the **temperature**, by the prescription (4.2) with  $G_q \equiv T$  and  $k_T = 0$ . Hence it is defined at every state on an arbitrary part  $p_i \in \mathcal{P}$ .

The temperature, however, may not be defined on their compositions (see the example at the end of Section 3.7). That is why we define a subset of  $\Sigma$ ,

$$\Sigma_{eq} \subset \Sigma,$$

so that at every state  $s \in \Sigma_{eq}$  all simple units have the same value of the temperature.

The system  $\mathfrak{S} = (\mathfrak{S}_1, \dots, \mathfrak{S}_n)$  whose state set equals  $\Sigma_{eq}$  is called the *simple system*.

For simple systems, the temperature has the averaged property because it is defined at any composition,  $T_s((p_i, p_j)) = T_s(p_i) = T_s(p_j)$ , compare (3.30).

The scaled copy of a simple system is then defined as a composition of scaled copies of individual simple units. Since parts of individual simple units are atomic, the state space of a simple system is given by (3.16), i.e.

$$\Gamma = \Gamma(p_1) \times \dots \times \Gamma(p_n). \quad (4.5)$$

We denote by  $\Gamma^{(r)}$  the state space of the scaled copy and  $rX \in \Gamma^{(r)}$  a point corresponding to  $X \in \Gamma$ .

Variables of the simple system may be always chosen so that the temperature  $\hat{T}$  is among them. The other variables are called **work variables** because their values may be always defined by a mechanical (or electric or magnetic) manipulation with the surrounding. Such a kind of manipulation is usually connected with a performed work, see Fig. 4.2. The work variables are deeply connected with the definition of parts  $p_1, \dots, p_n$  of the simple system. For example, a typical work variable is the volume that defines the size of (externally defined) parts  $p_i$  (see [21], p. 38).

Sometime, an extensive real variable  $\hat{E}$ , called the **energy**<sup>1</sup>, is chosen instead of the temperature (see again the end of Section 3.7). The relation of the energy and temperature is studied in equilibrium thermodynamics. We outline the essence of this relation in Section 6.5.

---

<sup>1</sup>The energy as a global variable is defined at the end of Chapter 5.

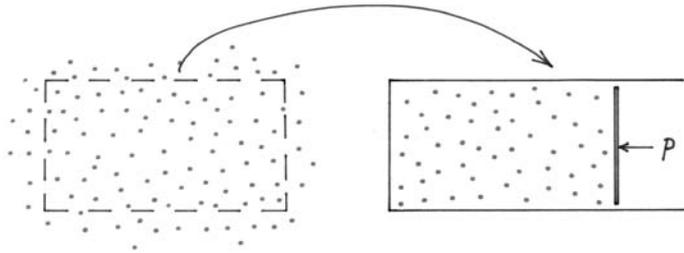


Figure 4.2: When taking off a sample of gas, we have to guarantee the same conditions as if it were in the vessel with the gas. It means to exert a pressure on walls. Then any change of the volume  $V$  of the sample is connected with a work,  $\delta W = p dV$ .

## 4.2 Thermomechanical systems

The discipline describing thermal, mechanical, hydrodynamical and electromagnetical processes in macroscopic media is called the **continuum thermomechanics** (see e.g. [13, 22, 23, 26, 27, 40, 43]). The important feature of the thermomechanical way of description is a rich use of many macroscopic subsystems that may be thought as parts of the system. The most important step, however, is an introduction of parts that are not material physical systems – *boundaries* of subsystems. It has an origin in classical thermodynamics where systems are always in a container (e.g. a box with gas endowed by a piston). The container does not play only the role of an external “limiter”, there are various variables that are assigned rather to the container than to the system alone. For example, the volume or the heat flow. Boundaries in continuum thermomechanics have, however, a more general meaning.

To define the structure of a thermomechanical system we return to the thought experiment in which a simple unit is formed by taking a sample of water from a water reservoir. The sample has all properties defined by the surrounding water because of the assumption of equilibrium. The only “free parameter” is the size of the sample – it occurs in the relation (4.2) as a “scaling parameter”  $v(p)$ . This size is, however, also defined externally by a wall forming the sample. Namely, the surrounding is not only the surrounding water but also the wall because it has to be present at the moment when the sample as a physical system arises<sup>2</sup>.

In this section, we study the situation when only *this aspect* of an external definition of a sample plays the role because its surrounding is not in equilibrium with it. The part  $p \in \mathcal{P}_{ch}$  represents the boundary delimiting the sample but a physical description of the sample cannot be done by assigning variables *only* to this boundary. Namely, there is no direct relation between states (and properties) of the surrounding and states (and properties) of the sample. Therefore, the description of such a sample cannot be done by the use of only this part.

That is why we add a part  $\sigma \in \mathcal{P}_{sub}$  which represents the sample as a physical system  $\mathfrak{S}_{sample}$ . Using (3.4), we may write

$$\sigma = F_s(\mathfrak{S}_{sample}). \quad (4.6)$$

The part  $p$  may be understood as a boundary of  $\sigma$  and therefore it is called the **oriented boundary**<sup>3</sup> and denoted by  $\partial\sigma$ . Thus

---

<sup>2</sup>Walls are supposed to be inert so that properties of the water are not influenced by them.

<sup>3</sup>“Oriented” means that the boundary defines always a uniquely given interior and exterior.

we obtain the structure

$$\mathcal{P} = \{\sigma, \partial\sigma\}. \quad (4.7)$$

The system of this structure is called the **discrete system**.

Remark: It may seem strange that we add the part  $\partial\sigma$  into the structure while  $\sigma$  represents the *whole* system (a sample). The structure, however, is chosen to provide the best description of the system by a set of variables assigned to its parts. The part  $\partial\sigma$  is crucial in description of discrete systems. In fact, it replaces partially the “lost” equilibrium because variables assigned to this part are definable on the surrounding.

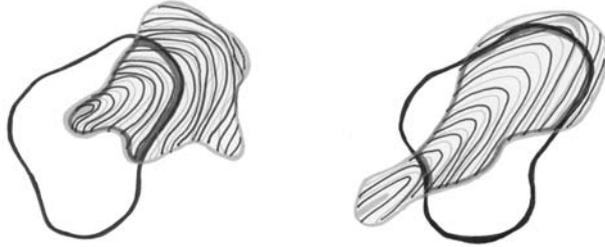


Figure 4.3: The discrete system need not be defined by a given piece of mass. The mass may flow through the boundary and its value “inside” then plays the role of a variable of the system.

In what follows, we will study discrete systems that are various macroscopic bodies described in common language by their arrangement, phase, chemical constitution of various components and so on. In this concrete meaning, they are also called **Schottky systems** (see e.g. [23, 26]). Each discrete system has a state set which corresponds to various states that are

taken into account in the study of the body (the body may or may not change the phase or chemical constitution, mass, momentum, energy, etc.), see Fig. 4.3. We also accept various manipulations with the body that are caused by a chosen surrounding of the discrete system. The set of manipulations,  $\mathcal{M}$ , is supposed to be rich enough so that it forms a large set of processes.

Variables of a discrete system split into two categories. The group of variables defined only on the part  $\sigma$  and the group defined only on the part  $\partial\sigma$ . Variables from the first group are called the **c-variables**. Typical c-variables are mass, energy, momentum, etc. Variables from the second group are called the **b-variables** (they are defined only on oriented boundaries). Within this category, special process variables play an important role. They describe an exchange of a physical property with surrounding. A typical example is the **heat**. Another example of an eminent importance is the **stress** that describes the exchange of momentum between the system and surrounding. This kind of b-variables will be called the **process b-variables**.

The great advantage of simple systems (defined in the last section) consists in a full definability of their states externally. Namely, the equilibrium guarantees that we always know ‘what is inside’. It is not the case of discrete systems. Values of c-variables are *not definable* externally (see Fig. 4.4). Therefore, the crucial role plays a special relation between certain c-variables and process b-variables, whose values are definable externally:

*A c-variable  $\hat{q}$  is said to have the **balance property** if for any process  $\pi$  there is such a class of parameterizations that if  $\pi_{\tau_1, \tau_2}$  belongs to this class then the derivative,*

$$\frac{d}{d\tau} (q_{s(\tau)}(\sigma)) \equiv \dot{q}_{s(\tau)}(\sigma), \quad (4.8)$$

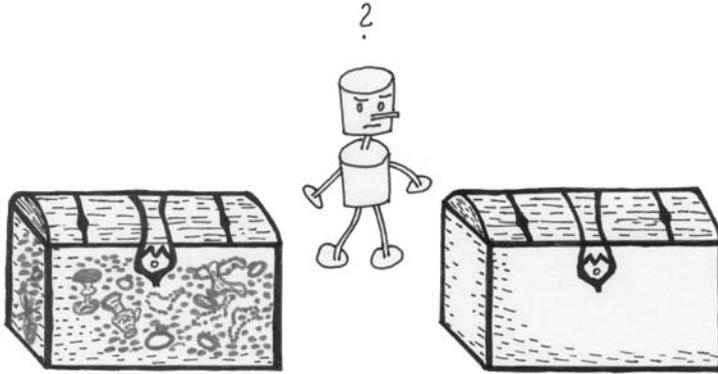


Figure 4.4: Where is a treasure? From an external point of view, some properties of the system are difficult to be determined.

exists for every<sup>4</sup>  $\tau \in (\tau_1, \tau_2)$ , and there is a process b-variable  $\hat{b}^q$  and c-variable  $\hat{\delta}^q$  so that the equality

$$\dot{q}_{s(\tau)}(\sigma) = b_{s(\tau)}^q(\partial\sigma) + \delta_{s(\tau)}^q(\sigma) \quad (4.9)$$

is valid for every  $\tau \in (\tau_1, \tau_2)$ .

The process b-variable  $\hat{b}^q$  is called the **flux** of the variable  $\hat{q}$ . The c-variable  $\hat{\delta}^q$  represents either an external **supply** of  $\hat{q}$  that is not connected with processes on the boundary (e.g. a volume force, like gravity or centrifugal force, is an external supply of the momentum) or the **production** of  $q$  (see e.g. [26]). The equality (4.9) presents in a general form the famous balance equality.

<sup>4</sup>This condition may be weaker by assuming only existence of the derivative almost everywhere within the interval  $(\tau_1, \tau_2)$ , see [40].

A typical example of a variable with balance property is the **mass**  $\hat{m}$ . It has the balance property so that

$$\delta_s^m(\sigma) = 0. \quad (4.10)$$

Remark: Variables with zero supply and production (such as the mass) have an origin in global variables having conservative properties. It is worth noticing a remarkable relation between simple and discrete systems. Variables assigned to simple systems are defined (measured) on an external surrounding by the global variables defined via equilibrium. Several variables assigned to discrete systems are defined (measured) on an external surrounding by the global variables defined via conservative properties of the equilibrium.

If a  $c$ -variable  $\hat{q}$  has the balance property and  $\hat{\delta}^q$  is always zero or it is an externally defined supply (e.g. defined by an external field) then any change of its value during any process is externally controllable. The problem is if a quantity has a nonzero production.

Let us have a macroscopic body that is understood as a physical system  $\mathfrak{S}$ . The main idea of the thermomechanical description of the body consists in imaging it as a composition of an uncountable number of discrete systems. The description is thus based on a special non-atomic structure. The structure is defined as follows: If  $p \in \mathcal{P}_{sub}$  then either there is a discrete system, so that  $p = \sigma$ , or there are  $n$  disjunct discrete systems and  $p = (\sigma_1, \dots, \sigma_n)$ . For every  $p \in \mathcal{P}_{sub}$ , there is exactly one element  $\text{Bd}(p) \in \mathcal{P}_{ch}$  so that if  $p = \sigma$  then  $\text{Bd}(\sigma) = \partial\sigma$ . These elements are called **oriented boundaries**<sup>5</sup>. If  $p \in \mathcal{P}_{ch}$  then

---

<sup>5</sup>Oriented boundaries are defined externally by the use of a suitable discrete system (or a composition of discrete systems). This system is somewhere “inside” the studied body but, whenever we understand it as

it is either an oriented boundary or its subpart. The **thermomechanical system** is described by a non-atomic structure of parts from  $\mathcal{P}_{sub}$  (discrete systems and their compositions) and  $\mathcal{P}_{ch}$  (oriented boundaries and their subparts).

Variables of a thermomechanical system arise as a natural enlarging of variables defined for discrete systems. The set of c-variables remains the same, only their properties may be characterized in more detail. For example, the fact that there is no only one part in  $\mathcal{P}_{sub}$  allows us to specify if they have averaged properties or are additive.

Concerning the b-variables, the situation is slightly different. Namely, we have b-variables that are defined only on oriented boundaries (and not on their subparts). Thanks to the unique relation between oriented boundaries and parts from  $\mathcal{P}_{sub}$ , the value of such a variable on an oriented boundary  $p'$  can be assigned also to the corresponding part  $p = \text{Bd}^{-1}(p')$ . A typical example of such a variable is ‘volume’.

Let us imagine, for a while, an oriented boundary  $\partial\sigma$  of a discrete system as a wall of a closed container that is occupied by a piece of the matter (it does not matter how “fully”, see Fig. 4.3). The volume of the space inside the container is determined by this boundary and not by the substance inside. Therefore the volume has to be understood as a b-variable  $\hat{V}$  whose definition domain covers all oriented boundaries. It points out its origin in thermodynamics – the volume is defined by an external box and not by matter inside. On the other

---

a discrete system, the resting part of the body belongs to the surrounding of *this* discrete system. The boundary is thus no real interface but it is rather defined by a mental construction. It resembles the composition of simple units – surroundings that define individual simple units are not really present but we ever may add them mentally.

hand, the volume behaves as an additive c-variable. Namely, if disjunct parts  $\sigma_1, \sigma_2$  have volumes  $V_1$  and  $V_2$ , respectively, then the volume of their composition is  $V_1 + V_2$ . Since the volume is b-variable we have to express its additivity as

$$V_s(\text{Bd}((\sigma_1, \sigma_2))) = V_s(\text{Bd}(\sigma_1)) + V_s(\text{Bd}(\sigma_2)). \quad (4.11)$$

This kind of b-variables may be understood as c-variables as well (the volume is usually understood as a c-variable and not a b-variable!). Notice, however, a great difference between b-variables that are understood as c-variables (as volume) and “pure” c-variables (such as energy). The value of b-variable on a concrete discrete system is determined by the situation on its boundary. Hence its value is fully determined by an external procedure definable on the surrounding of this discrete system. For example, the exchanging heat or the stress on the boundary are determined from outside. Typical c-variables, however, describe properties that *cannot* be determined at a state by only analyzing a “momentary situation” of the surrounding<sup>6</sup>.

### 4.3 Continuum

Parts defining the structure of a physical system are basic objects, building blocks, having no further detail specification (the detail structure of a part is given only by its subparts, i.e. the way by which they are composed from “smaller” ones). The only specification is given by various variables assigning to individual parts (notice that we do not specify the part corresponding to a ‘mass point’, for instance, otherwise then by values of

---

<sup>6</sup>A current energy of a body, for example, cannot be determined from a current state of the surrounding, see Fig 4.4.

variables assigned to it: mass, position, momentum, etc.). The problem is, however, if the structure is non-atomic. Then there are no basic blocks that enable us to describe the whole system by variables assigned to them. A typical non-atomic structure is a set of parts forming continuum.

The **continuum** means that smaller and smaller parts may be always chosen when forming the set of parts. The relation (3.2) expresses a typical situation of the structure describing the continuum. An example of the continuum is the structure of *thermomechanical system* whose parts from  $\mathcal{P}_{sub}$  form continuum because there are no smallest macroscopic pieces of matter. When modeling boundaries, it is also suitable to assume that subparts of any part from  $\partial\mathcal{P}_{sub}$  form the continuum.

The crucial idea consists in *modeling* parts of a non-atomic structure by some sets<sup>7</sup>. These sets are formed by collections of some abstract elements whose representative is denoted by the symbol  $u$ . These elements, however, have no physical meaning and serve only to form various sets of the class. Let  $\mathcal{U}$  be a “universe” of these elements, i.e. a set including all these elements.

Now, let us model parts from  $\mathcal{P}$  as sets formed by elements of  $\mathcal{U}$ . For every state  $s$ , let us choose a set of subsets of  $\mathcal{U}$ , i.e. the set

$$\mathcal{X}_s \subset \text{Pow}(\mathcal{U}),$$

( $\text{Pow}(\mathcal{U})$  is the set of all subsets of  $\mathcal{U}$ ) and postulate the existence of a one-to-one mapping,

$$\mathfrak{M}_s : \mathcal{P} \rightarrow \mathcal{X}_s,$$

that fulfills the following rules:

---

<sup>7</sup>Our approach thus differs from majority of others defining the continuum directly as a set of subsets of a set, see e.g. [10].

- (i)  $p' \stackrel{sp}{\subset} p \Leftrightarrow \mathfrak{M}_s(p') \subset \mathfrak{M}_s(p)$ ,
- (ii)  $[p, p'] \in D_{comp}$  implies  $\mathfrak{M}_s(p) \cap \mathfrak{M}_s(p') = \emptyset$ .

The set  $X_s = \mathfrak{M}_s(p)$  is said to be the **model of  $p$  at  $s$**  and  $\{\mathfrak{M}_s, s \in \Sigma\}$  the **model** of  $\mathcal{P}$ . We suppose that the structure includes the total part  $p_0$ , and denote

$$X_s^0 \equiv \mathfrak{M}_s(p_0).$$

If we model a non-atomic structure then any one-element subset of  $\mathcal{U}$  cannot be model of a part (the property (i) implies that the part  $\mathfrak{M}_s^{-1}(\{u\})$  then would be atomic). Therefore we claim

$$\forall s \in \Sigma, \forall u \in \mathcal{U} : \{u\} \notin \mathcal{X}_s. \quad (4.12)$$

The model of a structure describing continuum needs an additional assumption. Namely, there is no minimal size of pieces of a continuum structure – every piece may be always divided into smaller ones. It implies that if different elements  $u_1, u_2$  belong to a set modeling a part then there is another part whose model includes  $u_1$  but not  $u_2$ , i.e. it separates the element  $u_1$  from  $u_2$ . That is, we assume that for every  $s$  and every couple of different elements,  $u_1, u_2 \in X_s^0$ ,

$$\exists p \in \mathcal{P} : u_1 \in \mathfrak{M}_s(p) \wedge u_2 \notin \mathfrak{M}_s(p). \quad (4.13)$$

Elements of  $\mathcal{U}$  thus play the role of an auxiliary scaffold, by which the continuum description is build up, but whose concrete elements have no physical meaning. The model of  $\mathcal{P}$  for which (4.12) and (4.13) are valid is called the **continuum model** of the non-atomic structure  $\mathcal{P}$ .

Elements of  $\mathcal{U}$  alone have no physical meaning (as models of some parts) but the condition (4.13) strongly restricts an

arbitrariness in their assigning to parts at various states. We describe it by introducing the concept of configuration. Let  $\Upsilon$  be an index set so that it may be mapped on the whole  $\mathcal{U}$ .

We define a set of mappings,  $\mathfrak{C}$ , that includes for every state  $s$  the mapping  $\eta_s$ ,

$$\eta_s : \Upsilon \rightarrow \mathcal{U},$$

so that for every  $\xi \in \Upsilon$  and every  $p \in \mathcal{P}$  the condition

$$\eta_s(\xi) \in \mathfrak{M}_s(p) \Rightarrow \eta_{s'}(\xi) \in \mathfrak{M}_{s'}(p) \quad (4.14)$$

holds for any  $s, s' \in \Sigma$ . Mappings from  $\mathfrak{C}$  are called **configurations**.

The definition of configurations is sketched at Fig. 4.5.

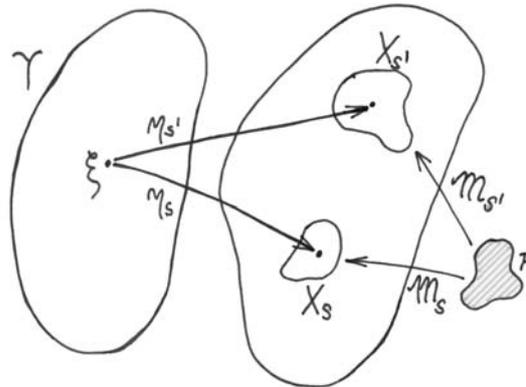


Figure 4.5: Parts are modeled as subsets of a set  $\mathcal{U}$ . A configuration means a mapping from an index set  $\Upsilon$  into  $\mathcal{U}$ , which keeps the “identity” of parts at individual states.

Obviously, we may choose an infinite amount of different set of configurations,  $\mathfrak{C}$ , for one continuum model. A nontrivial

question is, however, how big is this “liberty” if we fix the configuration for one state, say  $s_0$ . Namely, other configurations have to be related to  $\eta_{s_0}$  by the relation (4.14). The answer is surprisingly simple:

**Lemma 1:** If the mapping  $\mathfrak{M}_s$  is a continuum model then the choice of the configuration for *one* state from  $\Sigma$  defines *uniquely* a configuration for *any* state from  $s$ .

**Proof:** The state for which the configuration is fixed is denoted by  $s_0$ . Let there be two different sets  $\mathfrak{C}^1$  and  $\mathfrak{C}^2$ . It implies that there is a state  $s \in \Sigma$  so that there are two different configurations,  $\eta_s^1 \in \mathfrak{C}^1$  and  $\eta_s^2 \in \mathfrak{C}^2$ . It means that there is  $\xi \in \Upsilon$  so that  $\eta_s^1(\xi) \neq \eta_s^2(\xi)$ . Hence we have two different elements of  $\mathcal{U}$ , namely  $u_1 \equiv \eta_s^1(\xi)$  and  $u_2 \equiv \eta_s^2(\xi)$ .  $u_1, u_2 \in X_s^0$  and thus (4.12) implies that there is  $p \in \mathcal{P}$  so that  $u_1 \in \mathfrak{M}_s(p)$  and  $u_2 \notin \mathfrak{M}_s(p)$ . Next, we continue by following consideration. First, we use (4.14) for  $\mathfrak{C}^1$  by identifying  $s' \rightarrow s_0$  and obtain that  $\eta_{s_0}(\xi) \in \mathfrak{M}_{s_0}(p)$ . Second, we use (4.14) for  $\mathfrak{C}^2$  but in opposite direction, i.e. by identifying  $s \rightarrow s_0$  and  $s' \rightarrow s$ . Using the fact that at  $s_0$   $\mathfrak{C}^1$  merges with  $\mathfrak{C}^2$ , we get that  $u_2 \equiv \eta_s(\xi) \in \mathfrak{M}_s(p)$  which is a contradiction.  $\square$

The set of configurations is thus fixed by a choice of the configuration at one state. Moreover, a choice of this configuration as a one-to-one mapping (it may be always done by a suitable choice of the index set) guarantees that *all* configurations are one-to-one mappings:

**Lemma 2:** If the mapping  $\mathfrak{M}_s$  realizes a continuum model and the configuration  $\eta_{s_0}$  at a state  $s_0$  is a one-to-one mapping then the configuration at every state is a one-to-one mapping.

**Proof:** Let us suppose that there is a state  $s$ , so that there are two different elements  $\xi_1, \xi_2 \in \Upsilon$  for which  $\eta_s(\xi_1) = \eta_s(\xi_2)$ . At  $s_0$ , however,  $\eta_{s_0}(\xi_1) \neq \eta_{s_0}(\xi_2)$  that implies that there are two different

elements  $u_1 = \eta_s(\xi_1)$  and  $u_2 = \eta_s(\xi_2)$ . Consequently, there is a part  $p \in \mathcal{P}$  whose model at  $s_0$  includes  $u_1$  but does not include  $u_2$ . The same consideration as in the previous proof shows that it contradicts to the definition (4.14).  $\square$

*The couple  $[\Upsilon, \eta_{s_0}]$  where  $\eta_{s_0}$  is a one-to-one mapping is called the **referential configuration**.*

In what follows, let us suppose that the referential configuration is firmly chosen. For every part  $p \in \mathcal{P}$ , we may define a subset  $\mathcal{Z}(p)$  of  $\Upsilon$  by the prescription

$$\mathcal{Z}(p) \equiv \{\xi \in \Upsilon, \eta_{s_0}(\xi) \in \mathfrak{M}_{s_0}(p)\}. \quad (4.15)$$

If we have a continuum model then every state defines a uniquely given configuration,

$$s \rightarrow \eta_s \in \mathfrak{C}.$$

It does not mean, however, that the set  $\mathfrak{C}$  fully characterizes the state set  $\Sigma$  because there may be different states from  $\Sigma$  that are represented by the same configuration. Nevertheless, we may define a class of variables for which the specification of a configuration is sufficient.

*We say that a variable  $\hat{q}$  is **configuration dependent** if the value of this variable at any state  $s$  and any part  $p \in \mathcal{D}_{q,s}$  depends only on the configuration  $\eta_s$  on  $\mathcal{Z}(p)$ , i.e.*

$$q_s(p) = F(\eta_s(\xi)|_{\xi \in \mathcal{Z}(p)}). \quad (4.16)$$

If only configuration dependent variables are used in the description of a thermomechanical system, we speak about **continuum mechanics**.

Let  $\pi$  be a process realized by a set of manipulations with the studied thermomechanical system. We choose a parameterization  $\pi_{\tau_1, \tau_2}$  and suppose that we have a continuum model

of the studied system. Then the process in the parameterization  $\pi_{\tau_1, \tau_2}$  may be understood as a continuous succession of configurations  $\eta_{s(\tau)}$ . The mapping

$$\chi : \langle \tau_1, \tau_2 \rangle \rightarrow \mathfrak{C},$$

defined by the prescription

$$\chi(\tau) = \eta_{s(\tau)}, \quad (4.17)$$

is called the **motion**. The set of elements from  $\mathcal{U}$  that have in the process the same index  $\xi$  is called the **trajectory** during the process. It is worth reminding that both the concept of motion and trajectory are defined on an abstract set  $\mathcal{U}$ , and no assumption about a location of the system in space is necessary.

#### 4.4 State space of a continuum model

The concept of state space of the whole system uses essentially the occurrence of some “smallest” basic parts for which values of a complete set of variables are given at every state. These parts are clearly defined if the structure is atomic. If it is non-atomic, there are no such parts. Nevertheless, when forming continuum models we define the structure by means of elements of a set  $\mathcal{U}$ . Although these elements are not models of parts from  $\mathcal{P}$ , their large collections form *any* part of  $\mathcal{P}$  and the choice of a referential configuration determines *uniquely* elements of  $\mathcal{U}$  that have the same “label”  $\xi$  at various states, namely

$$u_s(\xi) = \eta_s(\xi). \quad (4.18)$$

May the set  $\Omega_0 \equiv \mathcal{Z}(p_0) \subset \Upsilon$ , that includes all “labels”, be identified with a set of some atomic parts? The problem is

that variables are defined only on parts from  $\mathcal{P}$  and thus they *are not defined* on individual elements of  $\mathcal{U}$ .

In continuum mechanics, where only configuration dependent variables are used, we may enlarge the set of parts  $\mathcal{P}$  to  $\mathcal{P}' = \mathcal{P} \cup \mathcal{P}_{atom}$ , with

$$\mathcal{P}_{atom} = \Omega_0, \quad (4.19)$$

and define a special variable  $\hat{u}$  by the prescription (4.18)<sup>8</sup>. Since the value of *every* variable is derivable from values of  $\eta_s(\xi)$  by (4.16), we may choose

$$Q = \{\hat{u}\} \quad (4.20)$$

that forms a complete set of local variables. The state space (3.18) is then identical with the set of all configurations, namely

$$\Gamma = \mathfrak{C}. \quad (4.21)$$

If there is a variable that is not configuration dependent (e.g. inner energy if thermal phenomena are taken into account), the situation looks hopelessly. Namely, values of such a variable cannot be assigned to atomic parts (4.19) and the knowledge of a concrete configuration cannot determine values of this variable for individual parts from  $\mathcal{P}$ .

Nevertheless, if this variable is *additive*, there is a possibility of defining its values for a suitable amount of parts from  $\mathcal{P}_{atom}$ , so that this information is sufficient for determining the value of the variable for *every* part from  $\mathcal{P}$ . To show it, we have to introduce firstly some important mathematical concepts.

*The so called  $\sigma$ -algebra  $\mathfrak{A}$  is a set of subsets of a given set  $X^0$  so that*

---

<sup>8</sup>“Values” of this variable are elements of an abstract set. But  $\mathcal{U}$  usually forms a Banach space – the Euclidean space  $\mathbb{E}_d$ .

- (i)  $X \in \mathfrak{A}$  implies that  $X^0 \setminus X \in \mathfrak{A}$ ,
- (ii) given a sequence (infinite or not)  $\{X_k\}$  of subsets of  $X^0$ ,  $X_k \in \mathfrak{A}$ , then  $\bigcup X_k \in \mathfrak{A}$ ,
- (iii)  $X^0 \in \mathfrak{A}$ .

If  $f$  is a mapping from  $X^0$  into a topological space  $\mathcal{T}$  then we say that  $f$  is **measurable** if for every open set  $Z \subset \mathcal{T}$  holds  $f^{-1}(Z) \in \mathfrak{A}$ .

Any complex-valued function  $\mu$  defined on a  $\sigma$ -algebra  $\mathfrak{A}$  is a (complex) **measure** if

- (i)  $\mu(\emptyset) = 0$ .
- (ii)  $\mu(\bigcup X_k) = \sum \mu(X_k)$  whenever  $\{X_k\}$  is a sequence (infinite or not) of pairwise disjoint subsets of  $\mathfrak{A}$ , i.e.  $X_i \cap X_j = \emptyset$  for  $i \neq j$ .

We say that a function is defined **almost everywhere** on  $X^0$  if for its definition domain  $X \subset X^0$  holds that  $\mu(X^0 \setminus X) = 0$ .

These definitions allow us to define the concept of an abstract integral of measurable functions on any  $\sigma$ -algebra. If  $\mu$  is a non-negative measure on a  $\sigma$ -algebra  $\mathfrak{A}$ ,  $X \in \mathfrak{A}$  and  $f$  is a complex-valued measurable function, we denote by

$$\int_X f \, d\mu$$

the integral of the function  $f$  over the set  $X$  (a precise definition of the integral is not quite simple and is not presented here, see e.g. [36]). The integral is a (finite or infinite) complex number that is defined uniquely by the function  $f$ , the set  $X$

and the measure  $\mu$ . We denote by  $L^1(\mu)$  the set of all complex measurable functions on  $X^0$  for which

$$\int_{X^0} |f| d\mu < \infty. \quad (4.22)$$

Let us go back to the study of a non-atomic structure represented by a continuum model. We suppose that there is a variable  $\hat{\mathcal{V}}$  whose values are positive real numbers and its definition domain covers the whole set  $\mathcal{P}$  at every state. If  $\mathcal{P}$  is a set  $\mathcal{P}_{sub}$  of a thermomechanical system then the volume is a good candidate for such a variable.

Remark: The set  $\mathcal{P}$  is not understood here as representing the whole structure. When studying a thermomechanical system, for example, it may represent the set  $\mathcal{P}_{sub}$  or a set of all subparts of an oriented boundary  $\partial\sigma$  (including  $\partial\sigma$  itself). Each such a substructure then may have different variables  $\hat{\mathcal{V}}$  (e.g. the volume for elements of  $\mathcal{P}_{sub}$  and the surface for  $\partial\sigma$  and its subparts).

Let us define for every  $s \in \Sigma$  a  $\sigma$ -algebra  $\mathfrak{A}^{\mathcal{X}_s}$  on  $X_s^0$  so that

$$\mathcal{X}_s \subset \mathfrak{A}^{\mathcal{X}_s}, \quad (4.23)$$

and choose a nonnegative measure  $\mu_{\mathcal{V}}$  on  $\mathfrak{A}^{\mathcal{X}_s}$  that coincides with  $\hat{\mathcal{V}}$  in the following sense. If  $X_s \in \mathcal{X}_s$  then

$$\mu_{\mathcal{V}}(X_s) = \mathcal{V}_s(\mathfrak{M}_s^{-1}(X_s)). \quad (4.24)$$

The measure  $\mu_{\mathcal{V}}$  is thus fixed only on sets that are models of parts. Consequently, measures of these sets are nonzero (because  $\mathcal{V}_s$  is always a positive number). What about the other sets in  $\mathfrak{A}^{\mathcal{X}_s}$ , i.e. sets that are not models of parts? May some of them have a nonzero measure, too? The important question is how wide is the group of such sets because the larger it is the

higher is the arbitrariness in modeling (realize that the value of  $\mu_{\mathcal{V}}(X)$  may be chosen arbitrary if  $X$  belongs into this group). Therefore we define:

We say that the model of  $\mathcal{P}$  is **scale complete** if the  $\sigma$ -algebra  $\mathfrak{A}^{\mathcal{X}_s}$  may be chosen in such way that the condition,

$$X \in \mathfrak{A}^{\mathcal{X}_s} \setminus \mathcal{X}_s \Rightarrow \mu_{\mathcal{V}}(X) = 0, \quad (4.25)$$

is valid for every state  $s$ .

**Remark:** The expression “scale complete” is chosen from the following reason. If  $\mathfrak{A}^{\mathcal{X}_s} \setminus \mathcal{X}_s$  includes some sets with a nonzero measure  $\mu_{\mathcal{V}}$  it may be interpreted as the absence of parts at certain size scales. Notice, that the scale completeness may not be defined only for continuum models. It is not clear, however, if an arbitrary continuum model may be supposed to be scale complete.

Let  $\hat{q}$  be a variable whose values are complex numbers<sup>9</sup> and

$$\hat{q} \ll \hat{\mathcal{V}} \quad (4.26)$$

(i.e.  $\hat{q}$  is absolutely continuous with respect to  $\hat{\mathcal{V}}$ ). Let us define a complex measure  $\lambda_s^q$  on  $\mathfrak{A}^{\mathcal{X}_s}$  corresponding to this variable by the prescription: If  $X_s \in \mathcal{X}_s$  then

$$\lambda_s^q(X_s) = q_s(\mathfrak{M}_s^{-1}(X_s)). \quad (4.27)$$

If  $X \in \mathfrak{A}^{\mathcal{X}_s} \setminus \mathcal{X}_s$ , we define its value arbitrary but in such a way that  $\lambda_s^q$  is absolutely continuous<sup>10</sup> with respect to  $\mu_{\mathcal{V}}$ . The Radon-Nikodym theorem [36] then says that there exists a function  $\phi_s^q \in L^1(\mu_{\mathcal{V}})$ ,

$$\phi_s^q : X_s^0 \rightarrow \mathbf{B}_q, \quad (4.28)$$

<sup>9</sup>This variable may be a component of a vector or tensor variable.

<sup>10</sup>The absolute continuity of one measure with respect to another is defined in the same way as for variables [36].

so that for every state  $s$  and every  $X \in \mathfrak{A}^{\mathcal{X}_s}$

$$\lambda_s^q(X) = \int_X \phi_s^q(u) d\mu_{\mathcal{V}}. \quad (4.29)$$

Especially, for sets from  $\mathcal{X}_s$  (i.e. sets that are models of parts), we can write

$$q_s(p) = \int_{\mathfrak{m}_s(p)} \phi_s^q(u) d\mu_{\mathcal{V}}. \quad (4.30)$$

The function  $\phi_s^q(u)$  is uniquely determined by the measure  $\lambda_s^q$  almost everywhere on  $X_s^0$ . That means that if we have two such functions then they may differ only on such a set  $X' \subset X_s^0$  whose measure  $\mu_{\mathcal{V}}(X') = 0$ . Such sets, however, cannot be models of parts from  $\mathcal{P}$  because any part from  $\mathcal{P}$  has a nonzero value of the variable  $\hat{\mathcal{V}}$ .

Nevertheless, the arbitrariness in the function  $\phi_s^q$  may be essentially higher if the model of the structure is not scale complete. Then there are sets in  $\mathfrak{A}^{\mathcal{X}_s} \setminus \mathcal{X}_s$  whose measure  $\mu_{\mathcal{V}}$  is nonzero. On these sets, the measure  $\lambda_s^q$  (which represents the variable  $\hat{q}$ ) may be chosen arbitrarily, which implies an additional arbitrariness in the choice of  $\phi_s^q(u)$  at a set .

Whatever the case, the set  $\{\phi_s^q \mid s \in \Sigma\}$  determines values of the variable  $\hat{q}$  on *all* parts from  $\mathcal{P}$  at arbitrary states. It depends on  $u$  but when using (4.18), we may construct a function  $\tilde{\phi}_s^q$  that depends on elements of the index set  $\Upsilon$ , i.e.

$$\tilde{\phi}_s^q(\xi) \equiv \phi_s^q(u_s(\xi)). \quad (4.31)$$

Now, we may understand  $\tilde{\phi}_s^q(\xi)$  as values of a variable on atomic parts from  $\mathcal{P}_{atom} = \Omega_0$ . This variable may not be defined on *all* atomic parts (since  $\phi_s^q$  also may not be defined everywhere) but it is defined on a sufficient amount of these parts so that

the value of variable  $\hat{q}$  is determined for any part at every state. When denoting

$$L^q \equiv \{\varphi_s^q, s \in \Sigma\},$$

we may enlarge the state space (4.21) into the state space

$$\Gamma' = \mathfrak{C} \cup L^{q_1} \cup L^{q_2} \cup \dots, \quad (4.32)$$

where  $L^{q_i}$  are additive variables that are absolutely continuous with respect to  $\hat{\mathcal{V}}$ .

However, if configuration dependent variables and additive variables do not form a complete set of variables then the state space  $\Gamma'$  does not represent fully the state set. It may happen when there is some essential variable, that is neither configuration dependent nor additive, and that cannot be replaced by some other configuration dependent or additive variables.

## Chapter 5

# Space

*Isn't space the machine that causes all those size-related miracles?*

V. Šedivý<sup>1</sup>

What is space? There are two main views about the nature of space in western thinking. Space is understood either as an *entity* of its own or as a *relation* between bodies ([35], p. 38). Though the second viewpoint was dominant in western culture (from Aristotle to Descartes, see [18]), physics, eventually, anchored its foundations in the idea of independent space which exists by itself. It was done by I. Newton in his “Principia”, [29]. This construction has secured physical thinking from an extreme relativity yielding by the idea of space which is just a relation between bodies (more in Chapter 6). The Newton’s conception of absolute space has shown itself as extraordinary successful. But what is space in the conceptual framework of our approach?

---

<sup>1</sup>[38], p. 222.

In this chapter, let us try to study space in the Newtonian spirit, i.e., as if it were a special entity. However, we have introduced no other entities than ‘physical systems’. If space is an entity then it has to be a special physical system. Newtonian space is a ‘container’ holding material objects and having an independent, absolute existence. Thus it may be understood as an absolutely indifferent surrounding. Therefore, we introduce a special class of systems that are studied *only* as surroundings. We postulate properties of these systems to model Newton’s space as well as possible. Especially the way in which Newtonian space saved physics from devastating relativity, plays eminent role in our “modeling”. Namely, space yields absolute *places*<sup>2</sup>– points of space. In our scheme, spatial points are understood as special parts of the description of a physical system. They are *defined* by its surrounding (space) in the spirit of the definition in Section 3.2. To describe points concretely, we use a suitable set of auxiliary physical systems called the coordinate systems. These systems also enable us to define *symmetries* of space as a physical indiscernibleness of descriptions in different coordinate systems.

The most important point is the relation of the system modeling space and an (arbitrary) system which is *in this space*. Namely, the first is the surroundings of the latter, and we know that surrounding is a source of all manipulations with the sys-

---

<sup>2</sup>Z. Neubauer gives an extraordinary deep contemplation about the role of ‘place’ (in a general sense) in the whole western thinking. He shows, for example, that the alphabet, as alphabetical ordering, is an order that allows constitution of a place as such. He emphasizes that “Aristotle claimed that if there truly were something like that (a place as such), that is if it existed in itself, it would be something so amazing, frightening and powerful, with so sovereign power that it would precede everything and be superior to everything - it would be omnipotent”, [28].

tem. What does it mean? Any manipulation is a set of changes. Any change of the system, which is in space only, is a manifestation of its inner dynamics. Hence, in our scheme, the inner dynamics of the system is *caused* by space. Though it may look rather strange, it only reflects the general line of thinking motivated by thermodynamics that never understands bodies as being “alone”. A ‘communication’ between the body and its surroundings is always present, whatever its meaning.

The important property of Newton’s space is its immobility. It is modeled in our approach by the surrounding, called the *Newton space*, that has only one state. It implies that it may cause only *one* manipulation. The inner dynamics, however, includes many changes of every state (in dependence on time interval during which the system is observed). It means that the manipulation caused by the Newton space is very *roughly* defined. The crucial result of this chapter is the proof that such a “roughness” is a result of properties of the Newton space (especially its symmetries). Moreover, the proof outlines *why* the inner dynamics has a form of an infinite amount of resulting states without an assumption about a passage of time. Thus ‘time’ may be understood, roughly speaking, as a mathematical result of properties of the Newton space.

As emphasized in Section 2.2, physicists try to construct highly sophisticated experimental device to be able to manipulate with studied physical systems very finely. They do not wish to get many various results when acting on the system by a defined action. From this point of view, the Newton space is a very bad ‘device’ resembling a “man’s hammer blow” in Fig. 2.2<sup>3</sup>. What about to improve this ‘device’? As well as an

---

<sup>3</sup>At this point, our concept of space approaches very closely to ideas of V. Šedivý [38], who compares space with a ‘machine’ (see the motto).

undefined hammer blow may be replaced by a precise device that can impact by a strictly defined way, we look for a more “sophisticated” space. It means to find a complete decomposition of the manipulation caused by the Newton space. When doing that, we obtain a modified space – the surrounding – that causes changes parameterized by a length parameter. We call it the *Mach space* and show its correspondence with ideas of Ernest Mach based on the criticism of the Newton’s absolute space.

## 5.1 Space as a surrounding

In our conception, there is nothing else than various physical systems characterized by their states. States may be changed as results of various manipulations with these systems. There is no spatial “background” in which systems are located and no independent flow of time in which manipulations obtain a temporal determination. If we want to speak about ‘space’ and ‘time’, we must use only concepts like ‘system’, ‘state’, ‘manipulation’, ‘change’, and so on.

Space in a common understanding is something in which any physical system is sited. This ‘something’ is not a part of the system – space is always an external entity. But we have no other entities than physical systems in our approach. That is why we introduce a special physical system called simply the **space** and denoted as  $\mathcal{E}$ . In what follows, we define properties of this system so that they model main features of space as an independent entity. These properties do not specify this system uniquely – there are many different spaces. Afterwards, we will specify some important types of spaces. Then we will use the symbol  $\mathcal{E}$  with indices.

The first property should model the idea that physical systems are “in space”. The idea of “being in space” is deeply related with the main property of space, namely, that it includes fixed places – spatial points. Physical systems may occupy these places by various manners while places themselves remain always the same. In our approach, only parts introduced in Chapter 3 are understood as fixed (firmly chosen) elements of the description. Recall that a special set of parts,  $\mathcal{P}_{ch}$ , has been introduced to model the situation when these fixed elements are defined by the surrounding (e.g., as an external wall that defines a container with a gas).

Space may be characterized as a “tool” defining special kind of parts – spatial points. The main idea of our approach is that the set of parts  $\mathcal{P}_{ch}$ , defined by  $\mathcal{E}$ , includes also parts that may be identified with spatial points. We, however, have no possibility of determining a concrete point in space otherwise than by means of another object located in space. Namely, this object (if it has some suitable properties) may serve as a referential frame that enables us to describe points by their coordinates. It motivates us to the following definition:

*Let  $\mathfrak{S}_{sur}$  be an arbitrary surrounding of  $\mathfrak{S}$ . We say that  $\mathfrak{S}_{sur}$  **defines spatial points for  $\mathfrak{S}$**  if there is an integer  $d$  and a set of physical systems,  $\mathcal{C}_d = \{a, b, \dots\}$ , so that the composition of each of them with  $\mathfrak{S}_{sur}$  (i.e.  $(\mathfrak{S}_{sur}, a)$ , etc.) is the surrounding of  $\mathfrak{S}$ , too, and each of these surroundings defines sets of parts,  $\mathcal{P}_{ch}((\mathfrak{S}_{sur}, a)) \subset \mathcal{P}(\mathfrak{S})$ ,  $\mathcal{P}_{ch}((\mathfrak{S}_{sur}, b)) \subset \mathcal{P}(\mathfrak{S})$ , etc., so that for every  $a \in \mathcal{C}_d$  there is a uniquely given one-to-one mapping,*

$$f_{\mathcal{E},a} : \mathbb{E}_d \rightarrow \mathcal{P}_{ch}((\mathfrak{S}_{sur}, a)),$$

*that assigns to every point  $\mathbf{x} \in \mathbb{E}_d$  an atomic part  $f_{\mathcal{E},a}(\mathbf{x})$  called the **spatial point**.*

Systems  $\{a, b, \dots\}$  are called the ***d*-dimensional coordinate systems**. The first property of  $\mathcal{E}$  thus may be postulated as follows:

(S1) *If  $\mathcal{E}$  is a surrounding of a physical system  $\mathfrak{S}$ , then  $\mathcal{E}$  defines spatial points for  $\mathfrak{S}$ .*

If we have a concrete *d*-dimensional coordinate system *a* we denote

$$\mathcal{P}_{ch}(a) \equiv f_{\mathcal{E},a}(\mathbb{E}_d) \subset \mathcal{P}_{ch}((\mathfrak{S}_{sur}, a)).$$

Remark: The *identity* of points in different coordinate systems is expressed by the relation

$$f_{\mathcal{E},a}(\mathbf{x}) = f_{\mathcal{E},b}(\mathbf{x}'). \quad (5.1)$$

It allows constructing various transformation laws interpreting the description in one coordinate system into another. The set of coordinate systems,  $\mathcal{C}_d$ , is usually supposed to define the same set of spatial points, i.e.  $\mathcal{P}_{ch}(a) = \mathcal{P}_{ch}(b)$ .

*If there is a maximal *d* for which  $\mathcal{C}_d$  exists, we call  $\mathcal{E}$  the ***d*-dimensional space**. We say that the system  $\mathfrak{S}$  is ***in (d-dimensional) space*** if there is a (*d*-dimensional) space  $\mathcal{E}$  that is the surrounding of  $\mathfrak{S}$ .*

Let us have a system  $\mathfrak{S}$  in *d*-dimensional space and a concrete *d*-dimensional coordinate system *a* defining the subset of parts  $\mathcal{P}_{ch}(a) \subset \mathcal{P}_{ch}$  whose elements are spatial points. The variable  $\hat{q} \in Q$  is called the **field in *a*** if for any state  $s \in \Sigma(\mathfrak{S})$ ,

$$\mathcal{D}_{q,s} = \mathcal{P}_{ch}(a). \quad (5.2)$$

The mapping  $q_s$ , given by (3.7), defines the **spatial field**,

$$q_s^a = q_s \circ f_{\mathcal{E},a}, \quad (5.3)$$

that assigns to every point  $\mathbf{x} \in \mathbb{E}_d$  the value(s)  $q_s^a(\mathbf{x})$  from  $\mathbf{B}_q$ , i.e.

$$q_s^a : \mathbb{E}_d \rightarrow \mathbf{B}_q.$$

We can have various coordinate systems within  $\mathcal{C}_d$  so that one variable is the field in one of them but it is not in another. That is why we choose a “reasonable” class  $\mathcal{C}_d$  defined as follows:

*For any couple of coordinate systems  $a, b \in \mathcal{C}_d$ , for each variable  $\hat{q} \in Q$  and any states  $s, s_n \in \Sigma$ ,  $n \in \mathbb{N}$ , the following conditions hold:*

- (i) if  $\hat{q}$  is a field in  $a$  then it is a field in  $b$ ,*
  - (ii) if  $q_s^a$  is a constant function then  $q_s^b$  is a constant function,*
  - (iii) if  $q_s^a$  is a continuous function then  $q_s^b$  is a continuous function, and*
  - (iv) if for every  $\mathbf{x}$ ,  $q_{s_n}^a(\mathbf{x}) \rightarrow q_s^a(\mathbf{x})$  then  $q_{s_n}^b(\mathbf{x}) \rightarrow q_s^b(\mathbf{x})$ ,*
- where “ $\rightarrow$ ” means the convergence in  $\mathbf{B}_q$ .

In what follows, we will use a firmly chosen reasonable class  $\mathcal{C}_d$  of coordinate systems and study only systems whose set of variables,  $Q$ , includes *only fields* (in an arbitrary coordinate system from  $\mathcal{C}_d$ ). It allows us to define the convergence of states from  $\Sigma$ :

*An infinite succession of states  $s_n$  **converges** to the state  $s$ ,*

$$s_n \rightarrow s,$$

*if  $q_{s_n}^a(\mathbf{x}) \rightarrow q_s^a(\mathbf{x})$  for each  $\hat{q} \in Q$  at every  $\mathbf{x} \in \mathbb{E}_d$  (the coordinate system  $a$  may be chosen arbitrarily from  $\mathcal{C}_d$ ).*

Remark: The convergence is point-like (it is defined at individual spatial points) which means that behavior of series of fields representing states  $s_n$  at very distant points ( $|\mathbf{x}| \rightarrow \infty$ ) does not play an essential role. Imagine, for example, the sequence of fields  $q_{s_n}^a(\mathbf{x})$  so

that  $q_{s_n}^a(\mathbf{x}) = q_s^a(\mathbf{x})$  within a succession of balls with a fixed center  $\mathbf{x}_0$  and increasing radius,  $R_n \rightarrow \infty$ . This sequence converges to  $q_s^a$  even if values of every field  $q_{s_n}^a$  outside the ball with the radius  $R_n$  are chosen arbitrarily.

This convergence enables us to characterize when a manipulation is not sensitive on behavior of fields at infinity:

*Let  $\mathfrak{S}$  be in space. We say that the state  $s \in \Sigma(\mathfrak{S})$  is **stable in the manipulation**  $\alpha \in \mathcal{M}(\mathfrak{S})$  if for any succession of couples,  $[s_n, s'_n] \in \alpha$ , holds that whenever  $s_n \rightarrow s \in \Sigma$  then there exists a state  $s' \in \Sigma$  so that  $s'_n \rightarrow s'$ , too.*

Let us suppose that the structure of parts is atomic and  $\mathcal{P}_{ch}(a) = \mathcal{P}_{atom}$  for every  $a \in \mathcal{C}_d$ . The state space of every part from  $\mathcal{P}_{ch}(a)$  is supposed to be the same (and independent on  $a$ ), namely  $\Gamma(P^a) \equiv \Gamma_0$ , and

$$\Gamma_0 = \mathbf{B} \equiv \mathbf{B}_{q_1} \times \dots \times \mathbf{B}_{q_n}, \quad (5.4)$$

where  $\mathbf{B}_{q_i}$  are Banach spaces of individual variables. Following (3.18), let us define the set of mappings  $\Gamma^a$  that represents various possibilities of assigning values of all variables to all spatial points<sup>4</sup>. Since spatial points are defined uniquely by the chosen coordinate system, the set  $\Gamma^a$  may be understood as a set of mappings from  $\mathbb{E}_d$  to  $\mathbf{B}$ . An element of  $\Gamma^a$  will be denoted simply by  $\psi$  (with omitting the specification of the coordinate system whenever there is no menace of misunderstanding). That means that  $\psi \in \Gamma^a$  is a mapping,

$$\psi : \mathbb{E}_d \rightarrow \mathbf{B}, \quad (5.5)$$

---

<sup>4</sup>These possibilities are thus characterized by a form of mappings in  $\Gamma^a$ . They may be continuous, differentiable, integrable, etc.

defined at every point  $\mathbf{x}$  by the prescription,

$$\psi(\mathbf{x}) = [q_s^a(\mathbf{x})]_{\hat{q} \in Q}. \quad (5.6)$$

We will always suppose that the set of variables  $Q$  (i.e. the set of fields) is *complete* with respect to the state space  $\Sigma$  of an arbitrary system under study. That means that any concrete  $\psi$  represents one state from  $\Sigma$ .

The set of parts  $\mathcal{P}_{ch}(a)$  consists of an uncountable number of spatial points. Every state  $s$  is represented by a set of spatial fields,  $\psi$ , whose values are defined at these points. We understand these fields as “distributions of physical properties in space”. Points, however, are not some separated parts – they form an infinitely dense continuum. That means that we should assume some “correlations” of values of fields at “nearby” points. For example, the continuity of fields  $\psi$  at a point  $\mathbf{x}$  means that values at points “infinitesimally near”  $\mathbf{x}$  cannot differ arbitrarily from  $\psi(\mathbf{x})$ . These correlations, however, are not only properties of individual states alone but also of their *changes*. If  $\psi$  changes to  $\tilde{\psi}$  (a set of fields with different values) the correlations at  $\mathbf{x}$  mean that  $\tilde{\psi}(\mathbf{x})$  depends also on values of  $\psi$  at some circumjacent points, i.e.  $\psi(\mathbf{x}')$ , and not only on  $\psi(\mathbf{x})$ . It motivates us to the following definition:

*Let  $\alpha \in \mathcal{M}$  be a manipulation with  $\mathfrak{S}$  caused by the surrounding  $\mathcal{E}$ . We say that  $\alpha$  is the **trivial manipulation** if there is a function  $f : \mathbf{B} \rightarrow \mathbf{B}$  so that for any couple  $\psi, \tilde{\psi} \in \Gamma^a$ , which represents states  $s, \tilde{s}$ , the condition*

$$[s, \tilde{s}] \in \alpha \Rightarrow \tilde{\psi}(\mathbf{x}) = f(\psi(\mathbf{x})), \quad (5.7)$$

*holds at every point where all fields in  $\psi$  are continuous.*

The trivial change thus represents situations when values of

changed fields at  $\boldsymbol{x}$ , i.e.  $\tilde{\psi}(\boldsymbol{x})$ , depend only on values of fields  $\psi$  at  $\boldsymbol{x}$ . It means that there is no influence of fields at surrounding points. The definition concerns only points at which all fields are continuous (at points of discontinuity, there is even no “initial” correlation of values of fields at near points). If all changes were trivial, spatial points would not be “cohesive” and space would disintegrate into an infinite number of incoherent elements. Therefore, we accept the following property of  $\mathcal{E}$ :

(S2) *Let  $\mathfrak{S}$  be in space  $\mathcal{E}$ . Then there exists a manipulation with  $\mathfrak{S}$ , caused by  $\mathcal{E}$ , that is not trivial.*

This property says more than that the relation (5.7) is not valid. Namely, if all states of the system  $\mathfrak{S}$  are represented by fields that are discontinuous at every point, then *any* manipulation is trivial. Therefore the property (S2) claims also that any physical system in space has a state at which all representing fields are continuous at least at one point.

In our approach, space is represented by a physical system,  $\mathcal{E}$ , though we do not introduce its “material” properties. Its nonmaterial character underlines the fact that it extends into infinity (in the sense that there are successions of spatial points, represented by coordinates  $\boldsymbol{x}_n$ , so that  $|\boldsymbol{x}_n| \rightarrow \infty$ ). The space  $\mathcal{E}$  is thus a little strange “entity” without a substantial specification.

How could space be approached by a material substance? Newton’s space, for example, may be *modeled* as an infinite, uniform substance (remember the idea of ether). But no infinite substance exists. To guarantee that a sufficiently large substance behaves like an infinite one, we accept this property of  $\mathcal{E}$ :

(S3) *If state  $s$  models the space  $\mathcal{E}$  then it is stable in any manipulation caused by  $\mathcal{E}$ .*

State  $s$  is an infinite (hypothetical) state of a material system that is in  $\mathcal{E}$  (e.g. an uniform substance). The stability means that this state may be approximated by a succession of states that are not infinite (they differ “at infinity”). The meaning of the fact that a state *models* the space will be defined in the next section by (5.10).

## 5.2 Symmetries of space

To characterize more various spaces, we define the crucial property of any space – its symmetries. We do it by using the concept of coordinate system. It is a material (or geometrical) structure that is not a subsystem of  $\mathcal{E}$ . The composition of a coordinate system, say  $a$ , and  $\mathcal{E}$  is a new physical system,  $(\mathcal{E}, a)$ . If the composition is done with another coordinate system  $b$ ,  $b \neq a$ , we obtain another physical system,  $(\mathcal{E}, b)$ . Though the systems  $(\mathcal{E}, a)$  and  $(\mathcal{E}, b)$  are different, they may be physically indistinguishable.

A model of such a situation is an infinite homogeneous environment that represents  $\mathcal{E}$ , and a system of three perpendicular infinite bars endowed with a sufficient set of scale lines. Let the system of bars be in a fixed position in the environment. Then it represents a coordinate system  $a$  because whenever a material object is in the environment, triplets of scale lines on individual bars determine localized parts of the system. If the same system of bars is fixed in another position in the environment, the same triplet of scale lines determines another part of the system. Thus the bars in another position represent another

coordinate system  $b$ . If we have only the environment and one of the coordinate systems, we cannot distinguish whether it is  $a$  or  $b$ . This fact represents a symmetry of the environment. If, for example, the environment is not homogeneous (i.e. it includes a salient place), the coordinate systems  $a$  and  $b$  can be simply distinguished because their positions may be determined with respect to this place. That is why we define:

*Coordinate systems  $a$  and  $b$  are said to be **equivalent** if there is no property (e.g. a number) having a source either in  $\mathcal{E}$  or in the coordinate systems alone which can distinguish them. For example, there is no function  $f$  so that*

$$f(\mathcal{E}, a) \neq f(\mathcal{E}, b). \quad (5.8)$$

A typical example of equivalent coordinate systems are the same but translated coordinate systems in the Euclidean space. On the other hand, the orthogonal and non-orthogonal ones are not equivalent because  $f(\mathcal{E}, a)$  from the relation (5.8) may represent the angle between two axis. If there is a salient (localized) place (point)  $P$  in  $\mathcal{E}$  then the translated coordinate systems may not be equivalent because  $f(\mathcal{E}, a)$  may be chosen as the distance of the origin of coordinate system  $a$  to  $P$ .

*Let  $\mathcal{G}$  be a mapping which transforms states of the system to other states, i.e.*

$$\mathcal{G} : \Sigma \rightarrow \Sigma.$$

*We say that  $\mathcal{G}$  is the **symmetry operation** of  $\mathcal{E}$  if there are equivalent coordinate systems  $a$  and  $b$ , and  $\mathcal{G}$  maps any state  $s$  to the state  $\mathcal{G}(s)$  so that for each  $\hat{q} \in Q$ ,*

$$q_s^a = q_{\mathcal{G}(s)}^b. \quad (5.9)$$

The concept of symmetry operation allows us to define states that model the space  $\mathcal{E}$ :

The state  $s$  of a system in  $\mathcal{E}$  is said to **model**  $\mathcal{E}$  if for any symmetry operation  $\mathcal{G}$  of  $\mathcal{E}$ ,

$$\mathcal{G}(s) = s. \quad (5.10)$$

Symmetry operations characterize symmetries of  $\mathcal{E}$ . We suppose that the domain of each symmetry operation is the whole state set  $\Sigma$ . Let  $\alpha \in \mathcal{M}(\mathfrak{S})$  be an arbitrary manipulation with  $\mathfrak{S}$  caused by the surrounding  $\mathcal{E}$ . Symmetries of space have to influence a concrete form of the manipulation  $\alpha$  because it is fully determined by this space. Namely, any couple  $[s, s'] \in \alpha$  has to respect all symmetries of the space. It may be formulated as the following statement:

**Proposition:** *If  $\mathcal{G}$  is a symmetry operation of  $\mathcal{E}$  then*

$$[s, s'] \in \alpha \Rightarrow [\mathcal{G}(s), \mathcal{G}(s')] \in \alpha. \quad (5.11)$$

The condition (5.11) is justified by the fact that the states  $s$  and  $\mathcal{G}(s)$  are equivalent from a physical point of view:  $s$  in  $(\mathcal{E}, a)$  is indistinguishable from  $\mathcal{G}(s)$  in  $(\mathcal{E}, b)$  and the couples  $(\mathcal{E}, a)$ ,  $(\mathcal{E}, b)$  alone cannot be differed anyhow.

Let a system  $\mathfrak{S}$  be in space  $\mathcal{E}$ . There is a deep relation between symmetries of the space  $\mathcal{E}$  and a possibility of finding global variables of  $\mathfrak{S}$  so that  $\mathcal{E}$  has conservative properties with respect to these variables. Namely, the condition (5.11) allows defining such variables. We, however, do not study these interesting questions here.

### 5.3 Special symmetries

Let us define for every  $\mathbf{y} \in \mathbb{E}_d$  the mapping

$$T_{\mathbf{y}} : \Sigma \rightarrow \Sigma$$

as follows:  $T_y(s) = s'$  if and only if for each  $\hat{q} \in Q$  and every  $a \in \mathcal{C}_d$ ,

$$q_{s'}^a(\mathbf{x}) = q_s^a(\mathbf{x} - \mathbf{y}). \quad (5.12)$$

If  $T_y$  is the symmetry operation then the equality (5.9) defines when coordinate systems  $a$  and  $b$  are equivalent. It has to be valid at every  $\mathbf{x} \in \mathbb{E}_d$ . Hence, (5.12) implies that

$$q_s^a(\mathbf{x}) = q_{T_y(s)}^b(\mathbf{x}) = q_s^b(\mathbf{x} - \mathbf{y}), \quad (5.13)$$

and thus the coordinate system  $b$  is translated by the vector  $\mathbf{y}$  with respect to  $a$ . The symmetry operation  $T_y$  is called the **translation**. If  $T_y$  is the symmetry operation for every  $\mathbf{y} \in \mathbb{E}_d$  then the surrounding  $\mathcal{E}$  is said to be **homogeneous**. It expresses the fact that there is no salient place within it. The homogeneity of space plays very important role in physics.

Similarly, we can define the symmetry operation called **rotation**. The definition is analogous to the previous one. Let  $\mathbf{Q}$  be a properly orthogonal matrix, i.e.

$$\mathbf{Q}^{-1} = \mathbf{Q}^\top, \quad \det \mathbf{Q} = 1. \quad (5.14)$$

For every matrix  $\mathbf{Q}$  fulfilling (5.14), the mapping

$$R_Q : \Sigma \rightarrow \Sigma$$

is defined as follows:  $R_Q(s) = s'$  if and only if for every  $\hat{q} \in Q$  and every  $a \in \mathcal{C}_d$ ,

$$q_{s'}^a(\mathbf{x}) = q_s^a(\mathbf{Q}\mathbf{x}). \quad (5.15)$$

If  $R_Q$  is the symmetry operation then it means that the coordinate systems  $a$  and  $b$ , fulfilling

$$q_s^a(\mathbf{x}) = q_s^b(\mathbf{Q}\mathbf{x}), \quad (5.16)$$

are equivalent. The symmetry operation  $R_Q$  is called the **rotation**. If  $R_Q$  is the symmetry operation for every  $\mathbf{Q}$  the surrounding  $\mathcal{E}$  is said to be **isotropic**. It means that there is no salient “direction” in  $\mathcal{E}$ .

The last symmetry studied in this section arises when there is no salient length in  $\mathcal{E}$ . Let us realize that coordinates of points  $\mathbf{x}$  express lengths in  $\mathcal{E}$  (coordinate systems endow  $\mathcal{E}$  with a metric). Therefore, if a new length unit is used we obtain a new inscription by “ $\mathbf{x}$ ”, i.e., we get a new coordinate system. Let us introduce the mapping  $\mathcal{O}_r : \Sigma \rightarrow \Sigma$  for any positive real number  $r \in \mathbb{R}_+$  as follows:  $\mathcal{O}_r(s) = s'$  if and only if

$$q_{s'}^a(\mathbf{x}) = q_s^a(r^{-1}\mathbf{x}). \quad (5.17)$$

There may be length units among dimensions of values from  $q_s^a$ . However, the change of values of fields caused by the change of length unit may be compensated by an appropriate change of other units. A condition allowing doing that is that dimensions of individual fields,  $[q_i]$ , where  $\hat{q}_i \in Q$ , cannot be combined to give the dimension of length [1]. That is, if for any real numbers  $p_1, \dots, p_n$ ,

$$[q_1]^{p_1} \dots [q_n]^{p_n} \neq \text{length}. \quad (5.18)$$

Then the relation (5.17) expresses a description of the field  $q_s^a$  in a new coordinate system using the  $r$  times smaller or larger length unit. Now, if there is no salient length in  $\mathcal{E}$  then there is no number coming from  $\mathcal{E}$  or the coordinate system alone which can distinguish coordinate systems  $a, a'$  using different length units (the used length unit cannot be compared with anything else to give a number which is different for individual couples  $(\mathcal{E}, a), (\mathcal{E}, a')$  alone). We say that  $\mathcal{E}$  is **scale symmetric** and the mapping  $\mathcal{O}_r$  is the corresponding symmetry operation.

## 5.4 Newton space

In Newton's "Principia" [29], space is defined as something external what remains always similar and immovable. The immobility of this space may be translated to our language as that  $\mathcal{E}$  can be only in one state, say  $s_0$ . This state has to have very special properties because space introduced by Newton includes no salient places, directions or lengths. Otherwise speaking, it is a one-state surrounding that is homogeneous, isotropic and scale symmetric. We will call it the **Newton space** and denote this space as  $\mathcal{E}_{Newt}$ .

The fact that the Newton space has only one state,  $s_0$ , implies that *any* change of this physical system is the *identity*,  $id = \{[s_0, s_0]\}$ . This space, as a surrounding, thus may caused only *one* manipulation with an arbitrary system being in this space because

$$\Sigma_{sur} \times \mathcal{M}_{sur} = \{[s_0, id]\},$$

and there is only one-element domain of the function  $\mathfrak{F}$  in the relation (2.6). Let us call this universal manipulation the **Newton manipulation** and denote it as  $\epsilon$ , i.e.

$$\epsilon = \mathfrak{F}([s_0, id]). \quad (5.19)$$

Notice that the relation (2.11) implies also that if the Newton space has the conservative property with respect to a global variable  $G$  then the value of this variable is the same at every state of the system because  $s_{sur} = s'_{sur} = s_0$ .

How does the manipulation  $\epsilon$  act on the state space of a physical system  $\mathfrak{S}$  for which the Newton space is the surrounding? That is, what is the subset of  $\Sigma \times \Sigma$  of various changes? In Newtonian physics, the system "located in empty space"

changes its initial state,  $s$ , into an infinite set of states parameterized by a parameter  $t$ . It expresses the inner dynamics of the system and the parameter  $t$  is called the **time**. In our scheme, the manipulation  $\epsilon$  thus includes, for any state  $s \in \Sigma$ , an infinite set of couples,

$$\text{Dyn}(s) \equiv \{[s, s(t)], t \in \mathbb{R}_{0,+}\} \subset \epsilon, \quad (5.20)$$

if we denote  $s(0) \equiv s$ .

Is this conclusion given only by empirical observations (that physical systems have inner dynamics changing their states into infinite sequences of states) or has it a deeper origin? Surprisingly, the fact that the Newton space manipulates with any system *in this way* is a consequence of its symmetries. To show it, let us proof the following theorem first.

**Theorem:** The manipulation  $\epsilon$  caused by the Newton space cannot be perfect for any system  $\mathfrak{G}$  which is in the Newton space, i.e., there are states  $s, s_1, s_2 \in \Sigma(\mathfrak{G})$ ,  $s_1 \neq s_2$ , so that  $[s, s_1], [s, s_2] \in \epsilon$ .

**Proof:** Let us suppose that the manipulation  $\epsilon$  is perfect, i.e., it is a mapping  $\epsilon : \Sigma \rightarrow \Sigma$  so that  $[s, \tilde{s}] \in \epsilon$  means that

$$\tilde{s} = \epsilon(s). \quad (5.21)$$

Due to the homogeneity and scale symmetry of the Newton space, (5.21) implies that also

$$T_y(\tilde{s}) = \epsilon(T_y(s)) \quad (5.22)$$

for every  $\mathbf{y} \in \mathbb{E}_d$ , and

$$\mathcal{O}_r(\tilde{s}) = \epsilon(\mathcal{O}_r(s)) \quad (5.23)$$

for every  $r > 0$ . In the field representation, states become sets of fields and the mapping  $\epsilon$  becomes a mapping  $\mathcal{T}$  between them. If we

represent  $s$  by  $\psi$  and  $\tilde{s}$  by  $\tilde{\psi}$  we obtain from (5.21) the definition of  $\mathcal{T}$ , namely

$$\tilde{\psi} = \mathcal{T}(\psi). \quad (5.24)$$

Let us have an arbitrary couple of states  $s, \tilde{s}$  fulfilling (5.21), i.e. an arbitrary couple of fields  $\psi$  and  $\tilde{\psi}$  fulfilling (5.24). Let  $\mathbf{x}_0$  be an arbitrary point at which fields  $\psi$  are continuous (such a point has to exist as implied by (S2)). We will study values of fields  $\tilde{\psi}$  at this point, i.e.  $\tilde{\psi}(\mathbf{x}_0)$ , in dependence on  $\psi$ . The relation (5.22) guarantees that if we choose a translated coordinate system, the mapping transforming the fields corresponding to  $s$  and  $\tilde{s}$  in the new coordinate system remains the same. That is why we may suppose that  $\mathbf{x}_0 = \mathbf{0}$  because we can ever choose a coordinate system with the center at an arbitrary point.

Now, let us choose a coordinate system using a smaller length unit. The scale symmetry of the Newton space guarantees that the mapping  $\mathcal{T}$  remains the same (as implied by (5.23)). Taking smaller and smaller units we construct a succession of fields,  $\psi_n$ , that converges at arbitrary point  $\mathbf{x} \in \mathbb{E}_d$  to  $\psi(\mathbf{0})$ ,

$$\lim_{n \rightarrow \infty} \psi_n(\mathbf{x}) = \psi(\mathbf{0}). \quad (5.25)$$

That means that the limit of the succession of states  $s_n$  corresponding to the fields  $\psi_n$  is a constant field,  $s_c$ , whose field representation is  $\psi_0$ , defined as

$$\psi_0(\mathbf{x}) \equiv \psi(\mathbf{0}). \quad (5.26)$$

Any state representing by constant fields is a model of  $\mathcal{E}_{Newt}$ . Hence (S3) implies that  $\epsilon(s_n)$  converges to  $\epsilon(s_c)$ . In the field representation,  $\mathcal{T}(\psi_n)$  converges to  $\mathcal{T}(\psi_0)$  at every point. On the other hand,

$$\lim_{n \rightarrow \infty} (\mathcal{T}(\psi_n))(\mathbf{0}) = \tilde{\psi}(\mathbf{0}), \quad (5.27)$$

because (5.23) guarantees that the re-scaled field  $\psi$  is mapped onto the re-scaled field  $\tilde{\psi}$  and all re-scaled fields have the same value at  $\mathbf{0}$ . As a result, we obtain that

$$\tilde{\psi}(\mathbf{0}) = \mathcal{T}(\psi_0)(\mathbf{0}). \quad (5.28)$$

As implied by (5.26),  $\psi_0$  is defined only by the value  $\psi(\mathbf{0})$ . Let us define a function  $f : \mathbf{B} \rightarrow \mathbf{B}$  so that

$$f(\mathbf{b}) = \mathcal{T}(\psi_{con}^b)(\mathbf{0}), \quad (5.29)$$

where  $\psi_{con}^b$  is a set of constant fields so that

$$\psi_{con}^b(\mathbf{x}) = \mathbf{b} \quad (5.30)$$

(notice that the scale invariance guarantees that the state corresponding to  $\psi_{con}^b$  has to exist in  $\Sigma$  if there is a state  $s_b \in \Sigma$  so that there exists a point  $\mathbf{x} \in \mathbb{E}_d$  at which fields representing  $s_b$  are continuous and reach the values  $\mathbf{b}$ ). The relation (5.28) then may be written as

$$\tilde{\psi}(\mathbf{0}) = f(\psi(\mathbf{0})). \quad (5.31)$$

The condition (5.31) can be written in the original coordinate system (because it depends only on *values* of the fields at the point  $\mathbf{x}_0$ ) and we obtain

$$\tilde{\psi}(\mathbf{x}_0) = f(\psi(\mathbf{x}_0)). \quad (5.32)$$

This condition is valid at all points of  $\mathbb{E}_d$  where fields  $\psi$  are continuous. It means that the manipulation  $\epsilon$  is trivial (see the definition at Section 5.1). Since there is no other manipulation than  $\epsilon$  we have obtained a contradiction with the property (S2). Therefore the mapping  $\mathcal{T}$  cannot exist.  $\square$

The core of the proof is the validity of (5.23) implied by the scale symmetry of  $\mathcal{E}_{Newt}$  (and by the assumption (5.18) which say, for example, that  $\psi$  cannot include a field whose values have the dimension of length). If  $\epsilon$  were dependent on a length  $\ell$ , i.e.  $\epsilon_\ell$ , the relation (5.23) could not be in force and  $\epsilon_\ell$  might be a mapping. An illustrative example is the gaussian transformation,

$$\tilde{\psi}(\mathbf{x}) = \pi^{-d/2} \ell^{-d} \int_{\mathbb{E}_d} e^{-\frac{(\mathbf{x}-\mathbf{x}')^2}{\ell^2}} \psi(\mathbf{x}') d^d \mathbf{x}', \quad (5.33)$$

that defines a perfect mapping between spatial fields whose domain is a broad class of fields  $\psi$  (if  $\psi$  includes more fields, (5.33) is defined for each component). Notice that the occurrence of the length parameter  $\ell$  is essential because the dimension of coordinates  $\mathbf{x}'$  has to be compensated by a length parameter (the homogeneity of space does not make possible using  $\mathbf{x}$  in place of  $\ell$ )<sup>5</sup>.

A length-dependent manipulation  $\epsilon_\ell$ , however, cannot be caused by the Newton space because it does not include a salient length  $\ell$ . The existence of such a manipulation would break its scale symmetry. The impossibility of occurrence of a *salient length* in  $\mathcal{E}_{Newt}$  may be also interpreted as the fact that  $\mathcal{E}_{Newt}$  includes *all lengths*. This expression may seem as rather unusual, but notice that the same interpretation is done in the case of homogeneity:  $\mathcal{E}_{Newt}$  cannot include a salient place, so it is supposed to include *all places*.

This small shift in locution, however, brings a remarkable result. Namely, if  $\mathcal{E}_{Newt}$  includes all lengths then it may caused (“simultaneously”) *all* mapping  $\epsilon_\ell$  parameterized by *all* length parameters. We may do it it formally by understanding  $\epsilon$  as an infinite “vector”,

$$\epsilon \equiv (\dots, \epsilon_\ell, \dots), \quad (5.34)$$

whose “components” are indexed by all lengths. Individual mappings,  $\epsilon_\ell$ , do not obey the condition (5.23) because the operator  $\mathcal{O}_r$  acts as  $\mathcal{O}_r(\epsilon_\ell(s)) = \epsilon_{r\ell}(\mathcal{O}_r(s))$ . That is,  $\mathcal{O}_r$  changes the length  $\ell$  into the length  $r\ell$  and thus realizes a “shift” of components in  $\epsilon$ . However, it does not change the “vector”  $\epsilon$  as a whole! The relation (5.23) thus may be formally understood

---

<sup>5</sup>When trying to “construct” a mapping  $\mathcal{T}$  without the use of a length parameter, we either disturb the homogeneity of the problem or obtain a trivial solution,  $\tilde{\psi}(\mathbf{x}) = f(\psi(\mathbf{x}))$ .

as remaining in force for the whole “vector”,

$$\epsilon(\mathcal{O}_r(s)) = \mathcal{O}_r(\epsilon(s)).$$

This interpretation of  $\epsilon$  leads to a “discovery” of time dynamics! Namely, the set of all lengths is isomorphic to the set of all positive real numbers. If we identify  $\ell \sim t \in \mathbb{R}_+$  and denote  $\epsilon_\ell(s) \equiv s(t)$  we have

$$\epsilon(s) = (\dots, \epsilon_\ell(s), \dots) \equiv \{s(t), t \in \mathbb{R}_+\}. \quad (5.35)$$

When adding the identity change,  $[s, s]$ , we can write (5.35) in the form (5.20). Thus the structure of time dynamics seem to appear as a result of symmetries of the Newton space.

## 5.5 Complete decomposition of $\epsilon$

The one-state Newton space can cause only one manipulation  $\epsilon$ . Since this space is homogeneous and scale symmetric,  $\epsilon$  cannot be a mapping. We have uncloaked a possible structure of  $\epsilon$  as consisting of infinite set of mappings parameterized by all length parameters. From now, we take this structure as given (taking into account also its “experimental evidence” in a form of a permanently observed time evolution).

The “vector” (5.34) resembles a parallel composition of the manipulation  $\epsilon$ . Since  $\epsilon_\ell$  are supposed to be mappings, (5.34) presents indeed a *complete decomposition* of  $\epsilon$ . In Section 2.2, we explained this concept as a result of searching for a better surrounding (e.g. a device) that can manipulate with the system much more finely to give only single valued results (see Fig. 2.2).

Let us try to introduce a “better” surrounding that causes the manipulations  $\epsilon_\ell$ . We will suppose that this surrounding is a

space, too, and denote it by  $\mathcal{E}_M$ . This space may be very similar to the Newton space (e.g. be homogeneous and isotropic) but it cannot be scale symmetric. Namely, (5.11) cannot be valid if  $\alpha \equiv \epsilon_{ell}$  and  $\mathcal{G} \equiv \mathcal{O}_r$ .

What should be its state set? For any length  $\ell$ ,  $\epsilon_\ell$  has to be defined via a given mapping  $\mathfrak{F}$  (see (2.6)), as

$$\epsilon_\ell = \mathfrak{F}(s_{sur}, \alpha_{sur}),$$

where  $s_{sur} \in \Sigma(\mathcal{E}_M)$  and  $\alpha_{sur} \in \mathcal{M}(\mathcal{E}_M)$ . It means that either the state of  $\mathcal{E}_M$  or the manipulation of  $\mathcal{E}_M$  has to depend on the parameter  $\ell$ . If states of  $\mathcal{E}_M$  were dependent on length parameters, then any couple of these states would depend on a *couple* of length parameters and hence any manipulation  $\alpha_{sur}$  would depend on the couple (or couples) of length parameters.

That is why we suppose that not states of  $\mathcal{E}_M$  alone but their couples,  $[s_{sur}, s'_{sur}]$ , depend on  $\ell$ . A simple example is the Newton space equipped with one material particle and a firmly chosen direction<sup>6</sup>. States of this space are given by shifting the position of the particle along the direction. While no length can be assigned to individual states (we have no possibility of defining the position of the particle in space alone<sup>7</sup>), any couple of different states of  $\mathcal{E}_M$  defines a length as a distance of the particle in these states. To be able to define a length also for the couple of identical states and differ shifts in opposite directions, we introduce the set  $\mathfrak{R}$ , whose elements are called the **generalized lengths**.  $\mathfrak{R}$  includes all lengths (i.e. objects denoted by  $\ell$ ), all minus lengths (denoted by  $-\ell$ ) and the object  $0_\ell$  called the zero length ( $0_\ell \equiv \ell - \ell$ ) ( $\mathfrak{R}$  is isomorphic to the set of real numbers).

<sup>6</sup>This space is neither isotropic nor homogeneous.

<sup>7</sup>Coordinate systems are not parts of space!

Being motivated by this example, we suppose that for any space  $\mathcal{E}_M$  there is a one-to-one mapping  $f_L$ ,

$$f_L : \Sigma_{sur} \times \mathfrak{R} \rightarrow \Sigma_{sur}, \quad (5.36)$$

so that for any  $s_{sur} \in \Sigma_{sur}$  the set  $\{f_L(s_{sur}, \ell), \ell \in \mathfrak{R}\} = \Sigma_{sur}$ , and

$$f_L(s_{sur}, 0\ell) = s_{sur}, \quad f_L(f_L(s_{sur}, \ell_1), \ell_2) = f_L(s_{sur}, \ell_1 + \ell_2). \quad (5.37)$$

It implies that for any couple of different states,  $[s_{sur}, s'_{sur}]$ , the equation  $f_L(s_{sur}, \ell) = s'_{sur}$  defines a unique length  $|\ell|$  (a positive generalized length).

## 5.6 Mach space

The space with a particle resembles the situation of celestial observers. For them, the Sun plays the role of such a material particle. Any couple of states forming a change defines a length – the distance of two positions of the Sun (measured in an angle measure multiplied by an arbitrarily chosen length unit). Changes of the Sun position, however, realize some basic clock for celestial observers<sup>8</sup>. Let us notice also that any clock may be understood as connecting a length (e.g. the high of sand column in a sand-glass, the length of a trajectory passed by a pendulum, etc.) with an arbitrary change<sup>9</sup>.

---

<sup>8</sup>Notice that other clocks were consequently derived from the movement of the Sun.

<sup>9</sup>The digital clock that shows only numbers uses the possibility of transforming any length to real numbers. We, however, have no experience with a “digital time”.

This noticing guides us to a simple model of the space  $\mathcal{E}_M$ . It is a surrounding having property of space and including special subsystems called **clocks**. Their important feature is the existence of a variable  $\hat{c}$ , called the **clock variable**, assigning a length to each clock at any state of  $\mathcal{E}_M$ . Values of clock variables, however, cannot be determined at individual states. Instead, the following property is supposed: in any couple of two different states of the system  $\mathcal{E}_M$ , say  $s_1^M, s_2^M$ , the clock variable of an arbitrary couple of clocks  $C_1, C_2$  fulfills

$$c_{s_1^M}(C_1) - c_{s_2^M}(C_1) = c_{s_1^M}(C_2) - c_{s_2^M}(C_2) \quad (5.38)$$

(a clock synchronization). It allows defining the mapping  $f_L$  as follows:  $f_L([s_1^M, s_2^M]) = c_{s_1^M}(C) - c_{s_2^M}(C)$  where  $C$  stands for an arbitrary clock.

Moreover, we assume that the identity manipulation with  $\mathcal{E}_M$ , denoted as  $id^M$ , may cause only the identity manipulation with the studied system. This is a natural idea connected with any clock: if there is no change on the clock, there is no change of anything. Formally, it is formulated as the demand that for every state  $s^M$  of the surrounding  $\mathcal{E}_M$  holds the condition

$$\mathfrak{F}([s^M, id^M]) = id, \quad (5.39)$$

where  $\mathfrak{F}$  is the mapping defined by (2.6) and  $id$  the identity manipulation with the studied system. This surrounding will be called the **clock space**. If we suppose that clocks are the same physical systems, negligibly small and located arbitrarily “densely”, the clock space is homogeneous and isotropic.

The clock space looks like the Newton space with an absolute flow of time. But there is no ‘flowing time’ in  $\mathcal{E}_M$ . We have only a set of states of a physical system,  $\mathcal{E}_M$ , whose important components are various subsystems called clocks. The

clocks belong inherently to this space. They cannot be looked apart from it. It resembles the situation in generally relativistic physics that shows that there cannot be some physical subsystems (e.g. clocks) that are independent of spacetime. But if clocks are indispensable physical systems belonging to the surrounding then we cannot exclude some physical influences of these subsystems on the studied system.

This is a very interesting point. In physics (excepting the theory of general relativity), clocks – as physical systems – play a negligible role, and if they are taken into consideration as a matter of fact, they are always supposed to have no influence on the system. Namely, the idea of an absolutely flowing time is so suggestive that there is even no reason to think about a necessary presence of some material systems called clocks. This attitude, however, is not supported by our experience. When going back to our main clock - the Sun in the sky - we cannot neglect its physical influence. But we can involve the Sun and the whole solar system into the studied system and look for a more distant clock with a negligible physical influence.

Such clocks are distant stars. We observe their regular motion in the sky and any change “here” can be related to their trajectory on the firmament. Moreover, a direct physical influence of them may be neglected. Nevertheless, the fact that distant stars move in the sky is in a way *related* to some physical phenomena observable on the Earth. Namely, there are gyroscopic effects that may be measured in rotating systems. The Earth rotates. But what does ‘rotation’ mean? In Newtonian physics, the rotation is defined towards (absolute) space<sup>10</sup>. But, as noticed by famous physicists and philosopher Ernest Mach,

---

<sup>10</sup>In fact, Newton proved the existence of absolute space by showing an absolute character of the rotation (in his famous bucket experiment).

the rotation of the Earth is not defined towards an absolute space but towards distant stars. The fact that we observe moving stars in the sky is a good explanation of some gyroscopic effects.

Though there is probably no causal influence of stars on the Earth, some noticeable relation between ‘clocks’ (stars) and some effects on the Earth cannot be omitted. We will not open some interesting questions connected with the problem of the relativity of rotation motion. The only result accepted by us is just the noticing that clocks may be interconnected with the studied system by some physical relations (whatever their physical meaning)<sup>11</sup>. Are then manipulations caused by such a clock space those that form a complete decomposition of the Newton manipulation  $\epsilon$ ?

The example with distant stars shows that the answer may be negative because certain gyroscopic effects cannot be observed in the Newton space. They may be understood as a “new” (non-Newtonian) *influence of the clock space* into a system which is in this space<sup>12</sup>. In this case, the space  $\mathcal{E}_M$  is characterized only by the existence of the mapping  $f_L$  and no by its “yield” of manipulations  $\epsilon_\ell$  giving a complete decomposition of the Newton manipulation  $\epsilon$ . The space  $\mathcal{E}_M$  understood in such a generalized sense will be called the **Mach space**.

The Mach space may or may not cause manipulations  $\epsilon_\ell$ . In any case it causes manipulations parameterized by all lengths,

---

<sup>11</sup>In the case of distant stars, this connection is principal because distant objects may work as good clocks only if there is some relative rotation. Namely, it is the only way in which they can work permanently.

<sup>12</sup>Notice also that ‘manipulation’ does not mean only a direct physical interaction. The manipulation caused by the Newton space means to “let system alone”, the manipulation caused by the clock space with distant stars means to “let rotating system alone”.

say  $(\dots, \epsilon_\ell^M \dots)$ . If  $\epsilon_\ell^M \neq \epsilon_\ell$  then their parallel composition,

$$\epsilon^M \equiv (\dots, \epsilon_\ell^M \dots),$$

defines a generalization of the Newton manipulation  $\epsilon$ . It is not clear, if it is worth speculating about a generalized of (one state) Newton space causing the manipulation  $\epsilon^M$  instead of  $\epsilon$ . Nevertheless, such a space would be more natural for celestial observers who can never study physical phenomena in the Universe without distant stars.

## 5.7 Adiabatic space

The Mach space is more general than the clock space. Especially the validity of the condition (5.39) may not be assumed. As a result, we get much more general situations. An important example of Mach space is used in thermodynamics to define the concept of adiabatic accessibility. This kind of Mach space presents an extremely difference of manipulations  $\epsilon_\ell^M$  from  $\epsilon_\ell$ . Namely, the important feature of this space is that it includes a weight in a homogeneous gravitational field which is a source or sink of energy of the system.

This space may be simply realized anywhere on the Earth's surface if the studied system is negligible small with respect to the Earth's radius<sup>13</sup>. Various states of this space are characterized by different positions of the weight. That means that any couple of states defines a plus or minus length (or zero length if there is no shift of the weight). Notice that not states alone,

---

<sup>13</sup>Though the space  $\mathcal{E}$  always produces parts forming the Euclidean space, it does not mean that it goes "through the whole Universe". For example, local inertial systems of the theory of general relativity (four-dimensional Minkowski spaces) are "relatively small".

but only their couples define the length (we have only one body in homogeneous field and its position cannot be absolutely determined). We will call this kind of Mach space the **adiabatic space**.

The adiabatic space is supposed to have the conservative property with respect to the global variable called the total energy. Usually, this variable is defined by the use of an adiabatic space:

*The **total energy** is a real global variable  $E$  so that the adiabatic space has conservative property with respect to this variable, and the relation,*

$$E(s_1^A) - E(s_2^A) = mg \ell([s_1^A, s_2^A]), \quad (5.40)$$

*holds for its values, where  $\ell([s_1^A, s_2^A])$  is the generalized length assigned to the couple of states  $[s_1^A, s_2^A]$  of the adiabatic space,  $m$  is a mass, and  $g$  is the intensity of the gravitational field.*

The equality (2.11) then allows us to define the value of total energy (up to a constant) for any system whose surrounding is the adiabatic space. Namely, (2.11) implies that

$$E(s_2) - E(s_1) = E(s_1^A) - E(s_2^A). \quad (5.41)$$

The difference of manipulations  $\epsilon_\ell$ , arising when the system is “let alone”, and  $\epsilon_\ell^M$ , caused by the adiabatic space, is without any doubts: the energy of the system is transmitted from the space to the system and vice versa. The difference between the adiabatic and clock space is also evident: there may be many nontrivial changes of the system while the weight does not change its position (i.e. the adiabatic space remains in the same state).

## Chapter 6

# Accessibility

*We need to free ourselves from the prejudices associated with the habit of thinking of the world as “inhabiting space” and “evolving in time”.*

**C. Rovelli**<sup>1</sup>

We have defined a special class of surroundings called generally the *space*. Each space defines special parts – spatial points – that have crucial meaning in physics. Namely, spatial points represent fixed *places* that allow us to define precise position (or distribution) of bodies as well as their motion<sup>2</sup>. The simplest

---

<sup>1</sup>[35], p. 7.

<sup>2</sup>Spatial points in our conception are always defined with respect to a concrete coordinate system from a class  $\mathcal{C}_d$ . A group of equivalent coordinate systems defines a certain “relativity” in the description. For instance, the galilean relativity is obtained by a group of coordinate systems in “rectilinear and uniform” motion. But this relativity does not annul an “absolute” character of spatial points. Realize that transformation laws between various coordinate systems are based on the *identity* of points (5.1) defined in different coordinate systems (in special relativity, it is the identity of points in Minkowski spacetime, called ‘events’).

kind of such a surrounding is the Newton space which models very faithfully the idea of an absolute, inert and immobile empty space. Another example is the Mach space that may be understood as space with clocks as well as space with a weight in gravitational field or something else<sup>3</sup>.

However, the assumption that the system is *in space* (whatever its concrete kind) is rather a metaphysical construction than a result of empirical observations. We observe things and various substances around us but no ‘space’ as independent entity *in which* are bodies. Remind that the dominant conception of space in western culture (till Newton) was its understanding as a *relation* between bodies (such as “contiguity” or “adjacency”, see [35], p. 38).

It is important to realize that the Newton’s predecessor, R. Descartes, laid foundations of physics (as a program of mathematical description of nature<sup>4</sup> [6]) *without* using the concept of space as an independent entity. Instead, Descartes identifies space with the *extension* of bodies, i.e. bodies are not *in* space but constitute space by *themselves* [20]. This extension is called “**res extensa**”. A perfect correspondence of the Descartes’ position with the central result of the general relativistic physics is startling. To demonstrate it, let us cite the leading physicists in the field of quantum gravity<sup>5</sup>, C. Rovelli, ([35], p. 7):

---

<sup>3</sup>The Mach space has also this spatial character (it defines spatial points) regardless some simple additional objects belonging to it (clocks, a weight in a gravitation field, etc.).

<sup>4</sup>Physics of R. Descartes represents a transitory stage in the mathematization of nature, connecting Galileo with Newton [17].

<sup>5</sup>The main point of the programm of quantum gravity is the formulation of physics without a given spatiotemporal background. Though the quantum gravity is intended to be a fundamental physical theory, its possible “applications” concern only extreme short scale structure of spacetime. At

*“In Newtonian and special relativistic physics, if we take away the dynamical entities – particles and fields – what remains is space and time. In general relativistic physics, if we take away the dynamical entities, nothing remains. The space and time of Newton and Minkowski are reinterpreted as configurations of one of the fields, the gravitational field. This implies that physical entities – particles and fields – are not immersed in space and moving in time. They do not live on spacetime. They live, so to say, on one another.”*

It is appealing to formulate physics without a given spatial (and temporal) background. And not only as the method of description at extreme spatiotemporal scales (where it hardly can be formulated otherwise) but as a central point of *any* physical description. In this chapter, we outline some possible features of such a (still hypothetical) approach in which the existence of space as a special entity (surrounding) is not supposed. We concentrate especially on the question how “inner dynamics” of the system could be studied. Namely, when studying the inner dynamics of a system, we have to leave the system alone, which means that no influence is supposed. But what means that the system is “let alone” if there is no possibility of “removing” surrounding matter? In Newtonian physics, it is space itself what remains after removing all bodies and material substances and physics is then formulated in this space only. However, when abandoning an independent space, the concept of “inner dynamics” becomes problematic.

We have to realize that an irremovable presence of surrounding substances (“res extensa”) means a permanent *interaction* of the system with something in its surrounding. How does

---

scales of everyday physics, the quantum gravity should coincide perfectly with the Newtonian physics.

the Newtonian conception avoid this problem? Namely, why may we suppose that space does not interact with studied systems, too? Here is an extraordinary subtle step in the Newton's thinking. Namely, Newton does not speculate *if* and *in which way* space affects matter. He demands only that space *is not affected* by matter by postulating that space *remains always the same*.

This is a very essential point of our contemplation on physics without a spatial background. Namely, the previous noticing allows removing various baggage related to the concept of space. Such as “voidness”, “impalpability”, “ethereality”, etc. Maybe these attributes are helpful in accepting (but not explaining!) the exclusiveness of rectilinear and uniform motion in space (e.g. by imaging motion with vanishing friction), but they are not essential in the main role that space plays in Newtonian physics. ‘Space’ could be represented by an arbitrary milieu that remains always the same. The problem, however, is that no such immobile, inert matter exists.

That is why we introduce the concept of *accessibility* that supersedes the demand of existence of an entity that remains always the same. The idea comes from the conception of adiabatic accessibility defined in thermodynamics. Its essence is as follows. Let us have a system and its surrounding which is an arbitrary (material) system (we have no ‘space’ around). A state  $s_2$  of the system is accessible from another state  $s_1$  if, roughly speaking, the change  $s_1 \rightarrow s_2$  is caused by a surrounding that *remains the same* during this change. That is, the surrounding behaves (when causing the change  $s_1 \rightarrow s_2$ ) *like* space. In the next section, we explain the importance of accessibility for definition of motion when no space yielding fixed places (spatial points) is supposed.

## 6.1 Relative motion

The Cartesian physics was not successful. One reason of its failure, among others, was the problem with definition of *motion*<sup>6</sup>. We say that a body moves from a place A to another place B. But if there is no background space, places can be determined only by a reference to bodies. Motion thus has to be constituted only by a shifting of parts of *res extensa* *relatively* to each other<sup>7</sup>. That implies that an evidence of motion needs a *distinction* of parts. But how can parts be distinguished without specifying their positions, i.e. places where they are located? It is difficult to avoid some circularity when trying to define motion by accepting such an utter relativity<sup>8</sup>. It became clear to Newton that the concept of motion could be defined by concept of place which is exempt from relativity. The idea of absolute space solves it perfectly because places are absolutely defined by space. However, it seems to be an exaggeratedly radical solution.

Notice the way in which we understand motion in everyday experience. We have no problems to explain that a car moves from a place A to B. The places A and B are *not* defined with respect to some absolute space. They are given parts of a

---

<sup>6</sup>The problem with the definition of motion in Descartes' conception is well described by I. Leclerk [18], pp. 192-194. In the next passage, we take briefly over his argumentation.

<sup>7</sup>It is both the Aristotle's and Descartes' conception of motion.

<sup>8</sup>In general relativistic physics, such a relativity is also accepted by possibility of formulating the problem in general coordinate systems (e.g. by defining motion with respect to a cluster of bees). But it is saved from problems with an "absolute" Cartesian relativity by the existence of very small referential systems (infinitesimal Minkowski spacetimes [7]) that behave as "islets with classical spacetime" and form a clue of the whole conception.

countryside that plays a role of a referential background toward which the car is moving. But there may be various changes in the countryside alone during the motion of car. How can such a movable countryside play the role of a referential system defining the movement of car? The way of avoiding an endless relativity is that accepted in everyday life: we do not try to define the motion in *any* situation<sup>9</sup>. If the majority of referential parts (including those representing places A and B) changes or is even destroyed then the motion from A to B is not defined at all.

The crucial idea is that we cannot define a ‘general motion’ but only concrete motions of concrete bodies with respect to other concrete bodies. It is enough to define the motion of our car if the countryside (or its chosen part) behaves *like* absolute space during the motion of car. Imagine the countryside and the car as two physical systems so that the first forms the surrounding of the latter. The “motion of car” may be understood as a change of one state of the car,  $s_1$ , into another state,  $s_2$ <sup>10</sup>. If the countryside remains during the motion of car in the *same state*, say  $s_0$ , then it behaves like a fixed background and may serve as a perfect referential system allowing to define (and quantify) the motion of car.

Let us try to formulate it only in the framework of abstract thermodynamics. In this approach, we cannot say that “during

---

<sup>9</sup>We do not live in *res extensa* where everything permanently moves with respect to anything else. Main feature of our “*res extensa*” is rather a permanent “supply” of a sufficient amount of things in rest. Thus the definition of motion is always possible and we do not realize that it is defined, in fact, only in special situations (during an earthquake, say, nobody takes care to define a movement of a car).

<sup>10</sup>We recall that the ‘state’ is a primitive concept of our approach. We need not define the states of car with a relation to the countryside!

a change of states of the system, its surrounding is in the same state”. Namely, there are no temporal relations of the system and its surrounding, there are no “simultaneous states”, etc. The surrounding is defined as a source of all manipulations with the system (see the definition at Section 2.4). In the spirit of this definition, the situation when motion may be defined, is described as follows: *The change  $s_1 \rightarrow s_2$  is caused by an identity manipulation with the surrounding.* Namely, the identity manipulation includes only such changes that the initial and final state of the surrounding are the same. Recall that the only manipulation with the Newton space is just the identity – the surrounding thus behaves *like* the Newton space.

Remark: The formulation may seem rather strange because of the word ‘caused’. In which way does the countryside cause the movement of the car? Here we again meet the uncommon feature of our approach. Namely, the causality is understood as a relation of the system and its surrounding. The fact that the car changes its state from  $s_1$  to  $s_2$  means two things. First, that such a change is possible due to an inner dynamics of the system, second, that it is “allowed” by the surrounding. In our approach, however, the both aspects are included in the concept of ‘manipulation’. The only source of manipulations is the surrounding (in the sense that it may also mean to “do nothing”).

Notice that the question if the surrounding (the countryside) is *permanently* at the same state is nonsensical in our approach. The manipulation is formed only by couples of states; no ‘intermediate’ states (of any change) cannot be defined<sup>11</sup>. The crucial question is only whether the change  $s_1 \rightarrow s_2$  can

---

<sup>11</sup>We may define, however, the process whose *any* couple of states  $[s(\tau_1), s(\tau_2)]$ , where  $\tau_1 < \tau_2$ , is caused by the identity manipulation.

actually be caused by an identity manipulation with the surrounding. If yes, we say that the state  $s_2$  is *accessible* from the state  $s_1$ .

The relation of accessibility then gives the restriction under which certain concepts (like the concept of motion<sup>12</sup>) may be defined. The definition of accessibility, however, depends on a concrete surrounding whose states occur in the definition (their couples define the identity manipulation). It is very far to the conception of ‘inner dynamics’ introduced in physics because that is supposed to be a property of the system alone, not depending on a concrete surrounding. The accessibility, however, seems to depend strongly on a concrete surrounding. The modification of the definition, that enables us to eliminate this dependence, will be done in the next section.

## 6.2 Definition of ‘accessibility’

The way of *eliminating* a concrete surrounding in the definition of accessibility is a remarkable step introduced in thermodynamics<sup>13</sup>. Namely, we may eliminate *any concrete* surrounding by a weakening of the definition:

*We say that the state  $s_2$  is **accessible** from the state  $s_1$  if there exists such a surrounding of the system, that the change  $s_1 \rightarrow s_2$  is caused by an identity manipulation with this surrounding.*

*We denote  $s_1 \overset{N}{\prec} s_2$ .*

---

<sup>12</sup>Notice an analogy between ‘motion’ in Cartesian physics and ‘work’ in thermodynamics.

<sup>13</sup>The work of E. H. Lieb and J. Yngvason [21] presents clearly this step when defining the adiabatic accessibility and, especially, when explaining its role in the axiomatic scheme of thermodynamics.

The letter “N” emphasizes the fact that the accessibility has the same basic property as the manipulation with the Newton space. Namely, that it is ‘identity’. Notice that the identity manipulation with the Newton space includes only one change,  $s_0 \rightarrow s_0$ , because the Newton space has only one state,  $s_0$ . In the definition of accessibility, the identity manipulation may be formed by an *arbitrary* couple of same states.

The definition of accessibility requires to study such manipulations with a system that are caused by the following “mechanism”. We find a physical system (e.g. a device) that plays the role of a surrounding of the system. It means that it is a source of all manipulations with the system<sup>14</sup>. The surrounding may be quite arbitrary, the system may interact with an arbitrary amount of objects but, when it reaches the state  $s_2$ , the surrounding has to be at the same state as when the system “was” at the state  $s_1$ .

Remark: The logic of the definition is following. We “try” various surroundings till find (or do not find) the “right” one. Any exchange of surrounding, however, may change the system (see the end of Section 2.4). Nevertheless, if a “testing” surrounding “destroyed” either the state  $s_1$  or  $s_2$  of the system then it could not be the “right” one.

The situation, when the system is surrounded by the “right” surrounding realizing the accessibility, corresponds to an arrangement when there is no “record” in the surrounding after the change from  $s_1$  to  $s_2$ . Surprisingly, it seems to have a connection with ideas of quantum mechanics: if the state  $s_1$  is pure then the state  $s_2$  has to be pure, too, because there is no entanglement with the surrounding.

---

<sup>14</sup>It also *defines* these manipulations – see the end of Section 2.4.

The situation, that is very close to that if the system *is* in the Newton space, arises when the surrounding has even the same conservative properties as the Newton space:

*We say that the state  $s_2$  is **conservatively accessible** from the state  $s_1$  if there exists such surrounding of the system that the change  $s_1 \rightarrow s_2$  is caused by an identity manipulation with the surrounding and this surrounding has conservative properties with respect to the same global variables as the Newton space.*

The surrounding that defines the accessibility can influence the system very heavily without changing its own state. An example is a massive elastic wall (a part of the surrounding) that changes momenta of light particles (being parts of the system). States differing by directions of momenta of particles thus may be accessible.

The conservative accessibility is a finer concept. Namely, the change  $s_1 \rightarrow s_2$  has to be realized without a change of global momenta and other global conservative quantities. There are no “massive walls” in the surrounding that causes “accessible” changes of states of the system. It is not so easy to manage such a change of the state “during which” the surrounding returns to the same state (e.g. an inverse of momenta of only one particle is not possible). It is an essential aspect of difference between the accessibility and conservative accessibility. It consists in a possibility or impossibility of occurrence of a *scale gap* between the system and the surrounding. The massive wall has to remain at the same state after reversing the momentum of a particle (*its* change of momenta is negligible). In other words, the description of the wall is done at much higher scales than that of the system<sup>15</sup>.

---

<sup>15</sup>There are some connotations to presence of macroscopic measuring

The “dynamics” of changes  $s_1 \rightarrow s_2$  that realize the conservative accessibility is described by conservative laws. A very nontrivial question is whether there is a general formulation of some additional conditions that guarantees the return of the surrounding to the initial state. Namely, the fulfillment of the conservative laws only, does not appear to be sufficient. In thermomechanics, for instance, there is a set of balance laws that describe conservative properties of the studied medium (Section 4.2). However, the balance laws are not sufficient to describe dynamics of the studied system. A set of the so-called **constitutive laws** has to be added. The search for such laws is the crucial problem of thermomechanics.

### 6.3 Time arrow

The Newton space is a very nice surrounding because its structure is as simple as be only possible. On the other hand, manipulations caused by this surrounding are very roughly defined. It behaves like a very bad measuring device that may give an infinite number of possible results. That is why we looked for a better “device” doing the same as the Newton space but manipulating with the system much more gentle, and giving exactly “targeted” results (it is the meaning of the complete decomposition of a rough manipulation). The Mach space is an example of such a surrounding. Its structure is more complicated though (as the structure of any sophisticated device), but it “produces” an infinite number of manipulations that cause uniquely defined results.

---

device in quantum mechanics. We, however, avoid detail discussions concerning quantum mechanics in the entire of the book.

The accessibility means the existence of surrounding that causes the change by the same “mechanisms” as the Newton space. Because the set of manipulation with Mach space is much more larger than a simple identity, we may radically generalize the definition of accessibility by replacing the Newton space by the Mach one. The definition is as follows:

*We say that the state  $s_2$  is **weakly accessible** from the state  $s_1$  if there exists such surrounding of the system that the change  $s_1 \rightarrow s_2$  is caused by such a manipulation  $\alpha$  with the surrounding, that any couple of states  $[s_{sur}^1, s_{sur}^2] \in \alpha$  defines uniquely a length  $\ell \in \mathfrak{R}$  via a mapping  $f_L$  (5.36). We denote  $s_1 \prec s_2$ .*

The term “weak” is justified by the fact that the accessibility implies the weak accessibility. Namely, the accessibility means that the manipulation  $\alpha$  is the identity, which defines the generalized length  $\ell = 0$ . The opposite implication, however, is not valid. The essence of the weak accessibility is the same: we look for an *arbitrary* surrounding which causes the change  $s_1 \rightarrow s_2$  so that its state may change into another that, however, may differ from the original one only by *something* that uniquely defines a certain plus, minus or zero length.

If that ‘something’ defining a generalized length is a clock (located within an arbitrary surrounding) then we call the relation  $\prec$  the weak accessibility in **clock realization**. It represents the **time arrow** in physics. Namely, the weak accessibility then simply means that the state  $s_1$  may be transformed into  $s_2$  by an arbitrary manipulation with a surrounding so that the surrounding is, at the end, at the same state as at the beginning, excepting the datum on the clock. The fact that there are states  $s_1, s_2$  so that  $s_2$  is not accessible from  $s_1$  expresses an impossibility of traveling to past. To say it in other words, everything (not belonging to the system) may be “tided

up” to the original conditions, only the ‘passed time’ cannot be “passed back”.

It offers an interesting possibility of defining the ‘time’ as a global variable:

*The global variable  $t$  is called the **time** if its values copy the relation  $\prec$  in the meaning that*

$$t(s_1) \leq t(s_2) \Leftrightarrow s_1 \prec s_2, \quad (6.1)$$

*where  $s_1 \prec s_2$  means the weak accessibility in clock realization.*

This definition indicates an interesting connection of motion and time. While the first may be defined by the use of accessibility motivated by the Newton space, the later is defined by a special realization of weak accessibility motivated by the Mach space.

## 6.4 Adiabatic accessibility

If that ‘something’, defining the length in the weak accessibility, is a weight in a gravitational field, we obtain the definition of the so-called **adiabatic accessibility** known from thermodynamics. We precede it here in the thermodynamic formulation taken from [21] (p. 21):

*A state  $s'$  is **adiabatically accessible** from state  $s$ , in symbols  $s \prec s'$ , if it is possible to change the state from  $s$  to  $s'$  by means of an interaction with some device (which may consist of mechanical and electrical parts as well as auxiliary thermodynamic system) and a weight, in such a way that the device returns to its initial state at the end of the process whereas the weight may have changed its position in a gravitational field.*

In our language, the ‘device’ represents an arbitrary surrounding. Any manipulation realizing the change in the definition is called the **adiabatic manipulation**. The crucial importance

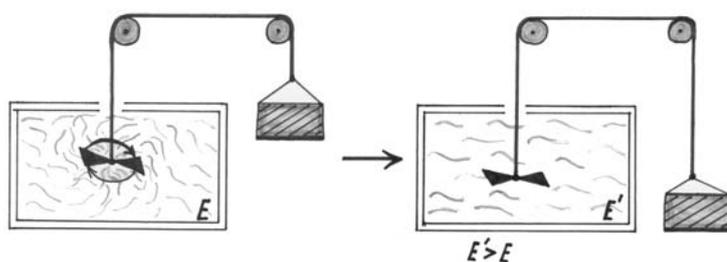


Figure 6.1: A typical process illustrating the adiabatic accessibility. A falling weight rotates a propeller which mixes a fluid, being isolated by a wall from a heat exchange between the surrounding. At the end of the process, the fluid is warmer than before and the weight is down. No other change has happened in the surrounding.

of the adiabatic accessibility is given by the fact that it defines the concept of entropy as a global variable:

*The global variable  $S$  is called the **entropy** if its values copy the relation  $\prec$  in the meaning that*

$$S(s_1) \leq S(s_2) \Leftrightarrow s_1 \prec s_2, \quad (6.2)$$

where  $s_1 \prec s_2$  means the adiabatic accessibility.

There are many questions concerning the entropy. For example, if it can be defined uniquely, if it can be understood as a (local) variable, and so on. The only situation when we know the positive answer on these questions, is the case when the system is a composition of simple systems defined in Chapter 4. Then the

entropy becomes the entropy of equilibrium thermodynamics. We outline the way of constructing the thermodynamic entropy in the next section.

Beforehand, we mention the following noticing. As everybody knows, entropy has a connection to time arrow. The presented approach reveals a possible explanation of this connection. When we compare the definitions of time and entropy, given by relations (6.1) and (6.2), we see that they are nearly identical. If we omit the specification of a concrete realization of weak accessibility, we obtain the *same* definitions. A provocative idea intrudes herewith. What if entropy and time are only different ways of *interpretation* of the same thing, the weak accessibility?

## 6.5 Thermodynamics and entropy

In the work of E. H. Lieb and J. Yngvason [21], it is solved the problem of a unique construction of the local variable  $\hat{S}$  that defines the global variable  $S$  (by the use of the rule (3.8)) that fulfills the condition (6.2). It is done only for special kind of systems – compositions of simple systems (see Section 4.1).

Let us have  $N$  simple systems,  $\mathfrak{S}_1, \dots, \mathfrak{S}_N$ . We will study the system

$$\mathfrak{S} = (\mathfrak{S}_1, \dots, \mathfrak{S}_N)$$

formed by their composition. The simple system  $\mathfrak{S}_i$  is formed by  $n_i$  parts and thus its composition is formed by a finite set of parts. Hence the structure of the system  $\mathfrak{S}$  is finitely atomic. Following (3.16), we define the state space of  $\mathfrak{S}$  as

$$\Gamma = \Gamma_1 \times \dots \times \Gamma_N, \quad (6.3)$$

where  $\Gamma_i$  is the state space of the  $i$ -th simple system. Elements of  $\Gamma$  will be denoted by  $X, Y, Z$ , elements of its scaled copy by  $rX, rY, rZ$ . The set of variables,  $Q$ , is supposed to be complete. It implies that elements of state space may be identified with states of the system,

$$X \sim s. \quad (6.4)$$

The state of a composition will be denoted by  $(X, Y)$  if  $X$  and  $Y$  are states of individual systems. The states  $X_1, X_2$  are said to be **comparable** if either  $X_1 \prec X_2$  or  $X_2 \prec X_1$  (the symbol  $\prec$  means only the adiabatic accessibility in this section). If the states  $X_1, X_2$  are in the relation  $X_1 \prec X_2$  and  $X_2 \prec X_1$  we write  $X_1 \overset{A}{\sim} X_2$ .

The composition of  $N$  simple systems is formed by  $N_c \equiv n_1 + \dots + n_N$  atomic parts,  $p_i$ , (see the definition of simple systems). We suppose that all these parts are mutually disjoint, i.e., the total part is defined as the composition of all parts,

$$p_0 = (p_1, \dots, p_{N_c}).$$

It implies that if we find the entropy as a local variable that has the averaged property, then it is defined on the part  $p_0$ , too, and  $S_s(p_0)$  defines the entropy as a global variable  $S(s)$ .

Being motivated by classical equilibrium thermodynamics, the wanted entropy is supposed not only to have the averaged property, but to be an *additive* variable<sup>16</sup> (i.e., it is defined for any composition as a sum of entropies of individual components). Since the entropy is supposed to be defined on simple

<sup>16</sup>The additivity is a very strong condition considering the relation (6.2). Namely, (as mentioned in [21], p. 20) the relation  $S(X) + S(Y) \leq S(X') + S(Y')$  says us *exactly* which adiabatic processes are allowed in the compound system, independent of any detailed knowledge of the manner in which the two systems interact.

systems, it has to be either intensive or extensive. However, it cannot be intensive because of its additivity. When summing up previous considerations, we look for an additive variable whose values at total part define a global variable  $S$  fulfilling (6.2). Otherwise speaking, we have to prove the following theorem, called the entropy principle (ibid., p. 19).

**Entropy principle:** *There is a real-valued variable  $\hat{S}$  defined at all states on all parts of the composition of simple systems called **entropy** such that*

- Monotonicity: *When  $X$  and  $Y$  are comparable states then*

$$X \prec Y \text{ if and only if } S(X) \leq S(Y). \quad (6.5)$$

- Additivity and extensivity: *If  $X$  and  $Y$  are states of some (possibly different) systems and if  $(X, Y)$  denotes the corresponding state in the composition of the two systems, then the entropy is additive for these states, i.e.,*

$$S((X, Y)) = S(X) + S(Y). \quad (6.6)$$

*$S$  is also extensive, i.e., for each  $r > 0$  and each state  $X$  and its scaled copy  $rX$ ,*

$$S(rX) = rS(X). \quad (6.7)$$

The crucial problem, however, is how to formalize the concept of adiabatic accessibility because its essence is that “there *exists* a surrounding (device) that causes such a manipulation with the system that realizes the change from  $s_1$  to  $s_2$ ”. But it is an absolutely hopeless task to check through *all* possible manipulations with the system. The approach of Lieb and Yngvason avoid this problem by an ingenious idea. Namely, they

do not study all possible surroundings realizing the adiabatic accessibility, but find *logical connections* that are supposed to be implied by the relation of adiabatic accessibility.

The logical connection means, for instance, that the existence of adiabatic manipulations realizing changes  $X \rightarrow Y$  and  $Y \rightarrow Z$  implies the existence of an adiabatic manipulation realizing  $X \rightarrow Z$ . These logical relations are then formulated in the form of axioms. The main idea then consists in finding a sufficient group of such axioms that implies the validity of the entropy principle. For illustration, let us present a first group of these axioms (ibid., p. 21):

- (A1)  $X \overset{A}{\sim} X$ .
- (A2)  $X \prec Y$  and  $Y \prec Z$  implies  $X \prec Z$ .
- (A3)  $X \prec X'$  and  $Y \prec Y'$  implies  $(X, Y) \prec (X', Y')$ .
- (A4) If  $X \prec Y$ , then  $rX \prec rY$  for all  $r \in \mathbb{R}_+$ .
- (A5) For  $0 < z < 1$ ,  $X \overset{A}{\sim} (zX, (1-z)X)$ .
- (A6) If  $(X, \varepsilon Z_0) \prec (Y, \varepsilon Z_1)$  holds for a sequence of  $\varepsilon$ 's tending to zero and some states  $Z_0, Z_1$  then  $X \prec Y$ .
- (A7) Assume  $X$  and  $Y$  are states in the same state space,  $\Gamma$ , that has a convex structure, i.e. if  $X, Y \in \Gamma$  then for any  $r \in (0, 1)$  the value  $rX + (1-r)Y$  belongs to  $\Gamma$  too. If  $r \in (0, 1)$  then  $(rX, (1-r)Y) \prec rX + (1-r)Y$ .

These axioms are justified by the definition of adiabatic accessibility and common physical notions. The validity of (A3), for instance, seems to be demonstrated by putting the individual systems side by side and “construct” an adiabatic manipulation based on those realizing  $X \rightarrow X'$  and  $Y \rightarrow Y'$  separately.

(A4), on the other hand, means that the possibility of realization of an adiabatic manipulation remains in force if the system is made smaller or larger (the question of a possibility of constructing an adequately smaller or larger device, as a “re-scaled surrounding”, may be relevant here).

The axiom (A7) evokes an idea of “mixing” of two separated systems at two different states. It says that we can always obtain, by an adiabatic manipulation, the state that is its “mixture”. It looks obviously if the both systems are gases (or liquids). Namely, we may simply put them side by side, remove a separating barrier and let them to mix mutually<sup>17</sup>. Why may we rely on a spontaneous mixing? It is a typical physical process that is justified by the **second law of thermodynamics**. At the first glance, it may seem strange that there is no axiom in the Lieb’s and Yngvason’s scheme called the “second law”. But we must not forget that the axioms *do not* “translate” thermodynamics into an axiomatic system, but *model* the relation of adiabatic accessibility as well as possible. The second law is thus an auxiliary statement justifying more or less the chosen axioms.

In thermodynamics, the second law expresses an impossibility of realizing certain processes. In the axiomatic scheme, this impossibility is formulated only for simple systems in the form of the following axiom (ibid., p. 42):

**Irreversibility.** *For each  $X \in \Gamma_{simple}$  there is a point  $Y \in \Gamma_{simple}$  such that  $X \prec Y$  but  $Y \prec X$  is not valid.*

Remark: When using the identification of states with points of state space, (6.4), we see that for states  $s_1 \sim X$ ,  $s_2 \sim Y$ , where  $X, Y$  occur in the previous axiom, the relation (6.2) implies that  $S(s_1) < S(s_2)$ .

---

<sup>17</sup>The meaning of (A7) in the case of solids is not so clear (ibid., p. 33).

If we compare it with (6.1), we see that an analogy of the irreversibility axiom for the weak accessibility in clock realization would be a statement that there is a “nonzero” flow of time.

If states are adiabatically accessible or not is deeply connected with the physical quantity called the **temperature**. We cannot, for instance, cool down a body only (without a change of work variables) by an adiabatic manipulation<sup>18</sup>. The concept of temperature is deeply related to a possibility of putting two systems into thermal contact. It is guaranteed by the following axiom (ibid., p. 55), where  $E$  denotes the value of energy of a simple system and  $\mathbf{w}$  values of its work coordinates:

**Thermal contact.** *For any two simple systems with state spaces  $\Gamma_1$  and  $\Gamma_2$  there is another simple system, the **thermal join** of  $\Gamma_1$  and  $\Gamma_2$ , with state space*

$$\Delta_{12} = \{(E, \mathbf{w}_1, \mathbf{w}_2) : E = E_1 + E_2, (E_1, \mathbf{w}_1) \in \Gamma_1, (E_2, \mathbf{w}_2) \in \Gamma_2\}. \text{ Moreover, } ((E_1, \mathbf{w}_1), (E_2, \mathbf{w}_2)) \prec (E_1 + E_2, \mathbf{w}_1, \mathbf{w}_2).$$

Notice that the last relation expresses another assumption concerning the adiabatic accessibility<sup>19</sup>. When splitting systems forming the thermal join, we obtain two systems in mutual thermal equilibrium, which enables us to define the temperature. It corresponds to the following axiom (ibid., p. 55):

**Thermal splitting.** *For any point  $(E, \mathbf{w}_1, \mathbf{w}_2) \in \Delta_{12}$  there is at least one pair of states,  $(E_1, \mathbf{w}_1) \in \Gamma_1$ ,  $(E_2, \mathbf{w}_2) \in \Gamma_2$ , with  $E = E_1 + E_2$ , such that  $(E, \mathbf{w}_1, \mathbf{w}_2) \overset{A}{\sim} ((E_1, \mathbf{w}_1), (E_2, \mathbf{w}_2))$ .*

<sup>18</sup>On the other hand, there is no problem to warm up a body by an adiabatic manipulation, see Fig. 6.1.

<sup>19</sup>In this case, the thermal join is realized without a change of the total energy,  $E = E_1 + E_2$ . That means (see (5.41)) that the weight does not move. Hence it fulfills the stronger condition of the accessibility,  $((E_1, \mathbf{w}_1), (E_2, \mathbf{w}_2)) \overset{N}{\prec} (E_1 + E_2, \mathbf{w}_1, \mathbf{w}_2)$ .

Writing the last relation as  $(E_1, \mathbf{w}_1) \overset{T}{\sim} (E_2, \mathbf{w}_2)$ , we may formulate the third axiom concerning the temperature (ibid., p. 56):

**Zerth law.** *If  $X \overset{T}{\sim} Y$  and if  $Y \overset{T}{\sim} Z$  then  $X \overset{T}{\sim} Z$ .*

When adding five more axioms (ibid., pp. 88,89), we obtain the group of axioms that implies the validity of the entropy principle, i.e., guarantees the existence of the entropy function fulfilling (6.2) for a special class of systems.

The logic of the construction open many interesting questions. For example, if there is a possibility of reformulating it for other classes of physical systems. Especially, the construction of the entropy function for thermomechanical systems, in *this* spirit, would be very inviting. The question, however, is whether it is possible. Namely, is the role of equilibrium (which defines simple systems and their compositions) crucial? It opens speculations concerning a relation of the entropy and time. The fact that time, as a global variable possessing properties of the generalized temperature, is also defined in “equilibrium” (see Section 2.6), may be relevant in this context. It would be also interesting to formulate an analogy of previous axioms for time ordering, i.e., to find axioms that model the weak accessibility in clock realization.



# Bibliography

- [1] G. I. Barenblatt, *Similarity, Self-Similarity, and Intermediate Asymptotics*, Consultants Bureau, New York (1979)
- [2] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, Wiley, New York (1985)
- [3] S. Carnot, *Réflexions sur la Puissance Motrice de Feu*, Bachelier, Paris (1824)
- [4] A. Carpinteri (Ed.), *Size-Scale Effects in the Failure Mechanisms of Materials and Structures*, E & FN SPON, London (1996)
- [5] W. A. Day, *A Commentary on Thermodynamics*, Springer, New York (1988)
- [6] R. Descartes, *Principles of philosophy* (1644), transl. by V.R. Miller and R.P. Miller, Reidel, Dordrech (1983)
- [7] A. Einstein, Grundlage der allgemeinen Relativitätstheorie, *Annalen der Physik* 49 (1916) 769-822
- [8] *Encyclopaedia of Mathematics*, Reidel, Kluwer Academic Publishers, Dordecht, Boston, Lancaster, Tokyo (1988)

- [9] R. Giles, *Mathematical foundations of thermodynamics*, Pergamon, Oxford (1964)
- [10] M. E. Gurtin, W.O. Williams, An Axiomatic Foundation for Continuum Thermodynamics, *Arch. Rat. Mech. Anal.*, 26 (1967) 83-117
- [11] I. M. Havel, Scale Dimensions in Nature *Int. J. General Systems* 24(3) (1996) 295
- [12] M. Holeček, Scale axis, fractal and some thermal phenomena, *J. Thermal Anal. Cal.* 60 (2000) 1093-1100
- [13] S. C. Hunter, *Mechanics of Continuous Media*, Ellis Horwood, New York (1976)
- [14] S. Kauffman, *Investigations*, Pergamon Press (2000)
- [15] D. Kondepudi, I. Prigogine, *Modern Thermodynamics*, Wiley, New York (1999)
- [16] J. Kvasnica, *Termodynamika*, SNTL, Praha (1965) (in Czech)
- [17] L. Kvasz, The Mathematisation of Nature and Cartesian Physics, *Philosophia Naturalis*, 40 (2003) 157-182
- [18] I. Leclerk, *The Nature of Physical Existence*, Washington D.C. (1972)
- [19] G. P. Lepage, What is renormalization?, *preprint* Cornell University, CLNS 89/970 (1989)
- [20] H. Lichtenegger, B. Mashhoon, Mach's Principle, <http://arxiv.org/abs/physics/0407078> (2004)

- [21] E. H. Lieb, J. Yngvason The physics and mathematics of the second law of thermodynamics, *Phys. Rep.* 310 (1999) 1 - 96
- [22] G. A. Maugin, *Material Inhomogeneities in Elasticity*, Chapman & Hall, London (1993)
- [23] G. A. Maugin, *The Thermomechanics of Nonlinear Irreversible Behaviour*, World Scientific, Singapore, New Jersey, London, Hong Kong (1999)
- [24] E. Mendoza (ed.) *Reflections on the Motive Force of Fire by Sadi Carnot and other Papers on the Second Law of Thermodynamics by E. Clapeyron and R. Clausius*, Glouster, MA: Peter Smith (1977)
- [25] W. Muschik, Empirical foundation and axiomatic treatment of Non-Equilibrium Temperature, *Arch. Rat. Mech. Anal.* 66 (1977) 379-401
- [26] W. Muschik, C. Papenfuss, H. Ehrentraut: A sketch of continuum thermodynamics, *J. Non-Newtonian Fluid Mech.* 96 (2001) 255-290
- [27] M. N. L. Narasimhan, *Principles of Continuum Mechanics*, John Willey & Sons, New York (1992)
- [28] Z. Neubauer, *God's Power and Computer Reason* (manuscript), Praha (2006)
- [29] I. Newton, *The Principia* (1687), transl. by I.B.. Cohen and A. Whitman, University of California Press, Berkley (1999)

- [30] L. Nottale, *Fractal Space-Time and Microphysics, Towards a Theory of Scale Relativity*, World Scientific (1993)
- [31] D. R. Owen, *A First Course in the Mathematical Foundations of Thermodynamics*, Springer, Heidelberg (1984)
- [32] M. Pitteri, On the axiomatic foundations of temperature, in: C. Truesdell, *Rational Thermodynamics*, Springer-Verlag, New York (1984)
- [33] J. Polchinski, Renormalization and effective lagrangians, *Nucl. Phys. B* 231 (1984) 269-295
- [34] A. Robinson, *Non-Standard Analysis*, Studies in Logic and the Foundations of Mathematics, North-Holland, 1966
- [35] C. Rovelli, *Quantum Gravity*, <http://www.cpt.univ-mrs.fr/~rovelli/book.pdf> (2003)
- [36] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, Inc. New York (1974)
- [37] A. Sochor, The alternative set theory and its approach to Cantor's set theory, in: H.J. Skala, S. Termini, and E. Trillas (eds.), *Aspects of Vagueness*, (1984) 161-203
- [38] V. Šedivý: Space, in *Man Bang*, Univers - Omnivers, o.p.s., Plzeň, (1998), (in Czech), translation to English: M. Kotišová (2005).
- [39] J. Šesták, *Science of Heat and Thermophysical Studies, A Generalized approach to Thermal Analysis*, Elsevier, The Netherlands (2005)

- [40] M. Šilhavý, *The Mechanics and Thermodynamics of Continuous Media*, Springer, Heidelberg (1997)
- [41] C. A. Truesdell, S. Bharata, *The Concepts and Logic of Classical Thermodynamics as a Theory of Heat Engines*, Springer, Heidelberg (1977)
- [42] C. A. Truesdell, *The Tragicomical History of Thermodynamics*, Springer, New York (1980) 1822-1854
- [43] C. A. Truesdell, *Rational Thermodynamics*, Springer, New York (1984)
- [44] A. Tzouvaras, Modeling vagueness by nonstandardness, *Fuzzy sets and systems*, 94 (1998) 385-396
- [45] P. Vopěnka, *Mathematics in the Alternative Set Theory*, Teubner Texte, Leipzig, 1979
- [46] K. G. Wilson, Renormalization group and critical phenomena *Phys. Rev. B* 4 No. 9 (1971) 3174-3205
- [47] K. G. Wilson, J. Kogut, The renormalization group and the  $\epsilon$  expansion, *Phys. Rep. C* 12 (1974) 75-200

# Index

- $\sigma$ -algebra, 107
- $d$ -dimensional space, 118
- absolute continuity, 71
- abstract thermodynamics, 10
- accessibility, 148
- additive variable, 70
- adiabatic accessibility, 153
- adiabatic manipulation, 154
- adiabatic space, 140
- atomic part, 59
- atomic structure, 59
- averaged property, 70
- averaging, 81
- b-variable, 96
- balance property, 96
- c-variable, 96
- causality, 44
- caused manipulation, 44
- clock, 136
- clock realization, 152
- clock space, 136
- clock variable, 136
- comparable states, 156
- complement part, 65
- complete decomposition, 38
- complete set of local variables, 74
- composed part, 59
- composition, 27, 36
- composition of parts, 57
- configuration, 103
- configuration dependent variable, 105
- conservative accessibility, 150
- conservative property, 48
- consistence of states and manipulations, 40
- constitutive laws, 151
- continuity rule, 76
- continuum, 101
- continuum mechanics, 105
- continuum model, 102
- continuum thermomechanics, 87, 93

- convergence of states, 119
- coordinate system, 118
- definition of spatial points, 117
- density of variable, 90
- discrete system, 95
- disjunct parts, 57
- energy, 92
- entropy, 5, 154, 157
- entropy principle, 157
- equilibrium, 51
- equivalent coordinate systems, 124
- extensive variable, 90
- field, 118
- finitely atomic structure, 59
- flux, 97
- full state space, 81
- generalized lengths, 134
- global variable, 46
- heat, 72, 96
- homogeneous surrounding, 126
- identity manipulation, 35
- inner state of the process, 42
- intensive variable, 90
- isotropic surrounding, 127
- l-space, 83
- local variable, 68
- Mach space, 138
- manipulation, 7, 11, 35
- mass, 98
- measurable mapping, 108
- measure, 108
- model, 102
- model of a part, 102
- model of space, 125
- motion, 106
- mutually isolated, 34
- Newton manipulation, 128
- Newton space, 128
- non-atomic structure, 59
- oriented boundary, 94, 98
- parallel composition, 37
- parameterization of the process, 41
- perfect manipulation, 37
- physical system, 22, 25
- process, 41
- process b-variable, 96
- process variable, 72
- production, 97
- referential configuration, 105
- representative set of subsystems, 26, 33
- res extensa, 142

- rotation, 126, 127
- scale, 55, 62, 68, 81
- scale complete model, 110
- scale symmetric space, 127
- scaled copy, 90
- Schottky system, 95
- second law, 4, 159
- set of global variables, 46
- set of manipulations, 35
- set of parts, 57
- set of variables, 68
- simple system, 92
- simple unit, 90
- simpler surroundings, 45
- space, 116
- spatial field, 118
- spatial point, 117
- stable state, 120
- state, 32
- state set, 32
- state space, 54
- state space of a part, 73
- state space of the system, 73
- stress, 96
- structure of mass points, 60
- structure of the system, 58
- subpart, 58
- subsystem, 26
- supply, 97
- surrounding, 43
- symmetry operation of space, 124
- system, 25
- tangent of the process, 42
- temperature, 91, 160
- thermal join, 160
- thermomechanical system, 99, 101
- time, 129, 153
- time arrow, 152
- total energy, 140
- total part, 58
- trajectory, 106
- translation, 126
- trivial manipulation, 121
- variable, 68
- weak accessibility, 152
- weakly atomic structure, 59
- work variables, 92

Miroslav Holeček  
ABSTRACT THERMODYNAMICS

Book work and composition in  $\text{\LaTeX}$  Michal V. Hanzelín.  
Cover design Zuzana Lebedová.  
Illustrations Lenka Kaplanová.  
O.P.S., Máchova 1492, 288 02 Nymburk (Czech republic);  
o-p-s@post.cz  
Printed in Typos - Digital Print, Plzeň (Czech Republic)  
Plzeň  $\equiv$  Nymburk 2006.  
First printing, 2006  
174 pages  
ISBN 80-239-6075-X